Prostate Cancer

A 17-gene Assay to Predict Prostate Cancer Aggressiveness in the Context of Gleason Grade Heterogeneity, Tumor Multifocality, and Biopsy Undersampling

Eric A. Klein [a,†,*], Matthew R. Cooperberg [b,c,†], Cristina Magi-Galluzzi [d], Jeffry P. Simko [b,e], Sara M. Falzarano [d], Tara Maddala [f], June M. Chan [b,c], Jianbo Li [f], Janet E. Cowan [b], Athanasios C. Tsiatis [f], Diana B. Cherbavaz [f], Robert J. Pelham [f], Imelda Tenggara-Hunter [b], Frederick L. Baehner [e,f], Dejan Knezevic [f], Phillip G. Febbo [f], Steven Shak [f], Michael W. Kattan [g], Mark Lee [f], Peter R. Carroll [b,**]

[a] Glickman Urological and Kidney Institute, Cleveland Clinic, Cleveland, OH, USA; [b] Department of Urology, University of California, San Francisco and UCSF Helen Diller Family Comprehensive Cancer Center, San Francisco, CA, USA; [c] Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, USA; [d] Robert J. Tomsich Pathology and Laboratory Medicine Institute, Cleveland Clinic, Cleveland, OH, USA; [e] Department of Pathology, University of California, San Francisco, San Francisco, CA, USA; [f] Genomic Health, Inc., Redwood City, CA, USA; [g] Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH, USA

Article info

Abstract

Background: Prostate tumor heterogeneity and biopsy undersampling pose challenges to accurate, individualized risk assessment for men with localized disease.
Objective: To identify and validate a biopsy-based gene expression signature that predicts clinical recurrence, prostate cancer (PCa) death, and adverse pathology.
Design, setting, and participants: Gene expression was quantified by reverse transcription–polymerase chain reaction for three studies—a discovery prostatectomy study (*n* = 441), a biopsy study (*n* = 167), and a prospectively designed, independent clinical validation study (*n* = 395)—testing retrospectively collected needle biopsies from contemporary (1997–2011) patients with low to intermediate clinical risk who were candidates for active surveillance (AS).
Outcome measures and statistical analysis: The main outcome measures defining aggressive PCa were clinical recurrence, PCa death, and adverse pathology at prostatectomy. Cox proportional hazards regression models were used to evaluate the association between gene expression and time to event end points. Results from the prostatectomy and biopsy studies were used to develop and lock a multigene-expression-based signature, called the *Genomic Prostate Score* (GPS); in the validation study, logistic regression was used to test the association between the GPS and pathologic stage and grade at prostatectomy. Decision-curve analysis and risk profiles were used together with clinical and pathologic characteristics to evaluate clinical utility.
Results and limitations: Of the 732 candidate genes analyzed, 288 (39%) were found to predict clinical recurrence despite heterogeneity and multifocality, and 198 (27%) were predictive of aggressive disease after adjustment for prostate-specific antigen, Gleason

† These authors contributed equally to this work.
* Corresponding author. Glickman Urological and Kidney Institute, Desk Q10-1, 9500 Euclid Avenue, Cleveland, OH 44195, USA. Tel. +1 216 444 5601.
** Corresponding author. University of California, San Francisco, Box 1695, 1600 Divisadero St., San Francisco, CA 94143-1695, USA. Tel. +1 415 353 7098; Fax: +1 415 353 7093.
E-mail address: pcarroll@urology.ucsf.edu (P.R. Carroll).

score, and clinical stage. Further analysis identified 17 genes representing multiple biological pathways that were combined into the GPS algorithm. In the validation study, GPS predicted high-grade (odds ratio [OR] per 20 GPS units: 2.3; 95% confidence interval [CI], 1.5–3.7; $p < 0.001$) and high-stage (OR per 20 GPS units: 1.9; 95% CI, 1.3–3.0; $p = 0.003$) at surgical pathology. GPS predicted high-grade and/or high-stage disease after controlling for established clinical factors ($p < 0.005$) such as an OR of 2.1 (95% CI, 1.4–3.2) when adjusting for Cancer of the Prostate Risk Assessment score. A limitation of the validation study was the inclusion of men with low-volume intermediate-risk PCa (Gleason score 3 + 4), for whom some providers would not consider AS.

Conclusions: Genes representing multiple biological pathways discriminate PCa aggressiveness in biopsy tissue despite tumor heterogeneity, multifocality, and limited sampling at time of biopsy. The biopsy-based 17-gene GPS improves prediction of the presence or absence of adverse pathology and may help men with PCa make more informed decisions between AS and immediate treatment.

Patient summary: Prostate cancer (PCa) is often present in multiple locations within the prostate and has variable characteristics. We identified genes with expression associated with aggressive PCa to develop a biopsy-based, multigene signature, the *Genomic Prostate Score* (GPS). GPS was validated for its ability to predict men who have high-grade or high-stage PCa at diagnosis and may help men diagnosed with PCa decide between active surveillance and immediate definitive treatment.

1. Introduction

Prostate-specific antigen (PSA) screening has been associated with a decline in PCa mortality but also has led to overdiagnosis and overtreatment of biologically insignificant disease [1]. Consequently, for selected patients with low-risk disease, active surveillance (AS)—expectant management with curative intervention only for those with tumor progression—is endorsed in practice guidelines as an alternative to immediate therapy [2]. Despite these recommendations, AS is underutilized, and >90% of men diagnosed with low-risk disease receive immediate treatment with surgery or radiation [3]. Overtreatment of biologically insignificant disease results in substantial cost and unnecessary morbidity [4], leading some agencies and professional organizations to question the value of routine screening [5,6].

An important impediment to the adoption of AS is the imperfect accuracy of conventional risk assessment at initial diagnosis [7]. Pretreatment risk-assessment tools [8,9] based on PSA, clinical stage, Gleason score, and other biopsy characteristics fare well in identifying patients at risk of aggressive disease but predict indolent disease for only a limited proportion of patients [10,11]. Moreover, for a substantial proportion (20–60%) of men classified as low-risk, current pretreatment assessment tools underestimate true tumor grade and, less commonly, true stage [12–14].

Molecular analyses of localized PCa have enabled the investigation of prognostic markers including tissue-based gene expression signatures, systems pathology profiles, and urine-based molecular markers [15,16]. Although many groups have demonstrated the potential of gene expression analysis to predict outcome in localized PCa [17–20], frequent genetic differences between regions of individual tumors and limited tumor sampling by needle biopsy pose challenges to molecular-based assays in PCa [21,22]. With these challenges in mind, we conducted two studies to identify genes for which expression in both prostatectomy and biopsy tissues consistently correlates with tumor aggressiveness regardless of multifocality, heterogeneity, or technical challenges associated with limited tumor obtained through biopsy. We then performed a third, independent, clinical validation study to determine whether a prespecified 17-gene signature can be measured in prostate biopsies to predict adverse pathology and improve risk stratification at diagnosis.

2. Materials and methods

2.1. Study design, patients, and specimens

Three studies were performed and are referred to as the *prostatectomy study*, the *biopsy study*, and the *validation study* (Fig. 1A; Supplement). The prostatectomy study sampled from a cohort of 2641 clinical T1/T2 PCa patients treated by radical prostatectomy at the Cleveland Clinic from 1987 to 2004. All patients with clinical recurrence (local recurrence or distant metastasis, $n = 127$) were selected, together with a random sampling of nonrecurrent patients, using an established stratified cohort sampling method ($n = 374$, with a 1:3 ratio of recurrent to non-recurrent patients) [23,24]. All samples analyzed were from fixed paraffin-embedded (FPE) prostatectomy specimens. The biopsy study included FPE prostate needle biopsy specimens from a separate cohort of 167 patients who had a diagnostic biopsy and underwent prostatectomy within 6 mo of diagnosis at the Cleveland Clinic between 1999 and 2007. Disease and vital status were determined from a database that was maintained prospectively, approved by an institutional review board (IRB), and compliant with the US Health Insurance Portability and Accountability Act, using data updated through October 2008.

The validation study conformed to the REMARK (Reporting Recommendations for Tumor Marker Prognostic Studies) guidelines for biomarker validation [25]. This prospectively designed protocol, including gene panel, algorithm, end points, analytical methods, and statistical methods, was agreed to by all investigators and locked prior to analyses. Consenting patients were identified from the IRB-approved University of California San Francisco (UCSF) Helen Diller Family Comprehensive Cancer Center Urologic Oncology Data Base (UODB). Men included were potential candidates for surveillance [26] but elected prostatectomy within 6 mo of their initial diagnostic biopsies (additional details are provided in the Supplement).
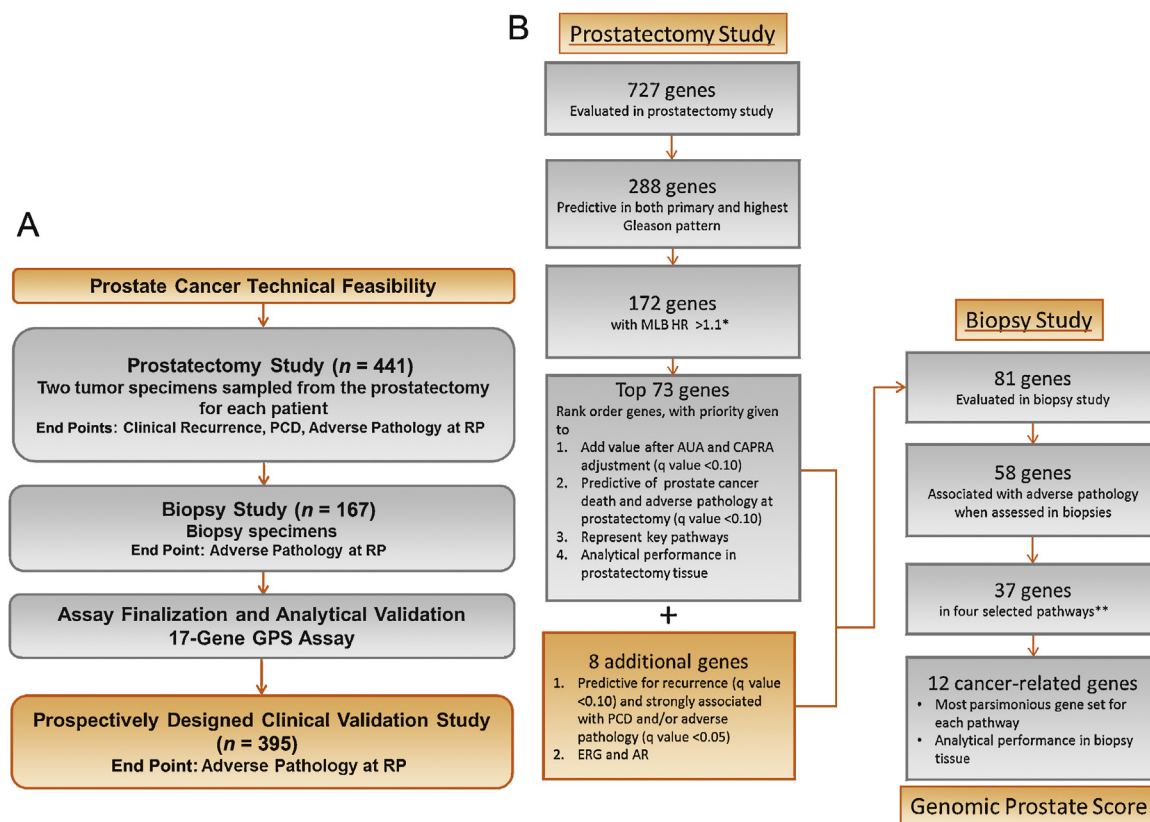
Fig. 1 – Identification of gene expression profiles predictive of prostate cancer aggressiveness: (A) Studies for development of the 17-gene Genomic Prostate Score (GPS) assay; (B) gene selection process for the GPS assay.
*Absolute maximum lower bound hazard ratio estimates in primary Gleason pattern, controlling false discovery at 10% (method of Crager [52]).
**Includes 14 stromal response, 17 cellular organization, 2 proliferation, and 4 androgen signaling genes.
AUA = American Urological Association; CAPRA = Cancer of the Prostate Risk Assessment; HR = hazard ratio; MLB = maximum lower bound; PCD = prostate cancer death; RP = radical prostatectomy.

2.2. Pathology

Pathology for all study specimens were centrally reviewed by expert pathologists using the 2005 International Society of Urological Pathology Consensus guidelines (S.M.F. and C.M.G. for the prostatectomy and biopsy studies, J.P.S. for the validation study) [27]. Three-dimensional mapping of prostatectomy specimens enabled enumeration of tumor foci and analysis of two spatially distinct tumor specimens representing (1) the most prevalent (primary) Gleason pattern and (2) the highest grade Gleason pattern. RNA from each sample was analyzed separately for gene expression. In patients whose primary Gleason pattern was also the highest Gleason pattern, two spatially distinct regions were selected for analysis. For the biopsy study, up to two representative needle biopsy FPE tissue blocks containing >2.0-mm involvement with tumor were selected for each patient. Twelve 5-μm unstained tissue sections were prepared for each biopsy specimen. The first and last sections were stained with hematoxylin and eosin to confirm presence of tumor and to guide manual microdissection to enrich for tumor-containing tissue for RNA analysis.

Biopsy specimens for the validation study were centrally reviewed to determine the primary and secondary Gleason score and to select a single, representative tumor-containing block with a minimum tumor length of 1 mm. The biopsy with the greatest volume of the highest grade tumor for each patient was used for analysis. Eight unstained, 5-μm, FPE biopsy sections were prepared from each case for histologic evaluation, manual microdissection of tumor-containing tissue, and RNA isolation. Prostatectomy specimens were centrally reviewed for primary, secondary, and

tertiary Gleason patterns, overall Gleason score, and pathologic T stage [27]. Biopsy reviews were performed blinded to prostatectomy pathology results and vice versa, and all pathology reviews were blinded to other clinical parameters and to gene expression data.

2.3. Assay methods and Genomic Prostate Score algorithm

For all studies, RNA was extracted, purified, and analyzed at Genomic Health, Inc., as described in the Supplement and previously published [28–30]. Prior to the prostatectomy study, 732 candidate genes were selected through a meta-analysis of publicly available microarray data sets (GSE3933, GSE10645, GSE5132, GSE3325) and additional comprehensive bioinformatic approaches including PCa-related genes enriched for those genes likely to have expression associated with outcome (Supplemental Table 1, Supplement). The expression of each gene was assayed either in singlet wells (prostatectomy study) or in triplicate wells (biopsy study and validation study) by TaqMan (Life Technologies, Carlsbad, CA, USA) quantitative reverse transcription–polymerase chain reaction (RT-PCR) assays and reported as crossing point values on a continuous scale. Eighty-one candidate genes selected from the prostatectomy study were assayed using the same methods in the biopsy study. Gene expression was normalized relative to reference genes to account for variation in RNA integrity due to tissue fixation and age of tissue blocks [30].

From the 81 genes evaluated in both studies, 12 genes associated with PCa aggressiveness and 5 reference genes were selected for inclusion in the final signature, called the *Genomic Prostate Score* (GPS)

(Fig. 1B; Supplement). For the validation study, GPS assessment in prostate biopsy tumor tissue was measured using a prespecified, analytically validated assay [30]. GPS uses preamplification to allow for lower RNA yield from prostate needle biopsies and is scaled between 0 and 100, with higher scores indicating more aggressive disease. All gene expression analyses were performed blinded to all clinical and pathologic data.

2.4. Statistical methods

For the prostatectomy study, the primary objective was to identify genes for which expression was associated with clinical recurrence-free interval (time from surgery to first local or distant metastatic recurrence detected by biopsy or imaging). Established methods including univariable and multivariable Cox proportional hazards regression models using weighted pseudo–partial-likelihood estimators evaluated the association between gene expression and clinical recurrence-free interval and other time to event end points [23,31,32]. Storey and Tibshirani's method [33] was used to control the false discovery rate (FDR) at 10% for gene identification (q < 0.10). For the biopsy study, the primary end point was presence of adverse pathology defined by high-grade disease (primary Gleason pattern 4 or any pattern 5) and/or non–organ-confined disease (pathologic stage T3) discovered at surgery and was analyzed using logistic regression.

Association of the GPS with adverse pathology was examined in the validation study using a six-cell multinomial logistic model representing combinations of pathologic grade and stage (Supplement). The prespecified primary end point was the ability of the GPS to predict prostatectomy grade and stage adjusting for biopsy Gleason score. Prespecified binary logistic regression models were also used to evaluate presence of adverse pathology in the prostatectomy specimen as well as high-grade disease and non–organ-confined disease, separately. Odds ratios [ORs] for GPS were calculated per 20-point increase in the score for ease of clinical interpretation, with 20 points approximating the difference between the average GPS of the highest 25th and lowest 25th percentiles of patients.

Favorable pathology was defined as freedom from high-grade and/or non–organ-confined disease. Area under the curve (AUC) for the receiver operating characteristic (ROC) and the likelihood of favorable pathology were calculated from multivariable binary logistic models of GPS adjusted for Cancer of the Prostate Risk Assessment (CAPRA) score [34], National Comprehensive Cancer Network (NCCN) risk groups [35], and a model including individual preoperative covariates. Based on these models, plots comparing the predicted probability of favorable pathology to the observed proportion of patients with favorable pathology for clinically defined subgroups were constructed. Decision-curve analysis [36] was used to compare a model combining CAPRA and GPS with the CAPRA score alone.

Analyses were independently performed by UCSF (J.E.C. and M.R.C.), the Cleveland Clinic (J.L.), and Genomic Health, Inc. (T.M.) using SAS v.9.2 (SAS Institute, Cary, NC, USA), R (R Foundation, Vienna Austria), and Stata v.12 (Stata Corp., College Station, TX, USA).

3. Results

3.1. Prostatectomy study results

The prostatectomy study included a final evaluable population of 441 patients with characteristics representative of the overall cohort (Supplemental Fig. 1, Supplement), 111 clinical recurrences (4% of original cohort), and 45 deaths (2%) due to PCa (Supplemental Table 2). As expected, baseline PSA, clinical T stage, biopsy Gleason score, pathologic T stage, surgical Gleason score, year of surgery, margin status, age,

and American Urological Association (AUA) risk group were significantly (p < 0.05) associated with clinical recurrence in univariable models; surgical Gleason score was the strongest factor in the multivariable model (Supplemental Table 3).

After controlling the FDR at 10%, 453 genes (62%) were associated with clinical recurrence in the samples containing either the primary (374 genes) or highest (367 genes) Gleason pattern, and 288 (40%) predicted recurrence regardless of the Gleason pattern (Fig. 2A; Supplemental Table 4). Importantly, consistent association between gene expression and outcome was observed across multiple features of aggressive PCa (eg, PCa death and the presence of adverse pathology at prostatectomy [37,38]) (Supplemental Table 4).

The 288 genes include multiple distinct gene groups defined by coexpression, biological function, and consistent association with clinical recurrence (Supplemental Fig. 2). Higher expression of stromal response and proliferation gene groups was associated with higher risk of clinical recurrence; higher expression of cellular organization, basal epithelial, androgen signaling, and stress response gene groups was associated with lower recurrence risk (Fig. 2B). Single-gene expression of androgen receptor (AR) and v-ets avian erythroblastosis virus E26 oncogene homolog (ERG) was not associated with clinical recurrence (q > 0.15 for both). Consistent findings were observed in the 258 patients for whom the primary Gleason pattern was of lower grade than the highest Gleason pattern (Fig. 2C) and in the subsets of patients with unifocal or multifocal disease (Fig. 2D). In multivariable analysis, after adjustment for AUA risk group (pretreatment PSA, biopsy Gleason score, and clinical stage) and controlling FDR at 10%, 198 (69%) of the 288 genes, including genes from the six gene groups identified in univariate analyses, remained strongly associated with clinical recurrence, suggesting that these genes provide prognostic value beyond conventional pretreatment clinical and pathologic factors (Supplemental Fig. 3).

3.2. Biopsy study results

The separate cohort of 167 AUA low- and intermediate-risk patients with biopsy tissue available from the Cleveland Clinic had characteristics similar to the overall cohort (Supplemental Table 6), and 58 (35%) had aggressive disease indicated by adverse pathologic features discovered at prostatectomy. Of the 81 genes evaluated in the biopsy study, quantitative RT-PCR analysis of tumor from FPE prostate needle biopsy tissue for each patient confirmed the association of 58 genes (72%) for adverse pathology at prostatectomy (Supplemental Table 7). Of the six biological pathways evaluated, the stromal response, androgen signaling, cellular organization, and proliferation groups were consistently associated with adverse pathology in both the prostatectomy and biopsy studies (Supplemental Fig. 4).

3.3. Development of the Genomic Prostate Score

A 17-gene GPS algorithm was developed from the 58 genes confirmed to be associated with aggressive disease in the
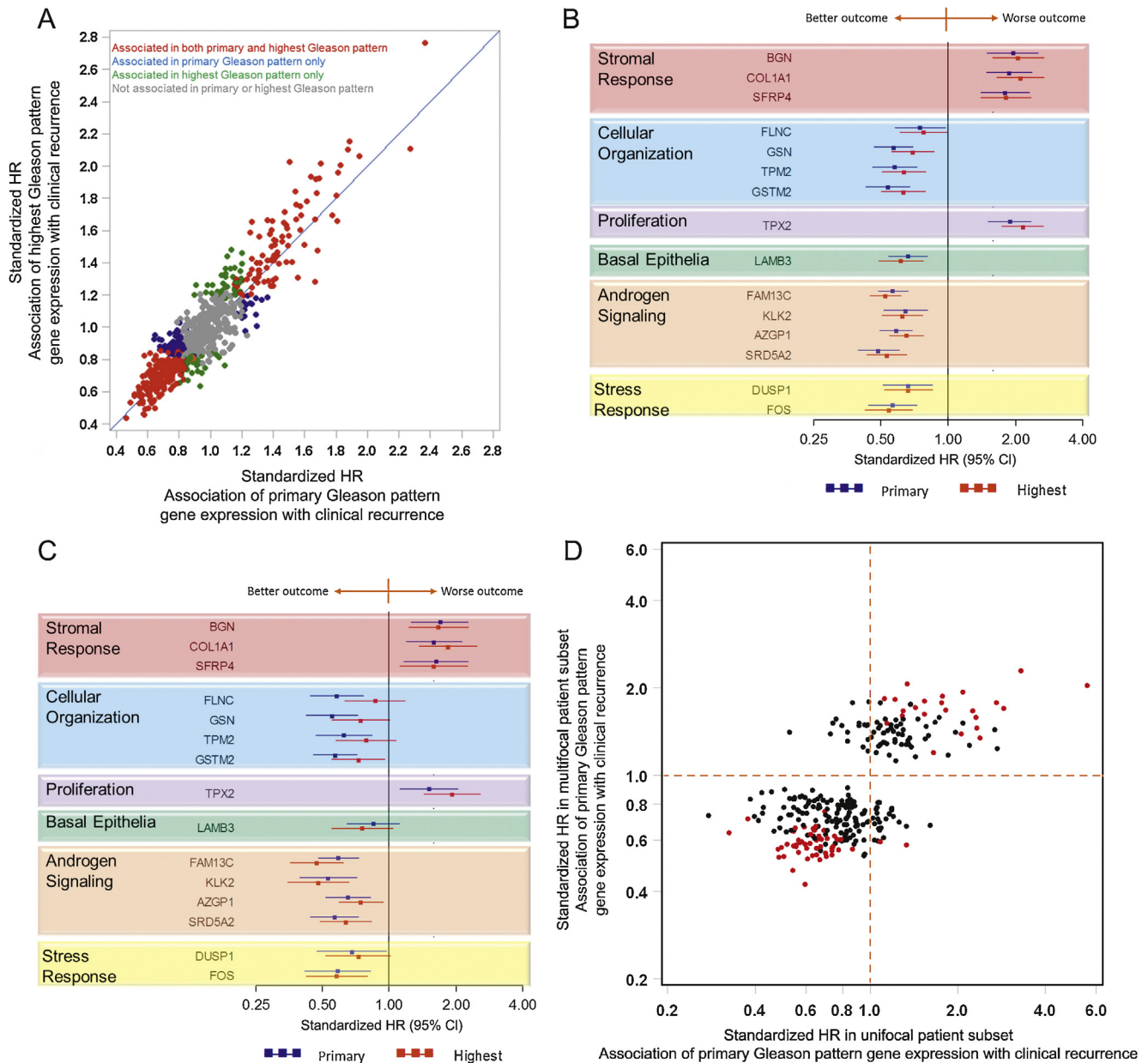
Fig. 2 – Association of gene expression profiles with outcome in the prostatectomy study. (A) Comparison of associations of gene expression with clinical recurrence in primary Gleason pattern and highest Gleason pattern tumor tissue. All 733 evaluable genes are depicted. (B) Analysis of gene expression from each pathway with clinical recurrence in primary and highest Gleason pattern tumor specimens from all 441 patients in the prostatectomy study. Forest plot shows standardized hazard ratios (HRs; estimate of relative risk per 1 standard deviation increase in gene expression measured on a log scale) and 95% confidence intervals. HR >1 indicates higher expression is associated with worse outcome. (C) Association of gene expression with clinical recurrence in primary and highest Gleason Pattern tumor tissue from 258 patients in whom primary and highest Gleason pattern are not the same. (D) The association of gene expression with clinical recurrence in patients with unifocal versus multifocal disease in the prostatectomy study. Red dots correspond to 77 of the 81 genes also evaluated in the biopsy study.
CI = confidence interval; HR = hazard ratio.

biopsy study, including 12 cancer-related genes from four gene groups and 5 reference genes (Supplement). Selection of the final set of cancer-related genes in GPS (Fig. 1B) was based on consistency across studies, representation of the four key pathways, and analytical performance (including reproducibility and higher mean expression with wide dynamic range). Strength of association with clinical recurrence was given preferential weighting for gene selection and algorithm

development (Supplemental Table 5). In a multivariable analysis adjusted for AUA risk group, the model consisting of the four gene groups of GPS was strongly associated with clinical recurrence (standard hazard ratio [HR]: 2.32; 95% confidence interval [CI], 1.81–3.00; q < 0.001) after regression-to-the-mean adjustment to partially account for overtraining. GPS provided additional risk discrimination beyond conventional pretreatment clinical risk assessment (Fig. 3A).
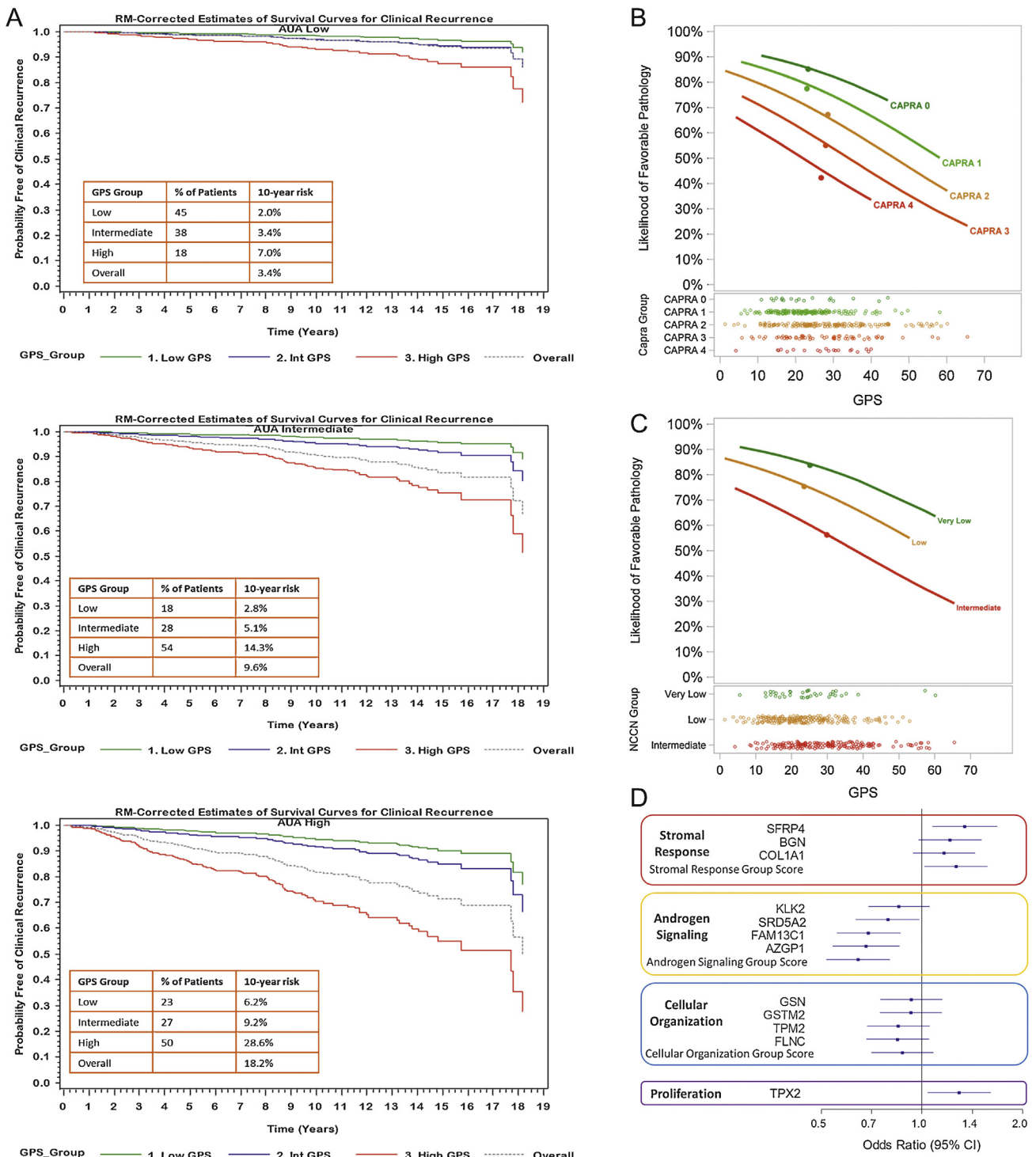
Fig. 3 – Improved risk discrimination with the addition of Genomic Prostate Score (GPS) to clinical risk assessment. (A) Risk discrimination by the GPS within American Urological Association (AUA) risk groups in the prostatectomy study. Regression-to-the-mean corrected estimated survival curves for clinical recurrence. Solid line curves represent low, intermediate, and high GPS groups defined by tertiles observed in the overall study. Dashed line curve represents the overall estimates for the given AUA risk group. (B, C) Solid circles correspond to the mean GPS on the x-axis and expected likelihood of favorable pathology on the y-axis, for each level of risk according to (B) CAPRA score or (C) NCCN risk group. Curves represent the probability of favorable pathology at each level of clinical risk, incorporating GPS. For example, for men with CAPRA score of 2, the average probability of favorable pathology is 67%; incorporating the GPS further refines risk, identifying patients with probabilities ranging from 37% to 85%. Distributions of the GPS values by clinical risk group are plotted as open circles beneath the curves. Assignment to CAPRA score or NCCN risk groups could be made for 386 and 388 patients, respectively. (D) Univariate odds ratios for individual genes and pathway subscores in GPS. The unscaled GPS is defined as: 0.735*Stromal Response group −0.368*Cellular Organization group −0.352*Androgen Signaling group +0.095*Proliferation. The unscaled score (GPS$_u$) is scaled between 0 and 100 (GPS) as follows: GPS = 13.4 × (GPS$_u$ + 10.5), with resulting values <0 set to GPS = 0 and values >100 set to GPS = 100. Forest plot shows odds ratios for prediction of adverse pathology at prostatectomy per 1 standard deviation increase in gene expression measured on a log$_2$ scale and 95% confidence intervals. Odds ratio >1 indicates higher expression is associated with worse outcome.

AUA = American Urological Association; CAPRA = Cancer of the Prostate Risk Assessment; CI = confidence interval; GPS = Genomic Prostate Score; NCCN = National Comprehensive Cancer Network; RM = regression to the mean.

3.4. University of California San Francisco Validation Study Cohort

As of August 2012, 514 prostatectomy patients who consented to participate in the UODB met clinical and pathology eligibility criteria for AS. Patients were excluded if the prostatectomy specimen was not available for re-review ($n = 28$, 5%), if the total length of tumor on biopsy was <1 mm ($n = 24$, 5%), or if the tumor was absent on central review of biopsy blocks ($n = 50$, 10%). Of 412 men whose biopsies were thus processed, the GPS successfully met all prespecified assay quality metrics in 395 (96%) [30], and this group formed the final evaluable population (Supplemental Fig. 5).

Participants had a median age of 58 yr and a mix of low- to low-intermediate clinical risk characteristics (Table 1). A total of 123 patients (31%) had high-grade or non–organ-confined disease at prostatectomy including 70 patients with high-grade and 87 with non–organ-confined disease. Scatter plots of association between GPS and clinical risk-stratification tools (CAPRA scores and NCCN groups) and a histogram of the GPS distribution (Fig. 3A and 3B; Supplemental Fig. 7) illustrate a broad range of GPSs at any given level of apparent clinical risk. GPS was weakly correlated with the CAPRA score ($r = 0.15$, $p = 0.002$).

3.5. Primary analysis of validation study

In the primary, prespecified multinomial analysis, GPS (assessed in biopsy tumor tissue) was a significant predictor of pathologic stage and grade at prostatectomy, adjusting for biopsy Gleason score ($p = 0.002$). In binary analyses adjusting for central biopsy Gleason score, each 20-point increase in GPS was associated with increased odds of high-grade disease (OR: 2.3; 95% CI, 1.5–3.7), non–organ-confined disease (OR: 1.9; 95% CI, 1.3–3.0), and high-grade and/or non–organ-confined disease (ie, adverse pathology; OR: 1.9; 95% CI, 1.3–2.9). In separate multivariable analyses adjusting for significant clinical covariates, the GPS was a consistent predictor of high-grade and/or non–organ-confined pathology, as were traditional clinical predictors (Table 2). The OR for each 20-point increase in GPS was 2.1 (95% CI, 1.4–3.2) adjusting for continuous CAPRA score, 1.9 (95% CI, 1.3–2.8) adjusting for NCCN risk group, and 1.9 (95% CI, 1.2–2.8) adjusting for age, PSA, clinical stage, and biopsy Gleason score.

3.6. Clinical utility of the Genomic Prostate Score

The association of individual genes and gene groups comprising GPS was consistent with the development studies. Downregulation of androgen-signaling genes most strongly predicted adverse pathology, followed by upregulation of stromal response and proliferation genes and downregulation of cellular organization genes (Fig. 3C). The addition of GPS to CAPRA improved the AUC for favorable pathology to 0.67 from 0.63 with CAPRA alone. However, in this population with a narrow range of risk, focused on men with low risk by traditional measures and thus suitable candidates for AS, AUC is limited in its ability to reflect clinically meaningful risk discrimination [39–41]. In this

Table 1 – Baseline characteristics of the prostatectomy, biopsy, and validation cohorts

Characteristic	Prostatectomy study[§] $n = 441$	Biopsy study $n = 167$	Validation study $n = 395$
Surgery year, range	1987–2004	1998–2007	1998–2011
Age			
Mean (SD)	61 (6.3)	60 (6.4)	58 (7.1)
Median (IQR)	62 (57–66)	61 (56–65)	58 (53–64)
Race			
White	83%	138 (83%)	359 (91%)
Black/Afro-Caribbean	12%	28 (17%)	13 (3%)
Asian/Hispanic/Other	5%	1 (<1%)	23 (6%)
Clinical tumor stage			
T1[†]	66%	149 (89%)	228 (58%)
T2	34%	18 (11%)	167 (42%)
Baseline PSA			
≤4	14%	23 (14%)	83 (21%)
>4–10	68%	133 (80%)	262 (66%)
>10–20	13%	11 (7%)	50 (13%)
>20	5%	0 (0%)	0 (0%)
Biopsy GS[*]			
≤6	70%	95 (57%)	301 (76%)
7[**]	25%	72 (43%)	94 (24%)
≥8	5%	0 (0%)	0 (0%)
Central surgical GS			
≤6	35%	29 (17%)	188 (48%)
7	61%	132 (79%)	198 (50%)
≥8	5%	6 (4%)	9 (2%)
Pathologic tumor stage			
T2	52%	97 (58%)	308 (78%)
T2+	7%	24 (14%)	6 (2%)
T3	41%	46 (28%)	81 (21%)
Adverse Pathology			
GS ≤3 + 4 and pT2	43%	107 (64%)	272 (69%)
GS ≤3 + 4 and pT3	25%	24 (14%)	53 (13%)
GS ≥4 + 3 and pT2	9%	14 (8%)	36 (9%)
GS ≥4 + 3 and pT3	23%	22 (13%)	34 (9%)
CAPRA score[‡]			
0	–	–	18 (5%)
1	–	–	141 (36%)
2	–	–	147 (38%)
3	–	–	60 (16%)
4	–	–	20 (5%)
Missing	–	–	9 (2%)
NCCN risk group[‡]			
Very low	–	–	37 (9%)
Low	–	–	191 (48%)
Intermediate	–	–	160 (41%)
Missing	–	–	7 (2%)

CAPRA = Cancer of the Prostate Risk Assessment; GS = Gleason score; IQR = interquartile range; NCCN = National Comprehensive Cancer Network; PSA = prostate-specific antigen; SD = standard deviation.

[§] For the prostatectomy study, only percentages are shown for categorical summaries because of the cohort sampling methodology.

[*] For the prostatectomy study, biopsies were not centrally graded.

[**] In the validation study, only GS 3 + 4 was included.

[†] Only T1c was included in the biopsy and validation studies.

[‡] NCCN risk group and CAPRA score could not be calculated for the prostatectomy and biopsy studies because of unavailable percentages of positive cores.

setting, decision-curve analysis and risk profiles have been shown to be more informative than ROC as measures of clinical utility [39,40].

In decision-curve analysis (Fig. 4), greater net benefit was realized for each end point through the combination of clinical (CAPRA) and genomic (GPS) information compared with clinical information alone. Over a range of threshold

Table 2 – Multivariable analyses of the Genomic Prostate Score and clinical/pathology covariates for prediction of adverse pathology at prostatectomy

Model	Variable	Odds ratio	95% CI	p value
1	GPS (per 20 units)	2.1	1.4–3.2	<0.001
	CAPRA (continuous)	1.6	1.2–2.0	<0.001
2	GPS (per 20 units)	1.9	1.3–2.8	0.001
	NCCN low vs very low	1.8	0.7–4.6	0.201
	NCCN intermediate vs very low	3.6	1.4–9.2	0.004
3	GPS (per 20 units)	1.9	1.2–2.8	0.003
	Age (continuous)	1.1	1.0–1.1	0.004
	PSA (continuous)	1.1	1.0–1.2	0.002
	Clinical stage T2 vs T1	1.6	1.0–2.5	0.059
	Biopsy Gleason score (3 + 4 vs 3 + 3)	1.7	1.0–2.9	0.050

CI = confidence interval; GPS = Genomic Prostate Score; NCCN = National Comprehensive Cancer Network; PSA = prostate-specific antigen; CAPRA = Cancer of the Prostate Risk Assessment.

probabilities, incorporation of GPS would be expected to lead to fewer treatments of patients who have favorable pathology at prostatectomy without increasing the number of patients with adverse pathology left untreated.

Within each level of clinical risk as ascertained by CAPRA score or NCCN risk group, GPS scores were widely distributed and further discriminated risk over a wide range (Fig. 3A and 3B). For example, an individual with a CAPRA score of 1 would have a 77% average likelihood of favorable pathology at prostatectomy; this estimate would be as high as 86% with a GPS of 10 or as low as 66% with a GPS of 40. To illustrate how well these models fit the observed pathology for subjects in the cohort, we compared the predicted probability versus observed proportion of favorable pathology for the CAPRA plus GPS and NCCN plus GPS models for clinically defined subgroups (Supplemental Fig. 6). For both CAPRA plus GPS and NCCN plus GPS, there was strong agreement between the predicted average and the observed proportion of favorable pathology at prostatectomy.

4. Discussion

Overtreatment of low-risk PCa is a widespread problem that has contributed to recommendations against routine use of PSA screening. Improved risk assessment at the time of diagnosis may help reduce overtreatment by discriminating
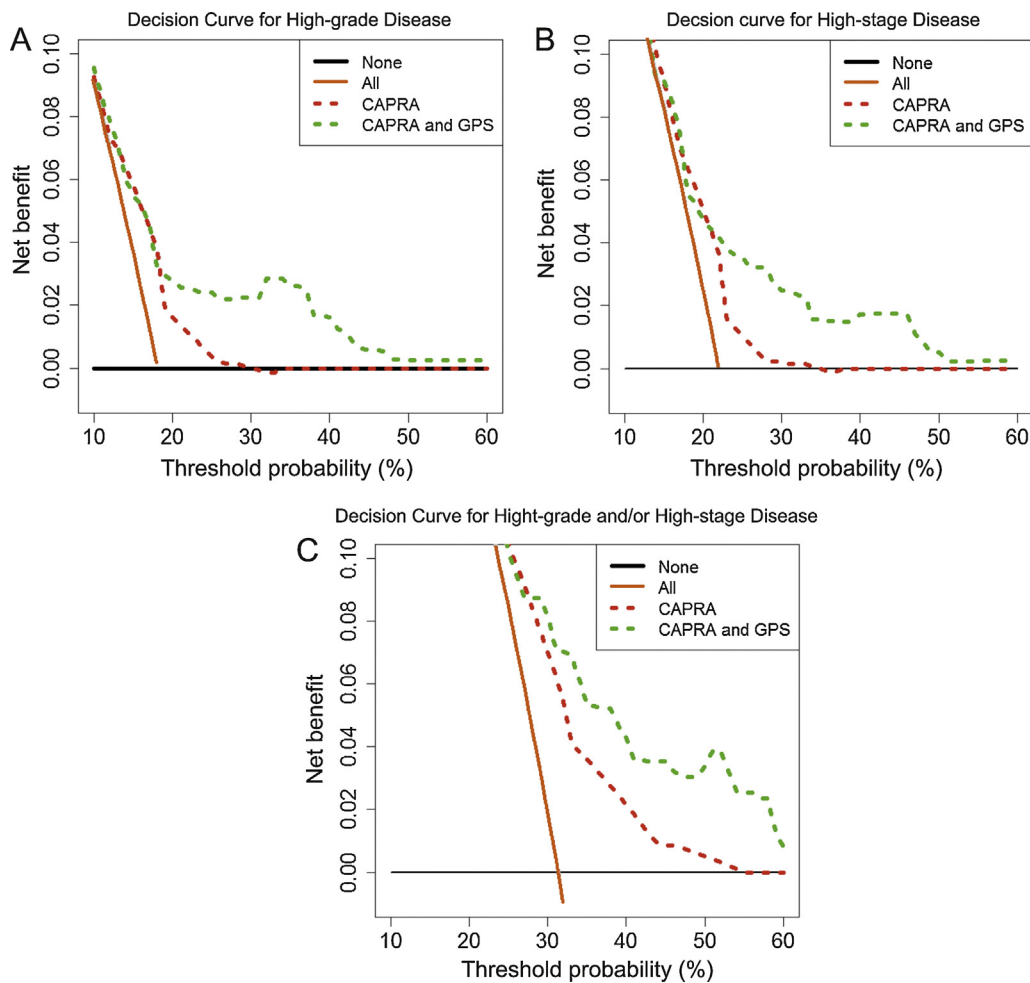
Fig. 4 – For the outcomes of (A) high-grade disease (Gleason pattern ≥4 + 3), (B) high-stage (pT3a) disease, and (C) high-grade and/or high-stage disease, decision curves are presented comparing a model based on clinical and pathologic characteristics summarized by the Cancer of the Prostate Risk Assessment (CAPRA) score to a model combining the CAPRA score with genomic information represented by the Genomic Prostate Score (GPS). For each analysis, the combined CAPRA plus GPS model yielded greater net benefit than the CAPRA score alone, indicating improved discrimination and calibration. CAPRA = Cancer of the Prostate Risk Assessment; GPS = Genomic Prostate Score.

indolent from aggressive disease more accurately, allowing more men (and their physicians) to pursue AS with greater confidence. Our studies were specifically designed to address the challenges of multifocality, heterogeneity, and limited biopsy sampling inherent in most PCa by evaluating whether a common underlying biology predictive of clinically aggressive disease could be identified in histologically heterogeneous prostate tumor foci and in tumor samples from needle biopsies. The 17-gene GPS, including genes representing multiple biological pathways, was shown to be an independent predictor of adverse pathology (p = 0.002) in a prospectively designed validation study of a large, contemporary cohort of men with low- to low-intermediate-risk PCa who were candidates for AS. Each 20-point increase in GPS was associated with a 2.3-fold increased risk of high-grade disease and a 1.9-fold increased risk of non–organ-confined disease. The GPS contributed additional prognostic information above and beyond existing, previously validated multivariable clinical risk-stratification tools. Furthermore, AUC, risk profiles, and decision-curve analysis all supported the added benefit of using GPS with standard risk stratification to identify those men with adverse pathology at time of diagnostic biopsy.

More accurate prediction of tumor grade and stage at diagnosis would address the key concerns of understaging and biopsy undersampling that have been barriers to the adoption of AS in contemporary practice. Men with low-risk PCa defined by conventional clinical and pathologic features already have a low probability of progression to lethal disease [42,43]. However, given a 30% rate of pathologic upgrading and/or upstaging between biopsy and prostatectomy [14] and the very low observed rates of surveillance uptake, even for low-risk tumors [3,44], more confident identification of men with very low likelihood of pathologic upstaging or upgrading should help support broader uptake of AS. Conversely, men with greater likelihood of adverse pathology may be encouraged to consider immediate definitive therapy or more intensive additional evaluation.

The genes composing the GPS were expressly chosen as those that were predictive when examined across different regions within individual patient tumors, including both the most common Gleason pattern and the highest grade cancer foci. In addition, GPS was developed to predict disease aggressiveness, with a primary emphasis on risk of long-term clinical progression but also the closely related near-term end point of adverse surgical pathology. Thus, successful validation suggests that the GPS reflects the underlying biology of a given tumor throughout the prostate including potential for invasion and distant metastasis.

Several limitations to this study should be acknowledged. First, our prostatectomy study analyzed 732 genes by polymerase chain reaction and did not use microarray or next-generation sequencing for comprehensive expression analysis. Although these alternative approaches may have identified different genes, use of the same technology (RT-PCR) for identification of genes, confirmation in biopsy specimens, and validation ensured equivalent analytic performance across all studies. Second, we focused on localized PCa specimens rather than metastatic samples to identify genes associated with aggressive disease. This decision was based on the objective of developing a "fit for purpose" assay for specific use in biopsies from clinically localized tumors. Finally, our validation study included men with biopsy Gleason 3 + 4 PCa, and in our primary analyses, we restricted our definition of high grade to primary Gleason pattern 4 or any pattern 5. The rationale for this decision was based on an explicit intention to include men whose tumors were expected to demonstrate a wider range of risk and an expectation that, given PCa epidemiology and policy trends, clinical entrance criteria for AS must eventually become broader [26,44]. Many Gleason pattern 3 + 4 tumors—especially those that are organ confined and/or those with a limited extent of pattern 4—may be marginally more aggressive, if at all, than Gleason 3 + 3 tumors. Early outcomes for these men on AS have been quite comparable to men with Gleason 3 + 3 tumors, justifying their inclusion in the validation study [45].

AS itself is not entirely benign. Serial prostate biopsies are uncomfortable and are associated with risks of bleeding, serious infection [46] and other adverse consequences [47]. Patients on surveillance can experience varying degrees of anxiety regarding the potential of their cancers to progress [48]. Additional studies are needed to determine whether GPS might also be used to identify men who may be candidates for less intense surveillance and perhaps some whose tumor characteristics are so indolent that they may be spared a diagnosis of "cancer" or converted to a program of watchful waiting, a less intensive alternative to AS [49,50].

5. Conclusions

The way forward for PCa must involve more personalized and evidence-based decision making regarding screening, diagnosis, and timing and intensity of treatment when required. Although multiple prior studies have related genomic markers to PCa clinical outcomes [17–20,51], the development of GPS specifically addresses the impact of tumor sampling in predicting aggressive PCa and included central pathology review and a large number of clinical recurrence events providing robust statistical power. By adding independent molecular information to established risk parameters, the GPS improves risk stratification at time of diagnosis and may favorably shift the balance of risks and benefits for men who are candidates for AS and facing challenging decisions regarding optimal management of localized PCa.

This work was presented in part at the 48th Annual Meeting of the American Society of Clinical Oncology, June 2012, Chicago, Illinois, USA, and at the Annual Meeting of the American Urological Association, May 2013, San Diego, California, USA.

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.2014.05.004.

References

[1] Strope SA, Andriole GL. Prostate cancer screening: current status and future perspectives. Nat Rev Urol 2010;7:487–93.

[2] Thompson I, Thrasher JB, Aus G, et al. Guideline for the management of clinically localized prostate cancer: 2007 update. J Urol 2007; 177:2106–31.

[3] Cooperberg MR, Broering JM, Carroll PR. Time trends and local variation in primary treatment of localized prostate cancer. J Clin Oncol 2010;28:1117–23.

[4] Sanda MG, Dunn RL, Michalski J, et al. Quality of life and satisfaction with outcome among prostate-cancer survivors. N Engl J Med 2008; 358:1250–61.

[5] Moyer VA, U.S. Preventive Services Task Force. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2012;157:120–34.

[6] Qaseem A, Barry MJ, Denberg TD, Owens DK, Shekelle P, Clinical Guidelines Committee of the American College of Physicians. Screening for prostate cancer: a guidance statement from the clinical guidelines committee of the American College of Physicians. Ann Intern Med 2013;158:761–9.

[7] Cooperberg MR, Carroll PR, Klotz L. Active surveillance for prostate cancer: progress and promise. J Clin Oncol 2011;29:3669–76.

[8] Cooperberg MR, Pasta DJ, Elkin EP, et al. The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. J Urol 2005;173:1938–42.

[9] Stephenson AJ, Kattan MW, Eastham JA, et al. Prostate cancer-specific mortality after radical prostatectomy for patients treated in the prostate-specific antigen era. J Clin Oncol 2009;27:4300–5.

[10] Dong F, Kattan MW, Steyerberg EW, et al. Validation of pretreatment nomograms for predicting indolent prostate cancer: efficacy in contemporary urological practice. J Urol 2008;180:150–4, discussion 154.

[11] Ploussard G, Epstein JI, Montironi R, et al. The contemporary concept of significant versus insignificant prostate cancer. Eur Urol 2011; 60:291–303.

[12] King CR, Long JP. Prostate biopsy grading errors: a sampling problem? Int J Cancer 2000;90:326–30.

[13] Conti SL, Dall'Era M, Fradet V, Cowan JE, Simko J, Carroll PR. Pathological outcomes of candidates for active surveillance of prostate cancer. J Urol 2009;181:1628–33, discussion 1633–4.

[14] Müntener M, Epstein JI, Hernandez DJ, et al. Prognostic significance of Gleason score discrepancies between needle biopsy and radical prostatectomy. Eur Urol 2008;53:767–76, discussion 775–6.

[15] Tomlins SA, Mehra R, Rhodes DR, et al. Integrative molecular concept modeling of prostate cancer progression. Nat Genet 2007;39:41–51.

[16] Makarov DV, Loeb S, Getzenberg RH, Partin AW. Biomarkers for prostate cancer. Annu Rev Med 2009;60:139–51.

[17] Singh D, Febbo PG, Ross K, et al. Gene expression correlates of clinical prostate cancer behavior. Cancer Cell 2002;1:203–9.

[18] Cheville JC, Karnes RJ, Therneau TM, et al. Gene panel model predictive of outcome in men at high-risk of systemic progression and death from prostate cancer after radical retropubic prostatectomy. J Clin Oncol 2008;26:3930–6.

[19] Ding Z, Wu CJ, Chu GC, et al. SMAD4-dependent barrier constrains prostate cancer growth and metastatic progression. Nature 2011; 470:269–73.

[20] Cuzick J, Berney DM, Fisher G, et al. Prognostic value of a cell cycle progression signature for prostate cancer death in a conservatively managed needle biopsy cohort. Br J Cancer 2012;106:1095–9.

[21] Berger MF, Lawrence MS, Demichelis F, et al. The genomic complexity of primary human prostate cancer. Nature 2011;470:214–20.

[22] Squire JA, Park PC, Yoshimoto M, et al. Prostate cancer as a model system for genetic diversity in tumors. Adv Cancer Res 2011;112: 183–216.

[23] Gray RJ. Weighted analyses for cohort sampling designs. Lifetime Data Anal 2009;15:24–40.

[24] Chen K, Lo S. Case-cohort and case-control analysis with Cox's model. Biometrika 1999;86:755–64.

[25] McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies. J Clin Oncol 2005;23: 9067–72.

[26] Cooperberg MR, Cowan JE, Hilton JF, et al. Outcomes of active surveillance for men with intermediate-risk prostate cancer. J Clin Oncol 2011;29:228–34.

[27] Epstein JI, Allsbrook Jr WC, Amin MB, Egevad LL. The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. Am J Surg Pathol 2005; 29:1228–42.

[28] Magi-Galluzzi C, Baehner F, Dutta D, et al. RNA yields and RT-PCR gene expression profiles obtained from manual-microdissected, fixed-paraffin-embedded tissue (FPET) from prostate adenocarcinomas. Mod Pathol 2007;20:352A.

[29] Magi-Galluzzi C, Millward CL, Maddala T, et al. RNA yields and RT-PCR gene expression profiles obtained from manual-microdissected fixed paraffin-embedded prostate cancer needle core biopsies. Mod Pathol 2010;23:205A.

[30] Knezevic D, Goddard AD, Natraj N, et al. Analytical validation of the Oncotype DX prostate cancer assay – a clinical RT-PCR assay optimized for prostate needle biopsies. BMC Genomics 2013;14: 690.

[31] Cox DR. Regression models and life tables. J R Stat Soc Series B Stat Methodol 1972;43:187–220.

[32] Goldstein LJ, Gray R, Badve S, et al. Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. J Clin Oncol 2008;26: 4063–71.

[33] Storey JD, Tibshirani R. Estimating the positive false discovery rate under dependence, with application to DNA microarrays. Technical Report No. 2001-28. Stanford, CA: Stanford University; 2001.

[34] Cooperberg MR, Broering JM, Carroll PR. Risk assessment for prostate cancer metastasis and mortality at the time of diagnosis. J Natl Cancer Inst 2009;101:878–87.

[35] NCCN clinical practice guidelines in oncology (NCCN Guidelines). Prostate cancer, version 2.2013. National Comprehensive Cancer Network Web site. http://www.nccn.org/professionals/physician_gls/f_guidelines.asp#prostate.

[36] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565–74.

[37] Eggener SE, Scardino PT, Walsh PC, et al. Predicting 15-year prostate cancer specific mortality after radical prostatectomy. J Urol 2011; 185:869–75.

[38] Hong SK, Sternberg IA, Keren Paz GE, et al. Definitive pathology at radical prostatectomy is commonly favorable in men following initial active surveillance. Eur Urol 2014;66:214–9.

[39] Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157–72, discussion 207–12.

[40] Vickers AJ, Elkin EB, Steyerberg E. Net reclassification improvement and decision theory. Stat Med 2009;28:525–6, author reply 526–8.

[41] Gu W, Pepe MS. Estimating the capacity for improvement in risk prediction with a marker. Biostatistics 2009;10:172–86.

[42] Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. N Engl J Med 2012; 367: 203–13.

[43] Donin NM, Laze J, Zhou M, Ren Q, Lepor H. Gleason 6 prostate tumors diagnosed in the PSA era do not demonstrate the capacity for metastatic spread at the time of radical prostatectomy. Urology 2013;82:148–52.

[44] Ganz PA, Barry JM, Burke W, et al. NIH State-of-the-Science Conference Statement: role of active surveillance in the management of men with localized prostate cancer. NIH Consens State Sci Statements 2011;28:1–27.

[45] Reese AC, Cowan JE, Brajtbord JS, Harris CR, Carroll PR, Cooperberg MR. The quantitative Gleason score improves prostate cancer risk assessment. Cancer 2012;118:6046–54.

[46] Nam RK, Saskin R, Lee Y, et al. Increasing hospital admission rates for urological complications after transrectal ultrasound guided prostate biopsy. J Urol 2010;183:963–8.

[47] Fujita K, Landis P, McNeil BK, Pavlovich CP. Serial prostate biopsies are associated with an increased risk of erectile dysfunction in men with prostate cancer on active surveillance. J Urol 2009;182:2664–9.

[48] Hayes JH, Ollendorf DA, Pearson SD, et al. Active surveillance compared with initial treatment for men with low-risk prostate cancer: a decision analysis. JAMA 2010;304:2373–80.

[49] Carter HB, Partin AW, Walsh PC, et al. Gleason score 6 adenocarcinoma: should it be labeled as cancer? J Clin Oncol 2012;30:4294–6.

[50] Esserman LJ, Thompson Jr IM, Reid B. Overdiagnosis and overtreatment in cancer: an opportunity for improvement. JAMA 2013;310: 797–8.

[51] Penney KL, Sinnott JA, Fall K, et al. mRNA expression signature of Gleason grade predicts lethal prostate cancer. J Clin Oncol 2011; 29:2391–6.

[52] Crager MR. Gene identification using true discovery rate degree of association sets and estimates corrected for regression to the mean. Stat Med 2010;29:33–45.