UCSF

UC San Francisco Previously Published Works

Title

ABILITY OF 2 PRETREATMENT RISK ASSESSMENT METHODS TO PREDICT PROSTATE CANCER RECURRENCE AFTER RADICAL PROSTATECTOMY

Permalink

https://escholarship.org/uc/item/83p6d576

Journal

Investigative Urology, 173(4)

ISSN

0021-0005

Authors

MITCHELL, JOSEPH A COOPERBERG, MATTHEW R ELKIN, ERIC P et al.

Publication Date

2005-04-01

DOI

10.1097/01.ju.0000155535.25971.de

Peer reviewed

ABILITY OF 2 PRETREATMENT RISK ASSESSMENT METHODS TO PREDICT PROSTATE CANCER RECURRENCE AFTER RADICAL PROSTATECTOMY: DATA FROM Capsure

JOSEPH A. MITCHELL, MATTHEW R. COOPERBERG, ERIC P. ELKIN, DEBORAH P. LUBECK, SHILPA S. MEHTA, CHRISTOPHER J. KANE AND PETER R. CARROLL

From the Department of Urology, Urology Outcomes Research Group, University of California, San Francisco, San Francisco, California, and TAP Pharmaceutical Products, Inc. (SSM), Lake Forest, Illinois

ABSTRACT

Purpose: Two methods widely used to predict the risk of treatment failure after radical prostatectomy for localized prostate cancer are the 3 level D'Amico risk classification and the Kattan nomogram. Although they have been previously validated, to our knowledge they have not been compared in a community based cohort. We tested the 2 instruments in the Cancer of the Prostate Strategic Urologic Research Endeavor (CaPSURE) database, a national registry of patients with prostate cancer, to assess their accuracy in a community based cohort.

Materials and Methods: Men were invited to join CaPSURE from 33 American urology practices, of which 30 were community based. A total of 1,701 men with localized prostate cancer (T1-3a) were treated with radical prostatectomy between 1989 and 2000. Patients who received neoadjuvant or adjuvant therapy were excluded. Recurrence was defined as 2 or more consecutive prostate specific antigen measurements of 0.2 ng/ml or greater, or a second treatment greater than 6 months after surgery. Freedom from progression (FFP) was based on life table estimates and Kaplan-Meier curves. Risk groups were compared using a Cox proportional hazards model and ANOVA.

Results: Based on the D'Amico classification 671 cases (39%) were classified as low risk, 446 (26%) were intermediate risk and 584 (34%) were high risk. Five-year FFP was 78%, 63% and 60% in the low, intermediate and high risk groups (HR 1.00, 1.87 and 2.32 respectively, p <0.0001). Mean 5-year FFP predicted by the Kattan nomogram in the same risk groups was 91%, 74% and 69%, respectively. Outcomes in the low risk group were tightly grouped about the mean but there was considerable dispersion of outcomes in the intermediate (30% to 98% FFP) and high (17% to 98%) risk groups.

Conclusions: Stratifying patients in CaPSURE into low, intermediate and high risk categories for disease as described by D'Amico or applying the Kattan nomogram resulted in statistically significant differences in predicted 5-year FFP. However, there was considerable overlap of outcomes between the intermediate and high risk groups. This analysis suggests that simply estimating disease recurrence by stratifying patients into low, intermediate and high risk groups may not provide sufficient information for predicting outcomes among individuals.

KEY WORDS: prostate, prostatic neoplasms, prostatectomy, risk

Several groups have proposed nomograms based on preoperative prostate specific antigen (PSA), biopsy Gleason sum and clinical stage to predict the likelihood of pathologically organ confined disease.^{1,2} However, these models may be of relatively limited value for treatment selection or planning because pathological stage may not consistently predict clinical outcome. A more relevant outcome to guide clinical decision making may be the likelihood of biochemical, local or metastatic disease recurrence after local therapy. An accurate tool to predict disease recurrence after treatment is needed to educate realistically patients about the odds of treatment success, determine the need for adjuvant or neoadjuvant therapy and guide clinical trials of new treatment strategies.

Two risk assessment schemes based on 3 preoperative disease characteristics, namely PSA, biopsy Gleason sum and American Joint Committee on Cancer (AJCC) clinical stage, were introduced separately in 1998 by D'Amico³ and

Submitted for publication June 9, 2003. Supported by TAP Pharmaceutical Products, Inc. (CaPSURE) and the National Institutes of Health. Kattan⁴ et al. The Kattan nomogram and D'Amico risk levels (low/intermediate/high) are now widely used to predict freedom from disease after radical prostatectomy (RP).^{3,4} These tools enjoy popularity because they are easy to use and simple to understand by patients and physicians alike. Although these 2 instruments have been successfully validated in the United States and more recently in Europe, they are based exclusively on patient populations from academic medical centers.^{3–5} To our knowledge they have yet to be directly tested in a large, community based cohort.

There are several perceived differences between community and academic settings. The volume of patients treated at academic centers is much larger than at community practices. At academic centers prostate cancer is usually treated by a small number of surgeons with a large annual surgical volume. In contrast, community urologists often perform fewer radical prostatectomies than academic urologists. Variations in surgeon volume have been shown to affect postoperative morbidity and they may also impact cancer outcomes. Unique screening and treatment practices may further skew academic populations. In addition, community

based practices may serve patients with different socioeconomic backgrounds than those treated at academic centers, which could also potentially influence practice patterns and outcomes.⁷

The Cancer of the Prostate Strategic Urologic Research Endeavor (CaPSURE), a disease registry developed and maintained at University of California-San Francisco, has collected prospective prostate cancer data on a large number of patients, mainly from community urology practices. We applied the Kattan nomogram and D'Amico risk levels to the CaPSURE data set to determine the relationship between the 2 tools in a community based cohort.

METHODS

CaPSURE is a longitudinal, observational disease registry of patients with prostate cancer and has been previously described in depth.^{8,9} At the time of this analysis there were 8,283 patients from 33 urology practices in the United States (30 community practices, and 3 academic and Veterans Administration centers) enrolled in the registry.

Of enrolled men 3,377 had undergone radical prostatectomy as of December 2002. Eligibility requirements for the current analysis were 1) RP performed between January 1, 1989 and December 21, 2002 (836 patients excluded), 2) no treatment before surgery or within 6 months of RP (400 excluded), 3) clinically localized (T1-T3a) disease (30 excluded), 4) no more than 1 of 3 pretreatment variables (baseline PSA, Gleason grade or clinical stage) missing (98 excluded) and 5) at least 2 followup PSA values recorded after RP (312 excluded), leaving 1,701 who met these eligibility criteria.

Treatment failure was defined by either of 2 events: 1) biochemical recurrence, defined as serum PSA 0.2 ng/ml or greater on 2 consecutive measurements after surgery, 10 or 2) a second prostate cancer treatment greater than 6 months after RP. We have previously reported that second treatment can be a surrogate marker of disease recurrence in this population. 11, 12 In contrast to this analysis, Kattan et al considered second treatment any time after surgery as treatment failure. 4 Date of disease recurrence was defined as the date of the first increased PSA or the initiation of second treatment.

The preoperative disease characteristics used by D'Amico³ and Kattan⁴ et al are serum PSA, primary and secondary biopsy Gleason grade, and clinical stage. Some patients in our study were missing data on 1 or more of these variables. In addition, Gleason sum was often reported in CaPSURE instead of individual primary and secondary scores, which are needed for the Kattan nomogram. Finally, our data were collected using the 1997 AJCC-UICC system instead of the 1992 system, which Kattan⁴ and D'Amico³ et al used. The 1997 stage can be readily converted to the 1992 stage except for T2a, which is divided into T2a or T2b in the 1992 system. Therefore, as Kattan et al have previously reported,⁴ we imputed missing or indeterminate values for 1 clinical variable using other available data. As noted, we allowed only 1 imputation for study inclusion. Patients with multiple missing variables were excluded from study. We imputed 5% of PSA, 9% of clinical stage (1997 to 1992 staging) and 15% of Gleason grade. Statistical analysis was performed separately on the entire cohort of 1,701 patients and on a subgroup of 1,205 (71%) without any missing data to confirm results.

Five-year rates of freedom from progression (FFP) were estimated by the Kattan nomogram for each patient. Additionally, each patient was categorized into the low, intermediate or high risk group according to criteria used by D'Amico et al³ with minor modifications to conform to the Kattan nomogram. High risk cases were defined as PSA greater than 20 ng/ml, Gleason sum 8 to 10, or clinical stage T2c or T3a. The original criteria of D'Amico et al did not include T3a.

Intermediate risk cases were defined as PSA 10.1 to 20 ng/ml, Gleason sum 7 or clinical stage T2b. Finally, patients were considered at low risk if PSA at diagnosis was 10 ng/ml or less, Gleason score was 2 to 6 and clinical T stage was T1 or T2a. The original criteria of D'Amico et al did not differentiate between T1c and T2a.

ANOVA was used to test for differences in mean nomogram score among the 3 risk groups. Actuarial survival in each group was estimated via Kaplan-Meier analysis and the HR for failure in each group was determined using Cox proportional hazards regression. We compared these actual 5-year FFP values to expected FFP values based on the mean nomogram score in each risk group. All statistical analysis was performed using SAS, version 8.2 (SAS Institute, Cary, North Carolina).

RESULTS

Among the 1,701 men included in the study median followup was 33.6 months. At surgery mean patient age \pm SD was 62.8 \pm 6.6 years (range 39 to 79), 11% of patients were black, 39% completed high school and 35% were college graduates (table 1).

Table 1. Patient demogra	aphics
Age at surgery:	
$Mean \pm SD$	62.8 ± 6.6
Range	39–79
No. younger than 60 (%)	500 (29)
No. 60–64 (%)	451 (27)
No. 65–69 (%)	511 (30)
No. 70 or older (%)	239 (14)
No. race/ethnicity (%):	
Black	185 (11)
White	1,455 (86)
Other/unknown	61 (4)
No. education (%):	
Less than high school	235 (14)
High school graduate	663 (39)
College graduate	593 (35)
Unknown	210 (12)
No. \$ income (%):	
Less than 30,000	376 (22)
30,000-50,000	359 (21)
Greater than 50,000	560 (33)
Unknown	406 (24)

	No. Pts (%)
D'Amico risk group at diagnosis:	
Low	671 (39)
Intermediate	446 (26)
High	584 (34)
Gleason total at diagnosis:	
2–6	1,371 (81)
7	269 (16)
8–10	61 (4)
Gleason primary/secondary at diagnosis:	
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)
PSA at diagnosis (ng/ml):*	
4 or Less	230 (14)
4.1–10	1,066 (63)
10.1–20	286 (17)
Greater than 20	119 (7)
1992 AJCC-UICC clinical stage at diagnosis:	
T1a/b	44 (3)
T1c	481 (28)
T2a	476 (28)
T2b	224 (13)
T2c	450 (26)
T3a	26 (2)

Based on the D'Amico preoperative risk classification 671 cases (39%) were classified as low risk, 446 (26%) were intermediate risk and 584 (34%) were high risk. Table 2 lists the preoperative clinical risk characteristics in our cohort. At diagnosis using 1992 AJCC-UICC prostate cancer staging criteria the majority of cases were T1c or T2a (56%) and had a Gleason grade at diagnosis of 2 to 6 (81%). Of the patients 77% presented with PSA less than 10 ng/ml (mean 9.3 \pm 10.0).

Local therapy failed in 24% of patients, of whom 248 (60%) experienced biochemical recurrence and 165 (40%) received secondary therapy. Overall mean actuarial FFP 5 years after radical prostatectomy was 67% (95% CI 63% to 71%). Figure 1 shows Kaplan-Meier survival curves for each of the 3 risk groups. The intermediate and high risk curves were statistically indistinct and indeed there was little difference between the 2 groups. The Kattan nomogram scores likewise diverged between the low and intermediate groups only with mean scores of 91%, 74% and 69%, respectively. Moreover, the range of nomogram scores increases across D'Amico groups and all 3 groups included patients with almost 100% predicted survival. Figure 2 shows the wide range of nomogram scores and the considerable overlap of the intermediate and high risk groups.

Subanalysis of recurrence-free survival by RP year did not significantly alter Kattan nomogram predictions. Similarly the imputation of missing preoperative characteristic data did not significantly alter survival results.

DISCUSSION

Application of the Kattan nomogram and D'Amico risk group categories to the CaPSURE database resulted in statistically significant differences in predicted 5-year FFP. However, this analysis demonstrated 2 potential shortcomings of these risk assessment schemes.

1) While outcomes in the low risk group were tightly grouped about the mean, there was considerable dispersion of nomogram outcomes in the intermediate (30% to 98% FFP) and high risk (17% to 98% FFP) groups. Likewise, Kaplan-Meier survival curves demonstrated little distinction between the intermediate and high risk groups. While there was a clinically significant separation between the low and nonlow risk groups (intermediate and high), it was increasingly evident that there was little difference in outcome between the high and intermediate risk groups.

Kattan⁴ and D'Amico³ et al also placed heavy emphasis on individual preoperative disease characteristics. For example, in the latter scheme a patient with Gleason 4+4 but otherwise favorable disease characteristics (ie T1c disease and PSA 5 ng/ml) is considered at high risk. In addition, the inability to account for other potentially promising predictors of disease recurrence after surgery, such as percent positive biopsy data, may have limited the ability of the Kattan nomogram and D'Amico risk categories to predict accurately disease recurrence.¹³

2) A potential limitation of the Kattan nomogram is that it overestimated actual 5-year FFP, as estimated by the Kaplan-Meier method (91% vs 78%, 74% vs 63% and 69% vs 60% for the nomogram vs the Kaplan-Meier life table, respectively, table 3). A potential explanation is differences in surgical technique, as discussed, which could result in different outcomes. In the current study we also used slightly different criteria for failure by second treatment (excluding adjuvant treatment given within 6 months of surgery) and by PSA recurrence (using 0.2 ng/ml rather than 0.4 ng/ml as a threshold) than in the original study of Kattan et al.4 With respect to failure by second treatment it

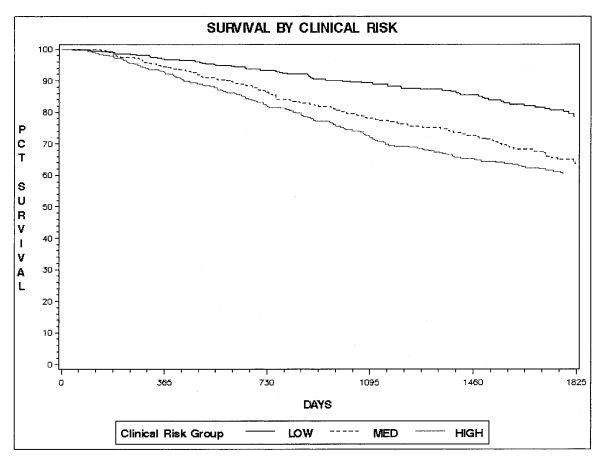


Fig. 1. Actuarial Kaplan-Meier CaPSURE recurrence-free survival curves for low, intermediate (MED) and high D'Amico risk groups. PCT, percent.

Kattan Nomogram Values by Clinical Risk Group

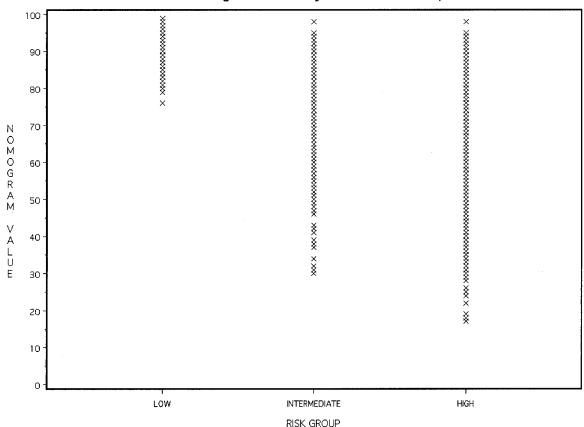


Fig. 2. Kattan nomogram predicted 5-year recurrence free survival for D'Amico risk groups. x, individual patient nomogram scores

	Risk Group		
	Low	Intermediate	High
D'Amico risk group (actual Kaplan-Meier FFP):			
No. pts	671	446	584
No. treatment failures	91	125	197
Mean % 5-yr FFP	78	63	60
95% CI*	71–83	55–70	53-66
No imputation % FFP*	78	62	57
HR†	1.00	1.87	2.32
% Kattan nomogram estimate:			
Mean 5-yr FFP ± SD	91 ± 4	74 ± 12	69 ± 17
Range	76–99	30–98	17-98

is also possible that for reasons not apparent from the available data community urologists participating in CaPSURE initiate second treatment more readily than the clinicians who participated in the study of Kattan et al, which would artificially inflate the failure rate.

However, differences in preoperative disease characteristics between the Kattan,⁴ D'Amico³ et al and CaPSURE patient populations may also account for the tendency of the nomogram to overestimate disease-free survival. The patients of Kattan⁴ and D'Amico³ et al tended to be at higher risk than patients in CaPSURE. Patients enrolled in CaPSURE tend to have lower PSA than those in the other 2 cohorts. While 33% of the patients of D'Amico et al³ and 29.9% of Kattan et al⁴ had preoperative PSA greater than 20, only 24% of the current cohort had PSA greater than 10 ng/ml. Patients in CaPSURE also tend to have lower clinical stage than those of Kattan et al. For example, in the data of

Kattan et al 49.4% of patients had T2b disease or greater compared to 41% in the current cohort. On further analysis almost twice as many patients in the series of Kattan et al had T2b disease (25% vs 13%) and 3 times more had T3a disease (5.9% vs 2%). Conversely almost twice as many patients in CAPSURE have T1c disease than in the study of Kattan et al (28% vs 15.1%). Finally, the patients of D'Amico et al³ were twice as likely as patients in CaPSURE to have Gleason sum 8 to 10.

These observed differences between the current patient population and those used to develop the Kattan nomogram and D'Amico classification are probably explained by several factors. 1) The CaPSURE database is largely community based compared to the exclusively academic data of the other 2 series (table 4). Differences between community and academic practice patterns, and patient characteristics may partly explain the lower PSA and clinical stage in CaPSURE.

Table 4. Study patient populations							
	CaPSURE Database D'Amico et al ³		Kattan et al ⁴				
Location	40 American Centers (33 community practices, 3 academic + 3 Veterans Affairs)		Hospital of University of Pennsylvania, Philadelphia, Pennsylvania	Methodist Hospital, Houston, Texas			
No. radical prostatectomy	1,701		888	983			
Mean age	62.8		Not available	63			
% White No. ng/ml PSA (%):	86		Not available	85			
Less than 4	230	(14)	85 (10)	217 (22.1)			
4–10	1,066	(63)	510 (57)	472 (48)			
10–20	286	(17)	210 (24)	187 (19)			
Greater than 20	119	(7)	83 (9)	107 (10.9)			
No. Gleason sum (%):							
2–6	1,371	(81)	681 (77)	Not available*			
7	269	(16)	133 (15)	Not available*			
8–10	61	(4)	74 (8)	Not available*			
No. 1992 AJCC stage (%):							
T1a/T1b	44	(3)	Not available	83 (8.4)			
T1c	481	(28)	256 (29)	148 (15.1)			
T2a	476	(28)	388 (44)	266 (27.1)			
T2b	224	(13)	93 (10)	246 (25)			
T2c	450	(26)	151 (17)	182 (18.5)			
T3a	26	(2)	Not available	58 (5.9)			
* Gleason grade reported by individual primary and secondary scores instead of sum.							

In addition, academic centers are tertiary referral centers, which may by nature may select for patients with more advanced or aggressive disease. 2) The trend toward lower clinical stage in CaPSURE may also be explained by a well documented under staging phenomenon. In a previous study we found that 24% of CaPSURE cases were clinically under staged. However, under staging in CaPSURE was less than at other academic institutions, casting doubt on this explanation. Another possible explanation is the continuing stage migration of prostate cancer toward lower stage disease. While our study included patients up to 2000 with 53% accrued since 1995, Kattan⁴ and D'Amico³ et al based their analyses on patients who underwent radical prostatectomy until 1996 and 1997, respectively.

The differences in pretreatment disease characteristics do not explain why Kattan et al consistently overestimated recurrence-free survival.4 The CaPSURE cohort with overall lower risk disease would be expected to have better recurrence-free survival than the nomogram prediction. A better explanation may be the treatment of disease recurrence of Kattan et al in their nomogram data. As mentioned, Kattan et al considered any secondary treatment after surgery as treatment failure. In contrast, we considered only secondary treatment greater than 6 months after surgery as failure. As a result, Kattan et al immediately considered patients who received adjuvant therapy for high risk disease (ie extracapsular extension, etc) to have treatment failure, possibly artificially inflating the number of true treatment failures. This definition of treatment failure may have contributed to the poor accuracy of nomogram predictions in this study.

CONCLUSIONS

Although stratifying patients into risk groups based on pretreatment variables (PSA, 1992 AJCC stage and Gleason grade) resulted in statistically significant differences in predicted and actual freedom from disease, there was considerable overlap between the high and intermediate risk groups. This analysis suggests that estimating disease recurrence using these criteria does not provide sufficient information to predict individual patient outcomes in these risk groups. New pretreatment assessment tools are needed, particularly to differentiate high and intermediate risk cases.

CaPSURE is managed by the Urology Outcomes Research Group, University of California-San Francisco.

REFERENCES

- Ross, P. L., Gerigk, C., Gonen, M., Yossepowitch, O., Cagiannos, I., Sogani, P. C. et al: Comparisons of nomograms and urologists' predictions in prostate cancer. Semin Urol Oncol, 20: 82, 2002
- Partin, A. W., Kattan, M. W., Subong, E. N., Walsh, P. C., Wojno, K. J., Oesterling, J. E. 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, 277: 1445, 1997
- D'Amico, A. V., Whittington, R., Malkowicz, S. B., Schultz, D., Blank, K., Broderick, G. A. et al: Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA, 280: 969, 1998
- 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
- 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
- Begg, C. B., Riedel, E. R., Bach, P. B., Kattan, M. W., Schrag, D., Warren, J. L. et al: Variations in morbidity after radical prostatectomy. N Engl J Med, 346: 1138, 2002
- Tarman, G. J., Kane, C. J., Moul, J. W., Thrasher, J. B., Foley, J. P., Wilhite, D. et al: Impact of socioeconomic status and race on clinical parameters of patients undergoing radical prostatectomy in an equal access health care system. Urology, 56: 1016, 2000
- 8. 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
- Cooperberg, M. R., Broering, J. M., Litwin, M. S., Lubeck, D. P., Mehta, S. S., Henning, J. M. et al: The contemporary management of prostate cancer in the United States: lessons from the Cancer of the Prostate Strategic Urologic Research Endeavor (CaPSURE), a national disease registry. J Urol, 171: 1393, 2004
- Freedland, S. J., Sutter, M. E., Dorey, F. and Aronson, W. J.:
 Defining the ideal cutpoint for determining PSA recurrence after radical prostatectomy. Prostate-specific antigen. Urology, 61: 365, 2003
- 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

- 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
- 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
- 14. Grossfeld, G. D., Chang, J. J., Broering, J. M., Li, Y.-P., Lubeck, D. P., Flanders, S. C. et al: Under staging and under grading in a contemporary series of patients under-
- going radical prostatectomy: results from the Cancer of the Prostate Strategic Urologic Research Endeavor database. J Urol, **165**: 851, 2001
- 15. Catalona, W. J. and Smith, D. S.: 5-Year tumor recurrence rates after anatomical radical retropubic prostatectomy for prostate cancer. J Urol, **152**: 1837, 1994
- Walsh, P. C., Partin, A. W. and Epstein, J. I.: Cancer control and quality of life following anatomical radical retropubic prostatectomy: results at 10 years. J Urol, 152: 1831, 1994
- tectomy: results at 10 years. J Urol, 152: 1831, 1994

 17. Zincke, H., Oesterling, J. E., Blute, M. L., Bergstralh, E. J., Myers, R. P. and Barrett, D. M.: Long-term (15 years) results after radical prostatectomy for clinically localized (stage T2c or lower) prostate cancer. J Urol, 152: 1850, 1994