THE JOURNAL of UROLOGY®
www.auajournals.org/journal/juro

OPEN

Prostate Cancer

JU Insight

Development and Validation of a Prostate Biopsy Risk Calculator in Black Men

Neil A. Mistry, Zequn Sun, Jamila Sweis, et al.

Correspondence: Adam B. Murphy (a-murphy2@northwestern.edu).

Full-length article available at https://doi.org/10.1097/JU.0000000000003774.

Study Need and Importance: Black men face higher incidence and mortality than non-Black men with respect to prostate cancer (PCa). Multivariable risk calculators (RCs) were developed to stratify patients undergoing prostate biopsy to reduce unnecessary biopsies and improve detection of clinically significant prostate cancer (csPCa). These calculators have historically relied on homogenous cohorts with few ethnic minorities, with concern for poorer calibration in low-risk Black men.

What We Found: We developed and validated 3 RCs tailored specifically to Black men using 2 cohorts of patients undergoing prostate biopsies in Chicago, Illinois. Using 2 modeling approaches, we created Mistry-Sun (MS) RCs 1 to 3 that use clinical variables likely available in (1) primary care office settings, (2) initial urologist consultation, and (3) subsequent urologist evaluation. We found that our models had an improved or similar area under the curve when compared against several available RCs. Our models showed better calibration for low-risk men resulting in fewer unnecessary biopsies (a combination of benign or Gleason Grade Group 1) than other available calculators with similar and appropriate risks of missed csPCa. The Table demonstrates how MS and other calculators perform in our external validation cohort with respect to missed csPCa and unnecessary biopsies. At lower risk thresholds, MS3 outperforms all other calculators since prostate volume addresses PSA elevation from benign prostatic hyperplasia. MS models 1, 2, and 3 reduce unnecessary biopsies at the 30% risk threshold compared to other RCs with a commensurate proportion of missed csPCa.

Limitations: Our study recruited a relatively small sample size and only within Chicago. We also did not incorporate more modern components of the PCa workup including MRI and biomarkers.

Interpretation for Patient Care: This is the first study to develop and validate a PCa risk-stratification tool specifically for Black men. Our tailored RCs can reduce unnecessary biopsies for Black men and addresses access disparities in precision oncology tools (Table).

Table. *Comparison of Biopsy Outcomes at 10% and 30% Risk Thresholds in the External Validation Cohort (N = 276)*

	Mistry-Sun 1	Mistry-Sun 2	Mistry-Sun 3	PCPT	PBCG	Kaiser
10% biopsy threshold						
Men below 10% threshold, No.	6	1	30	15	12	0
Total biopsies performed, No.	270	275	246	261	264	276
Missed GG2-5 among men below the 10% threshold, No. (%)	0 (0)	0 (0)	0 (0)	3 (20.0)	3 (25)	0 (0)
Unnecessary biopsies, No.	169	174	145	163	166	175
30% biopsy threshold						
Men below 30% threshold, No.	202	219	183	172	137	17
Total biopsies performed, No.	74	57	93	104	139	259
Missed GG2-5 among men below the 30% threshold, No. (%)	55 (27.2)	60 (27.4)	42 (23.0)	46 (26.7)	31 (22.6)	4 (23.5)
Unnecessary biopsies, No.	28	16	34	49	69	162

Abbreviations: GG, Gleason Grade Group; Kaiser, Kaiser Permanente Prostate Cancer risk calculator; PBCG, Prostate Biopsy Collaborative Group risk calculator; PCPT, Prostate Cancer Prevention Trial risk calculator version 2.0.
For men willing to undergo prostate biopsy if there is 10% or 30% risk of GG2 to GG5 prostate cancer, we list the number of men who fall below the threshold (tests negative), the total number of biopsies performed (tests positive), the number of missed GG2 to GG5 cancers (false-negatives, which should be $\leq$ 10% or $\leq$ 30% from the patient's perspective), and the number of unnecessary biopsies performed for men with GG1 prostate cancer or a negative biopsy. There were complete data in 276 out of the 292 men.

THE JOURNAL OF UROLOGY®

https://doi.org/10.1097/JU.0000000000003774
Vol. 211, 223-233, February 2024
Printed in U.S.A.

THE JOURNAL of UROLOGY®
www.auajournals.org/journal/juro

OPEN

Development and Validation of a Prostate Biopsy Risk Calculator in Black Men

Neil A. Mistry,[1]* Zequn Sun,[2]* Jamila Sweis,[1] Cordero McCall,[3] Norma Marshall,[1] Bernice Ofori,[1] Courtney M. P. Hollowell,[4] Rick A. Kittles,[5] Edward M. Schaeffer,[1] Michael Abern,[6] Peter Gann,[7] and Adam B. Murphy[1,4]

[1]Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, Illinois
[2]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois
[3]Medical College of Wisconsin, Milwaukee, Wisconsin
[4]Division of Urology, Department of Surgery, Cook County Health, Chicago, Illinois
[5]Morehouse School of Medicine, Atlanta, Georgia
[6]Department of Urology, University of Illinois at Chicago, Chicago, Illinois
[7]Department of Pathology, University of Illinois at Chicago, Chicago, Illinois

Purpose: We sought to develop and validate a prostate biopsy risk calculator for Black men and compare it with the Prostate Cancer Prevention Trial version 2.0, Prostate Biopsy Collaborative Group, and Kaiser Permanente Prostate Cancer Risk Calculators for the detection of Gleason Grade Group (GG) $\geq$ 2 prostate cancer (PCa).

Materials and Methods: We prospectively recruited 2 cohorts of men undergoing prostate biopsy from 5 facilities in Chicago. The first cohort was split into development (70%) and internal validation (30%) groups. The second was used for external validation. Iterative logistic regression was used to develop 3 models for predicting GG $\geq$ 2 PCa. Models were compared for discrimination using the C statistics, calibration curves, and net benefit curves. The frequency of unnecessary biopsies and missed PCas was compared at 10% and 30% risk thresholds.

Results: The 2 cohorts included 393 and 292 Black men, respectively. Our first model, Mistry-Sun 1, used serum PSA and prior negative biopsy. Mistry-Sun 2 added abnormal digital rectal exam (DRE) and an interaction term with abnormal DRE and PSA to Mistry-Sun 1. Mistry-Sun 3 added prostate volume, abnormal DRE, and age to Mistry-Sun 1. The C statistics were 0.74, 0.74, and 0.78, respectively, and were similar to or higher than established calculators. At the 10% and 30% risk thresholds our models had the fewest unnecessary biopsies and an appropriate proportion of missed GG $\geq$ 2 PCas.

Submitted February 5, 2023; accepted October 26, 2023; published November 2, 2023.

Support: This study was supported by the US Department of Defense (Prostate Cancer Research Program) Grants W81XWH-17-0608 and W81XWH-21-1-0850 (A.B.M., P.G.), Northwestern University Grant T37 MD014248 (C.M.), Exact Sciences (A.B.M.), and Biotechne/Exosome Diagnostics (A.B.M.).

Conflict of Interest Disclosures: E.M.S.: Pfizer, Astellas, Lantheus. All other authors declare that they have no competing interests as defined by *The Journal of Urology*®, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Ethics Statement: This study received institutional Review Board approval (IRB No. STU00205089).

Author Contributions: Data analysis and interpretation: A.B.M., P.G., E.S., N.A.M., N.M., Z.S., C.M.; Data acquisition: A.B.M., E.S., P.G., J.S., B.O., C.H., R.K., M.A.; Drafting the manuscript: N.A.M., N.M., C.M., Z.S., A.B.M.; Substantial contributions to the conception or design of the work: A.B.M., E.S., P.G.; Provided the edits: N.A.M., C.H., R.K., E.S., P.G., A.B.M. All authors have read and approved the final manuscript.

Data Availability: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Corresponding Author: Adam B. Murphy, MD, MBA, MSCI, Department of Urology, Northwestern University Feinberg School of Medicine, Arkes Pavilion, Suite 2300, 676 N Saint Clair St, Chicago, IL 60611 (a-murphy2@northwestern.edu).

* Co-first authors.

Conclusions: Tailoring a risk calculator to detect clinically significant PCa in Black men may improve biopsy decision-making and outcomes compared to tools developed in non-Black populations.

Key Words: prostate cancer, risk calculator, disparities, prostate biopsy, risk prediction

Prostate cancer (PCa) is the leading malignancy among US men and the second leading cause of cancer-related deaths in the US.[1] Black men have 70% higher PCa incidence and twice the mortality rate of White men.[2] PSA is the most widely used biomarker for PCa screening and reduces mortality by about 50%.[3,4] But its low specificity results in excessive unnecessary prostate biopsies and risk of pain, bleeding, urinary retention, and sepsis.[5,6] As a biomarker, PSA does not distinguish benign prostatic hyperplasia, prostatitis, high-grade prostatic intraepithelial neoplasia and PCa, especially in the diagnostic gray zone of 2 to 10 ng/mL. PSA cannot reliably discriminate clinically significant high-grade PCa Gleason Grade Group (GG) 2 to 5 from low-grade disease (GG1).

Multivariable risk calculators (RCs) have been developed to stratify patients undergoing prostate biopsy to improve the prediction of GG $\geq$ 2 PCa while reducing unnecessary biopsies and the overdiagnosis of GG1 PCa. These RCs are based on large European or American cohorts, such as the Prostate Cancer Prevention Trial version 2.0 (PCPT)[7] and the Prostate Biopsy Collaborative Group (PBCG),[8] which included few ethnic minorities. This homogeneity can lead to poor calibration in populations with differing PCa prevalence such as Chinese[9] or Korean[10] men. In Black men, our prior work has demonstrated the PBCG RC had poorer calibration in Black men in the lower range of risk, resulting in unnecessary biopsies.[6] Contemporary RCs have tried to mitigate these issues with diverse patient cohorts to better predict PCa in minority patients. One example is the Kaiser Permanente Prostate Cancer (Kaiser) RC, whose population includes 10.3% African Americans.[11] Another approach is to tailor RCs to men with differing prevalence to improve the prediction of GG $\geq$ 2 PCa while reducing unnecessary biopsies. This was demonstrated in the Korean[12] and the Huashan[13] RCs tailored to native Korean and Chinese populations, respectively.

Our primary objective was to develop and validate an RC for Black men undergoing prostate biopsy to predict GG $\geq$ 2 PCa while reducing unnecessary biopsies and compare it with established RCs. Given the proliferation of biomarkers and multiparametric MRI in biopsy decision making and accuracy, future iterations of the calculator will incorporate these tools.

MATERIALS AND METHODS

Patient Recruitment

This study used 2 separate prospectively recruited cohorts. One cohort was split into the development and internal validation groups, and a second cohort was used for external validation. Both were recruited from men referred for abnormal PSA or digital rectal exam (DRE) seen at Urology clinics in Chicago, Illinois.

The first cohort was recruited between 2009 and 2014 from the following 5 urology clinics: Northwestern Memorial Hospital, Cook County Health, University of Chicago, Jesse Brown Veteran's Administration Medical Center, and University of Illinois at Chicago. These men were recruited as part of an R01-funded study evaluating serum vitamin D levels and PCa risk (IRB No. STU00005398).

The second cohort was recruited using a DOD-funded protocol from 2017 - 2021 from 3 Chicago-institutions (Northwestern Memorial Hospital, Cook County Health, and University of Illinois at Chicago) to externally validate this RC (IRB No. STU00205089).

Inclusion and Exclusion Criteria

Eligible men must have a PSA < 50 ng/dL, self-report as Black race, and undergo a systematic $\geq$ 10 core transrectal ultrasound (TRUS)–guided prostate biopsy. The exclusion criteria included prior diagnosis of PCa or pelvic radiation. Since prostate MRI did not become available at our institution until after recruitment began and is underutilized in Black men nationally, patients who underwent prostate MRI were excluded from the study.[14,15]

Pathology

The pathology was reviewed by each institutions' uropathologist who assigned Gleason score using the 2005 and the 2014 International Society of Urological Pathology Guidelines.[16]

Clinical Data

Data were manually collected through electronic medical records and a coordinator administered the demographic and medical history questionnaire. Various mathematical forms were assessed for their association with GG2-5 PCa and the form with the strongest association by *P* value were preferred. The continuous variables included age (years), BMI (kg/m^2), prebiopsy PSA (ng/mL), prostate volume (cm^3) by TRUS, number of prior negative biopsies, and PSA density (ng/mL/cm^3).

The categorical predictors include history of prior negative prostate biopsy (0 vs $\geq$1), first-degree family history of PCa (yes vs no/unknown), abnormal DRE (yes vs no/not palpable), and use of 5-alpha reductase inhibitors in the past 6 months (yes vs no).

Missing Data

Complete case analysis was undertaken. Cases were excluded if data for any of the relevant predictor variables were missing.

Outcomes of Interest

The outcome of interest for the RC is the uropathologists' interpretation of the prostate biopsy modeled as GG $\geq$ 2 PCa vs GG1 PCa or negative biopsy/atypical small acinar proliferation. We followed the guidelines from the Transparent Reporting of Multivariable Prediction Models for Individual Prognosis or Diagnosis.[17] The models were developed to detect GG $\geq$ 2 PCa vs GG1 or negative biopsy/atypical small acinar proliferation. We calculated the predicted probabilities for GG $\geq$ 2 PCa using the formula provided in the supplemental data in the initial validation for the Kaiser RC and using the R code kindly provided by the first author for the initial validations of the PCPT and PBCG RCs.

Statistical Analysis

Clinical and Demographic Factors. For comparisons of clinical variables across biopsy outcomes, we used Kruskal-Wallis tests to compare medians for continuous variables and χ^2 tests for categorical variables. χ^2 trend tests were used to compare the distributions of nonbinary categorical variables.

Model Development and Internal Validation. We created 3 models similar to the Kaiser RC to mimic what is available to primary care providers, to urologists on the first visit, and to urologists with a prostate volume measured from prior imaging. We split the first patient cohort by randomly assigning participants into the development (n = 275, 70%) and internal validation (30%) groups. The variables of interest tested for use within our models were selected based on literature review and prior risk nomograms. PSA was log-transformed (base 2) as in the previous RCs. We created a binary logistic regression model for GG $\geq$ 2 PCa vs GG1 and negative biopsy. We then performed a univariate analysis of each independent variable, selecting the mathematical forms of the variables that had the strongest univariate associations with GG $\geq$ 2 PCa or were included in prior iterations of RCs. We also tested interaction terms between each predictor and used forward and backwards selection for the core predictive variables and tested all combinations for interaction terms in the Development cohort. The top 5 models that reported the highest C statistics for "GG $\geq$ 2 PCa" vs "GG1 PCa or negative biopsy" were tested in the internal validation cohort. Predictive discrimination was examined using C statistics; models with the lowest Akaike Information Criterion score were preferred in cases of ties.[18]

External Validation. The performance of our models was compared with the PCPT, PBCG, and Kaiser RCs using C statistics and corresponding 95% confidence intervals of "GG $\geq$ 2 PCa" vs "GG1 PCa or negative biopsy"; *P* values were calculated with DeLong's test[19] for external validation. We computed the individual-level predicted risk of GG $\geq$ 2 PCa in our 3 models and the PCPT, PBCG, and Kaiser RCs (model 2). We used the actual biopsy outcome of each participant to compute the discrimination (C statistic) and the corresponding 95% confidence intervals.

To produce output like the established RCs, the covariates from the binary logistic models were used to create multinomial logistic regression models with 3 prostate cancer outcome categories: GG $\geq$ 2 PCa, GG1 PCa, and negative biopsy with negative biopsy as the reference group. The predicted probabilities of GG $\geq$ 2 PCa for each patient in the multinomial and binomial versions of the models were directly compared to assure concordance.

Calibration. The biopsy outcomes were used to create calibration curves that graphed the RC models' predicted proportion of GG $\geq$ 2 PCa vs the actual proportion of GG $\geq$ 2 PCa in men in each group in 10% intervals. The Hosmer-Lemeshow test was used to evaluate goodness-of-fit between 0% to 30% to compare each model's calibration for low-risk men ($\leq$30%) to aid in avoidance of unnecessary biopsies. Thirty percent reflects real-world risk thresholds for biopsy and the average risk of GG $\geq$ 2 PCa on prostate biopsy.[20]

Net Benefit Curves. Net benefit curves were constructed from the calculated net benefit reported across probabilities by increments of 10% up to 100%, defined as the minimum probability of disease at which further intervention would be warranted, as net benefit = sensitivity $\times$ prevalence $-$ (1 $-$ specificity) $\times$ (1 $-$ prevalence) $\times$ w, where w is the odds ratio at the threshold probability.[21]

Biopsy Outcomes Comparison. Since patients have different risk thresholds for choosing to undergo biopsy with a median of 25% (IQR 10%-50%),[20,22] we evaluated the biopsy outcomes at the 10% and 30% risk thresholds.

RESULTS

Clinical and Demographic Factors

Black men in this study self-identified as African American, West African, Central African, and Afro-Caribbean. The development and internal validation cohort included 437 Black men and included 393 eligible men in the study. Among the 393 men, there were 123 negative biopsies, 148 GG1 PCas, and 122 GG $\geq$ 2 PCas. The external validation cohort included 402 Black men; 292 were eligible; 110 were excluded for prebiopsy MRI. Of 292 men in the external validation cohort, there were 121 negative biopsies, 66 GG1 PCas, and 105 GG $\geq$ 2 PCas. Tables 1 and 2 show the clinical characteristics of both cohorts. Supplemental Table 1 (https://www.jurology.com) compares both cohorts.

The Mistry-Sun Models

Three binary logistic models were developed, Mistry-Sun 1-3, with negative biopsy + GG1 PCa as the reference category for predicting GG $\geq$ 2 PCa.

The C statistics and 95% confidence intervals for the binary outcome of GG $\geq$ 2 PCa vs negative biopsy + GG1 PCa for the Mistry-Sun models and the PCPT, PBCG, and Kaiser RCs are displayed in Table 3. The covariates from the binary models were used to create multinomial versions with negative biopsy as the referent category for both GG1 and GG $\geq$ 2 PCa. The multinomial regression models' odds ratios with 95% confidence intervals for each model

Table 1. *Patient Characteristics of the Development and Internal Validation Cohort (n = 393)*

	Negative biopsy (N = 123)	GG1 PCa (N = 148)	GG2-5 PCa (N = 122)	*P* value[a]
Continuous variables				
Age				.69
Median (IQR), y	62.0 (56.0, 67.0)	60.0 (56.0, 66.0)	60.50 (58.0, 65.8)	
Missing	0	0	0	
BMI				.73
Median (IQR), kg/m^2	27.50 (24.6, 31.4)	28.30 (25.2, 32.3)	28.00 (25.0, 31.3)	
Missing	4	3	5	
PSA				**< .01**
Median (IQR), ng/mL	7.0 (5.1, 10.9)	6.5 (5.0, 10.1)	9.2 (5.6, 20.0)	
Missing	0	0	0	
Prostate volume[b]				**< .01**
Median (IQR), cm^3	51.1 (36.8, 76.1)	37.2 (29.4, 50.9)	35.00 (25.4, 50.0)	
Missing	4	4	7	
Total cores				.98
Median (IQR)	12.0 (12.0, 12.0)	12.0 (12.0, 12.0)	12.0 (12.0, 12.0)	
Missing	0	0	0	
Categorical variables				
Abnormal DRE[c]				**< .01**
Yes, No. (%)	36 (29.3)	25 (16.9)	50 (41.0)	
Missing	0	0	0	
PCa family history[d]				
Yes, No. (%)	18 (14.6)	35 (23.6)	21 (17.2)	.28
Missing	0	0	0	
Prior biopsies, No. (%)[e]				.23
0	86 (69.9)	129 (87.2)	110 (90.2)	
1	19 (15.4)	12 (8.1)	5 (4.1)	
2+	4 (3.3)	5 (3.4)	3 (2.5)	
Missing	9 (7.3)	1 (0.7)	3 (2.5)	
Publicly funded site, No. (%)				.35
Yes	102 (82.9)	120 (81.1)	91 (74.6)	
Missing	1 (0.8)	0 (0)	0 (0)	
5-ARI use, No. (%)				**.03**
Yes	20 (16.3)	10 (6.8)	8 (6.6)	
Missing	0	0	0	

Abbreviations: ARI, alpha reductase inhibitor; BMI, body mass index; DRE, digital rectal examination; GG, Gleason Grade Group; IQR, interquartile range; PCa, prostate cancer; PSA, prostate-specific antigen.
[a] Kruskal-Wallis tests were used to compare medians for continuous variables, and χ^2 tests were used for categorical variables; *P* values < .05 are bolded to highlight statistical significance.
[b] Transrectal ultrasound derived prostate volume using the ellipsoid formula.
[c] Abnormal DRE: digital rectal examination coded as abnormal vs normal/not palpable/not performed.
[d] Family history: patient has 1 or more first-degree relatives with prostate cancer vs no or unknown family history.
[e] Prior biopsies: 1 or more prior negative prostate biopsies.

for GG $\geq$ 2 PCa are in Table 4. When comparing the predicted probabilities for each patient in the multinomial and binomial versions of the models, the mean difference was less than 1% between the predicted probabilities for GG $\geq$ 2 PCa.

Internal Validation Cohort

The models were internally validated for discrimination in the Vitamin D cohort. The Mistry-Sun 1 model includes $\log_2$(PSA) and history of prior negative biopsy. The C statistic for this model was 0.68 for GG $\geq$ 2 PCa. This model excluded DRE and volume estimates to mimic data in the primary care setting. Our second model, Mistry-Sun 2, adds abnormal DRE and an interaction term between abnormal DRE and $\log_2$(PSA) to the Mistry-Sun 1 model to represent the first urology consultation for elevated PSA. The C statistic was 0.69. The Mistry-Sun 3 model adds prostate volume, age, and first-degree family history of PCa to the Mistry-Sun 1 model, with a C statistic of 0.72.

External Validation Cohort

We used C statistics to compare across the 3 Mistry-Sun and the established models in external validation. Mistry-Sun 3 performed the best with a C statistic of 0.78. For volume-free models, the C statistic for Mistry-Sun 2 (0.74) was higher than those for the other RCs such as PCPT (0.71), PBCG (0.73), and Kaiser (0.72); but discrimination was statistically significant compared against PCPT. Table 3 contains a comparison of the C statistics. Receiver Operative Characteristics Curves were generated to show that the Mistry-Sun RCs had a similar tradeoff between sensitivity and specificity relative to the established RCs for the binary logistic regression models (see Supplemental Figure 1, https://www.jurology.com).

Calibration

Using risk thresholds between 0% to 30% for risk of GG $\geq$ 2 PCa, which is consistent with patient risk thresholds and average risk of GG $\geq$ 2 PCa, the Mistry-Sun models demonstrate better calibration via the Hosmer-

Table 2. *Patient Characteristics of the External Validation Cohort (n = 292)*

	Negative biopsy (N = 121)	GG1 PCa (N = 66)	GG2-5 PCa (N = 105)	*P* value[a]
Continuous variables				
Age				.43
Median (IQR), y	61.0 (56.0, 65.0)	63.0 (59.0, 66.0)	61.0 (57.0, 67.0)	
Missing	0	0	0	
BMI				.37
Median (IQR), kg/m^2	29.0 (25.5, 32.9)	30.0 (26.9, 34.3)	29.4 (25.0, 31.9)	
Missing	0	0	1	
PSA				**< .01**
Median (IQR), ng/mL	5.3 (3.7, 7.6)	6.3 (4.6, 9.9)	9.1 (6.1, 14.7)	
Missing	0	0	0	
Prostate volume[b]				**< .01**
Median (IQR), cm^3	51.0 (31.8, 67.3)	40.2 (27.8, 50.6)	32.4 (22.0, 46.0)	
Missing	0	2	3	
Total cores				.66
Median (IQR)	12.0 (12.0, 12.0)	12.0 (12.0, 12.0)	12.0 (12.0, 12.0)	
Missing	0	0	0	
Categorical variables				
Abnormal DRE[c]				**.02**
Yes, No. (%)	20 (16.5)	11 (16.7)	35 (33.3)	
Missing	1	1	0	
PCa family history[d]				.33
Yes, No. (%)	23 (19.0)	19 (28.8)	19 (18.1)	
Missing	5	2	0	
Prior biopsies, No. (%)[e]				.25
0	106 (87.6)	62 (93.9)	99 (94.3)	
1	10 (8.3)	2 (3.0)	2 (1.9)	
2+	5 (4.1)	0 (0)	1 (1.0)	
Missing	0	2	3	
Publicly funded site, No. (%)				.11
Yes	14 (11.6)	6 (9.1)	22 (21.0)	
Missing	0	0	0	
5-ARI use, No. (%)				.07
Yes	10 (8.3)	1 (1.5)	2 (1.9)	
Missing	0	0	0	

Abbreviations: ARI, alpha reductase inhibitor; BMI, body mass index; DRE, digital rectal examination; GG, Gleason Grade Group; IQR, interquartile range; PCa, prostate cancer; PSA, prostate-specific antigen.
[a] Kruskal-Wallis tests were used to compare medians for continuous variables, and χ^2 tests were used for categorical variables; *P* values < .05 are bolded to highlight statistical significance.
[b] Transrectal ultrasound derived prostate volume using the ellipsoid formula.
[c] Abnormal DRE: digital rectal examination coded as abnormal vs normal/not palpable/not performed.
[d] Family history: patient has 1 or more first-degree relatives with prostate cancer vs no or unknown family history.
[e] Prior biopsies: 1 or more prior negative prostate biopsies.

Lemeshow Goodness-of-Fit test than the PCPT, PBCG, and Kaiser RCs (Figure 1). The *P* values for the Hosmer-Lemeshow test for Mistry-Sun 1 (0.092), Mistry-Sun 2 (0.166), and Mistry-Sun 3 (0.053) all exceed 0.05 at the predicted risk cutoff, which fails to reject the null hypothesis that the observed rates are the same as the predicted rates across 5 predicted probability intervals from 0% to 30%. The PCPT (0.021), PBCG (0.046), and Kaiser (0.003) RCs have *P* values $< .05$ and suggest that the predicted probabilities are different than the observed probabilities of GG $\geq$ 2 PCa.

Net Benefit Curves

The net benefit curves shows that the 3 Mistry-Sun models have similar net benefit to the other established RCs and are superior to the biopsy-all strategy (Figure 2).

Biopsy Outcomes Comparison

To compare the hypothetical outcomes of the different RC thresholds, we modeled the number of biopsies performed, missed GG $\geq$ 2 PCa diagnoses, and the number of unnecessary biopsies (negative or GG1) using the models' predicted probability of GG $\geq$ 2 PCa. Table 5 summarizes those findings. The Mistry-Sun models identify more men as low risk below both thresholds and has a lower number of unnecessary prostate biopsies than other RCs. Mistry-Sun models 1, 2, and 3 reduce unnecessary biopsies at the 30% risk threshold compared to all of the other RCs with a commensurate proportion of missed high-grade cancer. The percentage of men with missed GG $\geq$ 2 PCa was appropriate for men with thresholds of less than 10% (Mistry-Sun models' mean percentage of missed GG $\geq$ 2 PCa was 0%) and 30% (Mistry-Sun models' mean was 25.9%). Calculated sensitivity and specificity for all models is reported in Supplemental Table 2 (https://www.jurology.com).

DISCUSSION

Black men face a higher risk of being diagnosed with, and dying from PCa than their non-Black

Table 3. *Comparison of the 3 Mistry-Sun Models With the Established Risk Calculators by Their C Statistics in the External Validation Cohort*

Model	Internal validation	External validation						
	Mistry-Sun models C statistic (95% CI)	Mistry-Sun models C statistic (95% CI)	PCPT		PBCG		Kaiser	
			C statistic (95% CI)	*P* value[a]	C statistic (95% CI)	*P* value[a]	C statistic (95% CI)	*P* value[a]
Mistry-Sun 1: Log_2(PSA) + prior negative biopsy	0.677 (0.58, 0.79)	0.744 (0.69, 0.80)	0.711 (0.68, 0.80)	.048	0.728 (0.67, 0.79)	.22	0.716 (0.65, 0.78)	.18
Mistry-Sun 2: Log_2(PSA) + prior negative biopsy + abnormal DRE + log_2(PSA) × abnormal DRE	0.698 (0.62, 0.82)	0.743 (0.68, 0.80)	0.711 (0.68, 0.80)	.06	0.728 (0.67, 0.79)	.18	0.716 (0.65, 0.78)	.14
Mistry-Sun 3: Log_2(PSA) + prior negative biopsy + prostate volume + abnormal DRE + prostate cancer family history + age	0.715 (0.62, 0.86)	0.783 (0.73, 0.84)	0.711 (0.68, 0.80)	.003	0.728 (0.67, 0.79)	.006	0.716 (0.65, 0.78)	.006

Abbreviations: CI, confidence interval; DRE, digital rectal examination; Kaiser, Kaiser Permanente Prostate Cancer risk calculator; PBCG, Prostate Biopsy Collaborative Group risk calculator; PCPT, Prostate Cancer Prevention Trial risk calculator version 2.0; PSA, prostate-specific antigen.
Prior negative biopsy was coded as 0 vs 1 or more.
[a] *P* value of each established risk calculator compared to Mistry-Sun Models using the DeLong test.

counterparts, but the risk is not uniformly distributed. There is a demonstrable need for risk stratification tools to directly address the disparity and more accurately identify the men who may benefit most and least from a biopsy. The established RCs have poorer calibration in low-risk ranges for Black men. In our prior publication, we showed that the PBCG overestimated risk based on calibration. In our current publication, we see that PBCG and PCPT underestimated risk while Kaiser overestimated it using the calibration plots. This is likely due to the validation cohort having inherently higher risk as evidenced by higher rates of overall and high-grade prostate cancer.

To our knowledge, this is the first study to develop and validate a tailored risk-calculator in a cohort of Black men undergoing prostate biopsy. The most important clinical factors in Mistry-Sun 2 (lacking prostate volume) are PSA level, prior negative biopsy, abnormal DRE, and an interaction term between PSA and abnormal DRE. This model had better calibration and correspondingly identified more men who could avoid biopsy. All 3 Mistry-Sun models were similar or outperformed the established models in discrimination. This performance was only statistically significant for Mistry-Sun 1 and Mistry-Sun 2 vs PCPT. Mistry-Sun models were better calibrated, based on the Hosmer-Lemeshow test at lower risk thresholds, resulting in fewer unnecessary biopsies. While the Mistry-Sun models underpredict risk at predicted probabilities of 40% or higher, we believe this exceeds most urologists' and patients' probability thresholds and will lead to recommending a biopsy. The impact of the improved calibration becomes especially clear when modeling the numbers of unnecessary biopsies, where Mistry-Sun 2 has the lowest number of unnecessary biopsies at a higher

Table 4. *Odds Ratios and 95% Confidence Intervals for Each Covariate in the Mistry-Sun Risk Calculator Models*

Covariates	Mistry-Sun 1		Mistry-Sun 2		Mistry-Sun 3	
	GG1 vs no cancer	GG2-5 vs no cancer	GG1 vs no cancer	GG2-5 vs no cancer	GG1 vs no cancer	GG2-5 vs no cancer
Log_2(PSA)	1.027 (0.778, 1.357)	1.667 (1.233, 2.254)	0.922 (0.656, 1.297)	1.264 (0.856, 1.866)	1.270 (0.926, 1.741)	1.883 (1.331, 2.664)
Prior negative biopsy, yes vs no (ref)	0.613 (0.293, 1.284)	0.432 (0.178, 1.045)	0.571 (0.271, 1.205)	0.461 (0.189, 1.122)	0.504 (0.225, 1.125)	0.363 (0.136, 0.965)
Abnormal digital rectal exam, yes vs no or not palpable (ref)	-	-	0.160 (0.021, 1.194)	0.285 (0.031, 2.652)	0.460 (0.204, 1.038)	1.725 (0.792, 3.755)
Prostate cancer family history, yes vs no/ unknown (ref)	-	-	-	-	2.912 (1.281, 6.621)	1.548 (0.584, 4.101)
Age, y	-	-	-	-	1.036 (0.99, 1.083)	1.061 (1.008, 1.116)
Prostate volume,[a] cm^3	-	-	-	-	0.974 (0.963, 0.986)	0.965 (0.950, 0.980)
Log_2(PSA) × abnormal digital rectal exam	-	-	1.418 (0.762, 2.639)	1.738 (0.907, 3.332)	-	-

Abbreviations: GG, Gleason Grade Group; PSA, prostate-specific antigen; ref, reference group.
The models are for 3 different clinical settings with varied data availability. Category of "no cancer" includes men with atypical small acinar proliferation who were not biopsied again.
[a] Prostate volume is measured by the transrectal ultrasound ellipsoid formula.

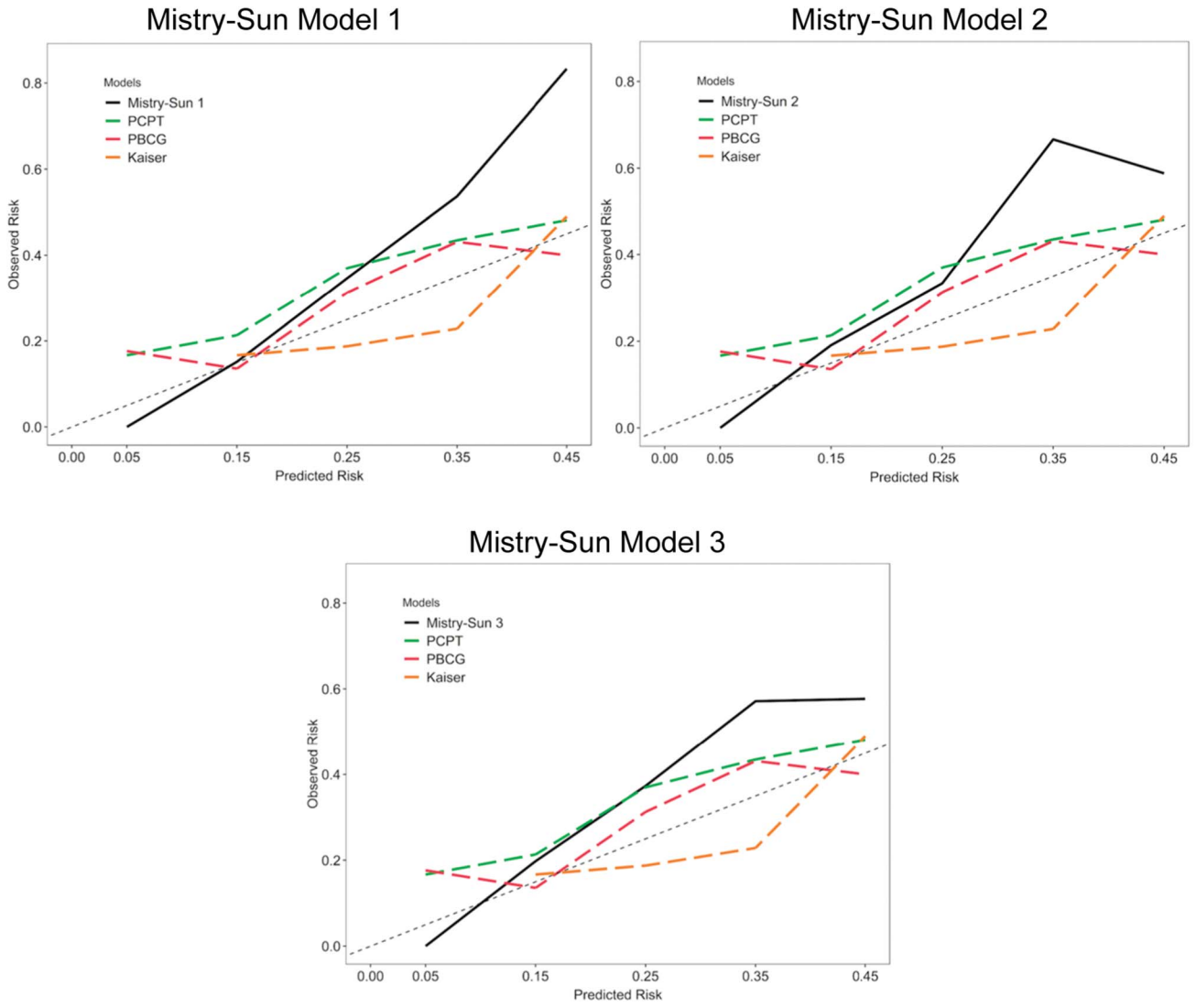

Figure 1. The calibration plots from 0% to 45% risk threshold for each of the Mistry-Sun models compared against the Prostate Cancer Prevention Trial risk calculator version 2.0 (PCPT), the Prostate Biopsy Collaborative Group risk calculator (PBCG), and the Kaiser Permanente Prostate Cancer risk calculator (Kaiser).

risk threshold for missed GG $\geq$ 2 PCa (30%) compared to all the other models while maintaining an appropriate number of missed GG $\geq$ 2 cancers.

Mistry-Sun 3 included volume and was anticipated to improve performance over the comparison models that did not include volume. Mistry-Sun model 3 reduces unnecessary biopsies at the 10% risk threshold when compared to Kaiser (another volume-based calculator), without missing any high-grade cancer. These volumes were derived from TRUS measurements at time of biopsy, which can differ from those on MRI especially at larger volumes and may not be available in biopsy-naïve men.[23] It is well documented that larger glands can produce more PSA, which is why PSA density can risk-stratify patients. Furthermore, MRI can provide a noninvasive measurement of volume for patients with high PSA values who otherwise have no prior known volume and would end up with a TRUS and prostate biopsy.[15] This may strengthen the rationale for prebiopsy prostate MRI for Black men.

The biggest strength of our study was the prospective cohorts of Black men. Our recruitment featured both publicly funded (University of Illinois at Chicago, Jesse Brown Veteran's Administration Medical Center, and Cook County Health) and private (Northwestern Medicine and University of Chicago) institutions, capturing the socioeconomic diversity of Chicago. Additionally, our cohorts reflect contemporary practice by selecting men referred to a urologist for abnormal PSA or DRE and biopsied using a standard 12-core approach.

Among our limitations are the relatively small sample size and geographic homogeneity. Another limitation is the lack of a centralized pathological

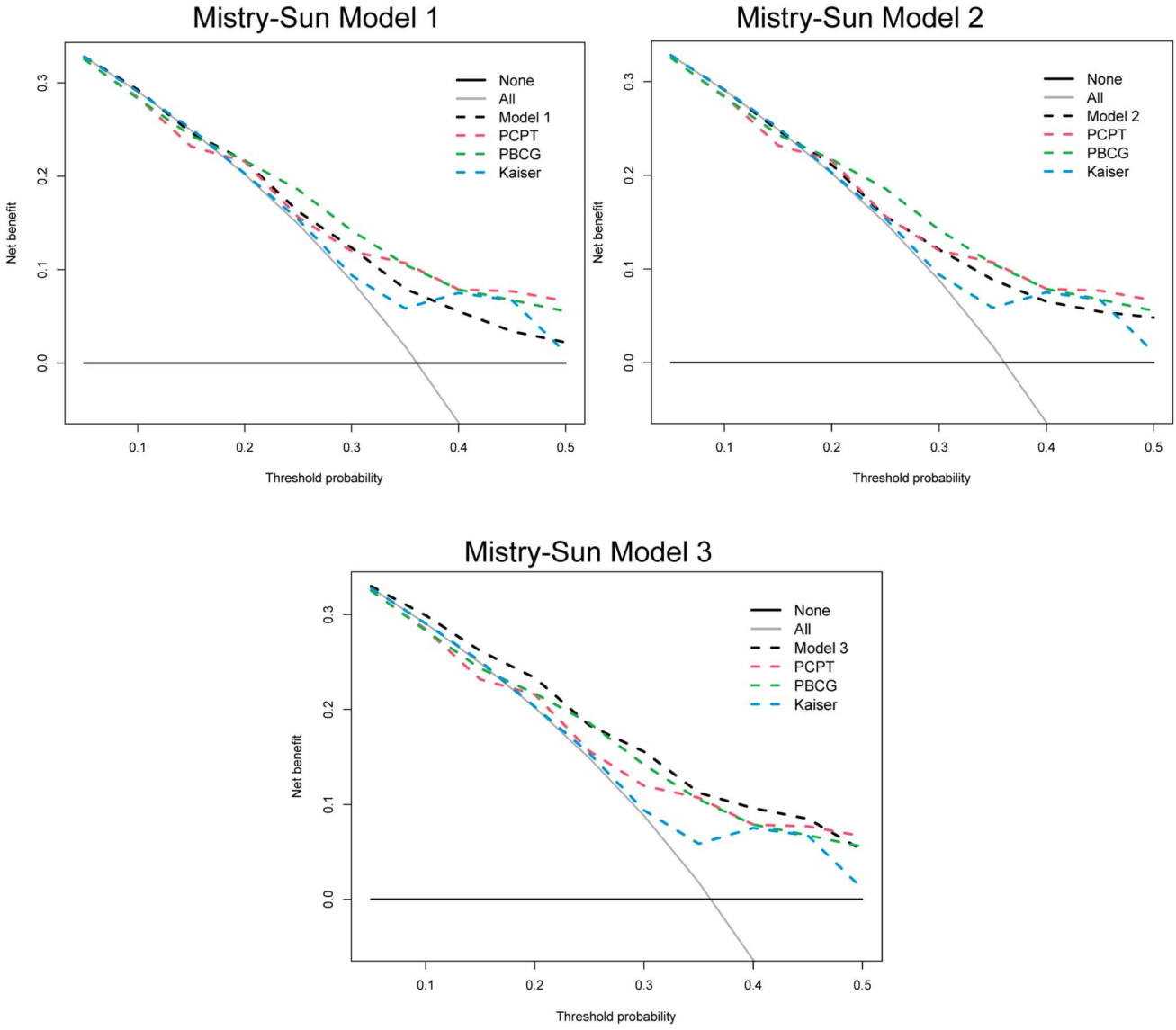

Figure 2. The net benefit curves from 0% to 50% risk threshold for each of the Mistry-Sun models compared against the Prostate Cancer Prevention Trial risk calculator version 2.0 (PCPT), the Prostate Biopsy Collaborative Group risk calculator (PBCG), and the Kaiser Permanente Prostate Cancer risk calculator (Kaiser).

Table 5. *Comparison of Biopsy Outcomes at 10% and 30% Risk Thresholds in the External Validation Cohort (N = 276)*

	Mistry-Sun 1	Mistry-Sun 2	Mistry-Sun 3	PCPT	PBCG	Kaiser
10% biopsy threshold						
Men below 10% threshold, No.	6	1	30	15	12	0
Total biopsies performed, No.	270	275	246	261	264	276
Missed GG2-5 among men below the 10% threshold, No. (%)	0 (0)	0 (0)	0 (0)	3 (20.0)	3 (25)	0 (0)
Unnecessary biopsies, No.	169	174	145	163	166	175
30% biopsy threshold						
Men below 30% threshold, No.	202	219	183	172	137	17
Total biopsies performed, No.	74	57	93	104	139	259
Missed GG2-5 among men below the 30% threshold, No. (%)	55 (27.2)	60 (27.4)	42 (23.0)	46 (26.7)	31 (22.6)	4 (23.5)
Unnecessary biopsies, No.	28	16	34	49	69	162

Abbreviations: GG, Gleason Grade Group; Kaiser, Kaiser Permanente Prostate Cancer risk calculator; PBCG, Prostate Biopsy Collaborative Group risk calculator; PCPT, Prostate Cancer Prevention Trial risk calculator version 2.0.

For men willing to undergo prostate biopsy if there is 10% or 30% risk of GG2-5 prostate cancer, we list the number of men who fall below the threshold (tests negative), the total number of biopsies performed (tests positive), the number of missed GG2-5 cancers (false-negatives, which should be $\leq$ 10% or $\leq$ 30% from the patient's perspective), and the number of unnecessary biopsies performed for men with GG1 prostate cancer or a negative biopsy. There were complete data in 276 out of the 292 men.

review and 2 different International Society of Urological Pathology Gleason grading standards, although we attempted to mitigate this by including uropathologists for review. Additionally, our nomogram did not include some modern components of the PCa workup such as biomarkers, prostate MRI, and fusion prostate biopsy. MRI was deliberately excluded to build a base model and to reflect the reality of our patient population. An analysis of the Surveillance, Epidemiology, and End Results–Medicare cohort shows that Black men have 60% lower odds of receiving a prostate MRI before biopsy.[15] Future iterations of the Mistry-Sun RCs will include additional tools and seek validation in diverse patient populations.

CONCLUSIONS

Specific gaps identified in prior RCs are based on the overestimation of high-grade PCa risk by adding Black race as a static model covariate, which exposes Black men to unnecessary biopsies.[6] The Mistry-Sun 2 RC features good discrimination and better calibration for low-risk men and highlights the predictive power of deriving an RC within a Black cohort.

Overall, by tailoring our nomogram to Black men, Mistry-Sun 2 can reduce unnecessary biopsies. While the goal of many RCs is to maximize utility across the US population, developing calculators in high-prevalence populations is reasonable for improving the performance and reducing potential harm.

REFERENCES

1. American Cancer Society. *Cancer Facts & Figures 2021*. American Cancer Society. Accessed January 17, 2022. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2021.html
2. 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(2):120-134.
3. Loeb S, Vellekoop A, Ahmed HU. Systematic review of complications of prostate biopsy. *Eur Urol*. 2013;64(6):876-892.
4. Bokhorst LP, Bangma CH, van Leenders GJLH. Prostate-specific antigen-based prostate cancer screening: reduction of prostate cancer mortality after correction for nonattendance and contamination in the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer. *Eur Urol*. 2014;65(2):329-336.
5. van Vugt HA, Roobol MJ, Kranse R. Prediction of prostate cancer in unscreened men: external validation of a risk calculator. *Eur J Cancer*. 2011;47(6):903-909.
6. Carbunaru S, Nettey OS, Gogana P. A comparative effectiveness analysis of the PBCG vs. PCPT risks calculators in a multi-ethnic cohort. *BMC Urol*. 2019;19(1):121.
7. Ankerst DP, Hoefler J, Bock S, et al. Prostate Cancer Prevention Trial risk calculator 2.0 for the prediction of low- vs high-grade prostate cancer. *Urology*. 2014;83(6):1362-1367.
8. Ankerst DP, Straubinger J, Selig K, et al. A contemporary prostate biopsy risk calculator based on multiple heterogeneous cohorts. *Eur Urol*. 2018;74(2):197-203.
9. Zhu Y, Wang JY, Shen YJ. External validation of the prostate cancer prevention trial and the European randomized study of screening for prostate cancer risk calculators in a Chinese cohort. *Asian J Androl*. 2012;14(5):738-744.
10. Yoon DK, Park JY, Yoon S. Can the prostate risk calculator based on Western population be applied to Asian population? *Prostate*. 2012;72(7):721-729.
11. Presti JC, Alexeeff S, Horton B, Prausnitz S, Avins AL. Prospective validation of the Kaiser Permanente Prostate Cancer Risk Calculator in a contemporary, racially diverse, referral population. *Urol Oncol*. 2021;39(11):783.e11-783.e19.
12. Park JY, Yoon S, Park MS. Initial biopsy outcome prediction in Korean patients-comparison of a noble web-based Korean prostate cancer risk calculator versus prostate-specific antigen testing. *J Korean Med Sci*. 2011;26(1):85-91.
13. Wu YS, Zhang N, Liu SH. The Huashan risk calculators performed better in prediction of prostate cancer in Chinese population: a training study followed by a validation study. *Asian J Androl*. 2016;18(6):925-929.
14. Abashidze N, Stecher C, Rosenkrantz AB, Duszak R Jr, Hughes DR. Racial and ethnic disparities in the use of prostate magnetic resonance imaging following an elevated prostate-specific antigen test. *JAMA Netw Open*. 2021;4(11):e2132388.
15. Gaffney CD, Cai P, Li D. Increasing utilization of MRI before prostate biopsy in Black and non-Black men: an analysis of the SEER-Medicare cohort. *AJR Am J Roentgenol*. 2021;217(2):389-394.
16. Epstein JI, Egevad L, Amin MB. The 2014 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol*. 2016;40(2):244-252.
17. Moons KG, Altman DG, Reitsma JB. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1-W73.
18. Pencina MJ, D'Agostino RB Sr. Evaluating discrimination of risk prediction models: the C statistic. *JAMA*. 2015;314(10):1063-1064.
19. 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(3):837-845.
20. Koo K, Hyams ES. Assessment of men's risk thresholds to proceed with prostate biopsy for the early detection of prostate cancer. *Int Urol Nephrol*. 2019;51(8):1297-1302.
21. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ*. 2016;352:i6.
22. Chandra Engel J, Palsdottir T, Ankerst D. External validation of the prostate biopsy collaborative group risk calculator and the Rotterdam prostate cancer risk calculator in a Swedish population-based screening cohort. *Eur Urol Open Sci*. 2022;41:1-7.
23. Helrich S, Pate W, Garg N, Barbosa P, Wason S. Comparison of magnetic resonance imaging to ultrasound for prostate sizing. *Can J Urol*. 2021;28(6):10889-10899.

EDITORIAL COMMENT

Mistry et al report the development and validation of a novel risk calculator for predicting clinically significant prostate cancer in Black men.[1] This study addresses an important need in a population underrepresented in the development of existing risk calculators. Importantly, the authors found that some well-known risk calculators, including one derived from the Prostate Cancer Prevention Trial whose study population included only 3.2% African Americans,[2] performed poorly in this population of Black men.

A notable omission, and one being addressed by the authors in future iterations of their risk calculator, is the absence of a prebiopsy MRI. Undoubtedly, the addition of MRI lesions and MRI-calculated volume will be of great value in reducing unnecessary biopsies and detecting clinically significant cancer.

So why is the development of this risk calculator relevant in the current MRI era? While it may seem to many of us that MRI has become ubiquitous, contemporary studies still show less than 20% utilization in both White and Black prostate cancer patients throughout the United States.[3] These types of instruments will, therefore, remain valuable to patients and providers with limited access to MRI. More broadly, this study highlights the limitations of widely used prostate cancer instruments developed from predominantly White populations.

Michael Feuerstein[1]
[1]*Lenox Hill Hospital*
New York, New York

REFERENCES

1. Mistry NA, Sun Z, Sweis J, et al. Development and validation of a prostate biopsy risk calculator in Black men. *J Urol.* 2024;211(2):223-233.
2. Thompson IM, Ankerst DP, Chi C, et al. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. *J Natl Cancer Inst.* 2006;98(8):529-534.
3. Leapman MS, Dinan M, Pasha S, et al. Mediators of racial disparity in the use of prostate magnetic resonance imaging among patients with prostate cancer. *JAMA Oncol.* 2022;8(5):687-696.