




# Polygenic risk of any, metastatic, and fatal prostate cancer in the Million Veteran Program

Meghana S. Pagadala , PhD,<sup>1,2,3</sup> Julie Lynch, PhD, RN, MBA,<sup>4,5</sup> Roshan Karunamuni, PhD,<sup>6</sup> Patrick R. Alba, MS,<sup>4,5</sup> Kyung Min Lee, PhD,<sup>4,5</sup> Fatai Y. Agiri, BS,<sup>4</sup> Tori Anglin, MHA,<sup>4,5</sup> Hannah Carter, PhD,<sup>7</sup> J. Michael Gaziano, MD, MPH,<sup>8,9</sup> Guneet Kaur Jasuja, MPH, PhD,<sup>10,11</sup> Rishi Deka, PhD,<sup>1,6,12</sup> Brent S. Rose, MD,<sup>1,6,11</sup> Matthew S. Panizzon, PhD,<sup>1,13</sup> Richard L. Hauger , MD,<sup>1,13,14</sup> Tyler M. Seibert , MD, PhD<sup>1,6,15,16\*</sup>

<sup>1</sup>Research Service, VA San Diego Healthcare System, San Diego, CA, USA

<sup>2</sup>Medical Scientist Training Program, University of California San Diego, La Jolla, CA, USA

<sup>3</sup>Biomedical Science Program, University of California San Diego, La Jolla, CA, USA

<sup>4</sup>VA Informatics and Computing Infrastructure, VA Salt Lake City Healthcare System (VINCI), Salt Lake City, UT, USA

<sup>5</sup>Department of Internal Medicine, Division of Epidemiology, University of Utah School of Medicine, Salt Lake City, UT, USA

<sup>6</sup>Department of Radiation Medicine and Applied Sciences, University of California San Diego, La Jolla, CA, USA

<sup>7</sup>Department of Medicine, University of California San Diego, La Jolla, CA, USA

<sup>8</sup>Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC), VA Boston Healthcare System, Boston, MA, USA

<sup>9</sup>Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA

<sup>10</sup>Center for Healthcare Organization and Implementation Research (CHOIR), VA Bedford Healthcare System, Bedford, MA, USA

<sup>11</sup>Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA; Department of Urology, University of California San Diego, La Jolla, CA, USA

<sup>12</sup>Department of Psychiatry, University of California San Diego, La Jolla, CA, USA

<sup>13</sup>Center for Behavioral Genetics of Aging, University of California San Diego, La Jolla, CA, USA

<sup>14</sup>Center of Excellence for Stress and Mental Health (CESAMH), VA San Diego Healthcare System, San Diego, CA, USA

<sup>15</sup>Department of Radiology, University of California San Diego, La Jolla, CA, USA

<sup>16</sup>Department of Bioengineering, University of California San Diego, La Jolla, CA, USA

\*Correspondence to: Tyler M. Seibert, MD, PhD, Department of Radiation Medicine and Applied Sciences, University of California San Diego, 9500 Gilman Dr, La Jolla, CA 92093-0861, USA (e-mail: tseibert@ucsd.edu).

**Background:** Genetic scores may provide an objective measure of prostate cancer risk and thus inform screening decisions. We evaluated whether a polygenic hazard score based on 290 genetic variants (PHS290) is associated with prostate cancer risk in a diverse population, including Black men, who have higher average risk of prostate cancer death but are often treated as a homogeneously high-risk group.

**Methods:** This was a retrospective analysis of the Million Veteran Program, a national, population-based cohort study of US military veterans conducted 2011–2021. Cox proportional hazards analyses tested for association of genetic and other risk factors (including self-reported race and ethnicity and family history) with age at death from prostate cancer, age at diagnosis of metastatic (nodal or distant) prostate cancer, and age at diagnosis of any prostate cancer.

**Results:** A total of 590 750 male participants were included. Median age at last follow-up was 69 years. PHS290 was associated with fatal prostate cancer in the full cohort and for each racial and ethnic group ( $P < .001$ ). Comparing men in the highest 20% of PHS290 with those in the lowest 20% (based on percentiles from an independent training cohort), the hazard ratio for fatal prostate cancer was 4.42 (95% confidence interval = 3.91 to 5.02). When accounting for guideline-recommended risk factors (family history, race, and ethnicity), PHS290 remained a strong independent predictor of any, metastatic, and fatal prostate cancer.

**Conclusions:** PHS290 stratified US veterans of diverse ancestry for lifetime risk of prostate cancer, including metastatic and fatal cancer. Predicting genetic risk of lethal prostate cancer with PHS290 might inform individualized decisions about prostate cancer screening.

Despite enormous mortality from prostate cancer (PCa), early detection remains controversial because prostate-specific antigen (PSA) testing results in false-positive results and frequent overdiagnosis of indolent PCa that may never have become symptomatic (1–3). As one of the most heritable cancers (4), genetic risk stratification is a promising approach for identifying individuals at higher risk of developing metastatic or fatal PCa (2,5,6). Genetic risk scores for PCa outperform family history and other clinical risk factors commonly available to inform

screening decisions (7–10). An ideal genetic test would focus on aggressive PCa and, given the strong age dependence of PCa (and of aggressive PCa, in particular), an ideal genetic score would estimate age-specific risk (11,12). Age-specific genetic risk could inform individualized decisions about PSA testing in the context of a given man's overall health and competing causes of mortality.

A major limitation of early studies of polygenic risk was an exclusive focus on men of European ancestry (13,14). Such

systematic bias may exacerbate existing health disparities in PCA incidence and health outcomes (15,16). This is particularly worrisome for men of African ancestry, who have a higher overall incidence of metastatic and fatal PCA (17,18). Recent efforts have incorporated data from more diverse populations, yielding improved performance (23).

Our group has developed a polygenic hazard score (PHS) that identifies men who are likely to develop aggressive PCA at younger ages. This score, which can be calculated from a single saliva sample at any point in a man's life, was strongly associated with age at diagnosis of aggressive PCA in large datasets (10,19,22). The score also improved the accuracy of conventional screening with PSA (7,10,12). We subsequently expanded the model to optimize performance in men of all ancestries, particularly men with African ancestry (19,22,24). Here, we seek to validate the ability of the PHS to identify men at risk of metastatic or fatal PCA within the Million Veteran Program (MVP) longitudinal cohort, one of the largest and most racially and ethnically diverse populations studied to date (25).

## Methods

### Participants

We retrospectively obtained data from the MVP, composed of individuals aged 19 years to over 100 years recruited from 63 Veterans Affairs Medical Centers across the United States, an ongoing study that began enrollment in 2011. Only men were included in this PCA study, comprising 590 750 individuals, over 100 000 of whom were Black or African American by self-report (Table 1). There were no inclusion or exclusion criteria for age. Median age at last follow-up was 69 years (interquartile range = 59-74 years).

### Genotype Data

All study participants provided blood samples for DNA extraction and genotyping. Blood samples were collected by phlebotomists and banked at the VA Central Biorepository in Boston, MA, where DNA was extracted and shipped to 2 external centers for genotyping. DNA extracted from buffy coat was genotyped using a custom Affymetrix Axiom biobank array. The MVP 1.0 genotyping array contains a total of 723 305 variants, enriched for low-frequency variants in African and Hispanic populations and variants associated with diseases common to the VA population (25). Details on quality control and imputation were previously described (26).

### Clinical Data Extraction

Each participant's electronic health record was integrated into the MVP biorepository, including standard diagnosis codes and laboratory values. A natural language processing tool was used to determine diagnosis of metastatic cancer. Further information is in the [Supplementary Methods](#) (available online).

### Polygenic Hazard Score (PHS290)

The most recent version of the PHS, called PHS290, was calculated as the vector product of participants' genotype dosage ( $X_i$ ) for 290 variants and the corresponding parameter estimates from Cox proportional hazards regression, as previously described (24). We calculated PHS290 for each MVP participant. Distributions were visualized using histograms for each race and ethnicity. Differences in mean PHS290 between race and ethnicity were assessed via ANOVA. Subgroup analyses with less than 100

**Table 1.** Participant characteristics for self-reported race and ethnicity groups, n = 590 750

	All	Asian	Black or African American	Hispanic White	Native American	Non-Hispanic White	Other <sup>a</sup>	Pacific Islander	Unknown
Participants, No. <sup>b</sup>	590 750 (378 366)	6644 (3876)	102 203 (47 838)	27 651 (13 407)	5835 (3927)	420 473 (299 266)	8226 (8224)	3246 (740)	16 472 (1088)
All participants	69 137 (49 400)	376 (287)	15 748 (8613)	2120 (1325)	504 (378)	48 339 (37 922)	720 (720)	236 (60)	1094 (95)
Prostate cancer	6413 (4274)	33 (24)	1604 (804)	213 (126)	48 (35)	4299 (3199)	68 (68)	18 (5)	130 (13)
Metastases from prostate cancer	1858 (1314)	5 (4)	363 (192)	59 (30)	10 (6)	1354 (1059)	19 (19)	4 (2)	44 (2)
Death from prostate cancer									
Age demographics, y									
Age at diagnosis, median [IQR], y	67 [62-72]	67 [61-74]	63 [58-68]	66 [60-71]	65 [61-69]	68 [63-73]	64 [60-69]	65 [60-70]	67 [61-73]
Age at last follow-up, median [IQR], y	66 [59-74]	56 [41-70]	62 [55-70]	59 [47-71]	63 [56-72]	68 [62-75]	63 [57-71]	59 [49-70]	56 [42-70]

<sup>a</sup> Participants who did not identify with any of the other race categories could indicate "Other."

<sup>b</sup> Numbers in parentheses indicate participants with family history information also.

events are not shown in the main text but are reported in [Supplementary Table 2](#) (available online) (27).

## Cox Proportional Hazards Analysis

We evaluated association of PHS290 with age at diagnosis of PCa and with 2 important clinical endpoints: age at diagnosis of nodal and/or distant metastases from PCa and age at death from PCa (ie, lifetime PCa-specific mortality). To visualize the association in the full dataset, we generated cause-specific cumulative incidence curves for each endpoint and each of several PHS290 risk groups. Cox proportional hazards models were used to assess these associations in the full dataset and in each racial and ethnic group. Individuals not meeting the endpoint of interest were censored at age at last follow-up.

Models used PHS290 as a continuous variable. Effect sizes were estimated using hazard ratios (HRs) between risk strata, as previously described (7,9,10,12,19,20,22,23), and with previously defined thresholds for PHS290 (determined in an independent dataset and using PHS290 percentiles among men unaffected by PCa and younger than 70 years old): 9.004659 (20th quantile), 9.123500 (30th quantile), 9.519703 (70th quantile), 9.639068 (80th quantile), 9.946332 (95th quantile) (24). Hazard ratios for each ancestry group were calculated to make the following comparisons: HR<sub>80/20</sub>, men in the highest 20% vs lowest 20%; HR<sub>95/50</sub>, men in the highest 5% of genetic risk vs those with average risk (30-70th percentile); and HR<sub>20/50</sub>, men in the lowest 20% vs those with average risk.

## Race and Ethnicity, Family History, and PHS290

To assess the independent predictive value of PHS290 beyond commonly used clinical risk factors, we tested a multivariable Cox proportional hazards model with self-reported race and ethnicity, family history, and PHS290 (7,9,22). Race and ethnicity were reported by the participants on a multiple-choice questionnaire, where race categories were Asian; Black or African American; Native American; Pacific Islander; White; or Other (for those who did not identify with any of the other race categories). Categories for ethnicity were Hispanic and Not Hispanic. For the present work, ethnicity was used only within White race to make 2 groups: Non-Hispanic White and Hispanic White. Participants who identified race as Black or African American were included in that group regardless of Hispanic or Not Hispanic self-reported ethnicity. Those who did not respond to the race and ethnicity questions were categorized in the present analysis as Unknown race and ethnicity. Family history was recorded as either the presence or absence of (one or more) first-degree relatives with PCa. Cox proportional hazards models tested associations with any, metastatic, or fatal PCa. For PHS290, the effect size was illustrated via the hazard ratio for the highest 20% vs lowest 20% of genetic risk and between other strata of PHS290. Hazard ratios for racial and ethnic groups were estimated using Non-Hispanic White as the reference.

A univariate Cox proportional hazards model was applied to test for association of race and ethnicity with PCa endpoints. Similarly, a univariate model tested for association of family history alone. The @anova function from the R “survival” package (version 3.2-13; Therneau 2021) was used to compare the nested Cox models (multivariable vs univariate) based on the log partial likelihood of the model fits. The impact of adding PHS290 to the base clinical model (comprised of self-reported race and ethnicity and family history) was assessed by testing for increase in Harrell’s concordance index (28). Statistical significance was set

at a 2-tailed alpha of 0.01 for the test of whether the multivariable model performed better than either univariate model alone.

## Statistical Analysis

In all statistical analyses, statistical significance for association with clinical endpoints was set at a 2-tailed alpha of .01. Per journal style guidelines, exact P values are reported if P is greater than or equal to .001; lower values are reported only as P less than .001. Confidence intervals were calculated from n=1000 bootstrapping with replacement.

## Additional Analyses

Several additional analyses are described in the [Supplementary Methods](#) (available online). These include metastatic disease on initial staging, medical care outside the VA, PSA testing, group-specific PHS290 thresholds, genetic ancestry, genetic principal components, and observed vs predicted risk.

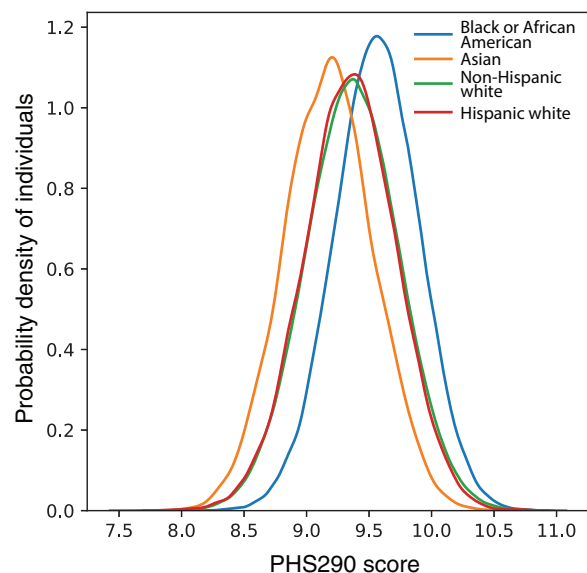
## Results

### PHS290 Score

The distribution of PHS290 in Non-Hispanic Whites was similar to that previously reported for men of European ancestry (mean=9.37, SD=0.37) (24). Mean PHS290 did vary by self-reported race and ethnicity, with statistically significant differences between all groups (ANOVA,  $P < .001$ ; all pairwise t tests,  $P < .001$ ). The distribution for the Hispanic ancestry group overlapped closely with that of the European group (mean=9.35, SD=0.37), whereas PHS290 tended to be lower among Asian men (mean=9.18, SD=0.35) and higher among Black men (mean=9.56, SD=0.34, t test) (Figure 1; [Supplementary Figure 1](#), available online). Only approximately 5% of Black men had PHS290 in the bottom quintile when using the fixed thresholds previously defined in an independent dataset of men of European ancestry.

### Association of PHS290 With PCa

PHS290 was associated with age at diagnosis of PCa, age at development of PCa nodal or distant metastases, and age at death from PCa ([Table 2](#); [Supplementary Tables 1-9](#), available online).



**Figure 1.** Polygenic hazard score based on 290 common variants (PHS290) score density plots in Million Veteran Program. PHS290 score density plot in select self-reported race and ethnicity groups.

**Table 2.** Association of PHS290 with any, metastatic, and fatal prostate cancer<sup>a</sup>

Clinical endpoint	No.	P <sup>b</sup>	HR (95% CI) <sup>c</sup>			
			HR <sub>80/20</sub>	HR <sub>20/50</sub>	HR <sub>80/50</sub>	HR <sub>95/50</sub>
Fatal prostate cancer						
All	590 750	<.001	4.42 (3.91 to 5.02)	0.48 (0.45 to 0.51)	2.12 (2.0 to 2.27)	3.0 (2.75 to 3.3)
Asian	6644	—	—	—	—	—
Black or African American	102 203	<.001	2.37 (1.73 to 3.29)	0.66 (0.56 to 0.77)	1.57 (1.33 to 1.85)	1.9 (1.5 to 2.42)
Hispanic White	27 651	—	—	—	—	—
Native American	5835	—	—	—	—	—
Non-Hispanic White	420 473	<.001	4.37 (3.77 to 5.05)	0.48 (0.45 to 0.52)	2.11 (1.96 to 2.27)	3.0 (2.69 to 3.34)
Other <sup>d</sup>	8226	—	—	—	—	—
Pacific Islander	3246	—	—	—	—	—
Unknown	16 472	—	—	—	—	—
Metastatic prostate cancer						
All	590 750	<.001	4.89 (4.57 to 5.21)	0.46 (0.44 to 0.47)	2.23 (2.16 to 2.31)	3.23 (3.07 to 3.39)
Asian	6644	—	—	—	—	—
Black or African American	102 203	<.001	3.03 (2.62 to 3.51)	0.59 (0.55 to 0.63)	1.78 (1.65 to 1.92)	2.28 (2.05 to 2.54)
Hispanic White	27 651	<.001	2.92 (1.96 to 4.43)	0.58 (0.47 to 0.71)	1.71 (1.4 to 2.11)	2.22 (1.65 to 3.01)
Native American	5835	—	—	—	—	—
Non-Hispanic White	420 473	<.001	4.62 (4.27 to 5.01)	0.47 (0.45 to 0.49)	2.17 (2.08 to 2.26)	3.13 (2.95 to 3.32)
Other <sup>d</sup>	8226	—	—	—	—	—
Pacific Islander	3246	—	—	—	—	—
Unknown	16 472	—	—	—	—	—
Prostate cancer						
All	590 750	<.001	5.2 (5.09 to 5.31)	0.44 (0.44 to 0.45)	2.31 (2.28 to 2.33)	3.39 (3.33 to 3.44)
Asian	6644	<.001	5.32 (4.04 to 7.09)	0.43 (0.37 to 0.49)	2.27 (1.98 to 2.61)	3.49 (2.83 to 4.34)
Black or African American	102 203	<.001	3.22 (3.06 to 3.39)	0.57 (0.56 to 0.58)	1.83 (1.79 to 1.88)	2.39 (2.3 to 2.48)
Hispanic White	27 651	<.001	4.24 (3.74 to 4.85)	0.49 (0.45 to 0.52)	2.06 (1.93 to 2.2)	2.92 (2.66 to 3.22)
Native American	5835	<.001	5.1 (4.04 to 6.55)	0.45 (0.4 to 0.51)	2.31 (2.05 to 2.62)	3.4 (2.85 to 4.12)
Non-Hispanic White	420 473	<.001	5.12 (4.99 to 5.26)	0.45 (0.44 to 0.45)	2.29 (2.26 to 2.32)	3.38 (3.31 to 3.44)
Other <sup>d</sup>	8226	<.001	5.52 (4.45 to 6.89)	0.43 (0.39 to 0.48)	2.36 (2.12 to 2.64)	3.59 (3.06 to 4.25)
Pacific Islander	3246	<.001	3.58 (2.5 to 5.37)	0.53 (0.43 to 0.63)	1.9 (1.57 to 2.33)	2.6 (1.98 to 3.57)
Unknown	16 472	<.001	5.27 (4.39 to 6.3)	0.44 (0.4 to 0.48)	2.32 (2.12 to 2.55)	3.45 (3.0 to 3.94)

<sup>a</sup> Cox proportional hazards model results from association with age at prostate cancer, metastatic prostate cancer (nodal or distant), and death from prostate cancer. Because the thresholds for a multivariable model must be constant across racial and ethnic groups and regardless of family history status—and it is potentially problematic to have race-specific thresholds in clinical settings—the thresholds for percentiles of PHS290 were taken from previously published values of an independent dataset of men with European genetic ancestry younger than 70 years and who did not have prostate cancer. For subgroup analyses with less than 100 events of the endpoint, the box is marked “—”; these statistically less reliable results are reported in [Supplementary Table 2](#) (available online). CI = confidence interval; HR = hazard ratio; PHS290 = polygenic hazard score based on 290 common variants.

<sup>b</sup> P values reported are from univariate models using PHS290 as the sole predictor variable.

<sup>c</sup> Hazard ratios compare men in various percentiles of genetic risk. HR<sub>80/20</sub>: highest 20% (>80th percentile of PHS290) vs average risk (30–70th percentile). HR<sub>20/50</sub>: lowest 20% (<20th percentile) vs average risk. HR<sub>80/50</sub>: highest 20% vs average risk. HR<sub>95/50</sub>: highest 5% (>95th percentile) vs average risk.

<sup>d</sup> Participants who did not identify with any of the other race categories could indicate “Other.”

These associations also held in all racial and ethnic subgroup analyses with more than 100 events. Comparing 80th and 20th percentiles of genetic risk in the full dataset (based on percentiles from an independent training cohort), men with higher PHS290 had an HR<sub>80/20</sub> of 5.20 (95% CI = 5.09 to 5.31) for any PCa, an HR<sub>80/20</sub> of 4.89 (95% CI = 4.57 to 5.21) for metastatic PCa, and an HR<sub>80/20</sub> of 4.42 (95% CI = 3.91 to 5.02) for fatal PCa. Cause-specific cumulative incidence curves for various PHS290 percentile groups demonstrated risk stratification ([Figure 2](#)). Consistent with previous reports, Black men had a higher average incidence of PCa than Non-Hispanic White men. However, the incidence among Black men with low PHS290 was comparable with that of the average among Non-Hispanic White men ([Figures 2 and 3](#); [Supplementary Figures 2, 3](#), available online).

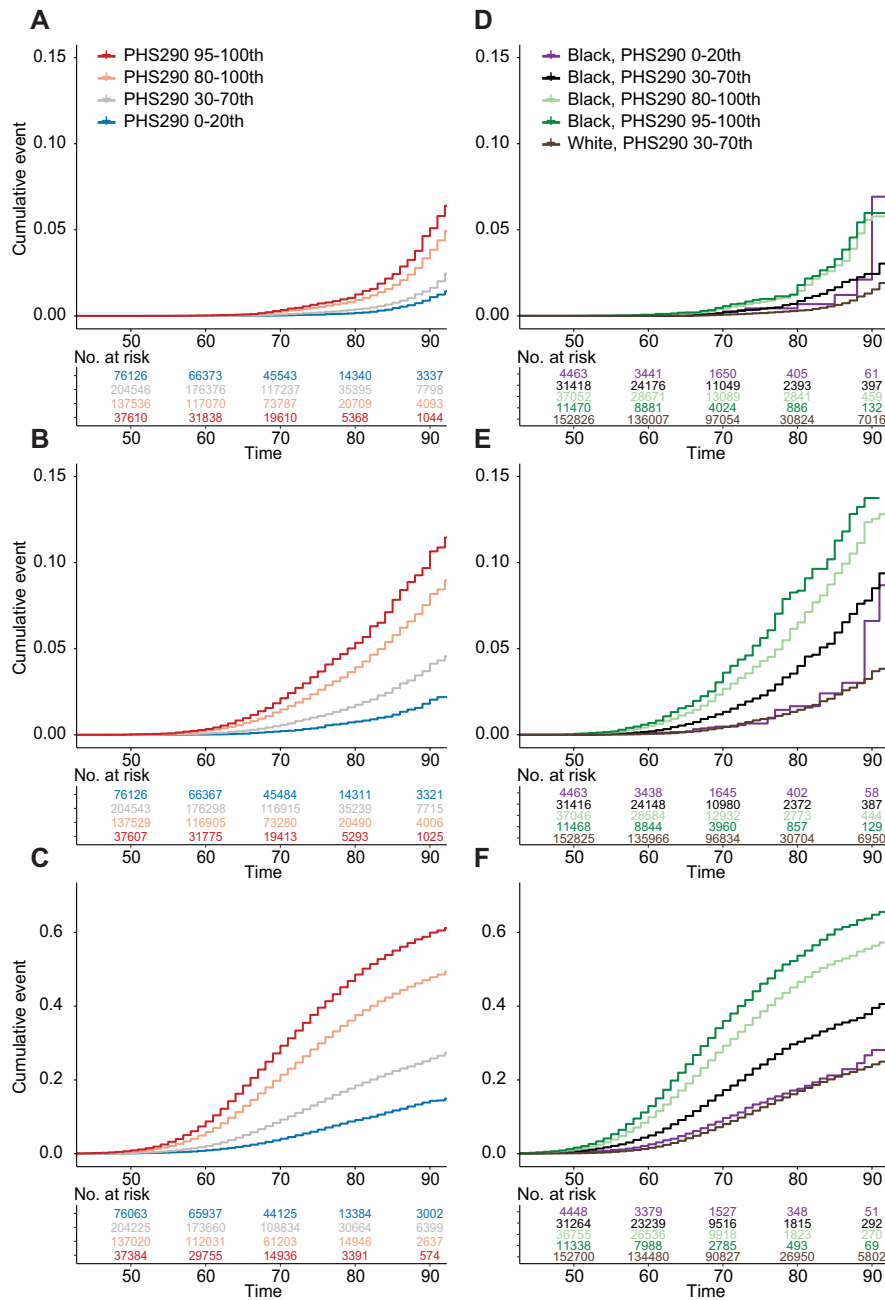
### Race and Ethnicity, Family History, and PHS290

Race and ethnicity and family history were each associated with every clinical endpoint in univariate models ([Supplementary Tables 9 and 10](#), available online). Men with a family history of PCa had a hazard ratio of 1.83 (95% CI = 1.53 to 2.17) for dying of PCa. The race and ethnicity associations were largely driven by an increased risk for Black men. Compared with Non-Hispanic White men, Black men had a hazard ratio of 2.53 (95% CI = 2.14 to 2.92) for dying of PCa.

PHS290 remained an independent predictor of PCa risk—including PCa death—when accounting for race and ethnicity and family history ([Table 3](#); [Supplementary Tables 11–19](#), available online). The multivariable model had stronger associations with each clinical endpoint than the univariate models (ANOVA  $P < .001$ ). Addition of PHS290 improved prediction for each clinical endpoint over the baseline multivariable clinical model of self-reported race and ethnicity and family history ([Table 4](#)). Independent of ancestry and family history, a high PHS290 (top 20%, based on percentiles from an independent training cohort) approximately quadrupled a man's risk of death from PCa compared with a low PHS290 (bottom 20%) ([Table 3](#); [Supplementary Table 20](#), available online).

### Discussion

In this large and diverse dataset, after accounting for current guideline-recommended risk factors (family history and race and ethnicity), PHS290 remained a strong independent predictor of dying from PCa. The genomic score was also associated with age at diagnosis of metastasis from PCa—a major driver of pain, disability, and aggressive medical therapy ([29](#))—and with age at diagnosis of any PCa. This study represents the largest and most racially and ethnically diverse independent validation of the



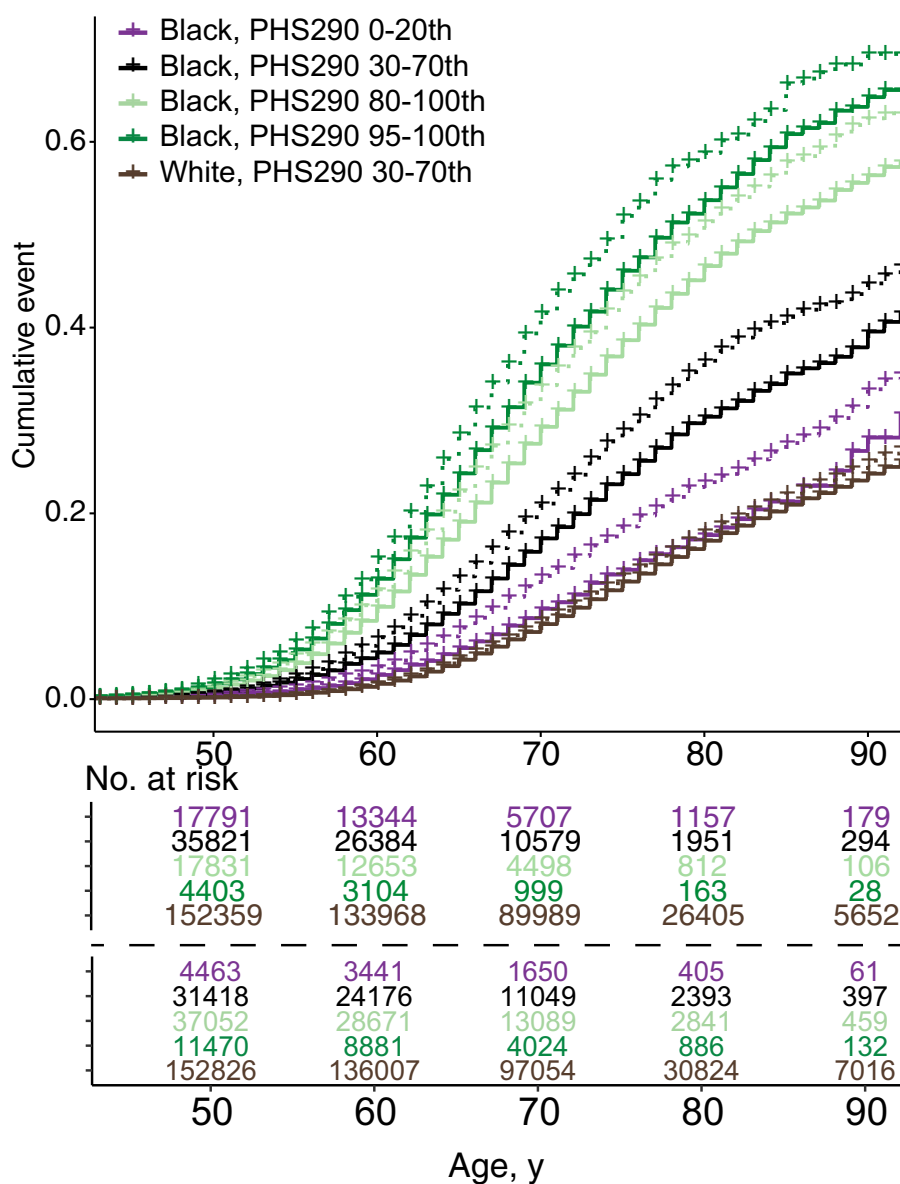
**Figure 2.** Million Veteran Program (MVP) cause-specific cumulative incidence. Cause-specific cumulative incidence within MVP, stratified by polygenic hazard score based on 290 common variants (PHS290), for (A) fatal prostate cancer, (B) metastatic prostate cancer, and (C) prostate cancer. PHS290 percentile groups shown for each endpoint: 0-20th, 30-70th, 80-100th, and 95-100th. Cumulative incidence for Black or African American (Black/AA) men in several PHS290 strata compared with average-risk Non-Hispanic White (PHS290 30-70th percentiles) for (D) fatal prostate cancer, (E) metastatic prostate cancer, and (F) prostate cancer. The y-axis scale was adjusted for (C) and (E) to show the higher incidence values for any prostate cancer.

association of a polygenic score with lifetime risk of metastatic and fatal PCa. Current clinical guidelines recommend individualized risk assessment with race and ethnicity and family history in deciding whether to pursue early detection of PCa (30-33). Through use of a diverse, longitudinal cohort, we demonstrate here that a polygenic score adds considerable information beyond race and ethnicity and family history for a man's individual risk of metastasis or death from PCa.

Guidelines recommend consideration of prostate screening at a younger age and at more frequent intervals in men of African ancestry (30) because Black men are substantially more likely to

die from PCa (18). The causes of this disparity are likely a combination of genetic, environmental, and social factors, including systemic racism (34-38). This study confirms a generally higher risk of PCa among Black men but also demonstrates that Black men have variable levels of lifetime risk. Black men with low PHS290 (based on percentiles from an independent training cohort) had a PCa risk comparable with average-risk Non-Hispanic White men, whereas Black men with high PHS290 had the highest risk of all subgroups.

Consistent with previous work, family history was independently associated with PCa risk when accounting for polygenic risk



**Figure 3.** Million Veteran Program (MVP) cause-specific cumulative incidence using MVP-specific and reference threshold. Cause-specific cumulative incidence for Black or African American (Black/AA) men compared with average-risk Non-Hispanic White (PHS290 30-70th percentiles) stratified by PHS290, for (A, D) fatal prostate cancer, (B, E) metastatic prostate cancer, and (C, F) prostate cancer. PHS290 percentile groups shown for each endpoint: 0-20th, 30-70th, 80-100th, and 95-100th. Incidence curves using MVP group-specific thresholds defined in MVP are shown in **dashed lines**, while reference thresholds are **solid lines**. For example, the purple solid line represents the 5% of Black men with PHS290 below the 20th percentile defined previously in an independent dataset of men with European ancestry. The purple dotted line represents the 20% of Black MVP participants with lowest PHS290 in the present MVP dataset.

in MVP, possibly by capturing yet unknown genetic factors and/or shared familial environmental factors (7,9,22,39,40). The relationship of environmental exposures, family history, and PCA risk merit further investigation (9), particularly in groups like veterans who may have been exposed to rare carcinogens (41).

One early detection strategy with strong evidence is early baseline PSA (eg, at age 40-49 years) (42-44). This strategy has not yet been widely adopted in the United States (45), but retrospective studies have demonstrated the predictive potential of obtaining an early baseline PSA. We found only a small proportion (14.7%) of MVP participants had an early baseline prediagnostic PSA test (Supplementary Methods, Supplementary Results, available online). Importantly, we lack information about why these men, and not the others, underwent early testing. Black race was

associated with higher likelihood of early baseline testing (consistent with guidelines), though early baseline PSA testing was still only conducted in 30.1% of Black men; family history did not have much apparent influence. Although acknowledging the small and likely biased nature of the subgroup analysis, we did find that early baseline PSA was independently associated with any, metastatic, and fatal PCA. On multivariable analysis in those with early baseline PSA, neither PHS290 nor Black race were statistically significantly associated with fatal PCA. This finding is consistent with another study, which found that a polygenic risk score was not independently associated with metastatic or fatal PCA after accounting for early baseline PSA (46). On the other hand, we found both PHS290 and Black race were independently associated with PCA and metastatic PCA after accounting for early

**Table 3.** Multivariable models combining self-reported race and ethnicity, family history, and PHS290 for 3 prostate cancer clinical endpoints<sup>a</sup>

Clinical endpoints	PHS290 <sup>b</sup>					HR (95% CI) <sup>c</sup>						
	HR <sub>20/50</sub>		HR <sub>80/50</sub>		HR <sub>95/50</sub>	Race and ethnicity						
	HR	CI	HR	CI	HR	Asian	Black or African American	Hispanic White	Native American	Other	Pacific Islander	Unknown
Fatal prostate cancer	0.49	4.17	2.06	2.88	1.67	0.68	1.97	1.04	0.98	1.9	1.74	0.73
	(0.46 to 0.53)	(3.59 to 4.88)	(1.91 to 2.23)	(2.58 to 3.23)	(1.4 to 1.96)	(0.15 to 1.43)	(1.69 to 2.31)	(0.69 to 1.42)	(0.32 to 1.85)	(1.12 to 2.81)	(0.0 to 4.43)	(0.0 to 1.95)
Metastatic prostate cancer	0.5	4.15	2.05	2.87	1.53	1.31	2.24	1.32	1.4	1.67	1.14	1.52
	(0.47 to 0.52)	(3.81 to 4.53)	(1.97 to 2.15)	(2.7 to 3.06)	(1.38 to 1.68)	(0.81 to 1.85)	(2.07 to 2.42)	(1.1 to 1.57)	(0.97 to 1.86)	(1.25 to 2.08)	(0.23 to 2.29)	(0.79 to 2.45)
Prostate cancer	0.47	4.69	2.19	3.14	1.78	1.28	1.83	1.1	1.04	1.2	0.98	0.9
	(0.46 to 0.47)	(4.57 to 4.81)	(2.16 to 2.21)	(3.08 to 3.2)	(1.73 to 1.83)	(1.13 to 1.43)	(1.78 to 1.87)	(1.04 to 1.16)	(0.95 to 1.15)	(1.1 to 1.29)	(0.73 to 1.25)	(0.72 to 1.08)

<sup>a</sup> Cox proportional hazards results for association with age at death from prostate cancer, at diagnosis of metastatic prostate cancer, and age at diagnosis with prostate cancer. This multivariable analysis was limited to the 378 366 participants who provided family history information in baseline survey data. CI = confidence interval; HR = hazard ratio; PHS290 = polygenic hazard score based on 290 common variants.

<sup>b</sup> PHS290 was included in the model as a continuous variable, and effect size is illustrated here via several hazard ratios.

<sup>c</sup> HR<sub>20/50</sub>: highest 20% (≥80th percentile of PHS290) vs average risk (30-70th percentile); HR<sub>80/50</sub>: lowest 20% (≤20th percentile) vs average risk. HR<sub>95/50</sub>: highest 5% (≥95th percentile) vs average risk. Because the thresholds for a multivariable model must be constant across racial and ethnic groups and regardless of family history status, the thresholds for percentiles of PHS290 were taken from previously published values of an independent dataset of men with European genetic ancestry younger than 70 years, and who did not have prostate cancer. Hazard ratios for race and ethnicity were estimated using Non-Hispanic White as the reference. Hazard ratios for family history were for 1 or more first-degree relatives diagnosed with prostate cancer.

**Table 4.** Harrell’s concordance index (95% confidence interval) for 3 prostate cancer clinical endpoints using race and ethnicity and family history, with or without PHS290<sup>a</sup>

Clinical endpoints	Race and ethnicity and family history	Race and ethnicity and family history and PHS290
Fatal prostate cancer	0.597 (0.579 to 0.618)	0.701 (0.684 to 0.721)
Metastatic prostate cancer	0.595 (0.587 to 0.606)	0.693 (0.684 to 0.703)
Prostate cancer	0.583 (0.581 to 0.586)	0.688 (0.685 to 0.690)

<sup>a</sup> This multivariable analysis was limited to the 378 366 participants who provided family history information in baseline survey data. Concordance index was greater with addition of PHS290 for each endpoint ( $P < .001$ ).

baseline PSA (with similar effect sizes to those found on multivariable analysis in the full MVP cohort).

Men at highest risk of metastatic or fatal PCa are potentially those most likely to benefit from screening, and we have shown that adding PHS290 to guideline-recommended risk factors improves stratification for these endpoints. Previous studies have further suggested genetic scores could also add value even after results from screening or diagnostic tests are already available, but this needs further investigation (47,48). Another compelling avenue for future studies is whether high genetic risk can be mitigated by lifestyle or other preventive intervention (49,50). Notwithstanding the possible advantages to risk-based screening, men should still be cautioned on the risks of screening: false-positive PSA testing results and overdiagnosis of nonthreatening cancer. Although men with high PHS290 have a higher absolute incidence of PCa metastasis or death, they also remain at risk of developing low-grade cancers likely to be detected with screening. Strategies to mitigate screening harms should be appropriately applied, including tests to reduce biopsies (multiparametric Magnetic Resonance Imaging and blood or urine tests) and active surveillance for cancer diagnoses with favorable prognosis (30).

Limitations of this study include heterogeneity of PCa screening and diagnostic pathways by clinicians across VA and other hospitals in the United States that could potentially introduce noise, although this heterogeneity likely leads to underestimation of associations with PCa. Medicare enrollment, for example, leads to more care outside the VA entirely, and our analyses do not suggest there were any major differences between participants who did or did not enroll in Medicare (Supplementary Methods, Supplementary Results, available online). Men on active-duty military service were not included in MVP. The natural language processing tool used to identify men with metastatic disease does not reliably distinguish regional nodal from distant metastases, so these were all considered as 1 endpoint. Finally, we acknowledge that although we have used race and ethnicity, genetic ancestry, and principal genetic components, none of these groups can account for—much less, disentangle—the complex web of biological and social factors associated with these categories. Further work will attempt to incorporate agnostic genetic ancestry groups and address impacts of admixture and local genetic ancestry on risk stratification with PHS (22).

We show that PHS290 stratified US men for lifetime risk of any, metastatic, and fatal PCa. Critically, this genetic risk stratification was successful within racial and ethnic subgroups in this diverse dataset. PHS290 was higher, on average, among Black men, who were also at higher risk from PCa. The combination of race and ethnicity, family history, and PHS290 performed better than any single risk factor in identifying men at highest risk of PCa metastasis and death. Predicting genetic risk of lethal PCa

with PHS290 might inform individualized decisions about screening and early cancer detection.

## Funding

RLH was funded by the VISN-22 VA Center of Excellence for Stress and Mental Health (CESAMH) and National Institute of Aging RO1 grant AG050595 (*The VETSA Longitudinal Twin Study of Cognition and Aging VETSA 4*). This research was supported by VA MVP022. MSP was supported by the National Institutes of Health (#1F30CA247168, #T32CA067754). TMS and RK were supported by the National Institutes of Health (NIH/NIBIB #K08EB026503), the Prostate Cancer Foundation, and the University of California (#C21CR2060).

## Notes

**Role of the funder:** The funders had no role in the design, the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

**Disclosures:** TMS reports honoraria from Varian Medical Systems and WebMD; he has an equity interest in CorTechs Labs, Inc and serves on its Scientific Advisory Board; he has received in-kind research support from GE Healthcare via a research agreement with the University of California San Diego. These companies might potentially benefit from the research results. The terms of this arrangement have been reviewed and approved by the University of California San Diego in accordance with its conflict-of-interest policies. MSP, JL, RK, PA, KL, FA, TA, JG, HC, GJ, RD, BR, MP and RLH report no disclosures.

**Author contributions:** Conceptualization: TMS; Data curation: MSP, JL, PA, KL, FA, TA, JG, GJ, RD; Formal Analysis: MSP; Funding acquisition: MSP, HC, RLH, TMS; Investigation: MSP, TMS; Methodology: MSP, RK, TMS; Project administration: JL, JG, MP, RLH, TMS; Resources: JL, TA, JG, GJ, BR, MP, RLH, TMS; Software: MSP, JL, RK, PA, KL, FA, TA, RD, TMS; Supervision: JL, RLH, TMS; Validation: MSP, RK, TMS; Visualization: MSP, TMS; Writing—original draft: MSP, RK, TMS; Writing—review and editing: MSP, JL, RK, PA, KL, FA, TA, JG, HC, GJ, RD, BR, MP, RLH, TMS.

**Acknowledgements:** This research used data from the Million Veteran Program, Office of Research and Development, Veterans Health Administration. This research was supported by the Million Veteran Program MVP022 award #I01 CX001727 (PI: RLH). This publication does not represent the views of the Department of Veterans Affairs or the United States Government.

**Prior presentations:** Results of this article were presented at GU ASCO 2022 (02/18/22).

## Data availability

It is not possible for the authors to directly share the individual-level data that were obtained from the MVP due to constraints stipulated in the informed consent. Anyone wishing to gain access to this data should inquire directly to MVP at MVPLOI@va.gov. The data generated from our analyses are included in the manuscript main text, tables, and figures and online [Supplementary Materials](#) (available online). The code used for analyses is available at <https://github.com/precimed/MVP-PCa-PHS>.

## References

- Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet*. 2014;384(9959):2027-2035. doi:10.1016/S0140-6736(14)60525-0
- Callender T, Emberton M, Morris S, et al. Polygenic risk-tailored screening for prostate cancer: a benefit-harm and cost-effectiveness modelling study. *PLoS Med*. 2019;16(12):e1002998. doi:10.1371/journal.pmed.1002998
- Loeb S, Bjurlin MA, Nicholson J, et al. Overdiagnosis and overtreatment of prostate cancer. *Eur Urol*. 2014;65(6):1046-1055. doi:10.1016/j.eururo.2013.12.062
- Mucci LA, Hjelmborg JB, Harris JR, et al., Nordic Twin Study of Cancer (NorTwinCan) Collaboration. Familial risk and heritability of cancer among twins in Nordic countries. *JAMA*. 2016;315(1):68-76. doi:10.1001/jama.2015.17703
- Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. *Nat Rev Genet*. 2018;19(9):581-590. doi:10.1038/s41576-018-0018-x
- Callender T, Emberton M, Morris S, Pharoah PDP, Pashayan N. Benefit, harm, and cost-effectiveness associated with magnetic resonance imaging before biopsy in age-based and risk-stratified screening for prostate cancer. *JAMA Netw Open*. 2021;4(3):e2037657. doi:10.1001/jamanetworkopen.2020.37657
- Seibert TM, Fan CC, Wang Y, et al. Polygenic hazard score to guide screening for aggressive prostate cancer: development and validation in large scale cohorts. *BMJ*. 2018;360:j5757. doi:10.1136/bmj.j5757
- Schumacher FR, Olama AAA, Berndt SI, et al.; Genetic Associations and Mechanisms in Oncology (GAME-ON)/Elucidating Loci Involved in Prostate Cancer Susceptibility (ELLIPSE) Consortium. Author correction: Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet*. 2019;51(2):363. doi:10.1038/s41588-018-0330-6
- Huynh-Le MP, Karunamuni R, Fan CC, et al.; on behalf of The PRACTICAL Consortium. Common genetic and clinical risk factors: association with fatal prostate cancer in the Cohort of Swedish Men. *Prostate Cancer Prostatic Dis*. 2021;24(3):845-851. doi:10.1038/s41391-021-00341-4.
- Karunamuni RA, Huynh-Le MP, Fan CC, et al.; UKGPCS orators. Additional SNPs improve risk stratification of a polygenic hazard score for prostate cancer. *Prostate Cancer Prostatic Dis*. 2021;24(2):532-541. doi:10.1038/s41391-020-00311-2
- Huynh-Le MP, Myklebust TÅ, Feng CH, et al. Age dependence of modern clinical risk groups for localized prostate cancer—a population-based study. *Cancer*. 2020;126(8):1691-1699. doi:10.1002/cncr.32702
- Huynh-Le MP, Fan CC, Karunamuni R, et al.; PRACTICAL Consortium. A genetic risk score to personalize prostate cancer screening, applied to population data. *Cancer Epidemiol Biomarkers Prev*. 2020;29(9):1731-1738. doi:10.1158/1055-9965.EPI-19-1527
- Duncan L, Shen H, Gelaye B, et al. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat Commun*. 2019;10(1):3328. doi:10.1038/s41467-019-11112-0.
- Petrovski S, Goldstein DB. Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine. *Genome Biol*. 2016;17(1):157. doi:10.1186/s13059-016-1016-y.



15. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet.* 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x
16. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature.* 2016;538(7624):161-164. doi:10.1038/538161a.
17. DeSantis CE, Miller KD, Goding Sauer A, Jemal A, Siegel RL. Cancer statistics for African Americans, 2019. *CA Cancer J Clin.* 2019;69(3):211-233. doi:10.3322/caac.21555.
18. Tsoodikov A, Gulati R, de Carvalho TM, et al. Is prostate cancer different in Black men? Answers from 3 natural history models. *Cancer.* 2017;123(12):2312-2319. doi:10.1002/cncr.30687.
19. Karunamuni RA, Huynh-Le MP, Fan CC, et al.; PRACTICAL Consortium. African-specific improvement of a polygenic hazard score for age at diagnosis of prostate cancer. *Int J Cancer.* 2021;148(1):99-105. doi:10.1002/ijc.33282.
20. Karunamuni RA, Huynh-Le MP, Fan CC, et al.; UKGPCS Collaborators. Performance of African-ancestry-specific polygenic hazard score varies according to local ancestry in 8q24. *Prostate Cancer Prostatic Dis.* 2022;25(2):229-237. doi:10.1038/s41391-021-00403-7.
21. Conti DV, Darst BF, Moss LC, et al. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. *Nat Genet.* 2021;53(1):65-75. doi:10.1038/s41588-020-00748-0.
22. Huynh-Le MP, Fan CC, Karunamuni R, et al.; PRACTICAL Consortium. Polygenic hazard score is associated with prostate cancer in multi-ethnic populations. *Nat Commun.* 2021;12(1):1236. doi:10.1038/s41467-021-21287-0.
23. Huynh-Le MP, Karunamuni R, Fan CC, et al.; UKGPCS Collaborators. Prostate cancer risk stratification improvement across multiple ancestries with new polygenic hazard score. *Prostate Cancer Prostatic Dis.* 2022. doi:10.1038/s41391-022-00497-7.
24. Huynh-Le MP, Karunamuni R, Fan CC, et al. Prostate cancer risk stratification improved across multiple ancestries with new polygenic hazard score. *Prostate Cancer Prostatic Dis.* 2021. doi:10.1101/2021.08.14.21261931.
25. Gaziano JM, Concato J, Brophy M, et al. Million Veteran Program: a mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol.* 2016;70:214-223. doi:10.1016/j.jclinepi.2015.09.016.
26. Hunter-Zinck H, Shi Y, Li M, et al.; VA Million Veteran Program. Genotyping array design and data quality control in the Million Veteran Program. *Am J Hum Genet.* 2020;106(4):535-548. doi:10.1016/j.ajhg.2020.03.004.
27. Karunamuni RA, Huynh-Le MP, Fan CC, et al.; PRACTICAL Consortium. The effect of sample size on polygenic hazard models for prostate cancer. *Eur J Hum Genet.* 2020;28(10):1467-1475. doi:10.1038/s41431-020-0664-2.
28. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA.* 1982;247(18):2543-2546. doi:10.1001/jama.1982.03320430047030.
29. Sartor O, de Bono JS. Metastatic prostate cancer. *N Engl J Med.* 2018;378(7):645-657. doi:10.1056/NEJMra1701695.
30. Moses KA, Sprenkle PC, Box G, et al. NCCN Guidelines Prostate Cancer Early Detection Version 1.2022. 2021. [https://www.nccn.org/professionals/physician\\_gls/pdf/prostate\\_detection.pdf](https://www.nccn.org/professionals/physician_gls/pdf/prostate_detection.pdf). Accessed August 22, 2022.
31. Wolf AMD, Wender RC, Etzioni RB, et al.; American Cancer Society Prostate Cancer Advisory Committee. American Cancer Society guideline for the early detection of prostate cancer: update 2010. *CA Cancer J Clin.* 2010;60(2):70-98. doi:10.3322/caac.20066.
32. Horwich A, Hugosson J, de Reijke T, et al.; European Society for Medical Oncology. Prostate cancer: ESMO Consensus Conference Guidelines 2012. *Ann Oncol.* 2013;24(5):1141-1162. doi:10.1093/annonc/mds624.
33. 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(10):761-769. doi:10.7326/0003-4819-158-10-201305210-00633.
34. Barocas DA, Grubb R 3rd, Black A, et al. Association between race and follow-up diagnostic care after a positive prostate cancer screening test in the prostate, lung, colorectal, and ovarian cancer screening trial. *Cancer.* 2013;119(12):2223-2229. doi:10.1002/cncr.28042.
35. Han Y, Rand KA, Hazelett DJ, et al. Prostate cancer susceptibility in men of African ancestry at 8q24. *J Natl Cancer Inst.* 2016;108(7):d1v431. doi:10.1093/jnci/djv431.
36. Mahal BA, Aizer AA, Ziehr DR, et al. Trends in disparate treatment of African American men with localized prostate cancer across National Comprehensive Cancer Network risk groups. *Urology.* 2014;84(2):386-392. doi:10.1016/j.urolgy.2014.05.009.
37. Yamoah K, Johnson MH, Choeurng V, et al. Novel biomarker signature that may predict aggressive disease in African American men with prostate cancer. *J Clin Oncol.* 2015;33(25):2789-2796. doi:10.1200/JCO.2014.59.8912.
38. Zhang H, Messing EM, Travis LB, et al. Age and racial differences among PSA-detected (AJCC stage T1cN0M0) prostate cancer in the U.S.: a population-based study of 70,345 men. *Front Oncol.* 2013;3:312. doi:10.3389/fonc.2013.00312.
39. Shi Z, Platz EA, Wei J, et al. Performance of three inherited risk measures for predicting prostate cancer incidence and mortality: a population-based prospective analysis. *Eur Urol.* 2021;79(3):419-426. doi:10.1016/j.eururo.2020.11.014.
40. Kachuri L, Graff RE, Smith-Byrne K, et al. Pan-cancer analysis demonstrates that integrating polygenic risk scores with modifiable risk factors improves risk prediction. *Nat Commun.* 2020;11(1):6084. doi:10.1038/s41467-020-19600-4.
41. Schecter A, Needham L, Pavuk M, et al. Agent Orange exposure, Vietnam war veterans, and the risk of prostate cancer. *Cancer.* 2009;115(14):3369-3371. doi:10.1002/cncr.24365.
42. Vickers AJ, Ulmert D, Sjoberg DD, et al. Strategy for detection of prostate cancer based on relation between prostate specific antigen at age 40-55 and long term risk of metastasis: case-control study. *BMJ.* 2013;346:f2023. doi:10.1136/bmj.f2023.
43. Lilja H, Ulmert D, Björk T, et al. Long-term prediction of prostate cancer up to 25 years before diagnosis of prostate cancer using prostate kallikreins measured at age 44 to 50 years. *J Clin Oncol.* 2007;25(4):431-436. doi:10.1200/JCO.2006.06.9351.
44. Preston MA, Gerke T, Carlsson SV, et al. Baseline prostate-specific antigen level in midlife and aggressive prostate cancer in black men. *Eur Urol.* 2019;75(3):399-407. doi:10.1016/j.eururo.2018.08.032.
45. Grossman DC, Curry SJ, Owens DK, et al.; US Preventive Services Task Force. Screening for prostate cancer. *JAMA.* 2018;319(18):1901-1913. doi:10.1001/jama.2018.3710.
46. Klein RJ, Vertosick E, Sjoberg D, et al. Prostate cancer polygenic risk score and prediction of lethal prostate cancer. *NPJ Precis Oncol.* 2022;6(1):25. doi:10.1038/s41698-022-00266-8.

47. Eklund M, Jäderling F, Discacciati A, et al.; STHLM3 Consortium. MRI-targeted or standard biopsy in prostate cancer screening. *N Engl J Med.* 2021;385(10):908-920. doi:[10.1056/NEJMoa2100852](https://doi.org/10.1056/NEJMoa2100852).
48. Jiang Y, Meyers TJ, Emeka AA, et al. Genetic factors associated with prostate cancer conversion from active surveillance to treatment. *HGG Adv.* 2022;3(1):100070. doi:[10.1016/j.xhgg.2021.100070](https://doi.org/10.1016/j.xhgg.2021.100070).
49. Ahmed M, Goh C, Saunders E, et al.; PRACTICAL Consortium. Germline genetic variation in prostate susceptibility does not predict outcomes in the chemoprevention trials PCPT and SELECT. *Prostate Cancer Prostatic Dis.* 2020;23(2):333-342. doi:[10.1038/s41391-019-0181-y](https://doi.org/10.1038/s41391-019-0181-y).
50. Plym A, Zhang Y, Stopsack KH, et al. A healthy lifestyle in men at increased genetic risk for prostate cancer. *Eur Urol.* 2022. doi:[10.1016/j.eururo.2022.05.008](https://doi.org/10.1016/j.eururo.2022.05.008). doi:[10.1016/j.eururo.2022.05.008](https://doi.org/10.1016/j.eururo.2022.05.008).