Title
VALIDATION OF THE KATTAN PREOPERATIVE NOMOGRAM FOR PROSTATE CANCER RECURRENCE USING A COMMUNITY BASED COHORT: RESULTS FROM CANCER OF THE PROSTATE STRATEGIC UROLOGICAL RESEARCH ENDEAVOR (CAPSURE)

Authors
GREENE, KIRSTEN L
MENG, MAXWELL V
ELKIN, ERIC P
et al.

VALIDATION OF THE KATTAN PREOPERATIVE NOMOGRAM FOR PROSTATE CANCER RECURRENCE USING A COMMUNITY BASED COHORT: RESULTS FROM CANCER OF THE PROSTATE STRATEGIC UROLOGICAL RESEARCH ENDEAVOR (CaPSURE)

KIRSTEN L. GREENE, MAXWELL V. MENG,* ERIC P. ELKIN, MATTHEW R. COOPERBERG,
DAVID J. PASTA, MICHAEL W. KATTAN, KATRINE WALLACE AND PETER R. CARROLL

From the Department of Urology, Program in Urologic Oncology, Urologic Outcomes Research Group University of California-San Francisco/Mt. Zion Cancer Center, University of California-San Francisco (KLG, MVM, EPE, MRC, DJP, PRC), San Francisco, California, Departments of Epidemiology and Biostatistics, and Urology, Memorial Sloan-Kettering Cancer Center (MWK), New York, New York, and TAP Pharmaceutical Products, Inc. (KW), Lake Forest, Illinois

ABSTRACT

Purpose: The Kattan preoperative nomogram combines preoperative prostate specific antigen (PSA), biopsy Gleason grade and clinical stage to estimate disease recurrence after radical prostatectomy. Several studies using patient data from academic centers have validated the nomogram. We assessed the performance of the Kattan nomogram using the Cancer of the Prostate Strategic Urological Research Endeavor database, a national, largely community based observational disease registry.

Materials and Methods: From the Cancer of the Prostate Strategic Urological Research Endeavor database we identified 1,701 men with clinically localized prostate cancer undergoing radical prostatectomy with sufficient pretreatment information and PSA followup after surgery. Disease recurrence was defined as 2 consecutive PSA values 0.2 ng/ml or greater, or a second cancer treatment more than 6 months after prostatectomy. A concordance index was used to evaluate the performance of the nomogram compared to observed 5-year recurrence-free survival (Kaplan-Meier). Kattan nomogram scores were calculated for each patient and stratified into 6 groups for analysis.

Results: In our cohort of 1,701 men 413 (24%) had evidence of disease recurrence. Median followup in these patients was 2.3 years. Kattan nomogram scores were 17% to 99% (mean 79%). The overall concordance index was 0.68. Varying the definition of recurrent disease and excluding patients with imputed data did not substantially alter nomogram performance (concordance index 0.65 to 0.70). The Kattan nomogram tended to overestimate 5-year freedom from recurrence in patients with scores of 65% and higher.

Conclusions: We noted the reasonable performance of the Kattan nomogram for predicting cancer outcomes after radical prostatectomy using a community based population. Although concordance is lower than in previous validation studies and the nomogram overestimates recurrence-free survival in patients at lower risk, the model is fairly robust and it provides important information when counseling patients regarding treatment options in the community setting. Further refinements in pretreatment estimation of disease-free survival and ultimately overall survival are needed.

KEY WORDS: prostate, prostatic neoplasms, recurrence, prostatectomy, prostate-specific antigen

Considerable debate continues to surround the appropriate selection of treatment in men diagnosed with prostate cancer.[1,2] Radical prostatectomy (RP) remains a viable option in men with clinically localized disease who desire curative intervention. However, the selection of this modality must be based on an accurate estimation of cancer outcomes after surgery. In recent years efforts have been made to estimate the likelihood of disease-free survival after RP.[3–5] One of the most commonly used tools is the nomogram developed by Kattan et al based on a cohort of men treated by a single surgeon at a United States tertiary referral center.[6] This nomogram, incorporating clinical T stage (1992), Gleason grade on diagnostic biopsy and pretreatment serum prostate specific antigen (PSA), attempts to predict biochemical recurrence 5 years after RP in those with clinically localized prostate cancer.

The Kattan nomogram was internally validated in patients at the same institution treated by 5 other surgeons.[6] Subsequent external validation at academic centers within the United States as well as 2 other continents confirmed the performance of the model for predicting disease-free survival.[7–9] Nevertheless, the usefulness and applicability in the community setting remain unexplored. Thus, we applied the Kattan nomogram in a cohort of men followed in the Cancer of the Prostate Strategic Urological Research Endeavor (CaPSURE) database, drawing from more than 30 community based urology practices, to assess the validity of the model in this patient population undergoing RP.

MATERIALS AND METHODS

CaPSURE is an observational database of men with biopsy proven prostate adenocarcinoma. Patients are recruited consecutively by participating urologists who report clinical data

and followup information on diagnostic tests and treatments. Data have been collected at 40 urology practices across the United States, namely 34 community based sites, 3 academic medical centers and 3 Veterans Administration medical centers. Data collection began in 1995 when newly diagnosed and previously diagnosed but still followed patients were asked to participate. Since 1999, only newly diagnosed patients have been recruited into the study cohort. Patients are treated according to the usual practice of their physician and they are followed until death or study withdrawal. Additional details of the project methodology have been reported previously.[10]

We examined data on 2,111 patients within CaPSURE enrolled as of May 2002 who were diagnosed with clinically localized disease (cT3a or less and N0M0) undergoing RP between January 1, 1989 and December 31, 2000 who had received neither neoadjuvant nor adjuvant treatment. Of these men we included 1,701 (81%) in the final study cohort with no more than 1 missing baseline data item (PSA, Gleason grade or T stage) and at least 2 followup PSA values after RP.

Disease recurrence was defined by biochemical failure (2 consecutive PSA values 0.2 ng/ml or greater after RP) or a second cancer treatment. Because serial PSA data were not complete for all patients, treatment failure was assumed if the patient received a second prostate cancer treatment more than 6 months after RP. Previous analyses of CaPSURE data have demonstrated that second treatment is a surrogate marker of disease recurrence within this population.[11,12] Kattan et al considered second treatment any time after RP, even within 6 months, as disease recurrence when developing the nomogram.[6] We repeated our analysis using the Kattan et al definition of biochemical recurrence after RP (PSA 0.4 ng/ml or greater, followed by an increased PSA value (ie 0.4 ng/ml $\leq$ PSA 1 $\leq$ PSA 2) to compare results using the 2 definitions.

Pretreatment disease variables used in the Kattan nomogram are serum PSA, primary and secondary Gleason grade on prostate biopsy, and clinical T stage. Some patients in CaPSURE were missing data on 1 or more of these variables, while others only had Gleason sum reported without individual primary and secondary scores. In addition, the Kattan nomogram is based on the 1992 American Joint Committee on Cancer-TNM staging system, while clinical stage in CaPSURE was based on the 1997 staging system. Except for stage T2a in the 1997 system the 1997 system is readily converted to the 1992 system. Similar to the original analysis of Kattan et al,[6] we imputed missing or indeterminate values for 1 clinical variable using the other 2 available variables to estimate the missing variable without regard to study outcome, ie disease recurrence. If more than 1 variable was missing, the patient was excluded from the study population. Of the 1,701 men included in analysis 5%, 9% and 15% had PSA, clinical T stage (1992) and Gleason grade imputed, respectively. Statistical analyses were performed on the entire cohort with imputed data as well as on the 1,205 (71%) patients not requiring imputed data to confirm the results.

For each patient the probability of 5-year freedom from recurrence survival was calculated using the Kattan nomogram. Nomogram probability scores were categorized into 6 groups, namely less than 40%, 40% to 64%, 65% to 74%, 75% to 84%, 85% to 94% and 95% or greater, for convenience in comparing predicted and observed survival. We compared observed percent recurrence-free rate at 5 years based on Kaplan-Meier actuarial analysis to the average predicted percent recurrence-free rate at 5 years calculated from the Kattan nomogram. In addition, the continuous nomogram score was used in Cox proportional hazards analysis to determine the concordance index, that is the proportion of randomly paired patients for whom the patient with the higher probability of recurrence (lower nomogram score) also

had earlier disease recurrence.[7,13–16] The concordance index ranges from 0 to 1 with 1 indicating perfect concordance, 0.5 indicating no association (ie no better than flipping a coin) and 0 indicating perfect discordance.

All analyses were performed using SAS for Windows, version 8.2 (SAS Institute, Cary, North Carolina) except the concordance index, which was calculated using S-Plus for Windows, version 6.0 (Insightful Corp., Seattle, Washington).

RESULTS

Of the 1,701 patients undergoing RP 413 (24%) experienced disease recurrence, including 248 by PSA criteria and 165 by second treatment. Median followup in those with recurrence was 2.3 years (maximum 9.5), while the median followup in patients without recurrence was 2.9 years (maximum 12.1). Table 1 lists the clinical characteristics of the entire cohort included in the validation study.

For all patients observed Kaplan-Meier freedom from recurrence survival at 5 years was 67% (95% CI 63 to 71). For the 1,205 patients with complete pretreatment information available, ie no imputed data, actuarial recurrence-free survival at 5 years was 66% (95% CI 61 to 71), suggesting that disease outcomes in patients with and without imputed data were similar (Cox proportional hazards model p = 0.13).

In the CaPSURE population Kattan nomogram scores (ie the predicted probability of remaining recurrence-free 5 years after surgery) were 17% to 99% (mean $\pm$ SD 79% $\pm$ 16%, median 83%). Table 2 shows observed recurrence-free survival in the entire cohort grouped into the 6 nomogram predicted survival categories. As expected, the lower the nomogram estimate, the less likely that patients would remain free of recurrence 5 years after RP. For example, patients with a Kattan nomogram value of less than 40% had an observed recurrence-free survival of 28% (95% CI 15 to 43), while those with a nomogram value of 95% or greater had an observed recurrence-free survival of 90% (95% CI 77 to 96). Figure 1 shows Kaplan-Meier curves for each Kattan prognostic group.

Figure 2 shows plots of predicted freedom from recurrence survival from the Kattan nomogram against the observed recurrence-free survival observed in each prognostic group. Theoretically one would expect that about 90% of men in the 85% to 94% nomogram category would be recurrence-free at 5 years, while about 80% in the 75% to 84% group would be

TABLE 1. *Patient characteristics*

No. pts	1,701
No. ng/ml PSA (%):	
4 or Less	230 (14)
4.1–10	1,066 (63)
10.1–20	286 (17)
Greater than 20	119 (7)
Median PSA (ng/ml)	6.6
No. primary/secondary Gleason grade:	
1–2/1–2	265 (16)
1–2/3	225 (13)
3/1–2	97 (6)
3/3	754 (44)
1–3/4–5	243 (14)
4–5/1–5	117 (7)
No. clinical T stage (%):	
T1a/b	44 (3)
T1c	481 (28)
T2a	476 (28)
T2b	224 (13)
T2c	450 (26)
T3a	26 (2)
Mean age at surgery $\pm$ SD	63 $\pm$ 7
No. race/ethnicity (%):	
White	1,451 (85)
Black	184 (11)
Hispanic	33 (2)
Other/unknown	33 (2)

TABLE 2. *Observed recurrence-free survival at 5 years by nomogram predicted survival estimates*

% Kattan Nomogram Group	No. Pts	Kaplan-Meier	
		% Observed 5-Yr Recurrence Free	95% CI
Less than 40	61	28	15–43
40–64	224	44	35–53
65–74	204	57	45–68
75–84	470	73	65–80
85–94	664	78	71–83
95% or Greater	78	90	77–96

recurrence-free at 5 years and so on. The nomogram tended to overestimate disease-free survival for prognostic categories of 65% to 74% and higher. In these patients at lower risk observed survival was lower than that predicted by the Kattan nomogram.

Using the entire CaPSURE cohort the overall concordance index was 0.68. We confirmed our results by repeating analysis in the CaPSURE subgroup without imputed data, the entire CaPSURE cohort using the Kattan definition of biochemical recurrence and 3-year intervals based on surgery date. Concordance index values were 0.65 to 0.72 for these sensitivity analyses (table 3).

DISCUSSION

The high prevalence of prostate cancer in men, combined with significant disease mortality and treatment related morbidity, has made the selection of therapy a subject of considerable debate. Accurate assessment of biological tumor behavior is important to determine not only who requires treatment, but also which therapy is likely to improve cancer specific outcomes. Traditional factors used in this process are clinical tumor stage, serum PSA and Gleason grade on diagnostic biopsy. However, each variable alone is generally a poor indicator of outcome after treatment. The widely used Partin nomogram addresses this issue to some degree, allowing the prediction of pathological stage at RP based on the combination of stage, PSA and Gleason score.[17] Although pathological stage can predict eventual treatment failure, its use as an end point is imperfect. Moreover, the question most frequently asked by patients is, "What is my chance of being cured of cancer 5 years after surgery?" The Kattan nomogram, also combining clinical prognostic factors, allows the preoperative prediction of disease recurrence after RP using serum PSA as the end point, ie evidence of recurrence.[6]

The Kattan nomogram was modeled on a single surgeon experience with 983 patients at a tertiary referral center. It was then validated in a population from the same institution with multiple treating surgeons as well as in international patient cohorts. Overall the nomogram was accurate across these various populations with concordance indexes between 0.67 and 0.83.[6–8] Nevertheless, the patients in these studies were drawn from academic centers and treated by oncological urologists. Our study confirms the validity of the Kattan nomogram in a large, community based cohort of patients with prostate cancer undergoing RP by surgeons with varying experience.

The concordance index for the CaPSURE cohort was 0.68. Varying the exact population and definition of biochemical failure did not substantially alter the accuracy of the model (concordance indexes 0.65 to 0.72). Thus, although accuracy may be slightly lower than in previous validation studies, it remains reasonably applicable to a wide spectrum of patients and surgeons. This confirmation is critical, given that many urologists are currently using this tool to counsel patients prior to definitive therapy. Community urologists can now be more confident in citing predictions of 5-year disease recurrence based on the Kattan nomogram.

Several inherent differences were present in our community based patient population compared to the prior validation cohorts. In the report of Kattan et al a single expert pathologist graded the biopsies, minimizing variability in 1 predictor variable.[6] In addition, PSA was determined using 1 assay (Hybritech Tandem-R, Perkin Elmer, Foster City, California) and digital rectal examination was performed by only a single practitioner. Although the validation studies of Graefen et al demonstrate variability in many of these areas,[7] none compare with the diversity within the CaPSURE database. Due to the multiple sites of practice dispersed within the United States multiple pathologists graded prostate biopsies, serum PSA was determined using different assays at various laboratories and clinical stage was assessed by individual urologists. The majority of operations were performed at community hospitals rather than at the tertiary referral centers analyzed in previous reports. It is likely that many of these factors account for the lower concordance index of the model when applied to our population. Nevertheless, given the heterogeneity of the data, the nomogram is fairly robust with an accuracy comparable to that in prior validation studies in more homogeneous groups. Moreover, the diversity and reflection of real world practice in the CaPSURE database is precisely the reason why we applied the Kattan nomogram and assessed its accuracy in our population.

Our patients differed from the original validation set with respect to clinical T stage as well as Gleason grade distribution. Some ambiguity was introduced by the cT2a and cT2b groups because we had to apply the 1997 T stage to the 1992 definition, as used by Kattan et al.[6] We also analyzed CaPSURE data using 2 definitions of PSA recurrence (2 consecutive PSA values 0.2 ng/ml and 0.4 ng/ml ≤ PSA 1 ≤ PSA 2). This did not change the concordance index, suggesting that other variables in diagnosis, staging or treatment accounted for the decreased accuracy of the nomogram in our patients.

We applied appropriate selection criteria to identify a cohort with sufficient data to assess accuracy of the nomogram. For example, a minimum of 2 postoperative PSA values was required to determine disease status. In addition, we excluded patients receiving neoadjuvant and adjuvant treatment to assess the impact of prostatectomy alone on disease recurrence. Thus, we believe that the study cohort is an unbiased group of patients in the community setting with adequate information for application of the Kattan nomogram.

Another potential limitation of our analysis is the use of imputed data in some patients. We performed this in a fashion similar to that of the original report of Kattan et al.[6] Calculated recurrence-free survival at 5 years was almost identical in the entire cohort, including those with imputed values, and the group with complete data, suggesting that the imputation process and inclusion of men with imputed data did not introduce bias.

The nomogram tended to underestimate disease recurrence in patients with recurrence-free survival scores of 65% to 74% and higher. This is significant because 71% of patients within CaPSURE are in this range. Plausible explanations for this performance in lower risk cases are pathological under grading or clinical under staging, earlier initiation of secondary therapy, differences in the incidence of positive surgical margins, intensity of followup, surgical technique and/or patient selection for surgery.

Given the tremendous variability in patients, our validation study reinforces the robustness of the Kattan nomogram as a tool to predict disease recurrence after RP. Ultimately the individual practitioner must determine whether a tool providing a CI approaching 0.70 is sufficiently useful for clinical application and patient counseling. Further study is needed to enhance the performance of the nomogram beyond 5 years and help determine what is an acceptable probability of recurrence for a given patient, considering treatment morbidity and life expectancy. Models to estimate disease specific
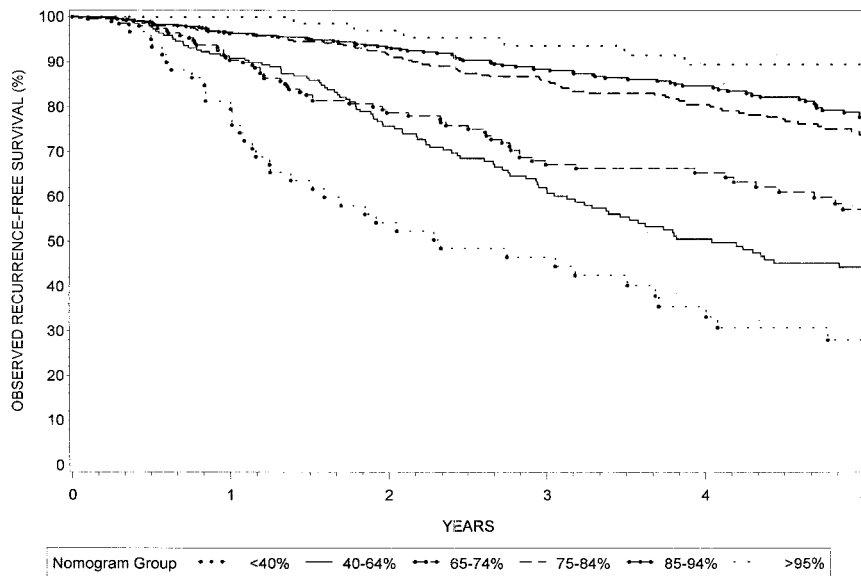
FIG. 1. Observed recurrence-free survival after RP in CaPSURE database stratified by Kattan nomogram category
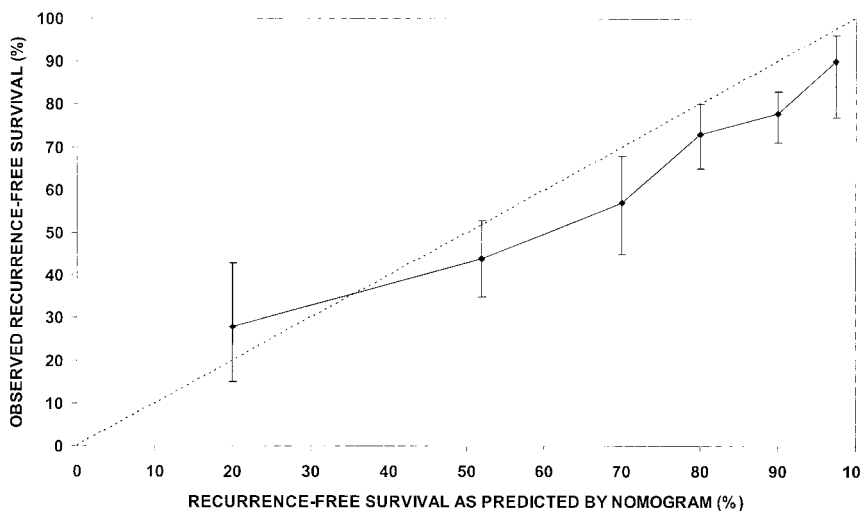

FIG. 2. Observed 5-year recurrence-free survival in CaPSURE cohort compared with predicted score calculated from Kattan nomogram. Dotted line indicates theoretical perfect agreement between prediction and observation. Bars indicate 95% CI.

TABLE 3. *Concordance index by study population and definition of biochemical recurrence*

Study Population	Recurrence Definition	No. Pts	Concordance Index	95% CI
All pts	2 Consecutive PSAs 0.2 ng/ml or greater	1,701	0.68	0.62–0.74
Nonimputed	2 Consecutive PSAs 0.2 ng/ml or greater	1,205	0.68	0.61–0.74
All 1989–1991	2 Consecutive PSAs 0.2 ng/ml or greater	164	0.65	0.53–0.77
All 1992–1994	2 Consecutive PSAs 0.2 ng/ml or greater	467	0.66	0.58–0.75
All 1995–1997	2 Consecutive PSAs 0.2 ng/ml or greater	671	0.70	0.61–0.80
All 1998–2000	2 Consecutive PSAs 0.2 ng/ml or greater	399	0.67	0.47–0.87
All pts	0.4 Ng/ml ≤ PSA 1 ≤ PSA 2	1,701	0.71	0.64–0.77
Nonimputed	0.4 Ng/ml ≤ PSA 1 ≤ PSA 2	1,205	0.72	0.65–0.79

and overall survival are necessary to help patients and physicians make informed decisions regarding prostate cancer treatment.

CONCLUSIONS

We report the performance of the Kattan nomogram for predicting disease recurrence after RP in a community based cohort of patients. The concordance index with observed disease recurrence was between 0.65 and 0.70 among various study populations and definitions of biochemical recurrence.

Although these values are slightly lower than in previous validation studies and the model tends to overestimate freedom from failure in lower risk categories (Kattan scores 65% or greater), the nomogram is robust and appropriate for use by most urologists in various practice settings.

REFERENCES

1. Holmberg, L., Bill-Axelson, A., Helgesen, F., Salo, J. O., Folmerz, P., Haggman, M. et al: A randomized trial comparing radical

prostatectomy with watchful waiting in early prostate cancer. N Engl J Med, **347:** 781, 2002

2. Jani, A. B. and Hellman, S.: Early prostate cancer: clinical decision-making. Lancet, **361:** 1045, 2003

3. D'Amico, A. V., Whittington, R., Malkowicz, S. B., Fondurulia, J., Chen, M. H., Kaplan, I. et al: Pretreatment nomogram for prostate-specific antigen recurrence after radical prostatectomy or external-beam radiation therapy for clinically localized prostate cancer. J Clin Oncol, **17:** 168, 1999

4. Graefen, M., Noldus, J., Pichlmeier, U., Haese, A., Hammerer, P., Fernandez, S. et al: Early prostate-specific antigen relapse after radical retropubic prostatectomy: prediction on the basis of preoperative and postoperative tumor characteristics. Eur Urol, **36:** 21, 1999

5. D'Amico, A. V., Whittington, R., Malkowicz, S. B., Schultz, D., Fondurulia, J., Chen, M. H. et al: Clinical utility of the percentage of positive prostate biopsies in defining biochemical outcome after radical prostatectomy for patients with clinically localized prostate cancer. J Clin Oncol, **18:** 1164, 2000

6. Kattan, M. W., Eastham, J. A., Stapleton, A. M., Wheeler, T. M. and Scardino, P. T.: A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. J Natl Cancer Inst, **90:** 766, 1998

7. Graefen, M., Karakiewicz, P. I., Cagiannos, I., Quinn, D. I., Henshall, S. M., Gyrgiel, J. J. et al: International validation of a preoperative nomogram for prostate cancer recurrence after radical prostatectomy. J Clin Oncol, **20:** 3206, 2002

8. Graefen, M., Karakiewicz, P. I., Cagiannos, I., Hammerer, P. G., Haese, A., Palisaar, J. et al: A validation of two preoperative nomograms predicting recurrence following radical prostatectomy in a cohort of European men. Urol Oncol, **7:** 141, 2002

9. Bianco, F. J., Jr., Kattan, M. W., Scardino, P. T., Powell, I. J., Pontes, J. E. and Wood, D. P., Jr.: Radical prostatectomy nomograms in black American men: accuracy and applicability. J Urol, **170:** 73, 2003

10. Lubeck, D. P., Litwin, M. S., Henning, J. M., Stier, D. M., Mazonson, P., Fisk, R. et al: The CaPSURE database: a methodology for clinical practice and research in prostate cancer. CaPSURE Research Panel. Cancer of the Prostate Strategic Urologic Research Endeavor. Urology, **48:** 773, 1996

11. Grossfeld, G. D., Chang, J. J., Broering, J. M., Miller, D. P., Yu, J., Flanders, S. C. et al: Does the completeness of prostate sampling predict outcome for patients undergoing radical prostatectomy?: data from the CAPSURE database. Urology, **56:** 430, 2000

12. Grossfeld, G. D., Chang, J. J., Broering, J. M., Miller, D. P., Yu, J., Flanders, S. C. et al: Impact of positive surgical margins on prostate cancer recurrence and the use of secondary cancer treatment: data from the CaPSURE database. J Urol, **163:** 1171, 2000

13. Kattan, M. W., Stapleton, A. M., Wheeler, T. M. and Scardino, P. T.: Evaluation of a nomogram used to predict the pathologic stage of clinically localized prostate carcinoma. Cancer, **79:** 528, 1997

14. Harrell, F. E., Jr., Lee, K. L. and Mark, D. B.: Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med, **15:** 361, 1996

15. Harrell, F. E., Jr., Califf, R. M., Pryor, D. B., Lee, K. L. and Rosati, R. A.: Evaluating the yield of medical tests. JAMA, **247:** 2543, 1982

16. Begg, C. B., Cramer, L. D., Venkatraman, E. S. and Rosai, J.: Comparing tumour staging and grading systems: a case study and a review of the issues, using thymoma as a model. Stat Med, **19:** 1997, 2000

17. Partin, A. W., Mangold, L. A., Lamm, D. M., Walsh, P. C., Epstein, J. I. and Pearson, J. D.: Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium. Urology, **58:** 843, 2001