Validation of a Genomic Classifier that Predicts Metastasis Following Radical Prostatectomy in an At Risk Patient Population

R. Jeffrey Karnes,* Eric J. Bergstralh, Elai Davicioni,† Mercedeh Ghadessi,† Christine Buerkı,† Anirban P. Mitra, Anamaria Crisan,† Nicholas Erho,† Ismael A. Vergara,† Lucia L. Lam,† Rachel Carlson, Darby J. S. Thompson, Zaid Haddad,† Benedikt Zimmermann,† Thomas Sierocinski,† Timothy J. Triche,† Thomas Kollmeyer, Karla V. Ballman,§ Peter C. Black,† George G. Klee and Robert B. Jenkins

From the Departments of Urology (RJK), Health Sciences Research (EJB, RC, KVJ) and Laboratory Medicine and Pathology (TK, GGK, RBJ), Mayo Clinic, Rochester, Minnesota, GenomeDx Biosciences, Inc. (ED, MG, CB, AC, NE, IAV, LLL, ZH, BZ, TS, TJT) and Department of Urologic Sciences, University of British Columbia (PCB), Vancouver and EMMES Canada (DJST), Burnaby, British Columbia, Canada, and Department of Pathology, University of Southern California (APM, TJT), Los Angeles, California

Purpose: Patients with locally advanced prostate cancer after radical prostatectomy are candidates for secondary therapy. However, this higher risk population is heterogeneous. Many cases do not metastasize even when conservatively managed. Given the limited specificity of pathological features to predict metastasis, newer risk prediction models are needed. We report a validation study of a genomic classifier that predicts metastasis after radical prostatectomy in a high risk population.

Materials and Methods: A case-cohort design was used to sample 1,010 patients after radical prostatectomy at high risk for recurrence who were treated from 2000 to 2006. Patients had preoperative prostate specific antigen greater than 20 ng/ml, Gleason 8 or greater, pT3b or a Mayo Clinic nomogram score of 10 or greater. Patients with metastasis at diagnosis or any prior treatment for prostate cancer were excluded from analysis. A 20% random sampling created a subcohort that included all patients with metastasis. We generated 22-marker genomic classifier scores for 219 patients with available genomic data. ROC and decision curves, competing risk and weighted regression models were used to assess genomic classifier performance.

Results: The genomic classifier AUC was 0.79 for predicting 5-year metastasis after radical prostatectomy. Decision curves showed that the genomic classifier net benefit exceeded that of clinical only models. The genomic classifier was the predominant predictor of metastasis on multivariable analysis. The cumulative incidence of metastasis 5 years after radical prostatectomy was 2.4%, 6.0%

Abbreviations and Acronyms
AUC = area under ROC curve
BCR = biochemical recurrence
CC = clinical only classifier
ECE = extracapsular extension
GC = genomic classifier
GPSM = Gleason score, preoperative PSA, SVI, SM
GS = Gleason score
MVA = multivariable analysis
ncRNA = noncoding RNA
N+ = lymph node involvement
PSA = prostate specific antigen
RP = radical prostatectomy
SM+ = positive surgical margin
SVI = seminal vesicle invasion

Accepted for publication June 5, 2013.
Study received Mayo Clinic institutional review board approval.
Supported by Mayo Prostate Cancer SPORE Grant NIH P50 CA81956, the Richard M. Schulze Family Foundation, National Research Council of Canada Industrial Research Assistance Program and GenomeDx Biosciences, Inc.
* Correspondence: Department of Urology, Mayo Clinic, 200 First St. Southwest, Rochester, Minnesota 55905; (e-mail: karnes.r@mayo.edu).
† Financial interest and/or other relationship with GenomeDx Biosciences.
‡ Financial interest and/or other relationship with GenomeDx Biosciences.
§ Current address: Center for Personalized Medicine, Department of Pathology and Laboratory Medicine, Children’s Hospital Los Angeles, Department of Pathology and Pediatrics, Keck School of Medicine at University of Southern California, Los Angeles, California.

http://dx.doi.org/10.1016/j.juro.2013.06.017
Vol. 190, 2047-2053, December 2013
© 2013 by AMERICAN UROLOGICAL ASSOCIATION EDUCATION AND RESEARCH, INC.
Printed in U.S.A.
Prostate cancer is the most common cancer in men. The global burden is expected to increase to 1.7 million new cases and 499,000 deaths annually by 2030. In the approximately 2.5 million men with prostate cancer in the United States the 5-year relative survival of locoregional disease, representing about 96% of diagnoses, is almost 100%. However, this overall favorable statistic hides the fact that 29,000 men died of prostate cancer in 2012. Most men who die of prostate cancer have localized disease on initial biopsy but after RP tumors are found with 1 or more adverse pathological features. These men at high risk often had increasing PSA or BCR. After RP bone metastasis developed in some men at high risk and/or they died of disease, typically within 10 years of diagnosis. Unfortunately, in those with high risk prostate cancer the mortality rate has not effectively improved in the last 20 years.

These findings highlight the need for the identification at diagnosis of truly aggressive tumors with inherently greater potential for early metastasis. These men have the most to gain from early secondary therapies and clinical trials. Currently, adverse pathological features, such as Gleason pattern 4 or greater, SM+, SVI, ECE and preoperative PSA 20 ng/ml or greater, are used to identify men at high risk. These men are candidates for secondary therapy such as postoperative radiation but few will have metastasis and die of prostate cancer even when treated expectantly. Thus, clinicians are reluctant to recommend postoperative radiation despite evidence showing its efficacy in men at high risk. Metastatic progression can be delayed, if not prevented, by early secondary intervention. However, clinicians perceive that the sum of the costs and morbidity of secondary therapy will exceed its benefits if applied to all patients with adverse pathological features.

We previously reported the development and validation of a GC (Decipher\textsuperscript{TM}) that predicts early metastasis after RP using an oligonucleotide microarray to profile RNA from formalin fixed, paraffin embedded RP specimens. Using a retrospective, nested, case-control design with a median followup of 16.9 years, we found unique patterns of differential expression in 192 early metastasis cases, ie within 5 years of increasing PSA, in comparison to 271 controls irrespective of BCR. We applied these expression patterns to develop and validate a 22-marker GC to predict early clinical metastasis. The GC was a more specific predictor of aggressive disease than clinical variables or previously reported gene signatures. Thus, we report a blinded study of a prospectively designed cohort to evaluate GC for predicting clinical metastasis in a contemporary, high risk population of patients treated with RP.

**MATERIALS AND METHODS**

**Study Design**

Patients treated with RP between 2000 and 2006 were identified from the tumor registry at our institution for a case-cohort study design. This involved identifying all patients with metastasis as well as a group representative of the full cohort (supplementary Appendix, http://jurology.com/). Thus, men at high risk for recurrence after open/robotic RP with no prior neoadjuvant/prostate cancer treatment were selected based on preoperative PSA greater than 20 ng/ml, pathological GS 8 or greater, SVI or a Mayo Clinic nomogram score of 10 or greater. The cohort of 1,010 men included 73 with metastasis, as evidenced by computerized tomography or bone scan. A 20% random sample of 202 men was drawn from the cohort, including 19 of 73 with metastasis. To increase sampling of metastasis the remaining 54 men with metastasis were added, resulting in a final study set of 256 (supplementary table 1 and figure 1, http://jurology.com/). Patients who did not experience metastasis regardless of BCR, defined as followup PSA 0.4 ng/ml or greater 30 days after RP, were censored at last followup. The study was approved by the Mayo Clinic institutional review board.

**Tissue Processing and Prognostic Classifier Application**

After histopathological re-review the dominant Gleason lesion (highest grade) was macrodissected in 238 available patient samples for microarray analysis. GC score was calculated for the 219 samples that passed quality control based on the predefined 22-marker classifier. All study participants were blinded to outcomes and clinical data except the statisticians at our institution who selected the study population. Previously described CC, combined genomic-clinical classifiers and scores of 2 validated prediction models (GPSM and the Stephenson nomogram) were evaluated. GC was also assessed on 384 additional post-RP cases from 3 independent data sets (supplementary Appendix, http://jurology.com/).
Statistical Analysis
Discrimination was measured by the AUC for censored survival data (survival ROC). Net clinical benefit was estimated using extension of decision curve analysis for survival data. Univariable analysis and MVA Cox proportional hazards models were used for risk ratio estimation. Cumulative incidence curves were constructed using Fine-Gray competing risks analysis. Survival analysis was weighted to estimate cohort parameters (supplementary Appendix, http://jurology.com/).

RESULTS
We profiled RNA from formalin fixed, paraffin embedded primary prostate cancer specimens from patients treated with RP. After exclusion due to tissue unavailability and quality control the study consisted of 219 patients, including 69 with metastasis, with a median followup of 6.7 years (supplementary figure 1, http://jurology.com/). Rates of BCR 3 years (35%) and metastasis 5 years (6%) after RP were similar to those in the original cohort, indicating representative sampling (supplementary table 1, http://jurology.com/). Median patient age was 63 years (range 46 to 78). Of the tumors 93% were GS 7 or greater, 47% were pT3 and 56% were SM+. Interestingly, 55% of cases with adverse pathological features met the criteria for low-intermediate D’Amico risk groups before RP, suggesting that many were significantly upgraded and up-staged after RP. Median time to BCR and metastasis after RP was 1.2 and 3.1 years, respectively.

Survival ROC curves were used to assess classifier discrimination. The AUC for GC was 0.79 (95% CI 0.68–0.87), outperforming all clinical variables (AUC 0.49 to 0.65, fig. 1). Incorporating all clinical variables into GC marginally increased the AUC to 0.82 (95% CI 0.72–0.88, data not shown). Median GC scores were consistently higher in patients with metastasis compared to those without metastasis at last followup in this study and in the 3 independent data sets (supplementary figure 2, http://jurology.com/). Decision curves were used to compare the net clinical benefit (clinical implication of false-positive findings) of genomic based classifiers with that of clinical only models. Overall the higher net benefit of genomic based classifiers compared to clinical only risk models suggested that GC models had increased specificity (ie lower false-positive results) without sacrificing sensitivity (fig. 2).

On univariable analysis GC had the highest significant HR among the classifiers (1.58 for each 10% score increase). Of clinical variables ECE, GS and SVI were significant prognostic factors (supplementary table 2, http://jurology.com/). Node positivity, preoperative PSA and SM+ were not significant. Patients receiving adjuvant hormone therapy had an increased HR for metastasis.

Table 1. Clinical characteristics of study patients with GC score

<table>
<thead>
<tr>
<th>No. Pts (%)</th>
<th>No. with Metastasis (%)</th>
<th>No. without Metastasis (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>219 (100)</td>
<td>69 (32)</td>
</tr>
<tr>
<td>Preop PSA (ng/ml):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 10</td>
<td>119 (54)</td>
<td>33 (28)</td>
</tr>
<tr>
<td>10–20</td>
<td>59 (27)</td>
<td>20 (34)</td>
</tr>
<tr>
<td>Greater than 20</td>
<td>41 (19)</td>
<td>16 (39)</td>
</tr>
<tr>
<td>Preop D’Amico risk group:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>38 (17)</td>
<td>6 (16)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>83 (38)</td>
<td>23 (28)</td>
</tr>
<tr>
<td>High</td>
<td>98 (45)</td>
<td>40 (41)</td>
</tr>
<tr>
<td>Pathological GS:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 or Less</td>
<td>15 (7)</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>111 (51)</td>
<td>29 (26)</td>
</tr>
<tr>
<td>8 or Greater</td>
<td>93 (42)</td>
<td>40 (43)</td>
</tr>
<tr>
<td>Pathological stage:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pT2N0M0</td>
<td>85 (39)</td>
<td>14 (17)</td>
</tr>
<tr>
<td>pT3aN0M0</td>
<td>102 (47)</td>
<td>40 (39)</td>
</tr>
<tr>
<td>pT3bN+M0</td>
<td>32 (15)</td>
<td>15 (47)</td>
</tr>
<tr>
<td>SM+</td>
<td>123 (56)</td>
<td>39 (32)</td>
</tr>
<tr>
<td>ECE</td>
<td>95 (43)</td>
<td>43 (45)</td>
</tr>
<tr>
<td>SVI</td>
<td>81 (37)</td>
<td>36 (44)</td>
</tr>
<tr>
<td>Postop adjuvant treatment:*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiation</td>
<td>24 (11)</td>
<td>9 (38)</td>
</tr>
<tr>
<td>Hormone</td>
<td>74 (34)</td>
<td>32 (46)</td>
</tr>
<tr>
<td>Postop salvage treatment:t</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiation</td>
<td>88 (31)</td>
<td>34 (50)</td>
</tr>
<tr>
<td>Hormone</td>
<td>86 (39)</td>
<td>62 (72)</td>
</tr>
<tr>
<td>Biochemical recurrence event</td>
<td>110 (50)</td>
<td>69 (63)</td>
</tr>
<tr>
<td>Prostate Ca specific mortality event</td>
<td>28 (13)</td>
<td>26 (100)</td>
</tr>
</tbody>
</table>

* Administered within 90 days postoperatively.
† Administered any time after 90 days postoperatively.

Figure 1. Cumulative survival ROC curves comparing GC score and individual clinicopathological factors for predicting clinical metastasis 5 years after RP. GC showed noticeably higher discrimination than individual clinicopathological factors. Path, pathological. Preop, preoperative.
(1.97, p = 0.02), likely because they were perceived to be most at risk.

On MVA only GC retained a significant HR when adjusted for clinical variables and postoperative adjuvant therapy (p < 0.001, table 2). GS was alternatively parameterized, eg 3 + 4, 4 + 3, 8 and 9-10, but this did not change the significance of GC (supplementary table 3, http://jurology.com/). Three additional MVAs were performed to model GC with clinical only nomograms. Only the Stephenson nomogram retained a significant HR with GC as the dominant variable (p = 0.04, supplementary table 4, http://jurology.com/).

To investigate the HR magnitude by increments in GC score we evaluated the effect size of each 10% increase for predicting metastasis after adjusting for postoperative treatment (supplementary Appendix and table 5, http://jurology.com/). We observed a general trend toward an increasing HR and a decreasing probability of metastasis-free survival with increasing deciles (Cochran-Armitage trend independence test p < 0.035). However, differences between deciles were not statistically significant, likely due to fewer patients in the higher GC score deciles. Score deciles were then incrementally collapsed to create 3 GC risk groups (GC less than 0.4, 0.4 to 0.6 and greater than 0.6), which showed significant differences in HR and metastasis-free survival compared to the reference group and to the prior level (table 3). Progression-free probability estimates and cumulative incidence plots revealed that 60% of patients had GC less than 0.4 (number at risk at time = 0) with only a 2.4% 5-year cumulative incidence of metastasis (fig. 3). In contrast, the 20% of patients with GC greater than 0.6 had a 22.5% 5-year cumulative incidence (p < 0.001). Similar results were obtained when excluding patients receiving adjuvant treatment (supplementary figure 3, http://jurology.com/).

Since GS was the strongest clinical predictor of metastasis, we compared the distribution of GC scores across GS groups. As expected, median GC increased with higher GS groups (table 3 and fig. 4). No patient with GS 6 or less had GC greater than 0.6 or clinical metastasis during study followup. Of patients with GS 7 tumors 41% had GC 0.4 or less, 86% of men with GS 7 tumors had low GC (less than 0.4) did not experience metastasis and only 3% died of the disease. As expected, 40% of patients with GS 8 or greater had a high GC score (greater than 0.6), of whom 62% experienced metastasis, we compared the distribution of GC scores across GS groups. As expected, median GC increased with higher GS groups (table 3 and fig. 4). No patient with GS 6 or less had GC greater than 0.6 or clinical metastasis during study followup. Of patients with GS 7 tumors 41% had GC 0.4 or less, 86% of men with GS 7 tumors had low GC (less than 0.4) did not experience metastasis and only 3% died of the disease. As expected, 40% of patients with GS 8 or greater had a high GC score (greater than 0.6), of whom 62% experienced metastasis and 41% died of the disease. However, 36% of patients with GS 8 or greater had a low GC score and most did not have metastasis (77%) or die of prostate cancer (85%). This reclassification was significant (McNemar chi-square p = 2.2 × 10^-16). Data suggest that GC may identify a subset of men with high risk GS 8 or greater tumors in whom clinical metastasis may never develop and, conversely, among patients with intermediate risk GS 7 tumors it may identify a subset enriched for clinical metastasis events. Such interpretations may need to be made in the context of additional clinical data.

Table 2. Multivariable Cox proportional hazards modeling of GC scores after adjusting for clinicopathological factors and adjuvant treatments

<table>
<thead>
<tr>
<th>Classifier</th>
<th>HR (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GC + Clinical variables</td>
<td>1.51 (1.29–1.76)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GC</td>
<td>1.59 (0.77–3.30)</td>
<td>0.21</td>
</tr>
<tr>
<td>Preop PSA</td>
<td>1.28 (0.85–1.93)</td>
<td>0.24</td>
</tr>
<tr>
<td>SVI</td>
<td>2.16 (0.90–5.16)</td>
<td>0.08</td>
</tr>
<tr>
<td>SM+</td>
<td>1.26 (0.58–2.75)</td>
<td>0.56</td>
</tr>
<tr>
<td>ECE</td>
<td>1.35 (0.60–3.00)</td>
<td>0.47</td>
</tr>
<tr>
<td>Lymph node involvement</td>
<td>0.73 (0.20–2.52)</td>
<td>0.63</td>
</tr>
<tr>
<td>Adjuvant treatment:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiation</td>
<td>0.65 (0.17–2.47)</td>
<td>0.52</td>
</tr>
<tr>
<td>Hormone</td>
<td>0.80 (0.31–2.03)</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Figure 2. Survival decision curve analysis comparing net benefit of GC genomic based classifiers and GC combined with clinical variables with clinical only models (CC, GPSM and Stephenson nomogram). Model performance was compared to extreme of classifying all patients at risk for clinical metastasis, thus, warranting treatment in all (dotted gray line) vs classifying no patients at risk and, thus, treating none (horizontal black dashed line). Decision to treat threshold (probability of metastasis used to trigger decision to treat) varied from 0 to 1 with sensitivity and specificity of each prediction model calculated at each threshold to determine net benefit. Optimal classifier had high net benefit above line indicating all patients treated. Net benefit of GC based models was superior at wide range of decision to treat thresholds. Stephenson 5-year, Stephenson nomogram derived 5-year survival probability.
DISCUSSION

This blinded, prospectively designed study independently validates a novel GC for predicting prostate cancer metastasis after RP. Results show that GC has improved performance over any individual clinicopathological variable or multivariable prediction model. In this contemporary (2000 to 2006) population of 1,010 men at risk for recurrence after RP the cumulative incidence of metastasis was only 6% at 5 years and 7.2% at 10 years. The GC score restratified this population, such that the low score group had an almost 3 times lower cumulative incidence of metastasis than the average and the high score group had an almost 4 times higher cumulative incidence of metastasis. Even after adjusting for postoperative therapy and clinical variables GC provided an independent prediction of metastasis. When assessed in 384 patients from independent data sets, GC retained its prognostic potential with higher scores in those in whom metastasis developed (supplementary Appendix, http://jurology.com/). Furthermore, clinical use studies indicated that GC can potentially modify case management, which was also evident at the individual level.

Stephenson et al previously developed and validated a nomogram to predict the probability of prostate cancer specific mortality using standard clinical parameters. Few of the patients in that study had a predicted 15-year metastasis incidence greater than 5%. They concluded that it is difficult to identify patients at substantially increased risk based on clinical parameters alone and there is an evident need for novel markers specifically associated with the biology of lethal prostate cancer. Drawing from advances in expression profiling and genomics, many signatures of aggressive disease have been proposed but to our knowledge none has demonstrated significant improvement for predicting aggressive prostate cancer over that of risk prediction models such as the Stephenson nomogram. This is probably because in most studies of patients treated with RP the end point was BCR, a nonspecific surrogate for...
metastatic and lethal disease. In fact, our investigations using transcriptome-wide expression profiling to compare patients with BCR (but not metastasis on long-term followup) and those with no evidence of disease in the absence of secondary treatment revealed few differences in RNA expression between the groups. This finding contrasts with patients with metastasis where primary tumor shows many thousands of differentially expressed RNAs compared to tumors in patients with no evidence of disease or with BCR only. Furthermore, most previous studies represented a broader prostate cancer population, which in the PSA screening era consisted of RP cases that were not necessarily representative of current cancer presentations.

In our previous series we used high density (approximately 1.4 million features) expression arrays to discover and validate primary tumor differential expression patterns associated with prostate cancer death and/or metastasis as end points. The transcriptome-wide approach allowed for the interrogation of a richer genomic data set, including thousands of differently expressed RNAs compared to tumors in patients with no evidence of disease or with BCR only. Furthermore, most previous studies represented a broader prostate cancer population, which in the PSA screening era consisted of RP cases that were not necessarily representative of current cancer presentations.

In this cohort almost 15% of patients had N+ disease and 45% received adjuvant therapy, the overall validation of GC in an adjuvant therapy-rich cohort may represent a limitation. While adjusting for secondary therapy on MVA showed that GC remained an independent, significant predictor of metastasis, we could not determine whether GC predicted a benefit from local (ie radiation) or systemic (eg hormone) therapy because patients were not randomized to these treatments. Evaluating GC in randomized clinical trial data sets would help us better understand the relationship between GC and the benefit from specific therapies. In turn, for future trials it can also lead to better selection of patients who are truly at high risk.

Experience in the last decade suggests that RP is used more frequently to manage more aggressive disease. Recent PSA screening guidelines may also likely result in a greater proportion of men presenting with more aggressive disease features. As a result, many contemporary patients with adverse pathological features may require additional post-RP intervention since long-term cancer control may not be achieved by surgery alone. Three large, randomized trials showed improved outcomes in patients with adverse pathological features when treated with adjuvant radiotherapy. Initial reports from the Radiation Therapy Oncology Group (RTOG) 96-01 trial indicated that intensification with multimodal therapy after RP decreased the incidence of metastasis. Despite this evidence it remains challenging to apply postoperative secondary therapy in individuals. Concerns for overtreatment causing toxicity and morbidity must be weighed against the potential harm of disease progression. The alternative is often observation with intervention only when PSA increases, which may prevent overtreatment but also delay treatment until the disease has already disseminated. A better predictor of metastasis by analysis of the primary tumor long before metastasis manifests clinically would enable

![Figure 4. GC scores of each pathological GS group, including patients with (red ovals) and without (blue ovals) clinical metastasis at study followup. Median GC score increased with GS but GC discriminated patients with clinical metastasis in all GS groups. Horizontal lines indicate GC risk groups (fig. 3).](image-url)
more tailored application of multimodal therapy and enhanced clinical trial design for high risk prostate cancer.

CONCLUSIONS
Genomic information significantly improved the prediction of metastatic disease compared to established clinicopathological risk factors in a contemporary cohort of men at high risk for recurrence after RP. Furthermore, most prognostic information for predicting metastatic disease was captured by genomic variables measured in the primary tumor. These data suggest that genomic alterations in aggressive prostate cancer manifest early, many years before metastatic disease is detected clinically. Using genomic information to characterize the true biological potential of a tumor to metastasize may ultimately lead to improved treatment of prostate cancer.

ACKNOWLEDGMENTS
Stephanie Fink assisted with experiments.

REFERENCES