Title

Patient-specific Meta-analysis of 2 Clinical Validation Studies to Predict Pathologic Outcomes in Prostate Cancer Using the 17-Gene Genomic Prostate Score

Authors

Brand, Timothy C
Zhang, Nan
Crager, Michael R
et al.

Publication Date

2016-03-01

DOI

10.1016/j.urology.2015.12.008

Peer reviewed

Patient-specific Meta-analysis of 2 Clinical Validation Studies to Predict Pathologic Outcomes in Prostate Cancer Using the 17-Gene Genomic Prostate Score

Timothy C. Brand, Nan Zhang, Michael R. Crager, Tara Maddala, Anne Dee, Isabell A. Sesterhenn, Jeffry P. Simko, Matthew R. Cooperberg, Shiv Srivastava, Inger L. Rosner, June M. Chan, Phillip G. Febbo, Peter R. Carroll, Jennifer Cullen, and H. Jeffrey Lawrence

OBJECTIVE To perform patient-specific meta-analysis (MA) of two independent clinical validation studies of a 17-gene biopsy-based genomic assay as a predictor of favorable pathology at radical prostatectomy.

MATERIALS AND METHODS Patient-specific MA was performed on data from 2 studies (732 patients) using the Genomic Prostate Score (GPS; scale 0-100) together with Cancer of the Prostate Risk Assessment (CAPRA) score or National Comprehensive Cancer Network (NCCN) risk group as predictors of the likelihood of favorable pathology (LFP). Risk profile curves associating GPS with LFP by CAPRA score and NCCN risk group were generated. Decision curves and receiver operating characteristic curves were calculated using patient-specific MA risk estimates.

RESULTS Patient-specific MA-generated risk profiles ensure more precise estimates of LFP with narrower confidence intervals than either study alone. GPS added significant predictive value to each clinical classifier. A model utilizing GPS and CAPRA provided the most risk discrimination. In decision-curve analysis, greater net benefit was shown when combining GPS with each clinical classifier compared with the classifier alone. The area under the receiver operating characteristic curve improved from 0.68 to 0.73 by adding GPS to CAPRA, and 0.64 to 0.70 by adding GPS to NCCN risk group. The proportion of patients with LFP >80% increased from 11% using NCCN risk group alone to 23% using GPS with NCCN. Using GPS with CAPRA identified the highest proportion—31%—of patients with LFP >80%.

CONCLUSION Patient-specific MA provides more precise risk estimates that reflect the complete body of evidence. GPS adds predictive value to 3 widely used clinical classifiers, and identifies a larger proportion of low-risk patients than identified by clinical risk group alone. UROLOGY 89: 69–75, 2016.

Financial Disclosure: *Nan Zhang 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 are the following: Matthew R. Cooperberg's institution has received research support from GenomeDx Biosciences Corp, Genomic Health, Inc., and Myriad Genetics, and consulting fees from Myriad Genetics. Jeffry P. Simko's institution has received research support from GenomeDx Biosciences Corp, Genomic Health, Inc., and Myriad Genetics, and consulting fees from Genomic Health, Inc. Peter R. Carroll has received research support from Genomic Health, Inc. Nan Zhang, Michael R. Crager, Tara Maddala, Anne Dee, Phillip G. Febbo, and H. Jeffrey Lawrence are all employees of Genomic Health, Inc. and own Genomic Health, Inc. stock. The remaining authors declare that they have no relevant financial interests.*

Funding Support: *The two clinical validation studies were supported by Genomic Health, Inc., which was involved in the design and conduct of the study: collection, management, analysis, and interpretation of the data; and preparation, review, and approval of the manuscript. Genomic Health, Inc. performed the genomic analyses described and provided financial support to the University of California San Francisco, the Center for Prostate Disease Research, and the Uniformed Services University of the Health Sciences to help defray tissue and data processing costs. No non-Genomic Health author received any personal compensation for the study, financial or otherwise. This research was also supported by the U.S. Department of Defense grant W81XWH-13-2-0074 to the University of California San Francisco, and through the Center for Prostate Disease Research and the Uniformed Services University of the Health Sciences (HU0001-10-2-0002).*

From the Madigan Army Medical Center, Tacoma, WA; the Genomic Health, Inc., Redwood City, CA; the Joint Pathology Center, Silver Spring, MD; the University of California, San Francisco, San Francisco, CA; the Center for Prostate Disease Research, Uniformed Services University of the Health Sciences, Bethesda, MD; and the Walter Reed National Military Medical Center, Bethesda, MD

Address correspondence to Nan Zhang, Ph.D., Genomic Health, Inc., 301 Penobscot Dr., Redwood City, CA 94063. E-mail: nzhang@genomichealth.com

Submitted: October 7, 2015, accepted (with revisions): December 3, 2015

http://dx.doi.org/10.1016/j.urology.2015.12.008
0090-4295

A number of new tissue-based molecular assays have been developed to provide improved risk stratification for men with newly diagnosed, clinically localized prostate cancer to help guide therapeutic decisions, such as immediate definitive therapy vs active surveillance. These prognostic assays must meet a suitable level of clinical validation before gaining widespread adoption into clinical urological practice. Studies using archival specimens can provide level I evidence if a suitable "prospective-retrospective" study design is used.[1]

One such assay is the Onco*type* DX Prostate Cancer Assay, a biopsy-based test that measures the expression of 17 genes (12 cancer-related and 5 reference) to provide a Genomic Prostate Score (GPS; scaled 0-100) as a measure of biologic aggressiveness. Two "prospective-retrospective" clinical studies have validated GPS as a strong and independent predictor of adverse pathology (AP) at radical prostatectomy (RP), thus, providing level I evidence.[2,3]

Using GPS together with National Comprehensive Cancer Network (NCCN) clinical risk group, the test currently provides an estimate of the likelihood of favorable pathology (LFP), which is furnished in the patient report of a commercial-grade assay, based on data from the first validation study. The second validation study confirmed the association between GPS and LFP. Although the baseline characteristics of both cohorts were similar, there were some differences, including the inclusion criteria for intermediate risk patients, the percentage of African American patients, and the pathologist responsible for the central pathology review. To provide more broadly representative and precise estimates of LFP, we combined information from both studies and examined predictive models using GPS with other clinical risk-stratifying tools.

To take full advantage of the combined data from these studies, which comprise a total of 732 patients, a patient-specific meta-analysis (MA) was performed.[4] Patient-specific MA combines predictions for individual patients from multiple studies, with weighting based on the precision of each prediction, thereby permitting more precise risk estimates with narrower confidence intervals that reflect the complete body of evidence collected across these studies.[4] Patient-specific MA was applied previously to the prognostic information provided by Onco*type* DX Breast Cancer Assay Recurrence Score together with pathology and clinical covariates in 2 studies of early stage breast cancer.[5] An educational tool providing estimates of distant recurrence risk of breast cancer based on the patient-specific MA calculation is available online.

Using the patient-specific MA, we sought (1) to determine if GPS added predictive value to 3 widely used clinical risk stratification tools (Cancer of the Prostate Risk Assessment [CAPRA] score, NCCN risk group, and American Urology Association/European Association of Urology [AUA/EAU] risk group), and (2) to provide a better estimate of the proportion of patients identified for whom the risk of aggressive disease is very low than is identified by clinical risk group alone.

MATERIALS AND METHODS

Patients and Study Design

The patient selection criteria and study design for the University of California San Francisco (UCSF) and the Uniformed Services University Center for Prostate Disease Research (CPDR) clinical validation studies of GPS have been described previously.[2,3] Both studies included patients who were potential candidates for active surveillance under relatively broad inclusion criteria at diagnosis—(ie, biopsy Gleason score [GS] 3 + 3 or 3 + 4 disease [Gleason 4 + 3 disease was permitted in the CPDR study]), prostate-specific antigen (PSA) ≤20, and clinical stage ≤T2—but who elected RP within 6 months of diagnosis. Patients were identified from the institutional review board-approved clinical databases and biobanks at each institution. The UCSF study patients were diagnosed from 1997 to 2011 and the CPDR study patients were diagnosed from 1990 to 2011. Exclusion criteria included any neoadjuvant therapy, <1 mm total biopsy tumor length, missing prostatectomy or biopsy tissue for central pathology review, and inadequate ribonucleic acid (RNA) quality for analysis. All biopsies and prostatectomies were centrally reviewed by academic urologic pathologists (Jeffry P. Simko for UCSF study, Isabell A. Sesterhenn for CPDR study), following the 2005 International Society of Urological Pathology Consensus guidelines.[6] Pathology review of biopsies was performed blinded to the review of RP and vice versa. Patients from these validation studies were excluded from the MA if they were found to have biopsy GS 4 + 3 disease at diagnosis or pathologic pT2$^+$ (because capsular incision renders the true tumor status with regard to organ confinement indeterminate) at RP upon central pathology review.

GPS Assay and End Point

The Onco*type* DX Prostate Cancer Assay (GPS assay) has been described previously, and analyzes the expression levels of 17-genes (12 cancer-related and 5 reference) to provide the GPS, which is scaled from 0 to 100.[7] The end point of AP at RP for both studies was defined as high-grade (primary GS pattern 4 or any pattern 5) or non–organ-confined disease (pT3).[7] Favorable pathology (FP) at RP is defined as low-grade (surgical GS 3 + 3 or 3 + 4 with no pattern 5) and organ-confined disease (pT2).

Statistical Methods

The analyses performed included GPS plus each of 3 widely used clinical risk assessment tools: CAPRA score,[8] NCCN risk group,[9] and AUA/EAU risk group.[10,11] For each of the combinations, multivariable logistic regression models with factors for the clinical risk assessment tool and GPS were fit separately to each study, with AP as the end point. The CAPRA score, which ranged from 0 to 5 across the 2 study populations, was treated as a continuous numerical variable in the logistic regression models. Patients classified as having clinical T2c disease were excluded from the patient-specific MA of GPS plus AUA/EAU risk group, because these patients are considered to have high-risk disease using the AUA/EAU risk classification system.

Both studies enrolled patients with NCCN very low, low, and intermediate risk disease. However, the enrollment of intermediate-risk patients in the UCSF study was restricted to patients with low volume (≤3 positive biopsy cores or ≤33% of positive cores) Gleason score 3 + 4 disease[2]; the CPDR study enrolled all intermediate-risk patients, regardless of tumor volume.[3] The patient-specific MA method can accommodate the "special population" of NCCN intermediate-risk patients with high tumor volume, who were included only in the CPDR study, provided 2 assumptions are met[4]: (1) there is no interaction between NCCN

risk-tumor volume category and GPS in the prediction of AP, which can be tested using the study data; and (2) if NCCN intermediate-risk, high tumor volume patients had been included in the UCSF study, the relation of their risk of AP to other patients' would have been similar to the relationship in the CPDR study. The second assumption is not testable based on data, but is plausible based on the characteristics of the populations of the 2 studies.

An MA likelihood ratio test was used to assess whether GPS added predictive value for the LFP above the clinical risk assessment tools by summing across studies the likelihood ratio chi-square test statistic, comparing the full model (GPS and clinical risk assessment tool) with the reduced model (clinical risk assessment tool alone).

The LFP for a given patient with specified clinical risk group and GPS was estimated by combining information from both studies using fixed-effect patient-specific MA.[4] Following this method, the individual patient's log odds of FP was first estimated for each study using logistic regression models. The 2 estimates were then averaged, weighting by their patient-specific inverse variances, and transformed to the probability scale to get the LFP estimate and 95% confidence interval.

For NCCN risk group, the NCCN intermediate-risk group high tumor volume patients in the CPDR study were handled as a special population for the patient-specific MA because these patients were not included in the UCSF study. To obtain an overall estimate of NCCN intermediate-risk group patients regardless of tumor volume, a logistic regression model was fit to the CPDR study data to estimate the probability that an NCCN intermediate-risk patient has high tumor volume given the patient's GPS. The LFP for an NCCN intermediate-risk group patient with given GPS was then estimated by averaging the LFP estimates for low tumor volume and high tumor volume, weighting by the estimated probability of high volume.

Receiver operating characteristic (ROC) curves and decision curves[12] were calculated using the patient-specific MA estimates of the likelihood of AP and the AP status data combined across both studies. In these calculations, sensitivity was estimated by the proportion of patients with AP who had patient-specific MA estimate of the likelihood of AP greater than each cutoff; specificity was estimated as the proportion of patients without AP who had patient-specific MA estimate of the likelihood of AP less than or equal to each cutoff. The area under the ROC curve (AUC) for GPS plus each clinical risk assessment tool was compared with the AUC for the clinical risk assessment tool alone, using the method of DeLong et al.[13]

RESULTS

Study Population Characteristics

A total of 732 evaluable patients contributed data to the MA (389 from the UCSF study and 343 from the CPDR study) (Supplemental Fig. S1). Patients were excluded from the MA if the biopsy GS was 4 + 3 (n = 13), pathologic T-stage at RP was pT2$^+$ (n = 45) after central review, or RP specimen was unavailable for central review (n = 7).

Baseline characteristics are described in Table 1 and Supplemental Table S1. Patient age at diagnosis, PSA at diagnosis, CAPRA score, NCCN risk groups, and AUA/EAU risk groups were all similarly distributed in these studies. Of note, there were more African American patients in the CPDR cohort (20%) than in the UCSF cohort (3%). NCCN intermediate-risk patients with biopsy GS 3 + 4 and high volume (>3 positive biopsy cores and >33% of positive cores) were enrolled only in the CPDR study. A similar percentage of patients had high-grade at RP (primary GS pattern 4 or any pattern 5) in the 2 studies, whereas 6% more patients had non–organ-confined disease in the CPDR study than in the UCSF study. GPS ranged from 1 to 74, with a median of 27 in both studies. Compared to either individual study, the 2 studies collectively represent a broader spectrum of contemporarily managed prostate cancer patients.

Multivariable Logistic Regression Analysis

Odds ratios for AP derived from the multivariable logistic regression models for the 2 studies individually are presented in Supplemental Table S2. GPS was a statistically significant predictor of AP after adjustment of CAPRA score or NCCN risk group or AUA/EAU risk group, in each study. The directions of the association of GPS and the clinical risk assessment tools with AP are the same in the 2 studies, and the magnitudes of association are broadly consistent with overlapping confidence intervals. The odds ratios from the multivariable models are not significantly different between the 2 studies ($P \geq .20$). Wald tests for the interaction of GPS with NCCN risk group-tumor volume category (Very Low, Low, Low-volume Intermediate and High-volume Intermediate) were not statistically significant (UCSF study $\chi^2 = 2.42$, 2 d.f., $P = .30$; CPDR study, $\chi^2 = 5.52$, 3 d.f., $P = .14$; MA $\chi^2 = 7.94$, 5 d.f., $P = .16$). A test specifically comparing the GPS odds ratio in NCCN intermediate high tumor volume patients to the GPS odds ratio in NCCN intermediate low tumor volume and other NCCN group patients in the CPDR study also found no evidence of interaction ($\chi^2 = 1.06$, 2 d.f., $P = .30$). These tests support the validity of the patient-specific MA with special populations (see Materials and Methods).

Likelihood ratio tests showed that GPS adds significant predictive value for the LFP to each of the 3 clinical risk tools—CAPRA score, NCCN risk group, and AUA/EAU risk group—in each study individually (all $P \leq .002$) and in the MA ($P < .001$; Supplemental Table S3).

LFP Estimated by Patient-specific MA

Within each CAPRA score, NCCN risk group, or AUA/EAU risk group, GPS further discriminated patients' LFP, yielding a wide range of LFP within each group (Fig. 1 and Supplemental Fig. S2). Combining GPS with CAPRA score provided the most individualized LFP estimates and the widest range of LFP among individual patients (Fig. 1A).

Comparison of LFP From Patient-specific MA and Individual Studies

The weighting scheme used to combine the log odds estimates from the 2 studies ensures that patient-specific MA estimates of LFP are more precise than for either of the individual studies.[4] Based on the estimates of the log odds of AP for the patients in the UCSF and CPDR studies,

Table 1. Distribution of baseline characteristics

Characteristic		Study		
		UCSF N = 389	CPDR N = 343	All N = 732
GPS (0-100)	Median	24	30	27
	Min	1	2	1
	Max	65	74	74
Age at diagnosis	Median	58.0	61.7	60.0
	Min	38.0	40.8	38.0
	Max	77.0	75.6	77.0
Race, n (%)	Caucasian	354 (91.0)	260 (75.8)	614 (83.9)
	African American	12 (3.1)	69 (20.1)	81 (11.1)
	Other	23 (5.9)	14 (4.1)	37 (5.1)
Diagnostic PSA ng/mL, n (%)	<4	79 (20.3)	80 (23.3)	159 (21.7)
	4-9.99	261 (67.1)	232 (67.6)	493 (67.3)
	10-20	49 (12.6)	31 (9.0)	80 (10.9)
Clinical T2 vs T1, n (%)	T1	226 (58.1)	236 (68.8)	462 (63.1)
	T2	163 (41.9)	107 (31.2)	270 (36.9)
Biopsy Gleason Score, n (%)	≤3 + 3	297 (76.3)	262 (76.4)	559 (76.4)
	3 + 4	92 (23.7)	81 (23.6)	173 (23.6)
NCCN risk group*, n (%)	Very Low	37 (9.7)	41 (12.3)	78 (10.9)
	Low	189 (49.5)	186 (55.7)	375 (52.4)
	Intermediate	156 (40.8)	107 (32.0)	263 (36.7)
	Total	382	334	716
CAPRA†, n (%)	0	19 (4.9)	15 (4.4)	34 (4.7)
	1	141 (36.3)	129 (37.9)	270 (37.0)
	2	149 (38.3)	105 (30.9)	254 (34.8)
	3	60 (15.4)	62 (18.2)	122 (16.7)
	4	20 (5.1)	24 (7.1)	44 (6.0)
	5	0	5 (1.5)	5 (0.7)
	Total	389	340	729
RP Gleason Score, n (%)	3 + 3	185 (47.6)	195 (56.9)	380 (51.9)
	3 + 4	135 (34.7)	80 (23.3)	215 (29.4)
	Any major pattern 4 or any pattern 5	69 (17.4)	68 (19.8)	137 (18.7)
Pathologic T-Stage, n (%)	T2	308 (79.2)	252 (73.5)	560 (76.5)
	T3a	70 (18.0)	77 (22.4)	147 (20.1)
	T3b	11 (2.8)	14 (4.1)	25 (3.4)

CAPRA, Cancer of the Prostate Risk Assessment; CPDR, Center for Prostate Disease Research; GPS, Genomic Prostate Score; NCCN, National Comprehensive Cancer Network; PSA, prostate-specific antigen; RP, radical prostatectomy; UCSF, University of California at San Francisco.

* NCCN is missing for 16 patients (n = 9 from CPDR study, n = 7 from UCSF study).

† CAPRA is missing for 3 patients from CPDR study.

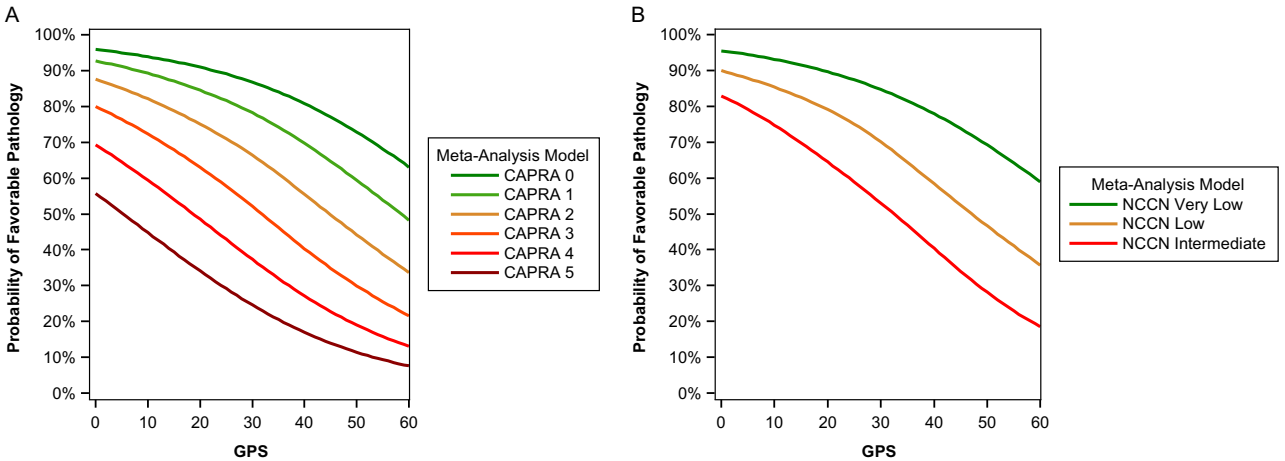

Figure 1. (A) Risk profile plot by patient-specific meta-analysis with GPS plus CAPRA. **(B)** Risk profile plot by patient-specific meta-analysis with GPS plus NCCN. CAPRA, Cancer of the Prostate Risk Assessment; NCCN, National Comprehensive Cancer Network; GPS, Genomic Prostate Score.

patient-specific MA confidence interval widths were narrower than the shortest of the confidence intervals from the 2 studies individually—a median of 24% (range 18%-29%) using GPS and CAPRA score, 23% (range 14%-29%) using GPS and NCCN risk group, and 24% (range 19%-29%) using GPS and AUA/EAU risk group.

The confidence intervals for LFP are asymmetric but have similar properties, with patient-specific MA intervals being tighter than individual study intervals. For illustration, LFP estimates and 95% confidence intervals for hypothetical patients based on both individual studies and the patient-specific MA are shown in Figure 2 and Supplemental

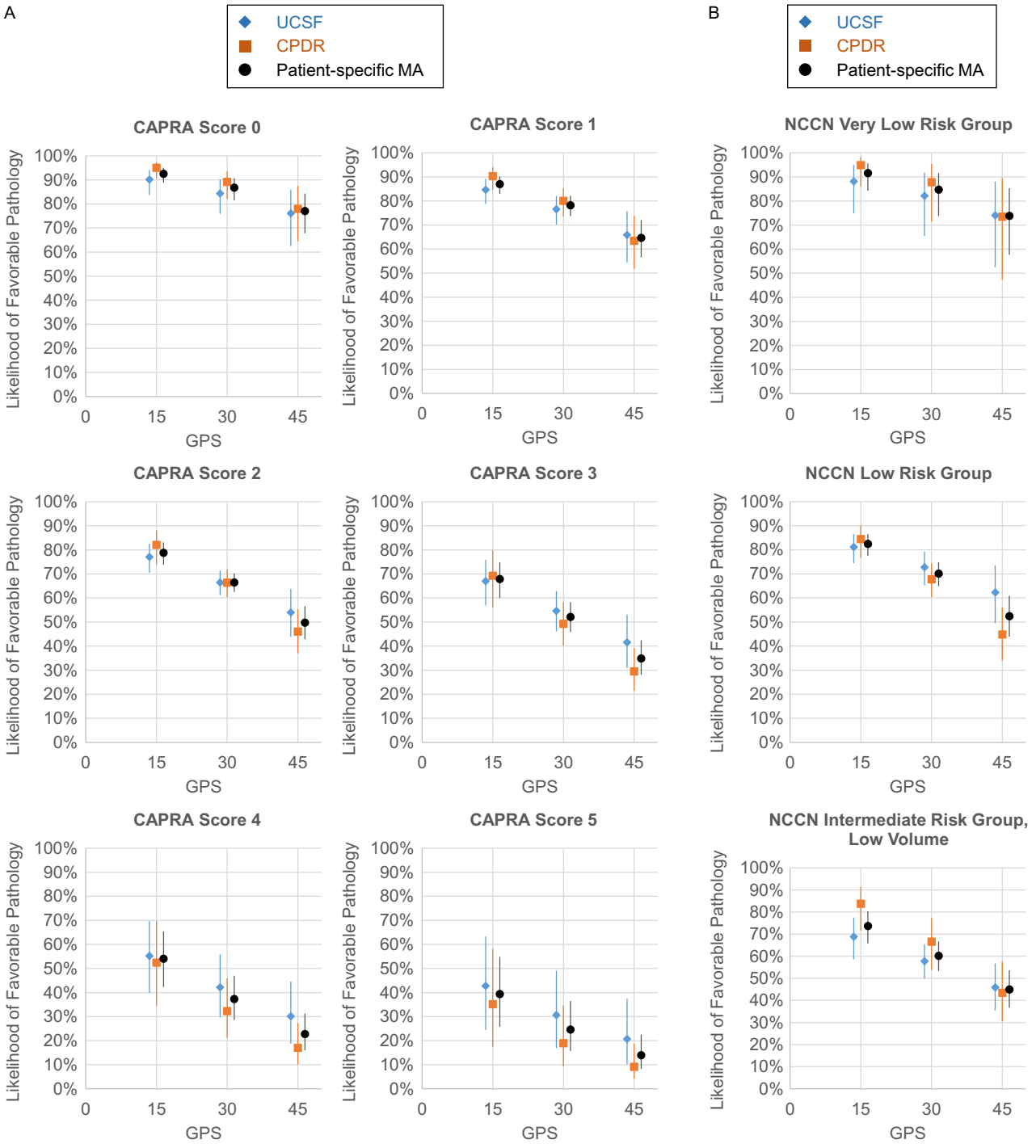

Figure 2. **(A)** Likelihood of favorable pathology for hypothetical individual patients with GPS values of 15, 30, and 45 and different clinical risk assessment tools from the 2 validation studies and the patient-specific meta-analysis for CAPRA. **(B)** Likelihood of favorable pathology for hypothetical individual patients with GPS values of 15, 30, and 45 and different clinical risk assessment tools from the 2 validation studies and the patient-specific meta-analysis for NCCN. CAPRA, Cancer of the Prostate Risk Assessment; CPDR, Center for Prostate Disease Research; GPS, Genomic Prostate Score; MA, meta-analysis; NCCN, National Comprehensive Cancer Network; UCSF, University of California at San Francisco.

Figure S3. Figure 2A compares the LFP estimates for hypothetical patients with low (15), intermediate (30), and high (45) GPS values using CAPRA score, whereas Figure 2B compares the LFP estimates for the same 3 hypothetical patients using the NCCN risk group populations common to the two studies (very low, low, and intermediate with low tumor volume). The comparison of the LFP estimates using AUA/EAU risk groups are in Supplemental Figure S2.

Clinical Significance

The CAPRA score used alone provided the highest ROC curve AUC among the 3 clinical risk assessment tools (Fig. 3C and Supplemental Fig. S4). Using GPS together with each clinical risk assessment tool significantly improved the AUC compared with the clinical risk assessment tool alone. AUC increased from 0.68 to 0.73 by adding GPS to CAPRA score, from 0.64 to 0.70 by adding GPS to NCCN risk group, and from 0.62 to 0.70 by adding GPS to AUA/EAU risk group (all $P < .001$).

In decision-curve analysis (Fig. 3), greater net benefit was shown when combining GPS with the clinical risk assessment tools compared with clinical risk tools alone. Over a wide range of threshold probabilities, incorporation of GPS would be expected to lead to fewer treatments of patients who have FP at RP without increasing the number of patients with AP left untreated.

As an illustration, in these 2 validation studies combined, 11% of all patients were clinically classified as NCCN very low risk, and 90% of these patients had an LFP >80%. In comparison, the proportion of all patients with an MA-estimated LFP >80% was 23% using GPS with NCCN. Using GPS with CAPRA identified the largest proportion—31%—of patients with an MA-estimated LFP >80%.

COMMENTS

The adoption of a prognostic biomarker into clinical practice requires a substantial body of evidence, including analytical validation, clinical validation, and evidence of clinical utility.[14] Clinical validation, in particular, requires that the assay be shown to be a strong predictor of 1 or more clinically meaningful end points in multiple studies using independent patient cohorts that collectively reflect a relevant group of patients with the disease. GPS was validated in 2 separate patient populations of men with localized prostate cancer who were candidates for active surveillance as a predictor of AP at surgery.[2,3] In both studies, the assay was a statistically significant predictor of outcome after adjustment for conventional clinical and pathologic features.

A major challenge of developing a refined risk prediction model based on these studies was how to incorporate the data from both patient cohorts into the prediction. Although the clinical characteristics of the 2 patient cohorts were similar, there were nonetheless some key differences in the representation of different racial groups and higher-risk patients. Not surprisingly, the risk estimates for AP, including the likelihood of high-grade and organ-confined disease, were numerically different in the 2 studies, al-

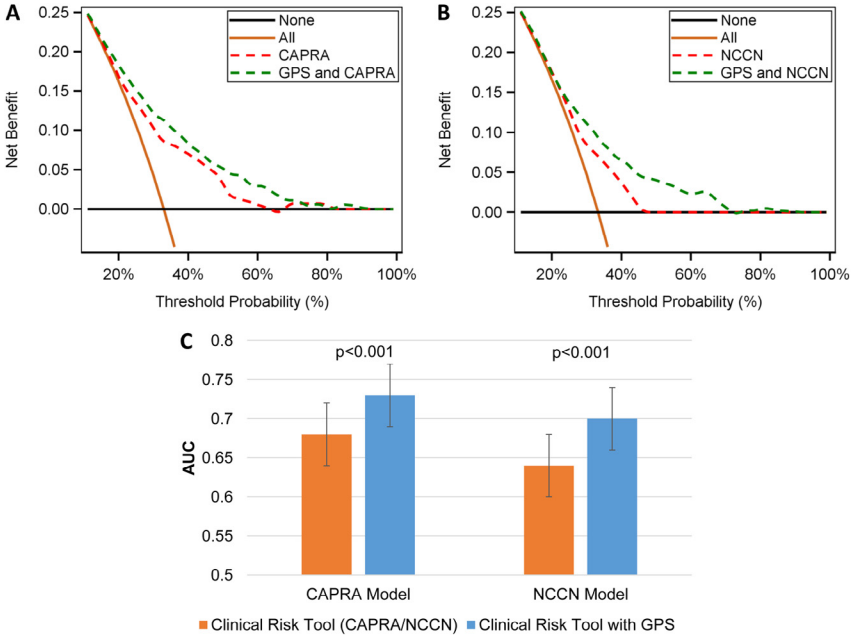

Figure 3. Decision curves from patient-specific meta-analysis based on models of GPS plus clinical risk assessment tools: **(A)** CAPRA, **(B)** NCCN, and **(C)** area under the curve (AUC) of receiver operating characteristics from patient-specific meta-analysis. CAPRA, Cancer of the Prostate Risk Assessment; GPS, Genomic Prostate Score; NCCN, National Comprehensive Cancer Network.

though the confidence intervals overlapped. Patient-specific MA permitted combining prediction information derived from more than 700 patients in the 2 studies to produce more precise risk estimates, with narrower confidence intervals than either study alone.

Another key issue was whether a predictive model could be developed that utilized different clinical risk stratification systems—for example, CAPRA, NCCN, and AUA/EAU—because clinicians use a variety of risk categorizations in their practice. In each of the 2 validation studies, a predictive model that incorporated GPS and NCCN risk groups improved risk stratification compared to NCCN risk assessment alone. In this MA, it was shown that the inclusion of GPS improved risk stratification for each of the 3 clinical risk tools. In addition, the level of discrimination for risk of AP was shown to improve with increasing granularity of the clinical risk tool used. GPS plus CAPRA provided better risk discrimination, followed by GPS plus NCCN risk, then GPS plus AUA/EAU risk group. Finally, the predictive model derived from this MA identifies a much higher fraction of patients with a high LFP (23% GPS + NCCN) than is identified by clinical risk group alone (11% NCCN alone).

CONCLUSION

Use of a patient-specific MA strategy, combining prediction information from 2 independent clinical validation studies, provided a more precise risk stratification for men with newly diagnosed, clinically localized prostate cancer, and showed that GPS adds predictive value to each of the 3 most widely used clinical risk-stratification tools. This new composite model should provide more individualized risk prediction and permit physicians and their patients to make decisions regarding active surveillance vs immediate treatment with greater confidence.

Acknowledgment. The authors thank David G. McLeod and Joel Moncur from Walter Reed National Military Medical Center; Yongmei Chen and Amina Ali from the Center for Prostate Disease Research; Janet E. Cowan and Imelda Tenggara-Hunter from the Department of Urology, University of California, San Francisco and UCSF Helen Diller Family Comprehensive Cancer Center; and the following individuals from Genomic Health: Chandra Arvizu, Athanasios C. Tsiatis, and Dejan Knezevic.

References

1. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst*. 2009;101:1446-1452.
2. Klein EA, Cooperberg MR, Magi-Galluzzi C, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. *Eur Urol*. 2014;66:550-560.
3. Cullen J, Rosner IL, Brand TC, et al. A Biopsy-based 17-gene Genomic Prostate Score predicts recurrence after radical prostatectomy and adverse surgical pathology in a racially diverse population of men with clinically low- and intermediate-risk prostate cancer. *Eur Urol*. 2015;68:123-131.
4. Crager MR, Tang G. Patient-specific meta-analysis for risk assessment using multivariate proportional hazards regression. *J Appl Stat*. 2014;41:2676-2695.
5. Tang G, Cuzick J, Costantino JP, et al. Risk of recurrence and chemotherapy benefit for patients with node-negative, estrogen receptor-positive breast cancer: recurrence score alone and integrated with pathologic and clinical factors. *J Clin Oncol*. 2011;29:4365-4372.
6. Epstein JI, Allsbrook WC Jr, Amin MB, Egevad LL, Committee IG. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma. *Am J Surg Pathol*. 2005;29:1228-1242.
7. 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.
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;173:1938-1942.
9. Prostate Cancer. NCCN Clinical Practice Guidelines for Oncology; 2015.
10. Mottet N, Bellmunt J, Bolla M, et al. EAU guidelines on prostate cancer. Part II: treatment of advanced, relapsing, and castration-resistant prostate cancer. *Eur Urol*. 2011;59:572-583.
11. Thompson I, Thrasher JB, Aus G, et al. Guideline for the management of clinically localized prostate cancer: 2007 update. *J Urol*. 2007;177:2106-2131.
12. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565-574.
13. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837-845.
14. Canfield SE, Kibel AS, Kemeter MJ, Febbo PG, Lawrence HJ, Moul JW. A guide for clinicians in the evaluation of emerging molecular diagnostics for newly diagnosed prostate cancer. *Rev Urol*. 2014;16:172-180.

APPENDIX

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.urology.2015.12.008.