

Improved prediction of prostate cancer recurrence through systems pathology

Carlos Cordon-Cardo, Angeliki Kotsianti, David A. Verbel, Mikhail Teverovskiy, Paola Capodieci, Stefan Hamann, Yusuf Jeffers, Mark Clayton, Faysal Elkhettabi, Faisal M. Khan, Marina Sapir, Valentina Bayer-Zubek, Yevgen Vengrenyuk, Stephen Fogarsi, Olivier Saidi, Victor E. Reuter, Howard I. Scher, Michael W. Kattan, Fernando J. Bianco Jr., Thomas M. Wheeler, Gustavo E. Ayala, Peter T. Scardino, Michael J. Donovan

J Clin Invest. 2007;117(7):1876-1883. <https://doi.org/10.1172/JCI31399>.

Research Article

Oncology

We have developed an integrated, multidisciplinary methodology, termed *systems pathology*, to generate highly accurate predictive tools for complex diseases, using prostate cancer for the prototype. To predict the recurrence of prostate cancer following radical prostatectomy, defined by rising serum prostate-specific antigen (PSA), we used machine learning to develop a model based on clinicopathologic variables, histologic tumor characteristics, and cell type-specific quantification of biomarkers. The initial study was based on a cohort of 323 patients and identified that high levels of the androgen receptor, as detected by immunohistochemistry, were associated with a reduced time to PSA recurrence. The model predicted recurrence with high accuracy, as indicated by a concordance index in the validation set of 0.82, sensitivity of 96%, and specificity of 72%. We extended this approach, employing quantitative multiplex immunofluorescence, on an expanded cohort of 682 patients. The model again predicted PSA recurrence with high accuracy, concordance index being 0.77, sensitivity of 77% and specificity of 72%. The androgen receptor was selected, along with 5 clinicopathologic features (seminal vesicle invasion, biopsy Gleason score, extracapsular extension, preoperative PSA, and dominant prostatectomy Gleason grade) as well as 2 histologic features (texture of epithelial nuclei and cytoplasm in tumor only regions). This robust platform has broad applications in patient diagnosis, treatment management, and prognostication.

Find the latest version:

<https://jci.me/31399/pdf>





Improved prediction of prostate cancer recurrence through systems pathology

Carlos Cordon-Cardo,¹ Angeliki Kotsianti,² David A. Verbel,² Mikhail Teverovskiy,² Paola Capodiecici,² Stefan Hamann,² Yusuf Jeffers,² Mark Clayton,² Faysal Elkhettabi,² Faisal M. Khan,² Marina Sapir,² Valentina Bayer-Zubek,² Yevgen Vengrenyuk,² Stephen Fogarsi,² Olivier Saidi,² Victor E. Reuter,¹ Howard I. Scher,¹ Michael W. Kattan,¹ Fernando J. Bianco Jr.,¹ Thomas M. Wheeler,³ Gustavo E. Ayala,³ Peter T. Scardino,¹ and Michael J. Donovan²

¹Memorial Sloan-Kettering Cancer Center, New York, New York, USA. ²Aureon Laboratories Inc., Yonkers, New York, USA.

³Department of Pathology, Baylor College of Medicine, Houston, Texas, USA.

We have developed an integrated, multidisciplinary methodology, termed *systems pathology*, to generate highly accurate predictive tools for complex diseases, using prostate cancer for the prototype. To predict the recurrence of prostate cancer following radical prostatectomy, defined by rising serum prostate-specific antigen (PSA), we used machine learning to develop a model based on clinicopathologic variables, histologic tumor characteristics, and cell type-specific quantification of biomarkers. The initial study was based on a cohort of 323 patients and identified that high levels of the androgen receptor, as detected by immunohistochemistry, were associated with a reduced time to PSA recurrence. The model predicted recurrence with high accuracy, as indicated by a concordance index in the validation set of 0.82, sensitivity of 96%, and specificity of 72%. We extended this approach, employing quantitative multiplex immunofluorescence, on an expanded cohort of 682 patients. The model again predicted PSA recurrence with high accuracy, concordance index being 0.77, sensitivity of 77% and specificity of 72%. The androgen receptor was selected, along with 5 clinicopathologic features (seminal vesicle invasion, biopsy Gleason score, extracapsular extension, preoperative PSA, and dominant prostatectomy Gleason grade) as well as 2 histologic features (texture of epithelial nuclei and cytoplasm in tumor only regions). This robust platform has broad applications in patient diagnosis, treatment management, and prognostication.

Introduction

A number of prostate cancer nomograms combine clinical and pathologic variables to predict the probability of disease recurrence or survival for individual patients (1–7). The postoperative nomogram, developed by Kattan et al. (5, 7), is widely used by physicians to estimate the probability of disease recurrence following radical prostatectomy, as signaled by a rising serum prostate-specific antigen (PSA) level. The prognostic variables in this nomogram are pretreatment serum PSA level; Gleason grade; and microscopic assessment of prostate capsular invasion, surgical margins, seminal vesicle invasion, and lymph node metastasis. The predictions appear to be accurate, with an area under the curve (AUC) ranging from 0.80 to 0.89 in different validation studies (2). However, for patients in the middle range, i.e., 7-year progression-free survival of 30% to 70%, the nomogram prediction is believed to be no more accurate than a coin toss (8). On further analysis, the concordance index of the nomogram indicates that it performs slightly better than midway between a model with perfect discrimination (1.0) and one with no discriminating ability (0.5) (8).

To develop a novel predictive model for PSA recurrence in prostate cancer patients treated with prostatectomy, we chose to apply a systems pathology methodology (9–11). Systems pathology makes use of novel technologies in object-oriented image analysis, pattern recognition, and quantitative biomarker multiplexing. The complex datasets obtained are analyzed by supervised machine learning algorithms, such as support vector regression for censored data (SVRc) (12–17). Our working hypothesis was that a more accurate tool for predicting patient outcome could be developed by using systems pathology. Specifically, our approach was to include, along with the conventional clinicopathologic variables, 2 types of features from the prostate tissue: morphometric features of specific cell types and immunohistochemical (IHC) analysis of biomarkers. We extended this approach by employing quantitative immunofluorescence in a second model, which further confirmed the standard IHC biomarker results and the utility of systems pathology in developing predictive tests.

Results

A model for predicting PSA recurrence using systems pathology. From a cohort of 539 patients treated at Baylor College of Medicine, 17 clinicopathologic features (see Supplemental Table 2; supplemental material available online with this article; doi:10.1172/JCI31399DS1) were retrospectively collected. Missing values for clinicopathologic features were imputed with regression models containing data for all of the features to estimate the value of the missing feature without reference to outcome. We then excluded patients with incomplete morphometric or molecular data or with missing outcome information. Of note, 271 of the 539 patients (50%) were excluded due to

Nonstandard abbreviations used: AR, androgen receptor; AUC, area under the curve; IHC, immunohistochemical/immunohistochemistry; MSKCC, Memorial Sloan-Kettering Cancer Center; PIN, prostatic intraepithelial neoplasia; PSA, prostate-specific antigen; SVRc, support vector regression for censored data; TMA, tissue microarray.

Conflict of interest: C. Cordon-Cardo, A. Kotsianti, M.J. Donovan, D.A. Verbel, M. Teverovskiy, P. Capodiecici, S. Hamann, Y. Jeffers, M. Clayton, F. Elkhettabi, F.M. Khan, M. Sapir, V. Bayer-Zubek, Y. Vengrenyuk, S. Fogarsi, and O. Saidi have financial interest in Aureon Laboratories Inc.

Citation for this article: *J. Clin. Invest.* 117:1876–1883 (2007). doi:10.1172/JCI31399.

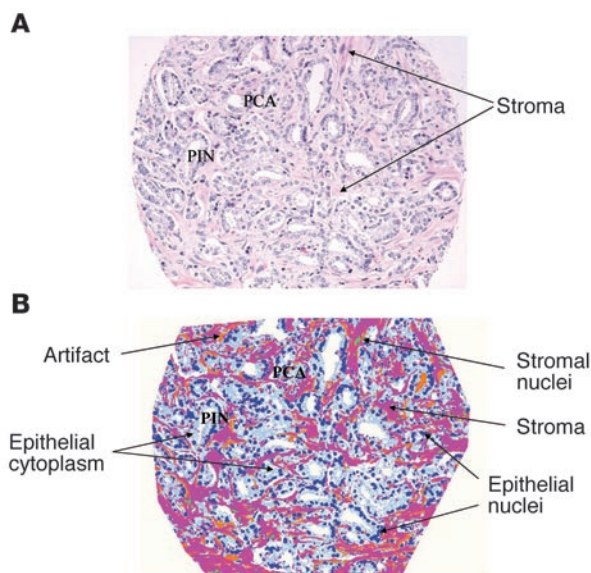


Figure 1

Image analysis of a prostate tissue sample. (A) H&E-stained section of a 0.6-mm TMA core demonstrating overall specimen heterogeneity, including the presence of prostate cancer (PCA), PIN, and stromal tissue with collagen, myofibroblasts, vessels, and scattered inflammatory cells. (B) Segmented H&E-stained image of PCA and PIN with epithelial nuclei in dark blue, cytoplasm in light blue, stromal nuclei in green, and remaining stromal elements in pink-purple. Artifacts have been annotated in orange.

insufficient tumor in the evaluable tissue microarray (TMA) cores (for additional details see “Image analysis and morphometry” in Supplemental Methods). In addition, those who had received neoadjuvant or adjuvant therapy (hormonal or radiation therapy) were considered nonevaluable and were removed from the cohort, leaving data from 262 evaluable patients to use in training the model (see Supplemental Table 1 for complete accountability information). Patient characteristics are detailed in Supplemental Table 2. The median age at diagnosis was 63 years (range, 38–81 years of age), and the median PSA before radical prostatectomy was 7.8 ng/ml (range, 0.9–81.9 ng/ml). In the prostatectomy samples, 40.8% of patients had a Gleason score of less than 7; 55% had a Gleason score of 7; and the remaining 4.2% had a Gleason score greater than 7. The 1992 tumor, node, and metastasis clinical stage was in T1c in 113 patients (43.1%), T2a–c in 130 patients (49.7%) and T3a–c in 17 patients (6.5%). In the pathologic analysis, cancer stage was T2N0M0 (tumor confined to prostate) in 158 patients (60%), T3aN0M0 (tumor extends through prostate capsule) in 70 patients (27%), and T3bN0M0 (tumor invades the seminal vesicle) or T2–3N+ (tumor either confined to prostate or with local extension and metastasis to regional lymph nodes) in the remaining 34 patients (13%). The clinical characteristics of the 262 patients were similar to those of the initial cohort of 539 patients (see Supplemental Table 2 for comparison of full cohort and evaluable patients).

A total of 37 patients (14%) experienced PSA recurrence, defined as 2 consecutive PSA measurements >0.2 ng/ml. Median follow-up of

patients with no recurrence observed was 59.1 months. The overall median time to recurrence was not reached.

Image analysis and morphometry. Digitized hematoxylin and eosin images of prostate sections containing at least 80% prostate cancer (Figure 1A) were processed using a custom-made image analysis system (Histology Labeling Tool; custom designed by Aureon Laboratories Inc.). Individual cellular elements, including epithelial and stromal nuclei, epithelial cytoplasm, and stroma, were segmented and then classified based on their location (stroma versus epithelium) and abundance within the tissue section (Figure 1B). A total of 496 morphometric features were quantified using the image software. Upon subsequent filtering, the set was reduced to 33 individual validated histologic features (Supplemental Table 4).

IHC biomarker analysis. Prostate tissue samples displayed in TMAs were assessed by traditional enzymatic IHC methods to establish biomarker antigen profiles. Six target antigens were selected based on their cellular distribution and 6 for their association with prostate cancer and/or progression as reported in published cDNA microarray studies and IHC analyses (18–25). From these 12 biomarkers, 43 features were recorded (see Supplemental Table 5). These features were derived from the analysis of each antibody within areas of a specific histologic type: prostate cancer, prostatic intraepithelial neoplasia

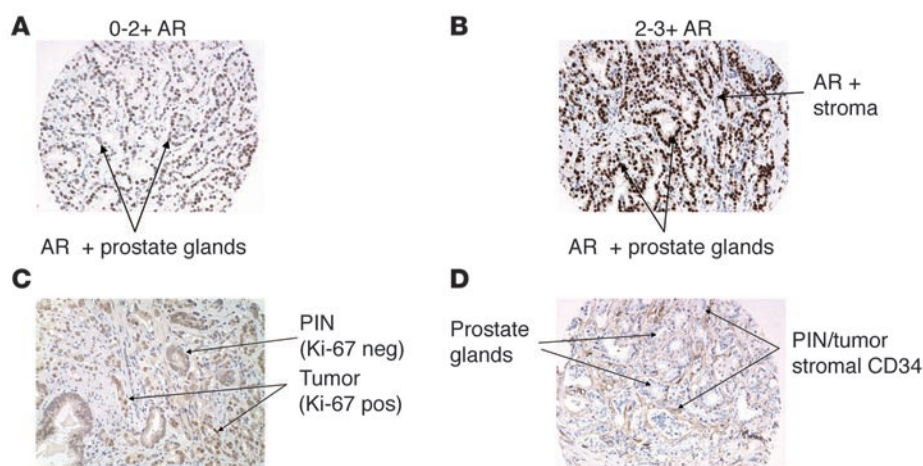


Figure 2

Standard IHC on selected TMA prostate tissue cores illustrates the variability of expression patterns and the subjective nature of the intensity scoring used by the pathologist. (A and B) Variability when scoring AR in tumor epithelial nuclei (intensity score of 100 cells counted, derived from the intensity value [0–3+] multiplied by the percentage of positive cells [0–100%] with a range of 0–300) is as follows: (A) 20% of the epithelial nuclei had 0 intensity, 50% had 1+ intensity, and 30% had 2+ intensity (staining index, 110/300); while (B) 30% of the epithelial nuclei had 3+ intensity and 70% had 2+ intensity (staining index, 230/300). Scoring additional markers (such as Ki67) also emphasized heterogeneity within a single section. (C) For Ki67, 97% of PIN had 0 intensity and 3% had 1+ intensity; 20% of prostate cancer had 0 intensity, 30% had 1+ intensity, and 50% had 2+ intensity. (D) For CD34 observed expression was identified with respect to its proximity to PCA and/or PIN.

**Table 1**

Features selected for PSA recurrence model

Feature	Weight
Lymph nodes	−23.32
Surgical margins	−11.73
Stromal area / total tissue area	−16.15
AR (tumor) staining index	−10.49
EpitNuc area / total tissue area	11.54
Stroma mean ± SD	−11.26
Biopsy Gleason score	−10.60
Seminal vesicle invasion	−6.40

EpitNuc, epithelial nuclear.

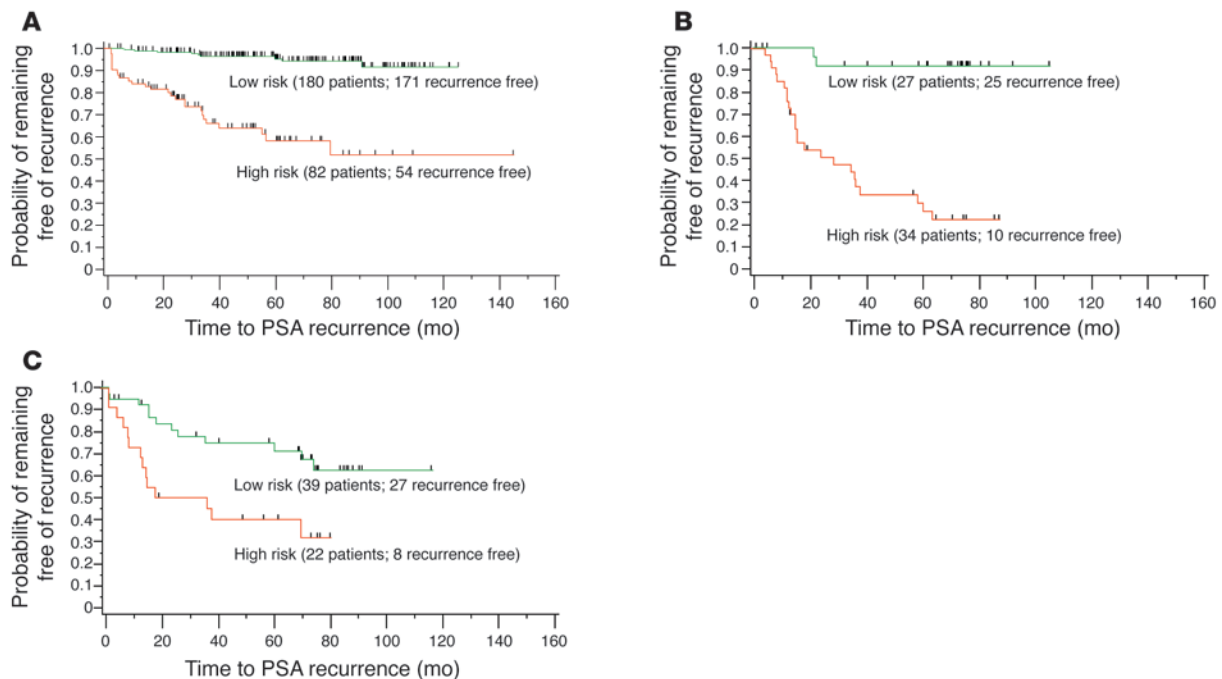
(PIN), and atrophic gland. For a summary of biomarker results, see Supplemental Table 6. Figure 2, A and B, illustrates the variability in staining intensity of the androgen receptor (AR) (scored on a scale of 0 to 3) observed across individual TMA cores. Figure 2, C and D, illustrates other associations observed for specific markers, such as a high level of Ki67 in the invasive tumor when compared with PIN and the proximity of tumor and PIN to CD34⁺ endothelial cells.

Feature selection and statistical analysis. We used domain expertise, which included knowledge of biomarker stratification and image analysis, combined with SVRc, to develop a model containing 8 features (4 clinicopathologic, 3 morphometric, and 1 molecular; Table 1) from the initial set of 93 features.

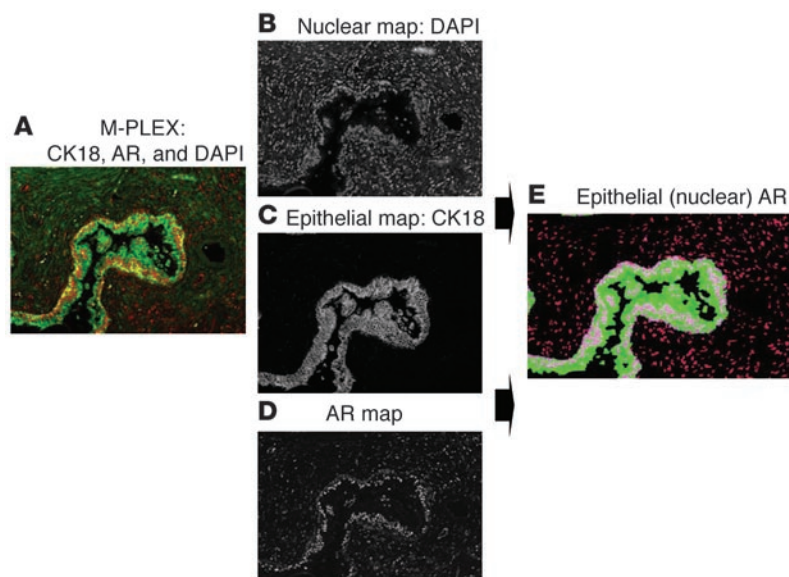
The 4 clinicopathologic features selected as being predictive of PSA recurrence (in order of their importance as indicated by their weights) were lymph node status (−23.32), surgical margin status

(−11.73), biopsy Gleason score (−10.60), and seminal vesicle invasion (−6.40) (Table 1). The negative weights indicate that the presence of each feature (or higher value of a continuous feature) was associated with a shorter time to PSA recurrence. These clinical features reflect both the biological potential of prostate cancer and the technical capabilities of the surgeon. In particular, positive surgical margins were recently identified as being an independent predictor of 10-year progression-free probability (26).

Of the 33 image features, only 3 were selected as improving the prediction of PSA recurrence: the relative area of stroma (weight, −16.15), the relative area of epithelial nuclei (weight, 11.54), and the variation in texture within stroma as expressed in the red channel (weight, −11.26). The first 2 features reflect the area occupied by stroma or epithelial cell nuclei in sections of prostate tissue composed predominantly of tumor elements but also of benign elements. The positive weight of relative epithelial nuclei area indicates that a larger area of epithelial nuclei was associated with a longer time to PSA recurrence; this feature may correspond to the “compactness,” i.e., the close packing of small, round nuclei that is seen in benign processes. Increasing amounts of epithelial nuclei are reflective of well differentiated prostate carcinoma, while increased stroma most likely reflects the presence of individual or isolated small clusters of epithelial cells (poorly differentiated). The quantitative measurements derived from these image patterns may represent more objective determinants of the traditional Gleason grading system. The final imaging feature, variation in stromal texture as evidenced by changes in the staining properties, most likely reflects the biochemical attributes of stroma associated with tumor versus benign elements. Additional morphologic attributes (27–38) derived from nuclear and stromal objects are currently under investigation.

**Figure 3**

Kaplan-Meier curves of freedom from PSA recurrence, according to SVRc model score. (A) Training cohort. (B) Validation cohort. Patients were stratified into low- and high-risk groups based on the sensitivity/specificity cut points. Tick marks indicate censored patients. (C) Validation cohort based on MSKCC 5-year PSA recurrence (PSAR) progression-free nomogram. Patients were stratified into high- and low-risk groups based on nomogram predictions (using an optimized log-rank χ^2 cut point) of probability of remaining free of recurrence.

**Figure 4**

Spectral image segmentation for quantifying nuclear AR in control prostate tissue. (A) Multiplex immunofluorescent image of prostate tissue stained with DAPI for nuclei (not shown), CK18 labeled with Alexa Fluor 488 for epithelial cells (green), and AR labeled with Alexa Fluor 568 for stroma and epithelial cells (red). By applying spectral optics to separate the respective fluorochromes, individual grayscale images were created for nuclei (B), epithelial cells (C), and AR (D). (E) Using algorithms based on pixel and object classification, a composite color image was rendered that segmented the AR that was present in DAPI-stained nuclei that were within CK18⁺ epithelial cells as well as the AR that was present in nuclei that were CK18 negative (stroma). The segmented image was the basis for quantifying the AR in epithelial nuclei as well as the AR within the stroma. Parameters calculated included overall nuclear area involved with AR, mean and maximum intensity, and distribution (% positive cells).

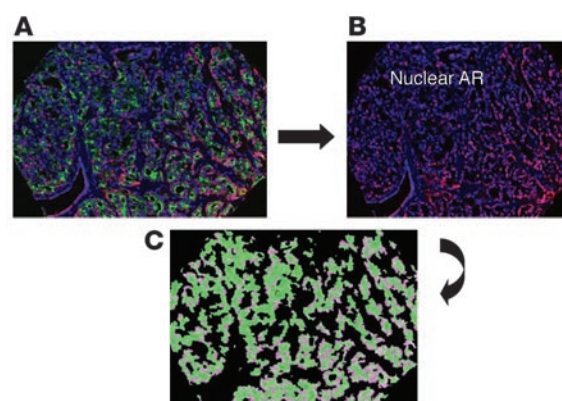
Fourteen of the biomarker features, encompassing 8 of the 12 biomarkers, were associated with PSA recurrence in univariate analysis. Of note, the presence of PSA in PIN was found to be the most statistically significant ($P = 0.001$). Only the staining index of nuclear AR (the activated form) within tumor epithelial cells, however, was identified by the model as being predictive of recurrence, with a weight of -10.49 . Two additional markers, CD34 and Ki67, although not selected by the model, demonstrated a trend toward association with PSA recurrence (data not shown).

On the training dataset, the model had a concordance index of 0.86. The sensitivity was 85%, and specificity was 81% for correctly predicting recurrence within 5 years of prostatectomy. Patients were divided into groups with low or high model-predicted probability of recurrence; the Kaplan-Meier estimates of freedom from recurrence for each group revealed their significant difference ($P < 0.0001$; log-rank test) and are illustrated in Figure 3A.

Validation study. The model was validated with an external cohort consisting of 366 patients from Memorial Sloan-Kettering Cancer Center (MSKCC). As in the training study, patients who had received neoadjuvant or adjuvant hormonal or radiation therapy were removed from the cohort. Complete record sets were obtained for 61 patients and were included in the analysis (see Supplemental Table 1 for complete accountability information). Of note, 301 of the 366 patients (82%) were excluded due to insufficient tumor in the evaluable TMA cores and/or whole sections from the prostatectomy samples (for additional details, see "Image analysis and morphometry" in Supplemental Methods). The median age at diagnosis was 62 years (range, 42–74), and the median preoperative PSA was 10.0 ng/ml (range, 2.0–69.5). Based on the prostatectomy samples, 15% had a Gleason score of less than 7, while 61% had a Gleason score of 7, and the remaining 24% had a Gleason score greater than 7 (see Supplemental Table 2). The characteristics of these 61 evaluable patients were similar to those of the initial cohort of 366 patients (see Supplemental Table 2 for comparison of full cohort and evaluable patients).

A total of 26 patients (43%) experienced PSA recurrence. For patients with no observed recurrence, the median follow-up time was 70.3 months. The overall median time to PSA recurrence was not reached.

The TMA cores from the 61 patients contained at least 50% prostate cancer cells as evaluated by H&E analysis. The acquired H&E images were processed by a histology labeling tool, and individual cores were subsequently evaluated for the presence of AR in tumor using IHC as described above. Application of the SVRc model to this validation cohort resulted in a concordance index of 0.82, sensitivity of 96%, and specificity of 72%. Figure 3B illustrates the Kaplan-Meier estimates of recurrence for records with low and high SVRc model scores. The groups showed a statistically significant difference in time to PSA recurrence ($P < 0.0001$; log-rank test). Of note, a model trained with just the clinicopathologic features achieved only a validation

**Figure 5**

Multiplex immunofluorescence of a representative TMA core utilized for quantifying AR. (A) Combined immunofluorescent image of prostate cancer stained with DAPI for nuclei (blue), CK18 for epithelial cells (green), and stromal and epithelial AR (dark pink). In B, the CK18 was removed and only the stromal and epithelial nuclear AR (dark pink) remained with DAPI-stained nuclei that were AR negative (blue). (C) After segmentation and classification, a composite image was generated that highlighted the nuclear AR (light pink) only in epithelial cells, which was then quantified based on the pixel intensity (brightness) of the AR.



concordance index of 0.70, sensitivity of 72%, and specificity of 72% on the same validation cohort, further demonstrating the value of systems pathology.

Figure 3C illustrates the Kaplan-Meier estimates of recurrence for the validation cohort using the MSKCC nomogram for estimating a 5-year PSA recurrence progression-free probability. An optimal cut point to separate high- and low-risk patients was identified by maximizing the χ^2 values in the log-rank test. Of note, although the nomogram cut point was optimized in the validation cohort and the SVRc risk stratification was optimized in the training cohort and then applied to the validation cohort, the Kaplan-Meier curve for the SVRc model in Figure 3B continues to illustrate a better stratification of low- and high-risk patients when compared with that in Figure 3C.

Quantitative immunofluorescence. Because of the relative subjectivity of scoring proteins in IHC analyses, we next aimed to develop a platform for quantitative assessment of multiple proteins using immunofluorescence combined with image analysis. To develop the assay, we analyzed multiple prostate tissue samples with DAPI to label nuclei, along with 2 fluorochrome-labeled antibodies: anti-cytokeratin 18 (CK18) to mark epithelial cells, and anti-AR. Based on the DAPI and CK18 staining, the Histology Labeling Tool software (custom designed by Aureon Laboratories Inc.) delineated specific tissue features (e.g., stromal versus glandular epithelium). Appropriate algorithms were then formulated for quantifying AR within these regions. The algorithms measure the mean, median, maximum, and standard deviation of AR (protein) intensity and distribution in epithelial nuclei and stromal nuclei. (See Supplemental Methods and Supplemental Table 7 for details regarding script development and a list of the 18 features quantified in the immunofluorescent image analysis scripts.) Of note, AR intensity represents the concentration of the antigen and is developed using a continuous pixel scale, as opposed to a 0–255 red/green/blue-limited (RGB-limited) value, thus allowing for an expanded dynamic range. The output from the scripts (i.e., the segmented images) was visually inspected by 2 pathologists for overall accuracy and performance. Figure 4A is a standard multi fluorochrome image that was deconstructed to produce 3 grayscale images (Figure 4, B–D), which represent discrete morphologic (i.e., DAPI, CK18) and AR maps. Figure 4E is a segmented image representing the nuclear AR within epithelial cells quantified by the script.

AR was then quantified in prostate samples from 59 patients in the validation cohort. We acquired 177 fluorescent images, representing each of the triplicate multiplex TMA cores, and grayscale images were segmented and classified to define AR present only within the DAPI-stained nuclei of epithelial cells. Figure 5 shows an example of a TMA core; Figure 5C is a color map illustrating AR within epithelial nuclei only. Among the 18 measurements for AR, one feature, the amount of AR in epithelial cell nuclei (including tumor and nontumor elements) relative to the total DAPI area of all epithelial nuclei was highly concordant with the IHC staining index for AR in tumor (Spearman rank correlation coefficient, 0.44; $P = 0.0011$). This feature was also independently associated with PSA recurrence when analyzed univariately. Therefore, the quantitative immunofluorescent data support the AR IHC biomarker results and provide a more objective means of assigning a value to AR expression in cellular compartments.

Extended analysis for predicting PSA recurrence using systems pathology. We conducted a second study using an extended cohort of 682 patients treated at MSKCC and comparable clinicopathologic fea-

tures as in the first study summarized above (see Supplemental Table 3). 101 patients (14.8%) experienced PSA recurrence, defined as 2 consecutive PSA measurements >0.2 ng/ml. Median follow-up of patients with no recurrence observed was 74 months. The overall median time to recurrence was not reached. The model was trained on 342 patients with a concordance index of 0.85 and validated on 340 patients with a concordance index of 0.77, sensitivity of 77% and a specificity of 72%. The Kaplan-Meier estimates of recurrence for records with low and high SVRc model score showed a statistically significant difference in time to PSA recurrence ($P < 0.0001$, log-rank test; see Supplemental Figure 1). For a complete list of selected features including order of importance and weight within the model see Supplemental Table 9.

Discussion

Prostate cancer is the second leading cause of cancer death among men in the United States, with an anticipated 234,000 newly diagnosed cases and nearly 30,000 deaths in 2006 (39). The majority of men with early-stage disease are cured with local therapy; however, approximately 15%–40% (dependent upon study cohort) will develop PSA recurrence (40). Furthermore, tumor progression for patients with prostate cancer is a slow process, the mean time from PSA recurrence to metastasis is 8 years, with a median of 5 years (41). The majority of tumors are indolent and require minimal intervention, but others are more aggressive and may be best treated early (i.e., by surgery, radiation therapy, hormonal therapy, systemic therapy, or clinical trial placement). These observations suggest that overall survival may depend on early identification of high-risk patients by predicting both the patients' time to PSA recurrence and their propensity to develop metastases. The current prognostic tools (i.e., Kattan nomogram [ref. 7], Partin Tables [ref. 41]) rely solely on clinical and pathologic variables. While they provide useful predictions of clinical states and outcomes, they need improvement in both accuracy and universality. The need for further refinement of risk stratification, especially for men, for whom the nomogram is indeterminate, has recently been acknowledged (8).

We utilized systems pathology to identify those patients who were at risk for PSA recurrence following a prostatectomy. Using this approach, we generated an integrated view of the disease, including quantitative assessment of cellular and microanatomic characteristics, molecular markers, and clinical variables to create an integrative and highly accurate prediction. The concordance index of the SVRc model, developed with 262 patients, was 0.82 when applied to the validation cohort; by comparison, when the Kattan nomogram (7), developed with 996 patients, was applied to this validation cohort, its concordance index was only 0.71. The SVRc algorithm is designed for training with many features and fewer events.

Applying our original feature selection technique, we have found a set of clinical variables, molecular biomarkers, and tissue morphometric features sufficient to create a clinically useful predictive test. The optimized model that was built during training was validated using an external cohort with a concordance index of 0.82, sensitivity of 96%, and specificity of 72%. We have broadened these analyses by developing an expanded SVRc model with 682 patients employing tumor-selected image analysis and quantitative immunofluorescence of the AR. The extended model validated with a concordance index of 0.77, sensitivity of 77%, and specificity of 72%. Both models exhibited comparable features with the addition of several clinical variables, including primary prostatectomy Gleason grade, preoperative PSA, and extracapsular extension



selected in the extended study. Of note, the sensitivity for both models was developed to identify recurrence within 5 years of the prostatectomy in order to influence potential therapeutic options, including PSA monitoring and use of early adjuvant therapy. Furthermore, AR expression data from the original model were independently supported using a novel quantitative approach for evaluating tumor antigens in tissue samples. By employing fluorescently labeled antibodies and spectral imaging coupled with image analysis, we successfully developed a quantitative measurement of the presence of AR in specific cellular compartments within prostate tissue and then evaluated the association of AR with PSA recurrence in an external cohort. We observed that elevated IHC AR staining in PIN was also associated with PSA recurrence in a univariate analysis (data not shown). Because PIN is considered to be a precursor of prostate cancer, this result suggests an even broader role for AR in prostate cancer. We also observed through quantitative immunofluorescence that the total AR content present within both tumor and PIN was independently associated with recurrence, supporting the IHC data and expanding on a role for AR in promoting tumor development. The ability to assess AR as a continuous variable (as opposed to more traditional nominal methods such as IHC assessment) allowed us to more accurately stratify patients with respect to their individual risk. Taken together, these results support quantitation of AR as part of the standard pathology evaluation of prostate cancer.

Although AR has been implicated in the progression of prostate cancer from androgen dependence to an androgen independence, its exact role in tumor recurrence has not been fully elucidated. Li et al. (42) and recently Inoue et al. (43) have demonstrated that high levels of AR protein were associated with treatment failure, and in both studies a high proliferative (Ki67) index was implicated in prostate cancer progression. Our findings of a trend toward association of Ki67 and CD34 with recurrence suggest that AR, either directly or indirectly, may mediate prostate cancer progression through mechanisms of proliferation (Ki67) and possibly angiogenesis (CD34).

We have successfully developed a systems pathology platform that integrates clinical features, tumor tissue morphometrics, and molecular analyses. Using SVRc, features were selected from the 3 domains and used to develop a predictive model for PSA recurrence. We believe this novel systems pathology approach has broad application in the field of personalized medicine, as it relates to tumor diagnostics, patient prognostication, and eventually to predicting response to specific therapeutics (44). Additional models are in development for predicting the probability of prostate cancer progression to bone metastasis and/or PSA rise following androgen deprivation therapy, either at the time of diagnostic needle biopsy or after radical prostatectomy.

Methods

Evaluation of recurrence. Between 1985 and 2003, patients with clinically-localized prostate cancer underwent radical prostatectomy without neoadjuvant therapy at MSKCC and/or by a single surgeon at Baylor College of Medicine, Houston, Texas, USA. Patients were included in the study only after informed consent and Institutional Review Board approval at both institutions. Time to recurrence was defined as the time (in months) from radical prostatectomy until the first of 2 consecutive PSA measurements each greater than 0.2 ng/ml. If a patient did not have recurrence as of his last visit, or if the patient outcome was unknown as of his most recent visit, then the outcome was considered censored.

Patient tissues. TMAs were constructed from paraffin blocks, which after review of the H&E-stained slide contained sufficient tumor cells and were representative of the prostatectomy specimen with respect to the reported Gleason grade and score. Three tissue cores with a diameter of 0.6 mm were taken from each specimen and randomly arrayed in the recipient paraffin block (Beecher Instruments). Sections (5 μ m) of the TMA block were placed on charged polylysine-coated slides and used for morphometric (H&E staining) and IHC analyses (see below). When tumor content was less than 80% in the available TMA cores for a given patient in the training set, the H&E whole sections from the patient's prostatectomy specimen were utilized for image analysis. The amount of tumor content was utilized as a filter to ensure optimal feature extraction.

Image analysis and morphometry studies. From the H&E-stained slides, representative areas of the prostate tissue from each patient, either from a single tissue core or a whole section, were imaged, digitized, and analyzed. Images were captured with an Olympus bright-field microscope at $\times 20$ magnification using a SPOT Insight QE Color Digital Camera (KAI2000; Diagnostic Instruments Inc.). Quantitative analyses of H&E and immunofluorescent images (see "Quantitative multiplex immunofluorescence" section below) was performed using a Histology Labeling Tool (software version 1.19; custom designed by Aureon Laboratories Inc.).

Image objects were classified into histopathological classes according to their spectra (e.g., color, channel values), generic shape (e.g., area, length), and spatial relationship properties, from which statistics were generated. For complete definitions of specific image object categories and the associated statistical measurements, see "Description of morphometric features" in Supplemental Methods and Supplemental Table 3, A and B, which summarize the bioimaging features.

IHC analysis. In TMA blocks from the training set, we analyzed a panel of 12 biomarkers: cytokeratin 18 (CK18; luminal cell marker), CK14 (basal cell marker), CD45 (lymphocyte marker), CD34 (endothelial cell marker), CD68 (macrophage marker), Ki67 (marker of proliferation), PSA (kallikrein), prostate-specific membrane antigen (PSMA; carboxypeptidase), cyclin D1 (cell cycle-related protein), p27^{Kip1} (cell cycle-related protein), AR (endocrine signaling protein), and Her-2/Neu (signaling protein). See Supplemental Table 8 for a list of the antibodies used. In samples from the validation cohort, only AR was analyzed. All markers were assessed with standard chromogenic IHC. For a complete description of IHC methods see Supplemental Methods.

Quantitative multiplex immunofluorescence. Multiplex immunofluorescence assay was performed utilizing CK18 combined with AR and DAPI (as a nuclear counter-stain) on a single TMA slide from the validation cohort. In brief, an antibody solution containing both anti-CK18 (1:7,000; Calbiochem) and anti-AR (1:5; LabVision) in 1% blocking reagent was applied for 60 minutes to a deparaffinized TMA tissue section containing 186 cores from 61 patients. After appropriate rinses, the slide was incubated with a mixture of Zenon Alexa fluor 488-conjugated anti-rabbit for CK18 and Alexa fluor 568-conjugated anti-mouse IgG1 for AR for 30 minutes, rinsed with phosphate-buffered saline, and subsequently imaged. For a complete description, including the method used to evaluate 5 antigens (CK18, AR, high molecular weight cytokeratin, p63, and alpha-methylacyl CoA racemase (AMACR)) on a single section of prostate tissue, see Supplemental Methods. The 5-antigen (quint-plex) multiplex assay was utilized in the extended cohort study.

Statistics. A version of the well known and highly regarded SVR (16, 17) machine learning algorithm was used. Our SVRc is an adaptation of traditional SVR that accommodates censored data (45). SVRc was developed in order to take advantage of the high-dimensional capability of SVR while adapting it for use with right-censored data (conditions prevalent in systems pathology). To accomplish this, a modified loss/penalty func-



tion was defined within the support vector machine that allows right and noncensored data to be processed. Our experience with SVRc has demonstrated that this approach can increase a model's predictive accuracy over that of the Cox model.

In conjunction with SVRc, we employed feature selection algorithms to identify the most important and predictive features in a prognostic model. In our original study, we employed "feature reduction," a feature selection algorithm developed specifically for SVRc. This approach relies upon the intrinsic capability of the linear kernel used in the SVR to determine the weight of individual features within a model, analogous to linear least-squares regression, allowing for the algorithm to compare the contributions of different features by a process that is computationally tractable and efficient. The contribution of a feature is evaluated as the product of its weight and variance. In the SVRc feature reduction algorithm, an initial SVRc model is constructed using all of the features in a training cohort. This model is then tested on the training cohort, and a fitness criterion (described below) is assessed. All the features in the model are ranked in order of the absolute value of their contribution (negative contributions imply negative correlation with time to recurrence). The feature with the lowest contribution to the model is dropped and a new model is constructed on the remaining features. This procedure is repeated until no more features are left. At this point, the model with the highest fitness is selected. In the case of multiple models with equal values of the fitness criterion, the model with the fewest features is selected.

The criterion used to assess fitness of each intermediate model during the feature reduction process was a combination of 3 evaluation metrics: the concordance index (46), sensitivity, and specificity. The concordance index (which is similar to the AUC [ref. 47]) estimates the probability that, of a pair of randomly chosen comparable patients, the patient with the higher predicted time to PSA recurrence from the model will experience recurrence within a shorter time than the other patient. The concordance index is based on pair-wise comparisons between 2 randomly selected patients who meet either of the following criteria: (a) both patients experienced the event and the event time of one patient is shorter than that of the other patient or (b)

only one patient experienced the event and his event time is shorter than the other patient's follow-up time. A concordance index of 0.5 would indicate that the model performs the same as a coin toss, while 1.0 would mean that the model has perfect ability to discriminate. To estimate sensitivity and specificity, typically evaluated for binary output, we first selected a clinically meaningful limit for time (PSA recurrence within 5 years) to separate early from late events. Patients whose outcome was censored before 5 years were excluded from this estimation. Thereafter every value of the output of the model, scaled between 0 and 100, was taken one after another as a potential cut point of the prediction. For each of these potential cut points, we evaluated the sensitivity and specificity of the classification. Sensitivity was defined as the percentage of patients who experienced PSA recurrence within 5 years that were correctly predicted; specificity was defined as the percentage of patients who did not experience a PSA recurrence within 5 years that were correctly predicted. Every cut point was evaluated by the product of sensitivity and specificity calculated for that cut point. The cut point with the highest value of the product was selected as the predictive cut point, and its sensitivity and specificity were considered to be the sensitivity and specificity of the model. Additional statistical methods used for evaluation of the extended cohort are described in Supplemental Methods (see "Analytical and statistical results").

Acknowledgments

We thank J. Edelson for assistance in developing the manuscript. We also thank Heidi Emerich and Nicole Roberts for technical assistance.

Received for publication January 3, 2007, and accepted in revised form April 9, 2007.

Address correspondence to: Michael J. Donovan, Aureon Laboratories Inc., 28 Wells Avenue, Yonkers, New York 10701, USA. Phone: (914) 377-4037; Fax: (914) 377-4001; E-mail: Michael.Donovan@aureon.com.

- Graefen, M., et al. 2002. A validation of two preoperative nomograms predicting recurrence following radical prostatectomy in a cohort of European men. *Urol. Oncol.* **7**:141-146.
- Graefen, M., et al. 2002. Validation study of the accuracy of a postoperative nomogram for recurrence after radical prostatectomy for localized prostate cancer. *J. Clin. Oncol.* **20**:951-956.
- Graefen, M., et al. 2002. International validation of a preoperative nomogram for prostate cancer recurrence after radical prostatectomy. *J. Clin. Oncol.* **20**:3206-3212.
- Kattan, M.W., Eastham, J.A., Stapleton, A.M., Wheeler, T.M., and Scardino, P.T. 1998. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *J. Natl. Cancer Inst.* **90**:766-771.
- Kattan, M.W., Wheeler, T.M., and Scardino, P.T. 1999. Postoperative nomogram for disease recurrence after radical prostatectomy for prostate cancer. *J. Clin. Oncol.* **17**:1499-1507.
- Smoletz, O., et al. 2002. Nomogram for overall survival of patients with progressive metastatic prostate cancer after castration. *J. Clin. Oncol.* **20**:3972-3982.
- Stephenson, A.J., et al. 2005. Postoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. *J. Clin. Oncol.* **23**:7005-7012.
- Stephenson, A.J., et al. 2005. Integration of gene expression profiling and clinical variables to predict prostate carcinoma recurrence after radical prostatectomy. *Cancer*. **104**:290-298.
- Hood, L. 2003. Systems biology: integrating technology, biology, and computation. *Mech. Ageing Dev.* **124**:9-16.
- Davidov, E., Holland, J., Marple, E., and Naylor, S. 2003. Advancing drug discovery through systems biology. *Drug Discov. Today*. **8**:175-183.
- Ideker, T., Galitski, T., and Hood, L. 2001. A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* **2**:343-372.
- Ye, Q.H., et al. 2003. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. *Nat. Med.* **9**:416-423.
- Su, A.I., et al. 2001. Molecular classification of human carcinomas by use of gene expression signatures. *Cancer Res.* **61**:7388-7393.
- Yeang, C.H., et al. 2001. Molecular classification of multiple tumor types. *Bioinformatics*. **17**(Suppl. 1):S316-S322.
- Brown, M.P., et al. 2000. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc. Natl. Acad. Sci. U. S. A.* **97**:262-267.
- Hastie, T., Tibshirani, R., and Friedman, J. 2001. *The elements of statistical learning: data mining, inference, and prediction*. Springer. New York, New York, USA. 552 pp.
- Cristianini, N., and Shawe-Taylor, J. 2000. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press. Cambridge, United Kingdom. 189 pp.
- Dhanasekaran, S.M., et al. 2001. Delineation of prognostic biomarkers in prostate cancer. *Nature*. **412**:822-826.
- Luo, J., et al. 2001. Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. *Cancer Res.* **61**:4683-4688.
- Rhodes, D.R., Barrette, T.R., Rubin, M.A., Ghosh, D., and Chinnaiyan, A.M. 2002. Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Res.* **62**:4427-4433.
- LaTulippe, E., et al. 2002. Comprehensive gene expression analysis of prostate cancer reveals distinct transcriptional programs associated with metastatic disease. *Cancer Res.* **62**:4499-4506.
- Ramaswamy, S., Ross, K.N., Lander, E.S., and Golub, T.R. 2003. A molecular signature of metastasis in primary solid tumors. *Nat. Genet.* **33**:49-54.
- Luo, J.H., et al. 2002. Gene expression analysis of prostate cancers. *Mol. Carcinog.* **33**:25-35.
- Welsh, J.B., et al. 2001. Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res.* **61**:5974-5978.
- Singh, D., et al. 2002. Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell*. **1**:203-209.
- Swindle, P., et al. 2005. Do margins matter? The prognostic significance of positive surgical margins in radical prostatectomy specimens. *J. Urol.* **174**:903-907.
- Veltri, R.W., Partin, A.W., and Miller, M.C. 2000. Quantitative nuclear grade (QNG): a new image analysis-based biomarker of clinically relevant nuclear structure alterations. *J. Cell. Biochem. Suppl.* **35**:151-157.



28. Hurwitz, M.D., DeWeese, T.L., Zinreich, E.S., Epstein, J.I., and Partin, A.W. 1999. Nuclear morphometry predicts disease-free interval for clinically localized adenocarcinoma of the prostate treated with definitive radiation therapy. *Int. J. Cancer*. **84**:594–597.
29. Olinici, C.D., Crisan, D., Olinici, C.I., and Vaida, M. 1997. Computer-based image analysis of nucleoli in prostate carcinoma. *Rom. J. Morphol. Embryol.* **43**:163–167.
30. Veltri, R.W., Miller, M.C., Partin, A.W., Coffey, D.S., and Epstein, J.I. 1996. Ability to predict biochemical progression using Gleason score and a computer-generated quantitative nuclear grade derived from cancer cell nuclei. *Urology*. **48**:685–691.
31. Veltri, R.W., et al. 1994. Quantitative nuclear morphometry, Markovian texture descriptors, and DNA content captured on a CAS-200 Image analysis system, combined with PCNA and HER-2/neu immunohistochemistry for prediction of prostate cancer progression. *J. Cell. Biochem. Suppl.* **19**:249–258.
32. Partin, A.W., et al. 1992. Use of nuclear morphometry, gleason histologic scoring, clinical stage, and age to predict disease-free survival among patients with prostate cancer. *Cancer*. **70**:161–168.
33. Wang, N., Stenkivist, B.G., and Tribukait, B. 1992. Morphometry of nuclei of the normal and malignant prostate in relation to DNA ploidy. *Anal. Quant. Cytol. Histol.* **14**:210–216.
34. Stephenson, R.A., Zahniser, D.J., Wong, K.L., and Hutchinson, M.L. 1991. An image analysis method for assessment of prognostic risk in prostate cancer: a pilot study. *Anal. Cell. Pathol.* **3**:243–248.
35. Eskelinen, M., Lipponen, P., Majapuro, R., and Syrjanen, K. 1991. Prognostic factors in prostatic adenocarcinoma assessed by means of quantitative histology. *Eur. Urol.* **19**:274–278.
36. Robutti, F., Pilato, F.P., and Betta, P.G. 1989. A new method of grading malignancy of prostate carcinoma using quantitative microscopic nuclear features. *Pathol. Res. Pract.* **185**:701–703.
37. Mohler, J.L., Partin, A.W., Lohr, W.D., and Coffey, D.S. 1988. Nuclear roundness factor measurement for assessment of prognosis of patients with prostatic carcinoma. I. Testing of a digitization system. *J. Urol.* **139**:1080–1084.
38. Freiha, F.S., McNeal, J.E., and Stamey, T.A. 1988. Selection criteria for radical prostatectomy based on morphometric studies in prostate carcinoma. *NCI Monogr.* **7**:107–108.
39. Jemal, A., et al. 2006. Cancer statistics, 2006. *CA Cancer J. Clin.* **56**:106–130.
40. Freedland, S.J., et al. 2005. Risk of prostate cancer specific mortality following biochemical recurrence after radical prostatectomy. *JAMA*. **294**:433–439.
41. Pound, C.R., et al. 1999. Natural history of progression after PSA elevation following radical prostatectomy. *JAMA*. **281**:1591–1597.
42. Li, R., et al. 2004. High level of androgen receptor is associated with aggressive clinicopathologic features and decreased biochemical recurrence-free survival in prostate: cancer patients treated with radical prostatectomy. *Am. J. Surg. Pathol.* **28**:928–934.
43. Inoue, T., et al. 2005. Androgen receptor, Ki67, and p53 expression in radical prostatectomy specimens predict treatment failure in Japanese population. *Urology*. **66**:332–337.
44. Saidi, O., Cordon-Cardo, C., and Costa, J. 2007. Technology insight: will systems pathology replace the pathologist? *Nat. Clin. Pract. Urol.* **4**:39–45.
45. Cordon-Cardo, C., et al. 2004. Improved prediction of PSA recurrence through systems pathology. *J. Clin. Oncol.* **22**(Suppl.):4591.
46. Harrell, F.E., Jr., Califf, R.M., Pryor, D.B., Lee, K.L., and Rosati, R.A. 1982. Evaluating the yield of medical tests. *JAMA*. **247**:2543–2546.
47. Stephenson, A.J., et al. 2006. Preoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. *J. Natl. Cancer Inst.* **98**:715–717.