Prostate Cancer

Combined Value of Validated Clinical and Genomic Risk Stratification Tools for Predicting Prostate Cancer Mortality in a High-risk Prostatectomy Cohort

Matthew R. Cooperberg a, n, Elai Davicconi b, n, Anamaria Crisan b, Robert B. Jenkins c, Mercedes Ghadessi b, R. Jeffrey Karnes d

a Departments of Urology and Epidemiology & Biostatistics, University of California, San Francisco, San Francisco, CA, USA; b GenomeDx Biosciences, Vancouver, British Columbia, Canada; c Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA; d Department of Urology, Mayo Clinic, Rochester, MN, USA

Abstract

Background: Risk prediction models that incorporate biomarkers and clinicopathologic variables may be used to improve decision making after radical prostatectomy (RP). We compared two previously validated post-RP classifiers—the Cancer of the Prostate Risk Assessment Postsurgical (CAPRA-S) and the Decipher genomic classifier (GC)—to predict prostate cancer–specific mortality (CSM) in a contemporary cohort of RP patients.

Objective: To evaluate the combined prognostic ability of CAPRA-S and GC to predict CSM.

Design, setting, and participants: A cohort of 1010 patients at high risk of recurrence after RP were treated at the Mayo Clinic between 2000 and 2006. High risk was defined by any of the following: preoperative prostate-specific antigen >20 ng/ml, pathologic Gleason score >8, or stage pT3b. A case-cohort random sample identified 225 patients (with cases defined as patients who experienced CSM), among whom CAPRA-S and GC could be determined for 185 patients.

Outcome measurements and statistical analysis: The scores were evaluated individually and in combination using concordance index (c-index), decision curve analysis, reclassification, cumulative incidence, and Cox regression for the prediction of CSM.

Results and limitations: Among 185 men, 28 experienced CSM. The c-indices for CAPRA-S and GC were 0.75 (95% confidence interval [CI], 0.55–0.84) and 0.78 (95% CI, 0.68–0.87), respectively. GC showed higher net benefit on decision curve analysis, but a score combining CAPRA-S and GC did not improve the area under the receiver-operating characteristic curve after optimism-adjusted bootstrapping. In 82 patients stratified to high risk based on CAPRA-S score ≥6, GC scores were likewise high risk for 33 patients, among whom 17 had CSM events. GC reclassified the remaining 49 men as low to intermediate risk; among these men, three CSM events were observed. In multivariable analysis, GC and CAPRA-S as continuous variables were independently prognostic of CSM, with hazard ratios (HRs) of 1.81 (p < 0.001 per 0.1-unit change in score) and 1.36 (p = 0.01 per 1-unit change in score). When categorized into risk groups, the multivariable HR for high CAPRA-S scores (≥6) was 2.36 (p = 0.04) and was 11.26 (p < 0.001) for high GC scores (≥6). For patients with both high GC and high CAPRA-S scores, the cumulative incidence of CSM was 45% at 10 yr. The study is limited by its retrospective design.

These authors made equal contributions.

* Corresponding author. Department of Urology, University of California, San Francisco, 1600 Divisadero St., Box 1695, San Francisco, CA 94143-1695, USA. Tel.: +1 415 515 3229.

E-mail address: mcooperberg@urology.ucsf.edu (M.R. Cooperberg).

http://dx.doi.org/10.1016/j.eururo.2014.05.039

0302-2838/© 2014 Published by Elsevier B.V. on behalf of European Association of Urology.
1. Introduction

Accurate risk stratification of prostate cancer (PCa), both at time of diagnosis and at other decision points, is essential to identify those patients at high risk of PCa-specific mortality (CSM). These patients are most likely to benefit from aggressive multimodal therapy, and it is important to distinguish them from the larger majority of patients who are cured by surgery or are otherwise at low risk of CSM, who may be spared the potential impact of additive treatments. The Cancer of the Prostate Risk Assessment Postsurgical (CAPRA-S) score was developed in a multi-institutional, community-based cohort to predict biochemical recurrence (BCR) and CSM following radical prostatectomy (RP) by incorporating preoperative prostate-specific antigen (PSA) levels and pathologic information into a straightforward, easy-to-use calculation of postoperative patient risk [1]. CAPRA-S has also been validated in another multi-institutional, sociodemographically and clinically diverse cohort, which confirmed its ability to predict both recurrence and CSM [2].

Over the last decade, many studies have tried to address the unmet clinical need for predicting aggressive PCa using genomic information [3–7]. The Decipher PCa genomic classifier (GC) risk prediction model was developed by investigators at the Mayo Clinic and GenomDx Biosciences to predict, with high specificity, early metastasis after RP [4]. Using oligonucleotide-microarray expression profiling of approximately 1.4 million markers in 545 tumors, machine learning algorithms were used to discover and validate a 22-marker gene expression signature of metastasis. The GC model measures the activity of genes implicated in proliferation, cell migration and adhesion, tumor motility, androgen-signaling, and immune system evasion [8]. In blinded validation studies in prospectively accrued cohorts [9], the GC model demonstrated improved performance over any individual clinicopathologic variable or clinical prediction model for clinical metastasis (confirmed by radiographic bone and computed tomography [CT] imaging) in post-RP [10] and post-BCR [11] patient cohorts.

In this study, we further examined the relationship between the CAPRA-S and GC scores for predicting CSM from the time of RP. We aimed to determine whether integrated genomic and clinical risk prediction models may further improve risk prediction compared with either model alone.

2. Materials and methods

2.1. Patient population

Subjects were identified from a population of 1010 men prospectively enrolled in the Mayo Clinic Department of Urology RP registry for PCa from 2000 to 2006. This population was clinically high risk, as defined by preoperative PSA level >20 ng/ml, pathologic Gleason score >8, or stage pT3b. Patients who received neo-adjuvant therapy or who were diagnosed with metastatic disease or failed to achieve PSA nadir after surgery were excluded. Clinical staging for patients with D’Amico high-risk disease or preoperative PSA >10 ng/ml underwent cross-sectional imaging with either CT or magnetic resonance imaging and bone scan to rule out the presence of metastatic disease before surgery. Data were collected from patients selected using a case-cohort approach, as this design allows inference measures (eg, survival estimates, hazards) about the whole cohort without requiring assessment of all 1010 patients [12,13]. The case-cohort design is most useful in analyzing time to failure in a large cohort in which the failure event is rare. The case-cohort design included all CSM events and a random sample of the full cohort. Of the 1010 men, 28 (3.0%) were documented to have died from PCa (at median follow-up of 6.9 yr). A 20% random sample of the entire cohort was selected for the analysis, including 11 patients with CSM (cases). The remaining 17 cases, who were not selected by random sampling, were also included for analyses (Supplemental Fig. 1).

2.2. Tissue and RNA processing

Following histopathologic review, total RNA was extracted and amplified from four to six 4-µm formalin-fixed, paraffin-embedded primary prostatic adenocarcinoma tissue sections from the nodule with the highest Gleason score. Macrodisssection was used to enrich for tumor cells. RNA was extracted and hybridized to Human Exon 1.0 ST GeneChips (Affymetrix, Santa Clara, CA, USA), which profile coding and nomenclature regions of the transcriptome, as described previously [10]. Following exclusion for tissue unavailability and microarray quality control (n = 38), 187 of the 225 patients sampled from the cohort remained with GC scores, of whom 185 had complete clinicopathologic data for estimating CAPRA-S scores.

2.3. Classifier assessment

We compared and integrated two previously validated post-RP classifiers: CAPRA-S and GC. CAPRA-S was developed using the Cancer of the Prostate Strategic Urologic Research Endeavor registry and was validated in the Shared Equal Access Regional Cancer Hospital database. CAPRA-S scores may be grouped into three validated groups: 0–2, 3–5, and ≥6 [14,15]. GC is a 0–1 score developed using clinical metastasis
after RP as the primary end point [8] and subsequently validated in an independent data set for prediction of clinical metastasis [10]. In some analyses, GC scores were categorized using previously defined cut-offs: < 0.4, 0.4–0.6, and >0.6 (indicating low, intermediate, and high risk, respectively) [10].

CAPRA-S and GC were combined into an integrated genomic-clinical classifier (GCC) using the parameter estimates derived from a Cox proportional hazards model for CSM. The predicted score is normalized to have values between 0 and 1, using the minimum and maximum predicted values that model could generate. It should be stressed that neither GC nor CAPRA-S was individually trained or refined in this patient population. The GCC model was validated using a bootstrap valuation method that estimates the optimism of the area under the receiver-operating characteristic curve (AUC) in the training set to adjust the AUC for overfitting and estimate the expected AUC in a potential validation set [16]. A total of 10 000 bootstrapped samples were run to determine the optimism, and the average optimism from those runs was used to adjust the AUC.

2.4. Statistical analysis

The primary end point in this analysis was CSM. Discrimination was measured by the concordance index. Survival analyses were weighted [12,17] to estimate parameters in the full cohort. Decision curve analysis was used to estimate the net benefit across a range of threshold probabilities for CSM at 5 yr after RP [18]. Univariable and multivariable Cox proportional hazards models for case-cohort study designs [19] were used to estimate hazard ratios (HRs) for both continuous and categorical predictive scores. An adaptation of the Firth method toward the Cox model to reduce the potential bias associated with small sample size was also used for univariable and multivariable analysis [20,21]. Interaction effects between prediction models and adjuvant therapy were evaluated by comparing the multivariable Cox model with interaction terms to the model without the interaction terms using the likelihood ratio test. Log-rank tests were used to assess the significance of differences in the Kaplan-Meier curves of patients in different risk groups. Cumulative incidence curves were constructed using Fine-Gray competing risks analysis to adjust for death by other causes [22]. Median follow-up times are estimates using the censoring distribution [23]. Statistical analyses were performed in R v.3.0, and all statistical tests were two-sided, using a 5% significance level.

3. Results

3.1. Performance of genomic and clinicopathologic risk models for predicting prostate cancer–specific mortality

GC scores were available for 187 patients (28 cases; median follow-up: 6.4 yr). Complete clinical data required to calculate CAPRA-S scores were available for 185 patients (Table 1). Patients in this high-risk cohort experienced CSM a median 4.8 yr after RP (interquartile range: 3.2–6.6). Medians and ranges for CAPRA-S and GC were 5 (2–12) and 0.37 (0.01–0.99), respectively.
The AUCs of CAPRA-S and GC as prediction models for CSM in comparison with individual clinicopathologic variables were compared using receiver-operating characteristic analysis (Supplemental Table 1). CAPRA-S and GC both have the highest AUCs: 0.75 (95% confidence interval [CI], 0.55–0.84) and 0.78 (95% CI, 0.68–0.87), respectively. In univariable analyses, CAPRA-S, GC, pathologic Gleason score, lymph node status, extracapsular extension, seminal vesicle invasion, and adjuvant androgen-deprivation therapy (ADT) were statistically significant predictors of CSM (Supplemental Table 3). In a multivariable analysis after adjusting for adjuvant treatment, the HR of CAPRA-S as a continuous variable was 1.36 for every 1-unit increase \( (p = 0.01) \), and the HR for GC was 1.81 for every 0.1-unit increase \( (p < 0.001) \). No significant interaction terms between the prediction models and adjuvant treatment were observed \( (p > 0.05) \), data not shown.

Use of a penalized approach to the Cox model appropriate for the small sample size and event rates in this study yielded similar results (Table 2). In exploratory analyses, when scores were dichotomized based on previously reported high-risk group cut points, multivariable analysis showed that patients with high CAPRA-S scores \( \geq 6 \) had an HR of 2.36 \( (p = 0.04) \), and patients with high GC scores \( \geq 0.6 \) had an HR of 12.2 \( (p < 0.001) \) for CSM (Supplemental Table 3). The corresponding HRs from the penalized model (Table 2) were 2.4 for high CAPRA-S scores and 11.3 for high GC scores. Kaplan-Meier plots using previously reported cut points for the models show significant differences in CSM-free survival for CAPRA-S and GC risk groups (Supplemental Fig. 2a and 2b). Survival differences were virtually unchanged when excluding patients who received any form of adjuvant therapy (Supplemental Fig. 2c and 2d).

### 3.2. Comparison of genomic and clinicopathologic model risk groups

As expected, some (although modest) correlation between GC and clinical risk factors captured by CAPRA-S is observed \( (r^2 = 0.38, p < 0.0001) \), but the trend shows that patients with both high GC score and high CAPRA-S score are most at risk to die from PCa (Fig. 1 and Supplemental Table 2). While patients with a higher CAPRA-S score had multiple adverse pathologic features, few who also had low GC scores \( (<0.4) \) developed metastasis or died of PCa at this follow-up despite a high frequency of BCR. For the group of patients with high-risk CAPRA-S scores \( \geq 6 \) or GC scores \( \geq 0.6 \), the cumulative incidence of CSM at 10 yr was estimated at...
Fig. 1 – Agreement between the genomic classifier (GC) and the Cancer of the Prostate Risk Assessment Postsurgical (CAPRA-S) scores. The distribution of CAPRA-S and GC scores for (A) biochemical recurrence (BCR), (B) clinical metastasis, and (C) prostate cancer–specific mortality (CSM). The dashed vertical lines show the boundaries for the low-risk (\( \leq 2 \)), intermediate-risk (3–5), and high-risk (\( \geq 6 \)) groups for CAPRA-S. The dashed horizontal lines mark the low (\(<0.4\)), intermediate (0.4–0.6), and high (\(\geq 0.6\)) GC scores, as described previously. The solid black lines and the surrounding gray shadows demonstrate the regression lines and their 95% confidence intervals.

CAPRA-S = Cancer of the Prostate Risk Assessment Postsurgical; GC = genomic classifier.

Fig. 2 – Cumulative incidence of prostate cancer–specific mortality (CSM) for the (A) Cancer of the Prostate Risk Assessment Postsurgical (CAPRA-S), (B) genomic classifier (GC), and (C) CAPRA-S high-risk stratified by GC. The cumulative probability of CSM increases with the CAPRA-S high risk when it is further stratified by GC.

CSM = prostate cancer–specific mortality; RP = radical prostatectomy.

13% and 30%, respectively (Fig. 2A and 2B). However, for patients with both high CAPRA-S and high GC scores, the cumulative incidence of CSM was 45% at 10 yr (Fig. 2C). Further evidence that the models provide complementary information was found using Kaplan-Meier analysis of each model’s higher-risk groups (Supplemental Fig. 3). Significant differences in CSM-free survival were observed when CAPRA-S risk groups were used to re-stratify GC > 0.6 (p = 0.005) and GC risk groups were used to re-stratify CAPRA-S > 5 (p < 0.001).

3.3. The integrated genomic and clinicopathologic risk model identifies a very high-risk subset of radical prostatectomy patients

Next, GC and CAPRA-S were combined using Cox regression to generate an integrated GCC. The formula for the GCC based on the Cox model coefficients was (0.21 × CAPRA-S) + (5.55 × GC). The AUC of the integrated model reached 0.80 (95% CI, 0.69–0.90). However, after cross-validation optimism adjustment, the AUC was 0.78, which was not significantly different from the individual models and was due to the limited number of events in this sample (the optimism-adjusted AUC was observed to vary between 0.6 and 1.0 over 10 000 bootstrapped iterations). Decision curve analysis, however, indicates that modest increases in AUC could be associated with larger gains in utility (Fig. 3). Compared with “treat none” and “treat all” scenarios (in which no risk prediction model is employed) to make treatment decisions across a range of CSM risk threshold probabilities, GCC had the highest net benefit compared with CAPRA-S or GC alone.

Finally, the 5-yr risk probability of CSM as calculated by each of the three risk prediction models for individual patient scores in this study population was evaluated (Fig. 4). For all three models, the probability of CSM at 5 yr is

---

virtually zero for patients with low-risk scores but rises dramatically for higher-risk scores. The 5-yr probability of CSM ranged between 0.0% and 17% for CAPRA-S (Fig. 4A), between 0.0% and 34% for GC (Fig. 4B), and between 0.0% and 53% for the integrated GCC model (Fig. 4C). These results further highlight the potential impact of combining genomic and clinical risk prediction models. A wider range of patient scores with the integrated GCC model had a lower predicted probability of CSM, and the exponential phase rose more steeply, indicating a subset of patients at markedly increased risk of CSM.

4. Discussion

The integration of tumor genomics into clinical practice for use in individualized patient risk prediction models holds great promise to improve management of high-risk PCa. This investigation follows previous reports on the validation of CAPRA-S [15] and GC [10], with results in this paper demonstrating that an integrated risk model combining GC and CAPRA-S provides improved risk prediction over either model alone. We have shown in this study that both GC and CAPRA-S can accurately predict CSM, and while these risk scores are only modestly correlated, they appear to provide complementary information that can be used for identification of a subset of post-RP patients at extremely high risk for death from PCa. Both CAPRA-S and GC can be determined immediately following RP, in contrast to measures based on PSA kinetics, and therefore these tools can be used to predict clinically significant events before disease progression can be detected clinically, radiographically, or in many cases even biochemically.

The results of this study demonstrate that in an at-risk population (ie, patients with adverse pathology), individuals with both high GC and high CAPRA-S scores represent a group of patients who are at markedly increased risk (>12-fold) of dying of their disease. The patient population sampled in this case-cohort study represents a group of surgical patients who are at considerably higher risk for treatment failure than the average man with PCa treated with RP. At 5 yr after RP, the population sampled for this analysis had a 33%, 8%, and 2% incidence of BCR, metastasis, and CSM, respectively.

GC did appear to “downclassify” a significant number of patients classified by CAPRA-S as high risk. Of these patients who were classified as high risk by CAPRA-S but as low risk by GC, many experienced BCR but not the rapid development of metastasis and CSM observed in those patients who were classified as high risk by both models. Given that the natural history of BCR is heterogeneous and relatively few men with BCR will experience early metastasis and CSM [24], these patients may represent a subset who perhaps could be observed and followed by tracking PSA kinetics.

After adjusting for adjuvant therapy in multivariable analysis, we found that CAPRA-S and GC remained independent and significant predictors of CSM. We did not find an interaction between CAPRA-S and GC with adjuvant therapy. While the use of adjuvant therapy was not universal and reflected inherent biases among the treating physicians, it does reflect current treatment practices for high-risk PCa. In the study cohort, nearly 12% of patients had positive lymph nodes, and 36% received adjuvant therapy; the validation of GC in an adjuvant therapy–rich cohort may represent a limitation. It is noteworthy that adjuvant ADT is associated with a negative effect on survival, which is almost certainly due to nonrandomized treatment assignments and resulting confounding by indication (ie, men receiving ADT likely have worse disease beyond what is captured by the variables in the adjusted analyses). This study is limited by its retrospective nature, and without randomization of these treatments, these results remain intriguing but hypothesis generating only.

5. Conclusions

For patients with adverse pathology after RP, outcomes vary greatly. Patients with CAPRA-S > 5 and GC > 0.6 were associated with a significantly increased risk of CSM in this cohort. As such, these men may benefit from additional secondary therapies, preferably in a clinical trial setting. Conversely, patients with both low CAPRA-S and low GC scores had excellent CSM-free survival even after adjusting for use of adjuvant therapy in this cohort. Future studies—ideally, randomized controlled clinical trials—will be required to determine whether the use of genomic and clinical risk prediction models actually improves patient cancer-specific and quality-of-life outcomes.

Author contributions: Matthew R. Cooperberg had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Cooperberg, Karnes, Davicioni.

Acquisition of data: Jenkins, Karnes, Davicioni, Ghadessi.

Analysis and interpretation of data: Crisan, Davicioni, Cooperberg.

Drafting of the manuscript: Cooperberg, Crisan.

Critical revision of the manuscript for important intellectual content: Cooperberg, Davicioni, Crisan, Karnes.

Statistical analysis: Crisan.

Obtaining funding: Davicioni.

Administrative, technical, or material support: Jenkins, Karnes.

Supervision: Karnes.

Other (specify): None.

Financial disclosures: Matthew R. Cooperberg certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Elai Davicioni, Anamaria Crisan, and Mercedes Ghadessi are employees of GenomeDX Biosciences. Matthew R. Cooperberg has been a consultant for GenomeDX but has received no financial or other compensation related to this paper. R. Jeffrey Karnes and Robert B. Jenkins have no direct financial incentive related to this paper, although their institution, the Mayo Clinic, does have a financial interest in the technology described in the paper. UCSF received a research grant from GenomeDX. GenomeDX has filed patents on prostate cancer biomarkers, including the Decipher panel, of which Robert B. Jenkins and Elai Davicioni are coinventors.
**Funding/Support and role of the sponsor:** NIH P50 CA91956 Mayo Prostate Cancer SPORE grant; Richard M. Schulze Family Foundation; National Research Council of Canada Industrial Research Assistance Program; Mayo Foundation for Medical Education and Research; and GenomeDx Biosciences Inc.

**Acknowledgment statement:** The authors acknowledge Dr. Christine Buerki, Kasra Yousefi, and Voleak Cheung for their assistance with the analysis and preparing of the manuscript (GenomeDx) and Dr. Darby Thompson (EMMES Canada) and Eric Bergstralh (Mayo Departments of Urology and Statistics) for statistical and methodological assistance. The authors appreciate the advice, encouragement, and support of Dr. Donald Tindall, PI of the Mayo Prostate Cancer SPORE, and Dr. Lesley Esford (National Research Council Canada).

**Appendix A. Supplementary data**

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eururo.2014.05.039.

**References**


