Title
P Mithal et al.

Journal
International Journal of Urology, 22(4)

ISSN
0919-8172

Authors
Mithal, Prabhakar
Howard, Lauren E
Aronson, William J
et al.

Publication Date
2015-04-01

DOI
10.1111/iju.12704

Peer reviewed

INTERNATIONAL JOURNAL OF UROLOGY
International Journal of Urology (2015) **22**, 363–367 doi: 10.1111/iju.12704

Original Article

Prostate-specific antigen level, stage or Gleason score: Which is best for predicting outcomes after radical prostatectomy, and does it vary by the outcome being measured? Results from Shared Equal Access Regional Cancer Hospital database

Prabhakar Mithal,[1] Lauren E Howard,[2,3] William J Aronson,[4,5] Christopher J Kane,[6] Matthew R Cooperberg,[7] Martha K Terris,[8,9] Christopher L Amling[10] and Stephen J Freedland[2,3]

[1]Department of Urology, University of Rochester Medical Center, Rochester, New York, [2]Duke Prostate Center, Division of Urologic Surgery, Department of Surgery, Duke University School of Medicine, [3]Urology Section, Department of Surgery, Veterans Affairs Medical Center, Durham, North Carolina, [4]Urology Section, Department of Surgery, Veterans Affairs Greater Los Angeles Healthcare System, [5]Department of Urology, University of California at Los Angeles School of Medicine, Los Angeles, [6]Department of Urology, University of California at San Diego Medical Center, San Diego, [7]Departments of Urology, and Epidemiology and Biostatistics, University of California, San Francisco and Urology Section, Department of Surgery, Veterans Affairs Medical Center, San Francisco, California, [8]Urology Section, Department of Surgery, Veterans Affairs Medical Center, [9]Division of Urologic Surgery, Department of Surgery, Medical College of Georgia, Augusta, Georgia and [10]Division of Urology, Department of Surgery, Oregon Health and Science University, Portland, Oregon, USA

Abbreviations & Acronyms
ADT = androgen deprivation therapy
BCR = biochemical recurrence
CRPC = castration-resistant prostate cancer
IQR = interquartile range
PC = prostate cancer
PCSM = prostate cancer-specific mortality
PSA = prostate-specific antigen
PSADT = prostate-specific antigen doubling time
RP = radical prostatectomy
SEARCH = Shared Equal Access Regional Cancer Hospital

Correspondence: Stephen J Freedland M.D., Cedars Sinai Medical Center, 8635 West 3rd, Suite 1070W, Los Angeles, CA 90048, USA. Email: stephen.freedland@cshs.org

Received 14 May 2014; accepted 11 November 2014. Online Publication 24 February 2015

Objectives: To assess the ability of preoperative prostate-specific antigen level, Gleason score and stage to predict prostate cancer outcomes beyond biochemical recurrence, specifically castration-resistant prostate cancer, metastases and prostate cancer-specific mortality in radical prostatectomy patients.

Methods: We carried out a retrospective study of 2735 men in the Shared Equal Access Regional Cancer Hospital database treated by radical prostatectomy from 1988 to 2011 with data available on pathological stage, grade and preoperative prostate-specific antigen. We used Cox hazards analyses to examine the predictive accuracy (c-index) of the preoperative prostate-specific antigen (log-transformed), path Gleason score (<7, 3 + 4, 4 + 3 and 8–10) and path stage grouping (pT2 negative margins; pT2 positive margins; pT3a negative margins; pT3a positive margins; pT3b; *vs* positive nodes) to predict biochemical recurrence, castration-resistant prostate cancer, metastases and prostate cancer-specific mortality.

Results: Median follow up was 8.7 years, during which, 937 (34%) had biochemical recurrence, 108 (4%) castration-resistant prostate cancer, 127 (5%) metastases and 68 (2%) prostate cancer-specific mortality. For the outcomes of biochemical recurrence, castration-resistant prostate cancer, metastases and prostate cancer-specific mortality, the c-indices were, respectively: prostate-specific antigen 0.65, 0.66, 0.64 and 0.69; Gleason score 0.66, 0.83, 0.76 and 0.85; and pathological stage group 0.69, 0.76, 0.72 and 0.80.

Conclusions: Gleason score can predict with very high accuracy prostate cancer-specific mortality in patients undergoing radical prostatectomy. Thus, Gleason score should be given more weight in nomograms to predict prostate cancer-specific mortality. Furthermore, men with a high Gleason score should be given special consideration for adjuvant treatment or referral to clinical trials because of a higher risk of prostate cancer-specific mortality.

Key words: disease progression, mortality, prostatectomy, prostatic neoplasms, risk factors.

Introduction

Preoperative PSA, Gleason score and stage are the classic triad used to predict PC outcomes. Multiple nomograms exist using these variables to predict BCR after various treatments with considerable accuracy.[1] Previously, we found these nomograms predict not just BCR, but aggressive BCR and PCSM with even greater accuracy.[2,3] These findings suggest that rather than

being completely redeveloped, existing BCR nomograms can likely be recalibrated when considering later and arguably more clinically relevant endpoints. It is not clear, however, how the weight of each variable should be adjusted if PCSM is the primary end-point rather than BCR.

The natural history of PC after RP is typically very long, with life expectancy measured in years if not decades. As such, it is very conceivable that some variables might better predict early events (i.e. BCR), whereas other variables better predict later events (i.e. PCSM). Specifically, although PSA and pathological stage primarily describe tumor size and extent, showing how advanced the PC is today, Gleason grade describes how aggressive a tumor is and might thus better approximate how rapidly it is likely to progress (i.e. how advanced it will be in the future). If the natural history of PC in a RP patient were to be viewed as a race, the PSA and stage would be indicators of the starting point and Gleason grade an indicator of speed. Both starting point and speed impact the time at which the end-point (BCR or PCSM) is reached; however, the contribution of each factor also depends on the distance to the end-point. More advanced starting points should confer a relatively greater advantage for closer (BCR) rather than distant (PCSM) end-points. Conversely, the greater speed will become most advantageous with more distant end-points, overcoming the advantage conferred by the starting point.

Based on this hypothesis, we predicted that at the time of RP, indicators of PC aggressiveness (i.e. Gleason score) should predict PCSM risk with greater accuracy than indictors of how advanced the PC is (i.e. PSA and stage). In keeping with this, we have previously shown that at the time of BCR, two measures of tumor aggressiveness, PSADT and time to BCR onset, predict PCSM with greater accuracy than either PSA or stage.[4,5] At the time of surgery, however, it is unknown whether a patient will develop a BCR, and thus PSADT is also unknown. We postulate that in this setting, Gleason score will be the best measure of PC aggressiveness. In summary, we hypothesized that although PSA, stage and Gleason score all are important predictors of castration resistance, metastasis and PCSM, as a biological measure of disease aggressiveness, Gleason score would be the strongest predictor of more delayed end-points, such as time to PCSM after RP.

Methods

Study population

We carried out a retrospective study using combined data from men undergoing RP from 1988 to 2011 at the Veterans Affairs Medical Centers in West Los Angeles, Palo Alto and San Diego, California; Durham, North Carolina; and Augusta, Georgia, which were collected in the SEARCH database after obtaining institutional review board approval from each institution to extract and combine data.[6] Patients who underwent neoadjuvant therapy are not included in SEARCH. Furthermore, patients who were missing data on pathological Gleason score ($n=379$), preoperative PSA ($n=90$) or pathological stage ($n=118$) were excluded, leaving 2735 patients for analysis.

Clinical and pathological variables

TNM scoring was carried out using the 2009 American Joint Committee on Cancer staging system. Biochemical recurrence was defined as a single PSA value >0.2 ng/mL, two values of 0.2 ng/mL or secondary treatment for an elevated PSA after RP. Distant metastases, defined as bone or visceral or non-pelvic adenopathy, were determined by review of radionuclide bone scans, magnetic resonance imaging scans, computed tomography scans, plain radiograph reports and clinical progress notes. The decision to carry out radiographic imaging was at the discretion of the attending physician. CRPC was defined using the Prostate Cancer Working Group 2 criteria: a 25% or greater increase in PSA and absolute increase of ≥2 ng/mL PSA from the nadir while on continuous ADT. PCSM was defined as death in any patient with metastasis showing PC progression after ADT without another obvious cause of death. All information relevant to study variables was ascertained from medical records.

Statistical analysis

Group characteristics were compared using either Spearman's test or Kruskall–Wallis for non-normally distributed continuous variables and the χ^2-test for categorical variables. We used Cox hazards analysis and c-index to examine accuracy of the preoperative PSA (log-transformed), pathological Gleason score (<7, 3 + 4, 4 + 3 and 8–10) and pathological stage grouping (pT2, negative margins; pT2, positive margins; pT3a, negative margins; pT3a, positive margins; pT3b; positive nodes) when predicting BCR, CRPC, metastases and PCSM. Results were graphed using Nelson–Aalen cumulative hazard curves. Equality of survival functions was tested using log–rank. *P*-values <0.05 were considered statistically significant. The STATA 11 software package (StataCorp, College Station, TX, USA) was used for all statistical analysis.

Results

The clinical characteristics of the study population are summarized in Table 1. Overall, there was a sizable percentage of black men (36.8%), and the majority of patients had a PSA <10 ng/mL, Gleason score <7 or 3 + 4, and 75% of the cohort had pT2 disease. There were 618 (23%) patients in the present study cohort who received salvage therapy. Of these, 203 (33%) received hormone therapy, 249 (40%) received radiation therapy and 166 (27%) received both hormone and radiation therapy with the median time from recurrence to secondary treatment being 5.0 months (IQR 0.9–18.7 months). The majority of patients underwent RP with no lymph node dissection (66%), bilateral nerve sparing (68%) and a retropubic approach (78%).

During a median follow up of 8.7 years after RP among men who were alive at last follow up, 937 men (34%) had BCR, 108 (4%) CRPC, 127 (5%) metastases and 68 (2%) PCSM. Among men who reached the end-point, the median time from RP to BCR, CRPC, metastases, and PCSM were 19, 78, 71 and 99 months, respectively (Fig. 1). All 68 patients who died from PC developed both CRPC and metastasis before death.

All three variables of the classic PC triad (i.e. PSA, stage and Gleason score) were found to be significantly, but weakly, correlated with one another, with no collinearity between

Table 1 Clinical and pathological characteristics of patients in the RP cohort

Total number	2735
Mean age at surgery (years)	62.0 ± 6.4
Race, *n* (%)	
Non-black	1728 (63.2%)
Black	1006 (36.8%)
Median PSA, ng/mL (IQR)	6.7 (4.8-10.2)
Median surgery year (IQR)	2003 (1998–2006)
Pathological Gleason score, *n* (%)	
<7	1067 (39.0%)
3 + 4	1059 (38.7%)
4 + 3	295 (10.8%)
8–10	314 (11.5%)
Pathological stage, *n* (%)	
pT2, negative margin	1295 (47.4%)
pT2, positive margin	756 (27.6%)
pT3a, negative margin	120 (4.4%)
pT3a, positive margin	251 (9.2%)
pT3b	245 (9.0%)
Positive nodes	68 (2.5%)
Extracapsular extension, *n* (%)	560 (20.5%)
Positive margins, *n* (%)	1225 (44.8%)
Seminal vesicle invasion, *n* (%)	280 (10.2%)
Lymph node dissection†, *n* (%)	
No	1806 (66.0%)
Yes	68 (2.5%)
Not done	861 (31.5%)
Nerve sparing†, *n* (%)	
None	321 (19.2%)
Unilateral	219 (13.1%)
Bilateral	1130 (67.7%)
Radical prostatectomy approach† *n* (%)	
Retropubic	2101 (78.3%)
Perineal	337 (12.6%)
Laparoscopic	66 (2.4%)
Robot assisted	179 (6.7%)
Median follow up, years (IQR)	8.7 (4.7–11.7)

†Reported among patients with data available

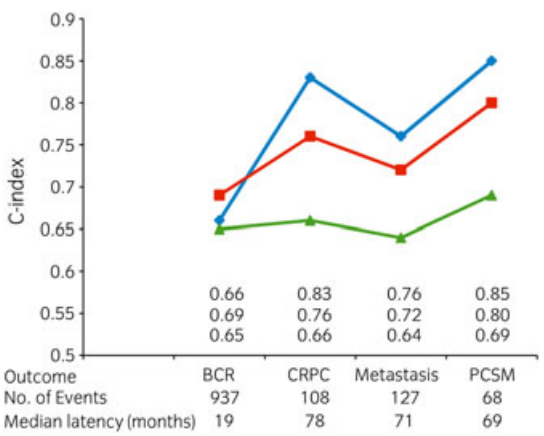

Fig. 1 Variable c-indices by clinical end-point. (—●—), Grade; (—▲—), PSA; (—■—), stage.

them. Spearman's rho values for PSA versus stage, PSA versus Gleason score, and stage versus Gleason score were 0.27, 0.24 and 0.36, respectively (all $P < 0.001$).

All three variables were significantly correlated with all outcomes examined (Table 2). In general, the predictive accuracy of all three predictor variables, as measured by c-index, increased as the end-point became farther from the time of surgery (Fig. 1). In other words, all variables tended to predict later events (i.e. PCSM) better than earlier events (i.e. BCR). However, although all three variables had similar predictive accuracy for BCR, the greatest improvement in accuracy over the end-points was seen in Gleason score. Specifically, for the outcomes of BCR, CRPC, metastasis and PCSM, the c-indices for Gleason score were 0.66, 0.83, 0.76 and 0.85, respectively. Gleason score also had the greatest accuracy amongst the three variables when predicting the later events of CRPC, metastasis and PCSM. The cumulative risks of BCR, CRPC, metastasis and PCSM stratified by Gleason score are shown in Figure 2.

Discussion

We found that as the clinical end-point becomes more distant from the time of surgery, while almost all features become more important, specifically Gleason score becomes a very strong predictor of poor outcome. Though Gleason score, PSA and pathological stage all predicted BCR with similar accuracy, Gleason score was the best predictor of PCSM, with very high accuracy (c-index = 0.85). These findings support our hypothesis that although PSA and stage are indicators of how advanced the cancer is and play important roles in predicting early outcomes, Gleason score better reflects how biologically aggressive the disease is and thus, how rapidly the cancer will grow over time. The clinical relevance is that men with high-grade Gleason score disease are at significantly increased risk of PCSM, and this information needs to be kept in mind when discussing adjuvant and early salvage treatments, as well as in designing clinical trials.

The present results are consistent with prior studies, which found that nomograms using PSA, Gleason score and stage, which were developed to predict BCR, actually predict aggressive BCR and PCSM with even greater accuracy than BCR itself.[2,3] If the relative contribution of PSA, grade and stage can be reassessed in direct relation to end-points later than BCR, this should lead to even further increases in the predictive power of existing nomograms.

The key prognostic importance of Gleason score in predicting distant end-points is evidenced by findings in contemporary studies of PCSM. In a large, multicenter study using Fine and Gray competing risk regression analysis, Eggener *et al.* reported that Gleason score played a central role in their postoperative nomogram, which was externally validated and achieved an overall c-index of 0.92 predicting 15-year PCSM risk. They concluded that Gleason score was the strongest determinant of PCSM.[7] Interestingly, in a preoperative nomogram predicting 15-year PCSM developed by the same group, a lower overall c-index of 0.82 was achieved.[8] As the authors acknowledged, most of those patients with low biopsy Gleason score who died were found to have been upgraded at the time of surgery.[8] In the present study, pathological Gleason score alone could predict PCSM with a c-index of 0.85, which is more accurate than most nomograms incorporating multiple variables. Although the present results were based on a modest number

Table 2 Correlations of variables with clinical outcomes

	BCR		CRPC		Metastasis		PCSM	
	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
PSA	2.1 (2.0–2.4)	<0.001	2.1 (1.6–2.7)	<0.001	1.7 (1.3–2.1)	<0.001	2.0 (1.5–2.8)	<0.001
Gleason grade								
<7	–	–	–	–	–	–	–	–
3 + 4	2.0 (1.7–2.4)	<0.001	2.9 (1.5–5.8)	0.002	1.1 (0.6–1.8)	0.807	2.2 (0.9–5.3)	0.069
4 + 3	3.4 (2.7–4.1)	<0.001	6.1 (2.9–12.8)	<0.001	3.4 (1.9–5.9)	<0.001	5.9 (2.3–14.8)	<0.001
8–10	4.5 (3.7–5.4)	<0.001	18.7 (10.1–35.4)	<0.001	8.4 (5.3–13.3)	<0.001	19.4 (9.0–41.9)	<0.001
Stage								
pT2, positive margin	2.7 (2.3–3.2)	<0.001	1.2 (0.7–2.3)	0.523	1.2 (0.7–2.0)	0.608	0.9 (0.5–2.4)	0.911
pT3a, negative margin	2.6 (1.9–3.4)	<0.001	1.9 (0.7–5.5)	0.230	2.5 (1.1–5.6)	0.031	2.7 (0.8–9.6)	0.116
pT3a, positive margin	3.4 (2.8–4.2)	<0.001	2.2 (1.1–4.4.3)	0.029	2.8 (1.6–5.0)	<0.001	2.9 (1.2–6.7)	0.014
pT3b	7.2 (5.9–8.7)	<0.001	8.0 (4.8–13.5)	<0.001	6.1 (3.8–9.8)	<0.001	12.1 (6.3–23.3)	<0.001
Positive nodes	10.6 (7.7–14.7)	<0.001	20.6 (11.0–38.5)	<0.001	9.5 (4.8–18.8)	<0.001	14.3 (5.3–37.1)	<0.001

of PC deaths, if validated in future studies, this would suggest that a greater weight should be assigned to pathological Gleason score when developing PCSM nomograms. This reinforces the need to give special consideration to patients with high-grade disease when carrying out adjuvant clinical trials, and also highlights the need to be able to better predict actual pathological grading before surgery.

The fact that PSA, grade and stage all become more important over time could be due to each being a measure of PC biology, and that all three variables are correlated with each other. In other words, a patient with an elevated PSA is more likely to have high-grade disease and advanced stage. As such, our findings suggest that all three factors remain significant predictors of later end-points, and just the relative weights of each need to be adjusted when predicting PCSM versus BCR. Furthermore, the present results highlight the work from others that found that not all BCR events progress to distant metastases.[4] Indeed, though the rate of BCR was high in the present study, our rates of distant end-points were low, also consistent with prior studies that oncological outcomes after RP are quite good.[8]

An important point from the present findings is that the risk of PCSM remained relatively low, even in patients with high-grade disease. Specifically, the 10-year risk of metastasis and PCSM was 19.5% and 13.8%, respectively, for men with Gleason 8 PC or higher (Fig. 2). Thus, even in the presence of high-grade disease, the vast majority of men (at least at 10 years) remain free of distant metastases. As such, the low positive predictive value even of the most robust variables used for PC risk assessment demands the development of novel molecular markers for more specific patient treatment selection. In addition, it argues that even among "high-risk" men, there is likely a need to use multiple variables to select the "highest-risk" men.

In the present study, we did not attempt to provide a tool for clinical use, as several highly accurate tools already exist, but rather to highlight the increased importance of Gleason score in contemporary PC risk assessment relative to other commonly

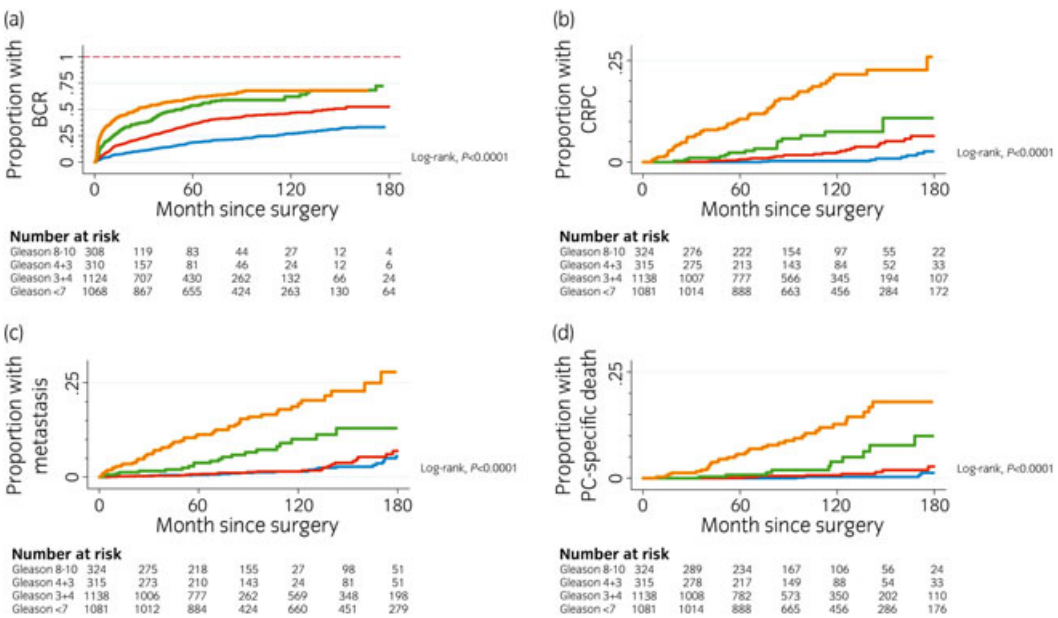

Fig. 2 Cumulative hazard curves after RP stratified by pathological Gleason score group for (a) BCR, (b) CRPC, (c) metastasis and (d) PCSM. (—◆—), Gleason <7; (—■—), Gleason 3 + 4; (—▲—), Gleason 4 + 3; (—●—), Gleason 8–10.

used variables.[8] Also, our goal was not to develop a multivariable model, and therefore we did not compare Gleason score with any existing multivariable models. Certainly, a multivariable model is ideal, and our data suggest that PSA, grade and stage should all be included in such models, with special emphasis placed on grade. Though we 100% agree that comprehensive tools should be used (indeed our group has developed many of these tools), they have not been adopted by all physicians currently managing prostate cancer.[9] As such, it is worthwhile to revisit the contributions of individual variables to risk assessment, as these are the pieces of information most likely to be accessible across all clinical settings.

The present study had several limitations. Gleason grading was not centralized, and therefore the exact Gleason scoring criteria used by each individual pathologist is not known. Despite this limitation, Gleason score was the strongest predictor of delayed outcomes. This suggests that if Gleason score had been graded in a systematic and uniform fashion, its predictive accuracy for late outcomes might have been even higher. The present study was carried out in a population of patients who underwent only RP within the Veterans Affairs health system. Thus, these findings might not be applicable to patients receiving other modes of definitive treatment and in other settings wherein the entire prostate is not examined. It was not possible to account for patient quality of life in the present study, a factor that must be incorporated into future analyses to assess the true impact of PC on the patient. Finally, given the long natural history of PC, longer follow-up times are required for more accurate risk analyses.

In conclusion, among men undergoing RP, pathological Gleason score was a better predictor of long-term outcomes, such as CRPC, metastasis and PCSM, than PSA and pathological stage. In nomograms to predict PCSM, Gleason score should be given more weight than PSA and stage. Men with high Gleason scores should be given special consideration for adjuvant clinical trials because of their increased risk of PCSM.

Conflict of interest

Matthew R Cooperberg is a paid consultant to Amgen, Abbott, Dendreon, Astellas.

References

1. Freedland SJ. Screening, risk assessment, and the approach to therapy in patients with prostate cancer. *Cancer* 2011; **117**: 1123–35.
2. Schroeck FR, Aronson WJ, Presti JC Jr *et al*. Do nomograms predict aggressive recurrence after radical prostatectomy more accurately than biochemical recurrence alone? *BJU Int.* 2009; **103**: 603–8.
3. Teeter AE, Presti JC Jr, Aronson WJ *et al*. Do Nomograms Designed to Predict Biochemical Recurrence (BCR) Do a Better Job of Predicting More Clinically Relevant Prostate Cancer Outcomes than BCR? A Report from the SEARCH Database Group. *Urology* 2013; **82**: 53–9.
4. Freedland SJ, Humphreys EB, Mangold LA *et al*. Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy. *JAMA* 2005; **294**: 433–9.
5. Freedland SJ, Humphreys EB, Mangold LA, Eisenberger M, Partin AW. Time to prostate specific antigen recurrence after radical prostatectomy and risk of prostate cancer specific mortality. *J. Urol.* 2006; **176**: 1404–8.
6. Allott EH, Abern MR, Gerber L *et al*. Metformin does not affect risk of biochemical recurrence following radical prostatectomy: results from the SEARCH database. *Prostate Cancer Prostatic Dis.* 2013; **16**: 391–7.
7. Eggener SE, Scardino PT, Walsh PC *et al*. Predicting 15-year prostate cancer specific mortality after radical prostatectomy. *J. Urol.* 2011; **185**: 869–75.
8. Stephenson AJ, Kattan MW, Eastham JA *et al*. Prostate cancer-specific mortality after radical prostatectomy for patients treated in the prostate-specific antigen era. *J. Clin. Oncol.* 2009; **27**: 4300–5.
9. Kim SP, Karnes RJ, Nguyen PL *et al*. Clinical Implementation of Quality of Life Instruments and Prediction Tools for Localized Prostate Cancer: Results from a National Survey of Radiation Oncologists and Urologists. *J. Urol.* 2013; **189**: 2092–8.