

Prostatic Diseases and Male Voiding Dysfunction

Models to Predict Positive Prostate Biopsies Using the Tyrol Screening Study

Prasanna Sooriakumaran, Majnu John, Paul Christos, Jasmin Bektic, Georg Bartsch, Robert Leung, Michael Herman, Douglas Scherr, and Ashutosh Tewari

OBJECTIVES	To describe two predictive models that predict for prostate cancer on biopsy derived from a large screening population. There are no published predictive models that predict prostate cancer in a screened population.
METHODS	The patients from the Tyrol screening study of known age, total prostate-specific antigen (PSA) level, digital rectal examination (DRE) findings, prostate volume, and percentage of free PSA, and who underwent an initial prostate biopsy from January 1992 to June 2004 were included (n = 2271). Multivariate logistic regression models were used to develop the biopsy positivity predictive models: nomogram 1, age, DRE, and total PSA; and nomogram 2, age, DRE, total PSA, and percentage of free PSA. The predictive accuracy of the models was assessed in terms of discrimination and calibration. External validation of the nomograms was performed using a urologically referred population of patients who underwent prostate biopsy (n = 599).
RESULTS	Both nomograms were well-calibrated internally and externally and discriminated well between patients with positive and negative biopsy findings for both the European and U.S. cohorts (model 2 better than model 1).
CONCLUSIONS	Our nomogram with age, total PSA, and DRE had good predictive ability to differentiate between screened patients with cancer on the initial prostate biopsy and those without. Adding the percentage of free PSA improves this predictive power further. These models might aid in clinical decision making regarding the need for biopsy in both European and U.S. populations. UROLOGY xx: xxx, xxxx. © 2011 Elsevier Inc.

Prostate cancer is the most common nondermatologic cancer affecting men in the Western world, with an estimated 217 730 new cases and 32 050 deaths occurring in 2010 in the United States alone.¹ Prostate cancer is also the third most common cause of death from cancer among men in many Western countries, including the United States and Austria after lung and colorectal cancer.¹⁻⁴ The diagnosis of prostate cancer is determined by the histologic findings; thus, a prostate biopsy remains the reference standard test. However, prostate biopsy is an invasive test and is associated with a number of adverse effects. These include hematuria, hematospermia, bleeding per rectum, and infection, which can range in severity from a mild urinary tract infection to severe, life-threatening septicemia.⁵ Patients also invariably experience some discomfort despite the use of local anesthesia.⁶ Hence, whether

to subject a man to prostate biopsy is determined by his risk of having prostate cancer. A patient deemed at high risk of a positive biopsy will be more likely to undergo biopsy than a patient at low risk of biopsy positivity. Thus, it is important to determine the risk with the greatest precision possible. There are established criteria for predicting prostate cancer risk according to age, digital rectal examination (DRE) findings, and prostate-specific antigen (PSA) level. However, it is known that most men with a PSA level greater than the age-specific reference range will not be found to have prostate cancer on biopsy. Adding DRE does not aid significantly in limiting the need for biopsy.⁷

More sophisticated methods of predicting whether a given patient will be diagnosed with prostate cancer on biopsy have been developed. Two artificial neural networks for the early detection of prostate cancer in men with a total PSA (tPSA) level of 2.5-4 ng/mL and 4-10 ng/mL were prospectively developed using a referral database from the European Prostate Cancer Detection Study.^{8,9} Two predictive models have also been published for the early detection of prostate cancer in men with abnormal DRE findings and tPSA values of <4 ng/mL.^{10,11} However, these methods were developed us-

From the Department of Urology, Weill Cornell Medical College, New York, New York; Division of Biostatistics and Epidemiology, Department of Public Health, Weill Cornell Medical College, New York, New York; and Department of Urology, Innsbruck Medical University, Innsbruck, Austria

Reprint requests: Ashutosh Tewari, M.D., Department of Urology, Weill Cornell Medical College, 525 East 68th Street, Starr 900, New York, NY 10065. E-mail: akt2002@med.cornell.edu

Submitted: September 17, 2010, accepted (with revisions): May 7, 2011

ing study populations consisting of men referred for early cancer detection or for lower urinary tract symptoms and, hence, might not be applicable to nonreferral patient populations. In the current era, screening for prostate cancer has become widespread, with established programs in the United States and many other countries. These programs identify a large number of potential candidates for biopsy if current criteria including DRE and age-specific tPSA are used. Approximately 70% of men with an increased PSA level identified in prostate cancer screening programs will not have prostate cancer.¹² Thus, it would be invaluable to the clinician to have a visual model to predict the risk of a prostate cancer diagnosis on biopsy. In the present report, we describe 2 predictive models created using the Tyrol cohort that can predict the risk of positive biopsy in a screened population.

MATERIAL AND METHODS

Patient Population

Since 1988, early prostate cancer detection has been promoted by the Department of Urology, University of Innsbruck (Innsbruck, Austria), using both PSA and DRE in the diagnostic evaluation of asymptomatic healthy men. Since 1993, an early prostate cancer detection program has been carried out in the Federal State of Tyrol (1 of the 9 Federal States of the Republic of Austria) with prospective data collection and documentation and the development of an associated biorepository. From September 1993 to September 1995, tPSA levels greater than the age-specific normal ranges¹³ (40-49 years, 0-2.5 ng/mL; 50-59 years, 0-3.5 ng/mL; 60-69 years, 0-4.5 ng/mL; and 70-79 years, 0-6.5 ng/mL) and a percentage of free PSA (%fPSA) level of <22% were used as criteria for recommending biopsy. Since October 1995, lower age-specific tPSA normal ranges¹⁴ (40-49 years, 0-1.25 ng/mL; 50-59 years, 0-1.75 ng/mL; 60-69 years, 0-2.25 ng/mL; and 70-79 years, 0-3.25 ng/mL) and the percentage of free PSA level of 18% were used.

The present study included a total number of 3838 men from the Tyrol Prostate Cancer Screening study, who underwent biopsy from January 1992 to June 2004. The data, including age and DRE results, were recorded for each patient. tPSA values were available for 3798 men (99.0%) and the %fPSA was available for 2402 men (62.5%). Because men with missing %fPSA data were older and to have greater tPSA levels, on average, than the rest of the cohort (data not shown) and no other variables were identified that could reliably be used to impute their values, these men were excluded.

Thus, the final study sample included 2271 patients (59.2%) with complete information for all variables of interest (age, DRE, tPSA, and %fPSA) and had a tPSA level of <100 ng/mL and prostate volume of <200 cm³.

Although the primary aim of the present study was to determine whether we could develop predictive models in a screened population, we also wanted to see whether our models based on the aforementioned screened population could be used to predict biopsy positivity in a urologically referred population. A total of 1180 patients underwent prostate biopsy at Weill Cornell Medical College (WCRC) for a tPSA elevated to greater than the age-specific normal levels and/or abnormal DRE findings from October 1990 and November 2005. However, the tPSA level and DRE findings were missing for 13 patients (1.1%), prostate volume data were missing for 37 patients (2.6%), and the %fPSA was missing for 564 patients (47.8%). Because this data set from WCRC was to be used for external validation, only patients with complete information for age, DRE, tPSA, prostate volume, and %fPSA were included; thus, the final data set used for external validation included 599 patients (52.5%).

Statistical Analysis

Multivariate logistic regression models were used to develop the biopsy positivity predictive models. We considered 2 different models: model 1 used the standard risk factors of age, DRE, and tPSA; and model 2 included those variables plus the %fPSA. Any possible nonlinear relationship between the predictive factors and biopsy outcome was evaluated using a multiple fractional polynomial method.¹⁵

The predictive accuracy of the models was assessed in terms of discrimination and calibration. Discrimination is the ability to differentiate between men with positive and negative biopsy findings and is measured using the receiver operating characteristic curve, summarized by the area under the curve (ie, concordance index). Bootstrap methods with 200 bootstrap samples were used to correct for overoptimism of each predictive model during internal validation. Calibration is the correlation between the number of those with positive biopsy findings predicted and the number found to be positive and was assessed by fitting locally weighted scatterplot smoothing curves to the observed proportion versus the predicted probability of biopsy positivity for prostate cancer. Confidence intervals were evaluated using the bootstrap method,¹⁶ with 95% coverage.

A decision analysis was performed to illustrate the clinical implications of performing biopsy according to the predicted risk of prostate cancer using the different predictive models developed from the Tyrol screening study sample and applying them separately to the Tyrol and WCRC cohorts.

Table 1. Summary of overall patient characteristics for Tyrol and WCRC cohorts

Characteristic	Tyrol	WCRC	P Value
Patients (n)	2271	599	
Age (y)	58.7 (9.6)	62.6 (8.8)	<.0001
tPSA (ng/mL)			<.0001
Median	3.8	5.2	
Interquartile range	2.5-6.1	4.1-7.2	
Patients with suspicious DRE findings (n)	208 (9.2)	211 (35.2)	<.0001
Prostate volume (cm ³)	33.0 ± 14.8	54.7 (31.3)	<.0001
%fPSA	14.0 ± 5.9	19.5 (10.4)	<.0001
Positive needle biopsy findings (n)	659 (29.0)	184 (30.7)	.4460

WCRC, Weill-Cornell Medical College; tPSA, total prostate-specific antigen; DRE, digital rectal examination; %fPSA, percentage of free PSA. Data in parentheses are percentages.

The normality of each continuous variable was assessed by plotting the histograms and using Shapiro-Wilks tests. *t* tests and Wilcoxon rank sum tests were used to compare normal and non-normal continuous variables, respectively; chi-square tests were used for the categorical variables. All analyses were done using R, version 2.9.1¹⁷ and the Regression Modeling Strategies package.¹⁸

RESULTS

The patient characteristics of the men undergoing biopsy from the Tyrol and WCMC cohorts are listed in Table 1. The Tyrol group was on average younger, had lower tPSA values, and had fewer patients with suspicious DRE findings. The Tyrol group also had smaller prostates and

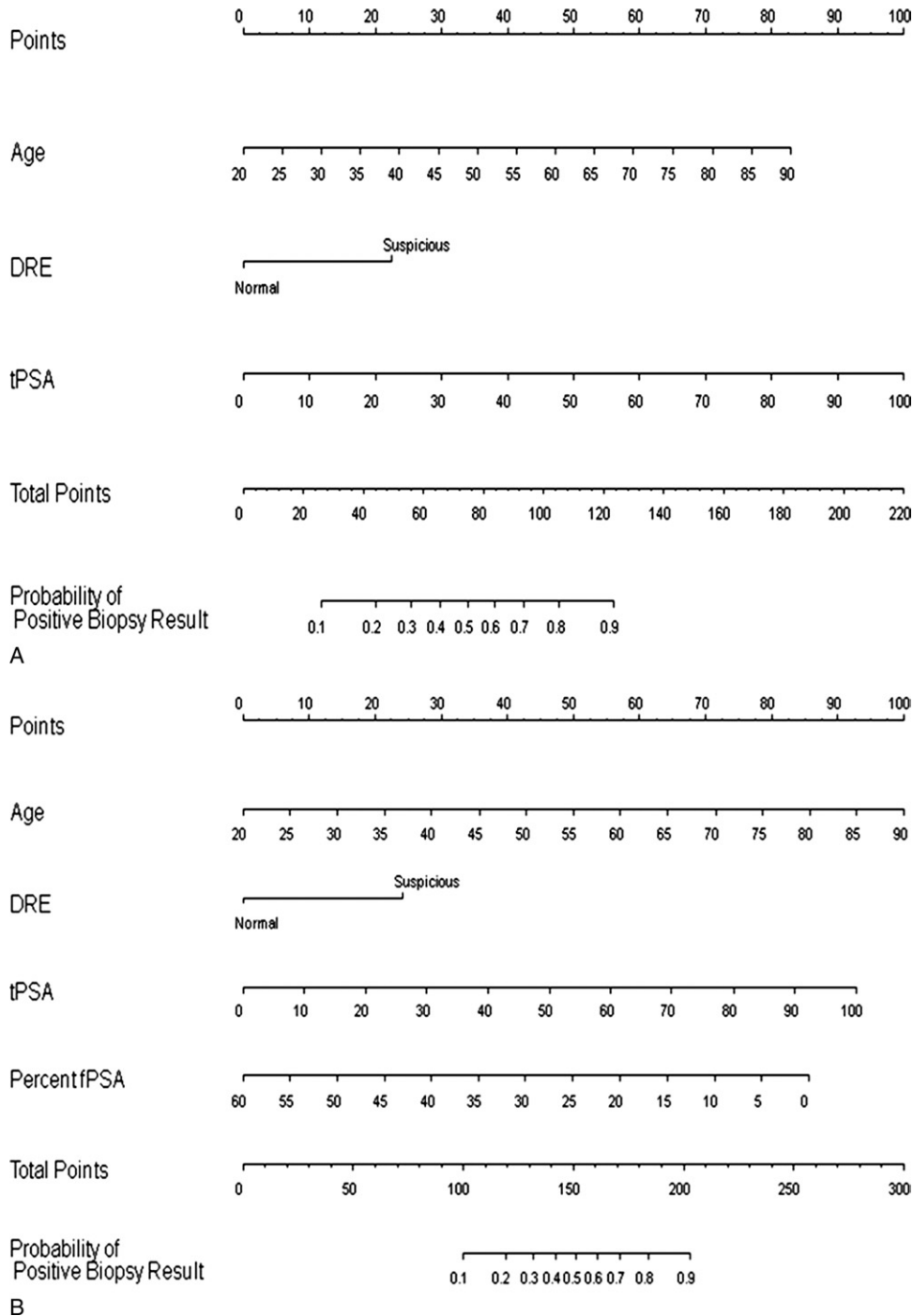


Figure 1. Predictive models to predict probability of positive prostate biopsy. To obtain model-predicted probability of positive biopsy result, locate patient values on each axis. Draw vertical line to points axis to determine how many points attributed for each variable’s value. Sum points for all variables. Locate sum on total points line to assess individual probability of positive biopsy result on probability of positive biopsy result line. **(A)** Model 1 (age, tPSA, and DRE). **(B)** Model 2 (age, tPSA, DRE, and %fPSA).

Table 2. Clinical implications of using prediction models for prostate cancer (Tyrol sample)

Variable	Biopsy (n)		Prostate Cancer (n)	
	Performed	Avoided	Found	Missed
Biopsy all	2271 (100)	—	659 (100)	—
Biopsy of those with $\geq 10\%$ risk of cancer				
Model with age, tPSA, and DRE	2241 (98.7)	30 (1.3)	657 (99.7)	2 (0.3)
Model with age, tPSA, DRE, and %fPSA	2176 (95.8)	95 (4.2)	653 (99.1)	6 (0.9)
Biopsy of those with $\geq 20\%$ risk of cancer				
Model with age, tPSA, and DRE	1554 (68.4)	717 (31.6)	540 (81.9)	119 (18.1)
Model with age, tPSA, DRE, and %fPSA	1484 (65.3)	787 (34.7)	531 (80.6)	128 (19.4)

Abbreviations as in Table 1.

Data in parentheses are percentages.

lower %fPSA values than the WCMC group. However, the proportions from both cohorts with positive prostate biopsies were similar. For both the Tyrol and the WCMC cohorts, biopsy-positive men were on average older and had greater tPSA levels; a greater proportion also had suspicious DRE findings. The biopsy-positive men in both cohorts also had smaller prostates and lower %fPSA values (data not shown).

The 2 biopsy positivity predictive models are shown in Figure 1. Model 1, which includes age, DRE, and total PSA, has a bootstrap corrected internally validated concordance index of 0.691 (95% confidence interval [CI] 0.667-0.716) and externally validated concordance index of 0.663 (95% CI 0.617-0.709). Adding the %fPSA to that model increased the discriminative ability, as evidenced by the improved concordance indexes for both internal (Tyrol) and external (WCMC) validation (concordance index 0.710, 95% CI 0.687-0.735; and concordance index 0.717, 95% CI 0.673-0.761, respectively). Both models were also well calibrated internally and externally on locally weighted scatterplot smoothing plots (data not shown).

The clinical implications of performing biopsy according to the predicted risk of prostate cancer of $\geq 10\%$ and $\geq 20\%$ using the 2 predictive models (Table 2). Only performing biopsy on those with a $\geq 10\%$ risk of prostate cancer from the full model (age, DRE, tPSA, and %fPSA) would have resulted in 95 fewer biopsies (4.2%), while only missing 6 cases of prostate cancer (0.9%). Increasing this threshold to 20% would spare 787 biopsies (34.7%) and miss 128 cases of cancer (19.4%).

Similarly, the clinical implications of performing biopsy according to the predicted risk of prostate cancer

of $\geq 10\%$ and $\geq 20\%$ using the 2 predictive models and applying them to the external WCMC study sample are listed in Table 3. Only performing biopsy on those with a $\geq 10\%$ risk of prostate cancer from the full model (age, DRE, tPSA, and %fPSA) would have resulted in 45 fewer biopsies (7.5%), while missing only 6 cases of prostate cancer (3.3%). Increasing this threshold to 20% would have spared 169 biopsies (28.2%) and miss 19 cases of prostate cancer (10.3%).

COMMENT

Predicting the risk of prostate cancer is a noble goal for the prostate cancer specialist. It can better determine the need for prostate biopsy, which is an invasive, unpleasant test with significant morbidity and, rarely, mortality. Although there are published models that predict biopsy positivity risk, none have been shown to be accurate enough without becoming too cumbersome to gain acceptance into widespread clinical practice. The Prostate Cancer Prevention Trial (PCPT) and European Randomized Study of Screening for Prostate Cancer cohorts have generated their own risk calculators.^{19,20} The PCPT-derived cancer risk calculator was created from data from 5519 men in the placebo arm of the PCPT aged >55 years with normal DRE findings and a PSA level of ≤ 3 ng/mL at study entry. This group underwent biopsy when the DRE findings were abnormal or the PSA level increased to ≥ 4 ng/mL and also at the end of the study (7 years). The European Randomized Study of Screening for Prostate Cancer risk calculator was created from the data from 6288 Dutch men in the screening arm

Table 3. Clinical implications of using prediction models for prostate cancer (WCMC sample)

Variable	Biopsy (n)		Prostate Cancer (n)	
	Performed	Avoided	Found	Missed
Biopsy all	599 (100)	—	184 (100)	—
Biopsy of those with $\geq 10\%$ risk of cancer				
Model with age, tPSA, and DRE	591 (98.7)	8 (1.3)	183 (99.5)	1 (0.5)
Model with age, tPSA, DRE, and %fPSA	554 (92.5)	45 (7.5)	178 (96.7)	6 (3.3)
Biopsy of those with $\geq 20\%$ risk of cancer				
Model with age, tPSA, and DRE	473 (79.0)	126 (21.0)	166 (90.2)	18 (9.8)
Model with age, tPSA, DRE, and %fPSA	430 (71.8)	169 (28.2)	165 (89.7)	19 (10.3)

Abbreviations as in Table 1.

Data in parentheses are percentages.

of the study and, unlike the PCPT cancer risk calculator, was based on the first round of screening results only. Although both tools have been shown to be superior to PSA alone in the prediction of biopsy positivity,²¹ the risk assessments show significant intermodel discrepancies. This is likely related to differences in study design that resulted from the different populations selected for entry (eg, the PCPT cancer risk calculator had a much larger proportion of black subjects than did the European Randomized Study of Screening for Prostate Cancer).²² The Prostate Biopsy Collaborative Group analyzed the prebiopsy PSA levels from 25 772 biopsies and 8503 cancer cases from 8 different cohorts in 6 different countries²³ and found that the European and U.S. populations had significant differences in the terms of risk of positive biopsy for the same baseline criteria. They attributed this to the more extensive workup of U.S. patients for other conditions before being selected for biopsy. In contrast to this, our study showed good external calibration for our models derived from a European cohort when validated in our U.S. sample.

Another population-based European screening trial was performed in Finland. In this, white men aged 55-67 years were screened during 1996 and 1997 and those with a serum PSA level of 4-20 $\mu\text{g/L}$ ($n = 758$) were used to establish the predictive risk models.¹² In the present study, the prostate cancer probability depended most strongly on the %fPSA. The tPSA level, prostate volume, and DRE findings also contributed to prostate cancer probability, although age and family history of prostate cancer did not. However, in the Tyrol screening study, the patients were referred for a biopsy according to their age-matched PSA level and %fPSA, with no restriction on age. In contrast to the Finnish study results, we found that age was a strong predictor of biopsy positivity.

In the present study, we developed 2 visual models that predict biopsy positivity in a population screened for prostate cancer. Our models are based on simple, easily available clinical parameters, such as age, DRE findings, tPSA, and %fPSA. We internally validated our models and used bootstrap samples to reduce overfit bias. This validation revealed that both our models generally had good discrimination according to concordance indexes of 0.691 and 0.710, respectively, for model 1 (age, DRE, tPSA) and model 2 (criteria in model 1 plus %fPSA). Thus, there is a 69.1% and 71.0% chance that a pair of biopsy-positive and biopsy-negative men will be correctly classified by model 1 and 2, respectively. Both predictive models were well calibrated internally and externally, meaning that the physician can give a reasonably accurate prediction of the probability that a screened patient will be positive at the initial biopsy test. The external validity of our models is reassuring, especially because other studies have found discrepancies in the risk assessment for models derived from European and U.S. cohorts. Our model appears of use in counseling men who

are screened for prostate cancer on the need to undergo prostate biopsy regardless of the locality of the population. For example, if model 2 was used and suggested a >20% risk of cancer, almost 35% of unnecessary biopsies would be avoided but at the expense of missing almost 20% of the cancers. What would be useful is to know whether how many of the missed cancer cases would be clinically significant (ie, of high grade); however, data were not available in our study and constitutes a limitation of this decision analysis. It should also be noted that our predictive models do not apply to patients who have already undergone a previous negative prostate biopsy, and we have not tried to validate our models for these difficult-to-treat patients.

The present study had are other limitations. Our predictive models were based on data only from 2271 (59.2%) of the 3838 patients who underwent biopsy, because we excluded those who had missing tPSA and %fPSA values (37.5%). The men with missing %fPSA values were older (mean age 67.0 vs 58.9 years, $P < .0001$) and also had greater total PSA values (median tPSA 6.6 vs 3.8 ng/mL, $P < .0001$) compared with the whole Tyrol cohort. Hence, only our first model, which included age, DRE, and tPSA, was representative of the entire Tyrol cohort undergoing biopsy. However, because the concordance index (0.691) of the first model was very similar to that of model 2 (concordance index 0.710), it could be used in the absence of information about the %fPSA.

CONCLUSIONS

We have developed 2 simple predictive models of increasing predictive power that appear reasonably reliable in predicting initial biopsy positivity in men screened for prostate cancer. These predictive models might aid clinical decision making about the need for prostate biopsy in screened men and might be equally applicable to U.S. and European populations.

References

1. Jemal A, Siegel R, Xu J, et al. Cancer Statistics. *CA Cancer J Clin*. 2010.
2. Vutuc C, Schernhammer ES, Haidinger G, et al. Prostate cancer and prostate-specific antigen (PSA) screening in Austria. *Wien Klin Wochenschr*. 2005;117:457-461.
3. Vutuc C, Waldhoer T, Haidinger G. Cancer mortality in Austria: 1970-2002. *Wien Klin Wochenschr*. 2004;116:669-675.
4. Jemal A, Bray F, Center MM, et al. Global cancer statistics. *CA Cancer J Clin* 2011;61:69-90.
5. Rodríguez LV, Terris MK. Risks and complications of transrectal ultrasound guided prostate needle biopsy: a prospective study and review of the literature. *J Urol*. 1998;160:2115-2120.
6. Djavan B, Waldert M, Zlotta A, et al. Safety and morbidity of first and repeat transrectal ultrasound guided prostate needle biopsies: results of a prospective European prostate cancer detection study. *J Urol*. 2001;166:856-860.
7. Djavan B, Remzi M, Marberger M. When to biopsy and when to stop biopsying. *Urol Clin North Am* 2003;30:253-262, viii.
8. Carlson GD, Calvanese CB, Partin AW. An algorithm combining age, total prostate-specific antigen (PSA), and percent free PSA to

- predict prostate cancer: results on 4298 cases. *Urology*. 1998;52:455-461.
9. Djavan B, Remzi M, Zlotta A, et al. Novel artificial neural network for early detection of prostate cancer. *J Clin Oncol*. 2002;20:921-929.
 10. Djavan B, Zlotta A, Kratzik C, et al. PSA, PSA density, PSA density of transition zone, free/total PSA ratio, and PSA velocity for early detection of prostate cancer in men with serum PSA 2.5 to 4.0 ng/mL. *Urology*. 1999;54:517-522.
 11. Eastham JA, May R, Robertson JL, et al. Development of a nomogram that predicts the probability of a positive prostate biopsy in men with an abnormal digital rectal examination and a prostate-specific antigen between 0 and 4 ng/mL. *Urology*. 1999;54:709-713.
 12. Finne P, Auvinen A, Aro J, et al. Estimation of prostate cancer risk on the basis of total and free prostate-specific antigen, prostate volume and digital rectal examination. *Eur Urol*. 2002;41:619-627.
 13. Oesterling JE, Jacobsen SJ, Chute CG, et al. Serum prostate-specific antigen in a community-based population of healthy men: establishment of age-specific reference ranges. *JAMA*. 1993;270:860-864.
 14. Horninger W, Berger A, Pelzer A, et al. Screening for prostate cancer: updated experience from the Tyrol study. *Can J Urol*. 2005;12(suppl 1):7-13, 92-93.
 15. Sauerbrei W, Cook MH, Benner A, et al. Multivariable regression model building by using fractional polynomials: description of SAS, Stata and R programs. *Comput Stat Data Anal*. 2006;50:21-39.
 16. Efron B. The bootstrap and modern statistics *J Am Stat Assoc*. 2000;95:3.
 17. R D Core Team. R: a language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing. Available from: www.R-project.org. Accessed August 7, 2010.
 18. Harrell F *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis*. New York: Springer-Verlag; 2001.
 19. Krane R, Roobol M, Schröder FH. A graphical device to represent the outcomes of a logistic regression analysis. *Prostate*. 2008;68:1674-1680.
 20. Thompson IM, Ankerst DP, Chi C, et al. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. *J Natl Cancer Inst*. 2006;98:529-534.
 21. Oliveira M, Marques V, Carvalho AP, et al. Head-to-head comparison of two online nomograms for prostate biopsy outcome prediction. *BJU Int*. 2010;107:1781-7830.
 22. van den Bergh RC, Roobol MJ, Wolters T, et al: The Prostate Cancer Prevention Trial and European Randomized Study of Screening for Prostate Cancer risk calculators indicating a positive prostate biopsy: a comparison. *BJU Int*. 2008;102:1068-1073.
 23. Vickers AJ, Cronin AM, Roobol MJ, et al. The relationship between prostate-specific antigen and prostate cancer risk: the Prostate Biopsy Collaborative Group. *Clin Cancer Res*;16:4374-4381.