The American Journal of Human Genetics, Volume *109*

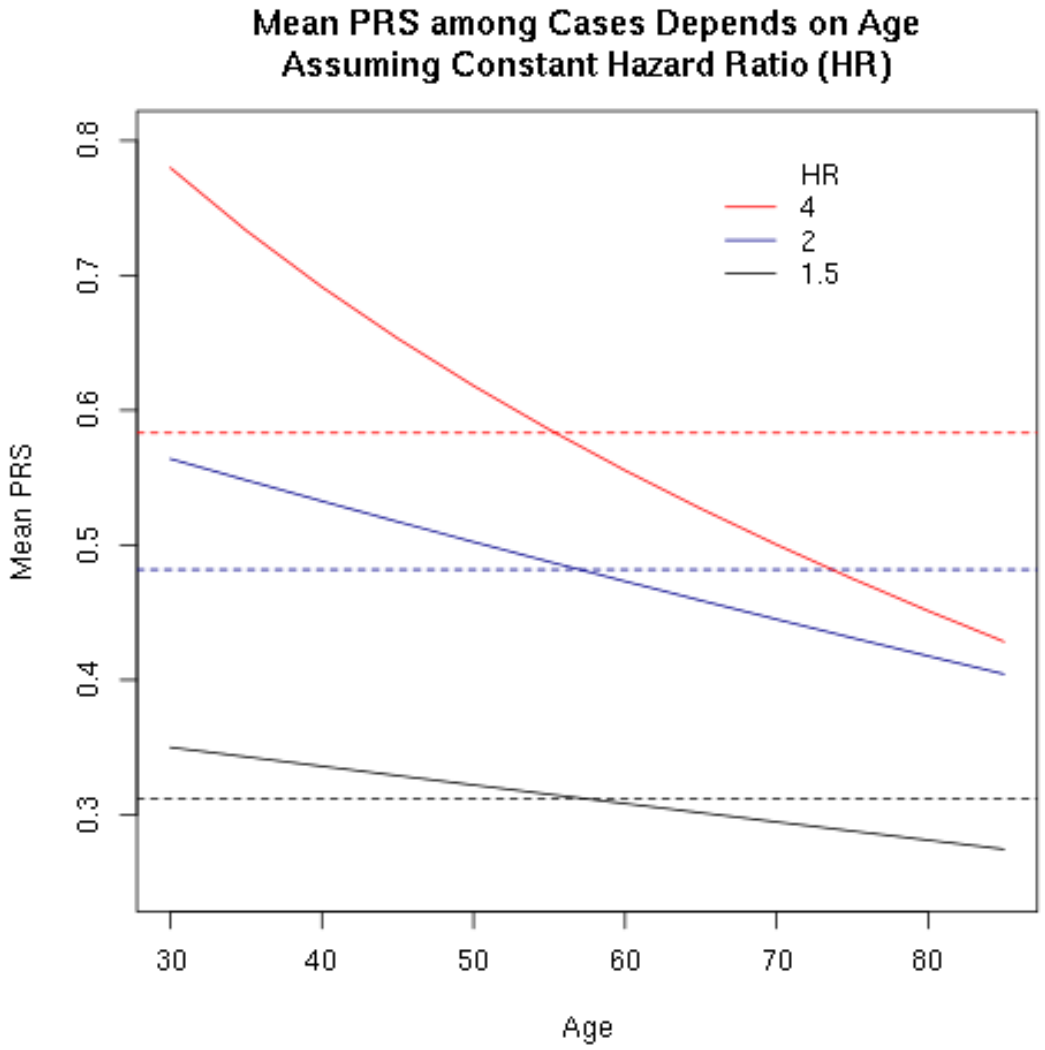
Supplemental information

Polygenic risk for prostate cancer: Decreasing relative risk with age but little impact on absolute risk

Daniel J. Schaid, Jason P. Sinnwell, Anthony Batzler, and Shannon K. McDonnell

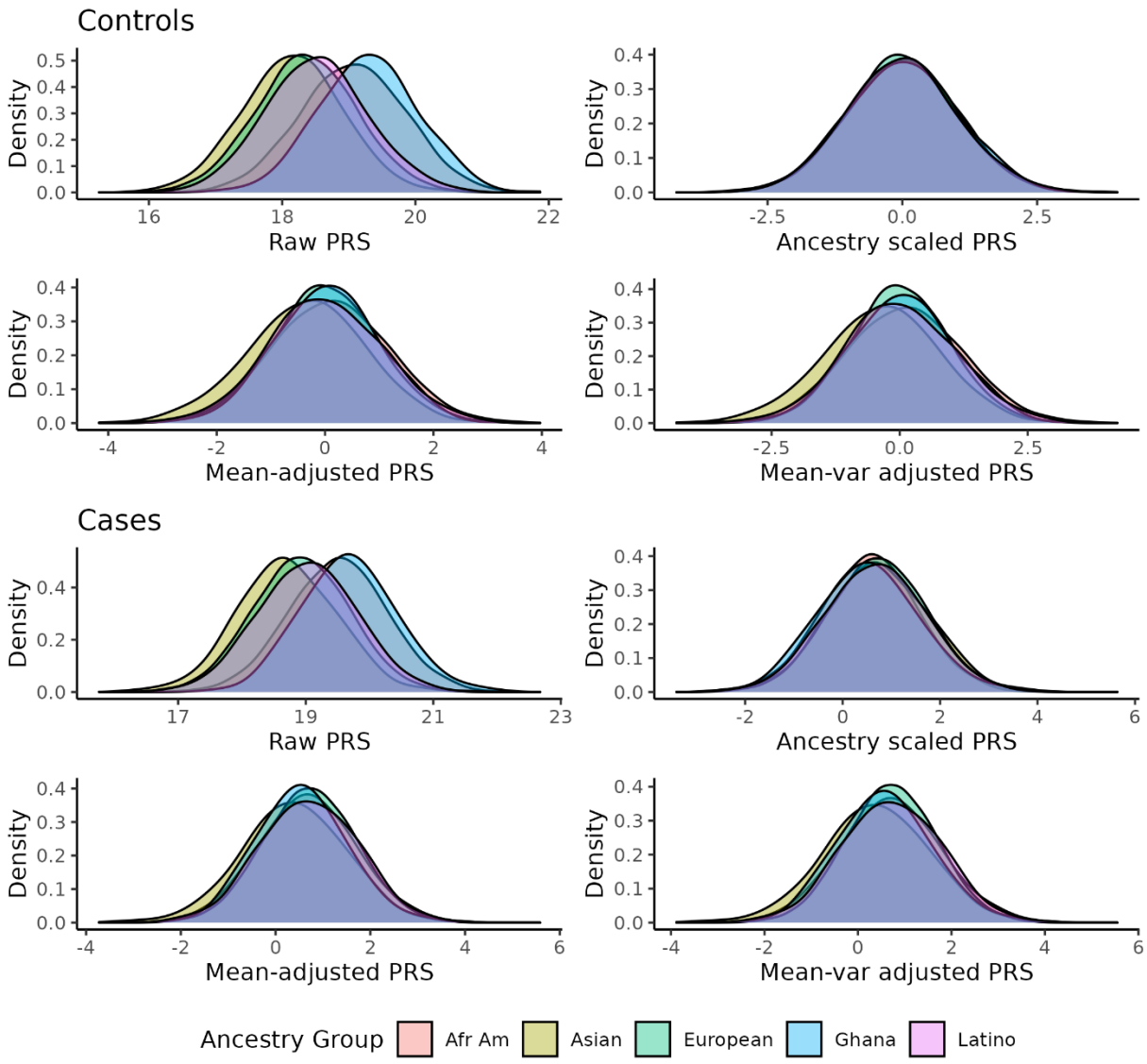
Supplemental Information

Figure S1.



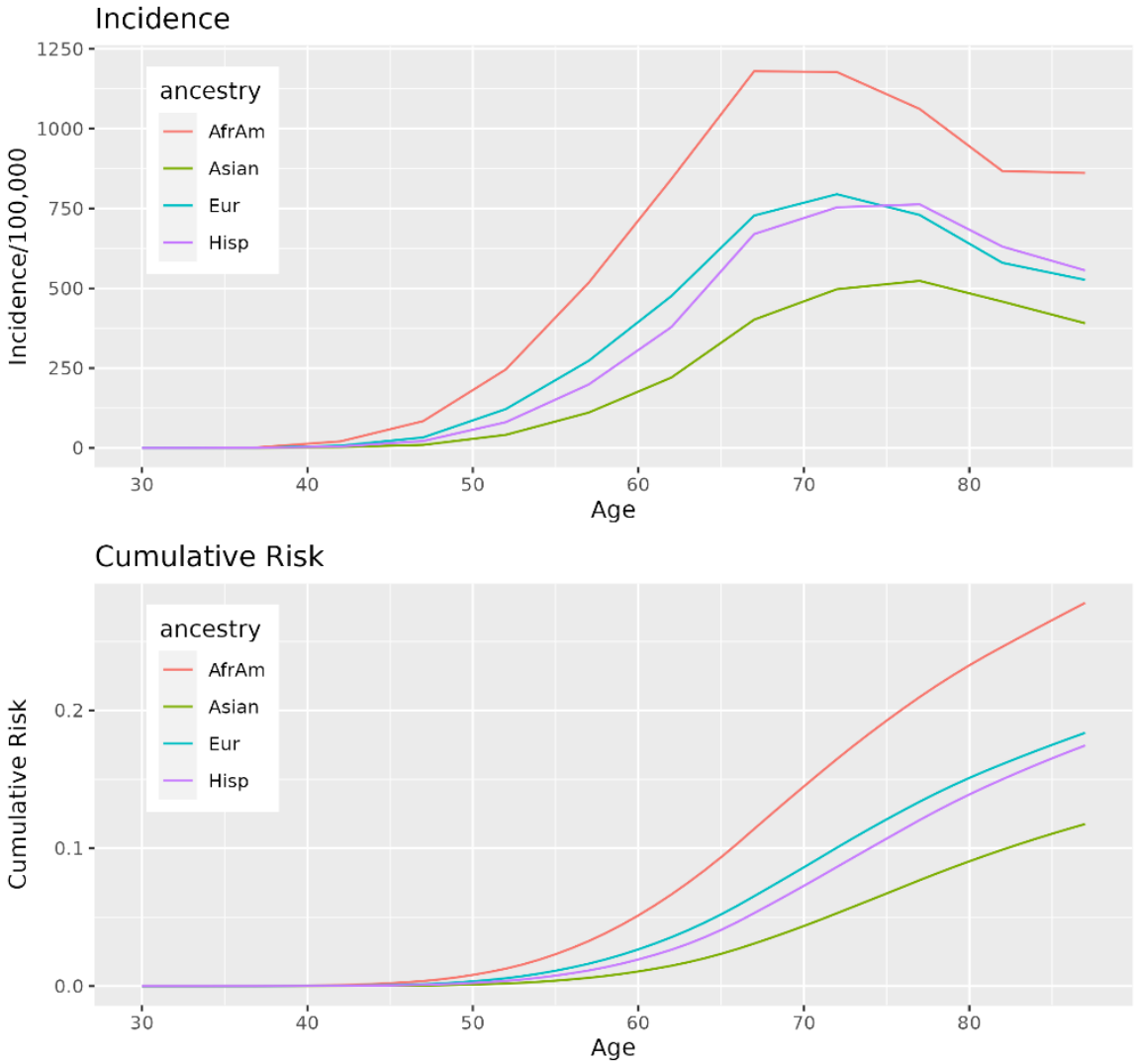
Mean PRS among cases over different ages, assuming a constant hazard ratio. The solid lines illustrate how the mean PRS is expected to decrease with age, and the horizontal dashed lines provide perspective on how the solid lines pivot from a constant value. See Methods section for Theoretical Mean PRS Among Cases & Age for derivations used to create Figure S1.

Figure S2



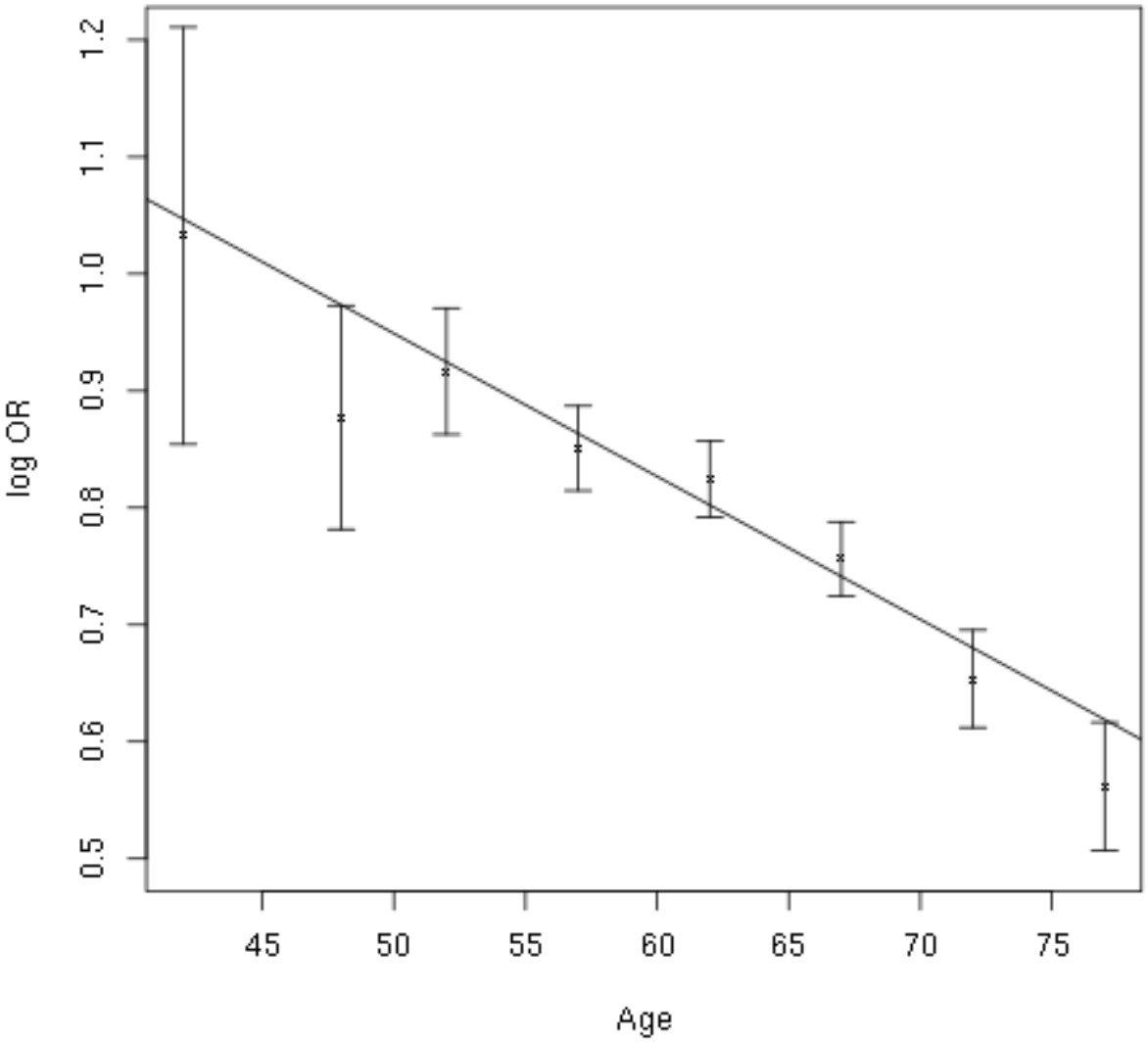
Distribution of PRS uncorrected (raw PRS) and corrected by various methods for controls and cases. See Methods section PRS Ancestry Correction by Projection onto 1,0000 Genome Reference Sample for details.

Figure S3.



Age-specific incidence of prostate cancer, per 100,000, for different ancestry groups (upper panel) and corresponding cumulative risk of prostate cancer (lower panel). See Methods section Age-Specific Incidence and Cumulative Risk by Ancestry for details.

Figure S4.



Piece-wise log-odds-ratios ("*" and their 95% confidence intervals represented as whiskers) and model assuming log-odds-ratio depends linearly on age (solid line). The piece-wise results are positioned on the x-axis at the median value of age within each age group. See Methods section on Linear Decrease in PRS Log-Odds-Ratio with Age for details.

Table S1. Studies obtained from dbGaP for Prostate Cancer Genome Wide Association Study Data. See Methods section Description of dbGaP Studies for more details

dbGaP Study Accession		Study Label	Study Full Name
phs000207.v1.p1		CGEMS	CGEMS Prostate Cancer GWAS - Stage 1 – PLCO (Embargo Release Date: December 22, 2009)
phs000306.v4.p1		GENEVA	GENEVA Prostate Cancer (Embargo Release Date: February 01, 2013)
phs000812.v1.p1		BPC3	Characterizing Genetic Susceptibility to Breast and Prostate Cancer - BPC3 (Embargo Release Date: June 11, 2015)
phs000838.v1.p1		GHANA	Ghana Prostate Study (Embargo Release Date: July 10, 2015)
phs000882.v1.p1		PEGASUS	National Cancer Institute (NCI) Prostate Cancer Genome-wide Association Study for Uncommon Susceptibility Loci (PEGASUS) (Embargo Release Date: January 24, 2017)
phs001391.v1.p1		ONCO	OncoArray: Prostate Cancer (Embargo Release Date: March 21, 2018)
phs000733.v1.p1		ICPCG	The International Consortium for Prostate Cancer Genetics Genome Wide Association Study of Familial Prostate Cancer (Embargo Release Date: March 10, 2015)

Table S2. Sample Size before and after quality control, merging, and removal of related subjects.

Study	Downloaded Data		After Initial QC (1)		After Merging and Removing Related Subjects (2)	
	Cases	Controls	Cases	Controls	Cases	Controls
BPC3	2782	4458	2775	4451	2179	3573
CGEMS	1151	1101	1140	1097	520	414
GENEVA	4304	4529	4264	4496	4115	4371
ICPCG	2568	1422	2520	1392	2106	1022
ONCO	52700	37751	52658	37723	51489	36910
PEGASUS	4599	2841	4595	2840	3670	2290
GHANA	474	458	474	458	461	452
Total	68578	52560	68426	52457	64540	49032

(1) Reasons for exclusion include low call rate (<80%), low heterozygosity (<0.4) on any chromosome, self-reported sex inconsistent with chromosome X and Y data.

(2) One subject was randomly removed from each related pair (kinship coefficient ≥ 0.0442)

Table S3: Exclusions by study.

	BPC3 (N=5752)	CGEMS (N=934)	GENEVA (N=8486)	ICPCG (N=3128)	ONCO (N=88399)	PEGASUS (N=5960)	WAFR (N=913)	Total (N=113572)
Missing Age	0	0	0	123	2677	0	0	2800
Age < 30	0	0	0	0	39	0	0	39
Missing Ancestry	0	0	0	27	0	0	0	27

Table S4. Description of studies included in analyses [a]

	BPC3 (No.=5752)	CGEMS (No.=934)	GENEVA (No.=8486)	ICPCG (No.=2978)	ONCO (No.=85683)	PEGASUS (No.=5960)	WAFR (No.=913)	Total (No.=110706)
Status								
Case	2179 (37.9%)	520 (55.7%)	4115 (48.5%)	2072 (69.6%)	51257 (59.8%)	3670 (61.6%)	461 (50.5%)	64274 (58.1%)
Control	3573 (62.1%)	414 (44.3%)	4371 (51.5%)	906 (30.4%)	34426 (40.2%)	2290 (38.4%)	452 (49.5%)	46432 (41.9%)
Ancestry								
Afr.Amer	0	0	4521 (53.3%)	0	6354 (7.4%)	0	0	10875 (9.8%)
Asian	0	0	1935 (22.8%)	0	1122 (1.3%)	0	0	3057 (2.8%)
European	5752 (100.0%)	934 (100.0%)	0	2978 (100.0%)	75999 (88.7%)	5960 (100.0%)	0	91623 (82.8%)
Ghana	0	0	0	0	0	0	913 (100.0%)	913 (0.8%)
Latino	0	0	2030 (23.9%)	0	2208 (2.6%)	0	0	4238 (3.8%)
Family History PrCa								
Unknown	1483	0	3024	0	29730	5960	913	41110
No	3826 (89.6%)	844 (90.4%)	4923 (90.1%)	906 (30.4%)	45896 (82.0%)	0	0	56395 (81.0%)
Yes	443 (10.4%)	90 (9.6%)	539 (9.9%)	2072 (69.6%)	10057 (18.0%)	0	0	13201 (19.0%)
Age Group								
[30,45)	0	0	77 (0.9%)	79 (2.7%)	1163 (1.4%)	0	2 (0.2%)	1321 (1.2%)
[45,50)	22 (0.4%)	0	206 (2.4%)	210 (7.1%)	2557 (3.0%)	0	2 (0.2%)	2997 (2.7%)
[50,55)	86 (1.5%)	0	444 (5.2%)	448 (15.0%)	8121 (9.5%)	0	134 (14.7%)	9233 (8.3%)
[55,60)	339 (5.9%)	121 (13.0%)	905 (10.7%)	628 (21.1%)	16423 (19.2%)	300 (5.0%)	146 (16.0%)	18862 (17.0%)
[60,65)	929 (16.2%)	0	1296 (15.3%)	620 (20.8%)	19219 (22.4%)	1212 (20.3%)	150 (16.4%)	23426 (21.2%)
[65,70)	1669 (29.0%)	497 (53.2%)	1816 (21.4%)	562 (18.9%)	18923 (22.1%)	1768 (29.7%)	169 (18.5%)	25404 (22.9%)
[70,75)	1413 (24.6%)	0	1796 (21.2%)	257 (8.6%)	11651 (13.6%)	1674 (28.1%)	175 (19.2%)	16966 (15.3%)
[75,88)	1294 (22.5%)	316 (33.8%)	1946 (22.9%)	174 (5.8%)	7626 (8.9%)	1006 (16.9%)	135 (14.8%)	12497 (11.3%)
Age,median (range)	69 (45,87)	67 (57, 77)	67 (44, 77)	61 (33, 87)	63 (30, 87)	67 (57, 77)	65 (42, 87)	64 (30, 87)

[a]Excluding men with missing age, age < 30 years, or missing ancestry

Table S5. Self-reported ancestry versus ancestry based on the maximum estimate of admixture probability. See Methods section Ancestry: Self—Reported and Genetically Informed for more details.

Genetic Admixture Max Probability	Self-reported Ancestry				
	African American	Asian	European	Latino	Ghana
African	10354	1	0	11	913
Amerindian	7	113	0	1637	0
Asian	33	2746	7	24	0
European	481	197	88638	2566	0

Table S6. Intercept and slope for models of log-risk of PRS as a function of age in years.

Model	Intercept (SE)	Slope (SE)
Weighted Cox	1.2504 (0.049)	-0.0138 (0.0015)
Logistic Regression	1.1935 (0.035)	-0.0122 (0.0010)

Table S7. Log-relative risk estimates (beta) and their standard errors (se) for Figure 2.

Ancestry	beta_persd	se_persd	beta_up90	se_up90
Afr Amer	0.7054	0.0309	1.3904	0.0681
Ghana	0.5132	0.0834	1.0916	0.2031
Latino	0.6858	0.0433	1.1945	0.1038
Asian	0.7743	0.0654	1.4182	0.1377
European	0.7666	0.0140	1.3720	0.0326

Table S8. Log-relative risk estimates (beta) and their standard errors (se) for Figure 3.

Ancestry	Age	beta_persd	se_persd	beta_up90	se_up90
Afr Am	[30, 55)	0.7867	0.0403	1.4177	0.0831
Afr Am	[55, 60)	0.7188	0.0408	1.3718	0.0888
Afr Am	[60, 65)	0.6699	0.0417	1.3081	0.0937
Afr Am	[65, 70)	0.7160	0.0473	1.4152	0.1084
Afr Am	[70, 88)	0.6454	0.0593	1.4397	0.1570
Ghana	[30, 55)	0.7179	0.1838	1.1702	0.4502
Ghana	[55, 60)	0.6935	0.1604	1.6039	0.4118
Ghana	[60, 65)	0.7092	0.1485	1.6726	0.3498
Ghana	[65, 70)	0.5790	0.1562	0.9123	0.4200
Ghana	[70, 88)	0.3404	0.0963	0.8264	0.2365
Latino	[30, 55)	0.7937	0.0848	1.4939	0.1839
Latino	[55, 60)	0.8194	0.0774	1.4036	0.1577
Latino	[60, 65)	0.7615	0.0622	1.2271	0.1374
Latino	[65, 70)	0.6782	0.0614	1.2064	0.1367
Latino	[70, 88)	0.5812	0.0651	1.0218	0.1673
Asian	[30, 55)	0.8122	0.1175	1.4138	0.2478
Asian	[55, 60)	0.7960	0.1041	1.5453	0.2087
Asian	[60, 65)	0.9118	0.1006	1.6689	0.1861
Asian	[65, 70)	0.8676	0.0942	1.4500	0.1744
Asian	[70, 88)	0.6744	0.0933	1.2949	0.2177
Eur	[30, 55)	0.9384	0.0177	1.6969	0.0332
Eur	[55, 60)	0.8715	0.0156	1.5575	0.0304
Eur	[60, 65)	0.8081	0.0157	1.4314	0.0326
Eur	[65, 70)	0.7178	0.0183	1.2914	0.0403
Eur	[70, 88)	0.6227	0.0296	1.1045	0.0784

Table S9. Tests of heterogeneity of relative risks across ages for parameters in Table S8.

	Test of Heterogeneity of Relative Risk across Ages	
Ancestry	Per SD PRS	Upper 90th Percentile of PRS
Afr Am	0.221	0.902
Ghana	0.112	0.228
Latino	0.099	0.321
Asian	0.473	0.746
Eur	0.000	0.000

Table S10. Log-relative risk estimates (beta) and their standard errors (se) for Figure 5.

FamHx	Age	beta_perSD	se_perSD
No	[30, 55)	0.8922	0.0236
No	[55, 60)	0.8427	0.0204
No	[60, 65)	0.7824	0.0221
No	[65, 70)	0.6713	0.0230
No	[70, 88)	0.6141	0.0382
Yes	[30, 55)	0.9803	0.0459
Yes	[55, 60)	0.9366	0.0464
Yes	[60, 65)	0.8072	0.0526
Yes	[65, 70)	0.7777	0.0617
Yes	[70, 88)	0.7432	0.1107

Methods

Theoretical Mean PRS Among Cases & Age

The mean PRS among cases depends on the strength of association of the PRS with disease and it is possible for the mean PRS to be greater among younger cases than older cases, even if the hazard ratio associated with a PRS is constant over all ages. This is because men who have greater values of PRS are at the greatest susceptibility for disease and are more likely to succumb at a younger age.

Below we derive the expected PRS among cases, and how this expectation depends on age, when assuming a constant hazard ratio (e.g., proportional hazards model). The derivation follows standard methods for survival analyses. Assume that the standardized PRS, z, has a standard normal density, $\phi(z)$. The probability of disease at age a, conditional on z, is

$P(a \mid z) = \lambda_o(a) e^{\beta z} S_o(a)^{\exp(\beta z)}$, where $\lambda_o(a)$ is the baseline hazard rate, $S_o(a) = \exp[-\sum_{t=0}^{a} \lambda_o(t)]$, and β is the log hazard ratio constant over age. From these, we determine the density of z conditional on disease at age a:

$$P(z \mid a) = \frac{\lambda_o(a) e^{\beta z} S_o(a)^{\exp(\beta z)} \phi(z)}{\int_{-\infty}^{\infty} \lambda_o(a) e^{\beta z} S_o(a)^{\exp(\beta z)} \phi(z) \partial z}. \quad (1)$$

The expected value of z among diseased cases at age a is then

$$E[z \mid a] = \int_{-\infty}^{\infty} zP(z \mid a)dP \qquad (2)$$

To illustrate this numerically, we assume that age of disease diagnosis has an exponential distribution (i.e., constant hazard rate of λ=.003, the mean baseline incidence for European ancestry), and a constant log hazard ratio of β, making it easy to numerically integrate equations (1) and (2). Figure S1 illustrates how the mean PRS among cases decreases with age, while the mean PRS among controls is expected to be approximately zero.

PRS Ancestry Correction by Projection onto 1,0000 Genome Reference Sample

Because the distribution of PRS differs across different ancestries due to SNP allele frequency differences, we evaluated three approaches to correct for population differences: 1) centering and scaling the PRS within each ancestry group, using the mean and standard deviation for controls within each ancestry group; 2) projection of data onto 1,000 Genome reference panel and correction of mean PRS;[1] 3) projection of data onto 1,000 Genome reference panel and correction of both mean PRS and variance of PRS. These latter two methods intend to provide a continuum of correction for men of different ancestries, some admixed.

The projection methods #2 and #3 (Christopher Kachulis, Broad Institute, personal communication), are based on projecting the study sample PRS onto a reference sample. This is accomplished by computing the PRS on the reference sample and using linear regression to regress the reference PRS on the top (maybe 10) principal components of the reference sample. The regression coefficients from this reference regression are used to predict the PRS in the study sample, by using the sample principal components. This predicted value is then subtracted from the sample PRS to adjust for ancestry. Because this approach only corrects for the mean of the distribution, it might not fully correct for ancestry if the variance of the PRS differs across ancestry. To adjust for the variance, one can create residuals from the linear regression in the reference sample, and then perform a second linear regression of the squared residuals on the principal components in the reference sample. This can then be used to predict the variance in the study sample, by using the regression coefficients with the sample principal components. To illustrate, the PRS adjusted score would be computed as

$$PRS_{adjusted} = \frac{PRS_{sample} - (\alpha_o + \sum \alpha_i PC_i)}{\sqrt{\beta_o + \sum \beta_i PC_i}},$$

where α coefficients are estimated by regression of the PRS on the principal components in the reference samples, β coefficients are estimated by regression of the squared residuals on the principal components in the reference samples**,** and PRS_{sample} and PC_i are from the study samples.

When computing the above PRS corrections, the genetic variants need to be available in both the study and reference samples. For the prostate cancer PRS, there were 220 variants available in our study sample, but only 212 of these were available in the 1000 Genome reference (8 variants were in our prostate cancer studies but not available in the 1000 Genome reference data). To compute the principal components, we removed the 212 risk variants and removed variants in linkage disequilibrium, resulting in approximately 100,000 variants.

The panels in Figure S2 below illustrate the distribution of the raw PRS, the PRS corrected by self-reported ancestry mean and standard deviation among controls ("Ancestry scaled PRS"), the PRS corrected by projection on the reference samples, correcting for the mean ("Mean-adjusted PRS"), and the PRS corrected by projection correcting for both mean and variance ("Mean-Var adjusted PRS"). For our prostate cancer study, it can be seen that the projection methods do not fully correct for ancestry, by viewing the non-overlapping distributions among controls. In contrast, the self-reported ancestry mean and standard deviation correction performed better.

Age-Specific Incidence and Cumulative Risk by Ancestry

The age-specific incidence rates of prostate cancer, as described in the main manuscript methods, are illustrated in the upper panel of Figure S3. This figure illustrates the greater incidence rate among African American ancestry across all ages, the lesser incidence among Asian ancestry and similar incidence among European and Latino (Hispanic) ancestries. The incidence rates, λ_t, can be used to compute the cumulative incidence of prostate cancer by age *a*, $F(a) = \exp[-\sum_{t=0}^{a} \lambda_t]$. The cumulative incidences in the ancestral populations are illustrated in the lower panel of Figure S3. By age 87, the life-time risk of prostate cancer is expected to be 28% among African American men, 18% among European ancestry, 17% among Latino ancestry, and 12% among Asian ancestry.

Linear Decrease in PRS Log-Odds-Ratio with Age

The results in the main manuscript, illustrated in Figure 4, were based on fitting Cox proportional hazards models with weights the inverse of population incidence rates for controls. Since men were enrolled based on case-control studies, we evaluated the sensitivity of our conclusions of log-relative-risk decreasing linearly with age among European ancestry by fitting piece-wise logistic regression models for age partitioned into 5 year intervals, from age 30 to age 88. In addition, a logistic regression model assuming linear change in PRS risk according to age was fit by the form of status ~ cohort + (age-30) + PRS + I((age-30)* PRS). The results from piece-wise fits and the linear model are illustrated in Figure S4. The linear decrease in log-odds-ratio fits the data well for ages 50-70. Table S6 below shows that parameter estimates and their standard errors for the weighted Cox model and the logistic regression model are consistent, and quite close for the linear decrease in log risk.

Description of dbGaP Studies

A brief description of each the studies obtained from dbGaP and listed Table S1 is provided below. Complete descriptions are available from the dbGaP web site (https://www.ncbi.nlm.nih.gov/gap/)

CGEMS

The Cancer Genetic Markers of Susceptibility (CGEMS) prostate cancer genome-wide association study (GWAS) included genotyping approximately 550,000 SNPs (Phase 1A with HumanHap300 and Phase 1B HumanHap240, both from Illumina, San Diego, CA) in 1,172 prostate cancer patients and 1,157 controls of European ancestry from the Prostate, Lung, Colon and Ovarian (PLCO) Cancer Screening Trial. Selected publications include:[2]

Acknowledgement**:** Data submitted to dbGaP by Lead Principal Investigator Stephen J. Chanock. Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute and Core Genotyping Facility, Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute (NCI), National Institutes of Health, Department (NIH), Department of Health and Human Services (DHHS), Bethesda, MD, USA. dbGaP accession phs000207.v1.p1.

GENEVA

This study is part of the Gene Environment Association Studies initiative (GENEVA, http://www.genevastudy.org) funded by the trans-NIH Genes, Environment, and Health Initiative (GEI). The version 1 release of this dataset included genotype data for the Japanese and Latino populations in the study. Genotyping was performed at the Broad Institute of MIT and Harvard, a GENEVA genotyping center and at the University of Southern California. Selected publications include:[3; 4]

Acknowledgement: Funding support for the GENEVA Prostate Cancer study was provided through the National Cancer Institute (R37CA54281, R01CA6364, P01CA33619, U01CA136792, and U01CA98758) and the National Human Genome Research Institute (U01HG004726). Assistance with phenotype harmonization, SNP selection, data cleaning, meta-analyses, data management and dissemination, and general study coordination, was provided by the GENEVA Coordinating Center (U01HG004789-01). dbGaP accession phs000306.v4.p1.

BPC3

The Breast and Prostate Cancer Cohort Consortium (BPC3) was established in 2003 to pool data and biospecimens from nine large prospective cohorts to conduct research on gene-environment interactions in cancer etiology. The BPC3 GWAS includes the following cohorts: the American Cancer Society Cancer Prevention Study-II (CPS-II); the European Prospective Investigation of Cancer (EPIC); the Physician's Health Study (PHS); the Nurses' Health Studies I and II (NHS and NHSII); the Health Professionals Follow-up Study (HPFS); the Multiethnic Cohort (MEC); the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial; and the Alpha-Tocopherol, Beta-Carotene (ATBC) Study. Selected publications include:[5-7]

Acknowledgement: The Breast and Prostate Cancer Cohort Consortium (BPC3) genome-wide association studies of advanced prostate cancer and estrogen-receptor negative breast cancer was supported by the National Cancer Institute under cooperative agreements U01-CA98233, U01-CA98710, U01-CA98216, and U01-CA98758 and the Intramural Research Program of the National Cancer Institute, Division of Cancer Epidemiology and Genetics. dbGaP accession phs000812.v1.p1

GHANA

Participants were recruited through the Ghana Prostate Study (a population-based component and a clinical component) between 2004 and 2006. Additional prostate cancer cases were recruited between 2008 and 2012. Selected publications include:[8-14]

Acknowledgement: The genome-wide association study of prostate cancer in West African men project was supported by the Intramural Research Program of the National Cancer Institute, National Institutes of Health, Department of Health and Human Services including Contract No.

HHSN261200800001E. The datasets have been accessed through the NIH database for Genotypes and Phenotypes (dbGaP). A full list of acknowledgements can be found in the supplementary note [8]. dbGaP accession phs000838.v1.p1

PEGASUS

This genome-wide association study was funded by the National Cancer Institute (NCI) to identify uncommon susceptibility loci for prostate cancer. A total of 7440 subjects of European ancestry from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial were genotyped using the Illumina HumanOmni2.5.

Acknowledgement: The National Cancer Institute (NCI) Prostate Cancer Genome-wide Association Study for Uncommon Susceptibility Loci (PEGASUS) was supported by the Intramural Research Program of the NCI. Please see publication number 14 dbGaP accession phs000882.v1.p1.[15]

ONCO

Original description of the study: From ELLIPSE (linked to the PRACTICAL consortium), ~78,000 SNPs were contributed to the OncoArray. A large fraction of the content was derived from the GWAS meta-analyses in European ancestry populations (overall and aggressive disease; ~27K SNPs). An additional just over 10,000 SNPs were selected from the meta-analyses in the non-European populations, with a majority of these SNPs coming from the analysis of overall prostate cancer in African ancestry populations as well as from the multiethnic meta-analysis. A substantial fraction of SNPs (~28,000) were also selected for fine-mapping of 53 loci not included in the common fine-mapping regions (tagging at r2>0.9 across ±500kb regions). A few thousand SNPs related with PSA levels and/or disease survival as well as SNPs from candidate lists provided by study collaborators, as well as from meta-analyses of exome SNP chip data from the Multiethnic Cohort and UK studies, were also selected. A large number of studies contributed to the total sample (99,622): see description at the web link https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001391.v1.p1.

Acknowledgement: dbGaP accession phs001391.v1.p1. See below for **OncoArray: Prostate Cancer Acknowledgements.**

ICPCG

The aim of this study was to perform a GWAS for prostate cancer cases that came from pedigrees with multiple men affected with prostate cancer. Pedigrees were identified that had 3 or more related prostate cancer cases and have an average age at diagnosis ≤ 75 years. Only one case was chosen from each pedigree; if more than one case was available in a pedigree, the most

aggressive case was chosen, or if no aggressive cases, the case with the earliest age of diagnosis was chosen. Male controls were selected such that they were unrelated to cases and to each other and had a distribution of race and birth year similar to the cases. A GWAS was performed based on genotyping with the Illumina 5M plus exome SNP set. Selected references include:[16-19]

Acknowledgement: Data was provided by principal investigator Lisa Cannon Albright, PhD. The University of Utah, UT, USA, and funding provided by R01 CA089600. National Institutes of Health, Bethesda, MD, USA. The genotyping data was generated and provided by the International Consortium for Prostate Cancer Genetics (ICPCG). The ICPCG was funded by a grant from the National Institutes of Health, U01 CA89600. dbGaP accession phs000733.v1.p1.

Ancestry: Self-Reported and Genetically Informed

The program ADMIXTURE[20] was used to estimate genetic admixture probabilities using a reference sample of 1000 Genome supplemented with data from the Human Genome Diversity Project.[21] We found the ADMIXURE software to provide odd results for very large sample sizes (~100,000), presumably due to numerical accuracy when computing the log-likelihood. For this reason, we partitioned samples into batches of size no greater than 10,000. The results in Table S5 below illustrate that potential misclassification was minimal among subjects self-reported as African American, European and Ghana. Asian subjects had a small number genetically classified as Amerindian, yet there is close ancestry among Amerindian and Asians, so these might represent historical ancestries. Latino ancestry is known to be admixed among Amerindian, European, and African. Note that the classification by maximum genetic admixture probability includes subjects that are bordering 50:50 admixture between two ancestries, such as 51% European and 49% African who self-report as African.

Model to Adjust Weights to Fit Age Distribution Among Cases

To evaluate the sensitivity of the weights in the Cox model, we modified the population incidence (hazard) rates to better fit the age distribution of the cases to allow for the possibility that the cases were sampled with preference to certain ages. Based on theory of survival analysis, the probability density of an event occurring at age t is

$$f(t) = \lambda_t S(t)$$

where λ_t is the population hazard rate and $S(t) = \exp(-\sum_i^t \lambda_t)$, assuming age is partitioned into one year increments. Assuming multiplicative changes to λ_t to adapt our weights to the age of diagnosis among the cases, we modeled the hazards as $\lambda_t e^{\beta_t}$, where β_t are parameters to estimate. The likelihood for the cases is

$$L = \prod_{t=30}^{87} \left(\lambda_t e^{\beta_t} \exp(-\sum_i^t \lambda_t e^{\beta_t}) \right)^{N_t}$$

Where $30 \le t \le 87$ for our cases and N_t is the number of cases at age of diagnosis t. We estimated the β_t parameters by the Newton-Raphson method, and use the revised hazard rates $\lambda_t e^{\beta_t}$ to weight the controls (weight of $1/(\lambda_t e^{\beta_t})$, and used ancestry-specific λ_t and β_t in these computations.

OncoArray: Prostate Cancer
Acknowledgments

Aarhus: This study was supported by the Danish Strategic Research Council (now Innovation Fund Denmark) and the Danish Cancer Society. The Danish Cancer Biobank (DCB) is acknowledged for biological material.

AHS: This work was supported by the Intramural Research Program of the NIH, National Cancer Institute, Division of Cancer Epidemiology and Genetics (Z01CP010119).

ATBC: This research was supported in part by the Intramural Research Program of the NIH and the National Cancer Institute. Additionally, this research was supported by U.S. Public Health Service contracts N01-CN-45165, N01-RC-45035, N01-RC-37004, HHSN261201000006C, and HHSN261201500005C from the National Cancer Institute, Department of Health and Human Services.

BioVu: The dataset(s) used for the analyses described were obtained from Vanderbilt University Medical Center's BioVU which is supported by institutional funding and by the National Center for Research Resources, Grant UL1 RR024975-01 (which is now at the National Center for Advancing Translational Sciences, Grant 2 UL1 TR000445-06).

Canary PASS: PASS was supported by Canary Foundation and the National Cancer Institute's Early Detection Research Network (U01 CA086402)

CCI: This work was awarded by Prostate Cancer Canada and is proudly funded by the Movember Foundation - Grant # D2013-36.The CCI group would like to thank David Murray, Razmik Mirzayans, and April Scott for their contribution to this work.

CeRePP: None reported

COH: SLN is partially supported by the Morris and Horowitz Families Endowed Professorship

COSM: The Swedish Research Council, the Swedish Cancer Foundation

CPCS1 & CPCS2: Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev Ringvej 75, DK-2730 Herlev, DenmarkCPCS1 would like to thank the participants and staff of the Copenhagen General Population Study for their important contributions.

CPDR: Uniformed Services University for the Health Sciences HU0001-10-2-0002 (PI: David G. McLeod, MD)

CPS-II: The American Cancer Society funds the creation, maintenance, and updating of the Cancer Prevention Study II cohort.CPS-II thanks the participants and Study Management Group for their invaluable contributions to this research. We would also like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries, and cancer registries supported by the National Cancer Institute Surveillance Epidemiology and End Results program.

EPIC: The coordination of EPIC is financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The national cohorts are supported by the Danish Cancer Society (Denmark); the Deutsche Krebshilfe, Deutsches Krebsforschungszentrum and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation, Greek Ministry of

Health; Greek Ministry of Education (Greece); the Italian Association for Research on Cancer (AIRC) and National Research Council (Italy); the Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF); the Statistics Netherlands (The Netherlands); the Health Research Fund (FIS), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, Spanish Ministry of Health ISCIII RETIC (RD06/0020), Red de Centros RCESP, C03/09 (Spain); the Swedish Cancer Society, Swedish Scientific Council and Regional Government of Skåne and Västerbotten, Fundacion Federico SA (Sweden); the Cancer Research UK, Medical Research Council (United Kingdom).

EPICAP: The EPICAP study was supported by grants from Ligue Nationale Contre le Cancer, Ligue départementale du Val de Marne; Fondation de France; Agence Nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES)The EPICAP study group would like to thank all urologists, Antoinette Anger and Hasina Randrianasolo (study monitors), Anne-Laure Astolfi, Coline Bernard, Oriane Noyer, Marie-Hélène De Campo, Sandrine Margaroline, Louise N'Diaye, Sabine Perrier-Bonnet (Clinical Research nurses)

ERSPC: This study was supported by the DutchCancerSociety(KWF94-869,98-1657,2002-277,2006-3518, 2010-4800); The Netherlands Organisation for HealthResearch and Development (ZonMW-002822820,22000106,50-50110-98-311, 62300035), The Dutch Cancer Research Foundation(SWOP), and an uncoditional grant from Beckman-Coulter-HybritechInc.

ESTHER: The ESTHER study was supported by a grant from the Baden Württemberg Ministry of Science, Research and Arts.The ESTHER group would like to thank Hartwig Ziegler, Sonja Wolf, Volker Hermann, Heiko Müller, Karina Dieffenbach, Katja Butterbach for valuable contributions to the study.

FHCRC: The FHCRC studies were supported by grants R01-CA056678, R01-CA082664, and R01-CA092579 from the US National Cancer Institute, National Institutes of Health, with additional support from the Fred Hutchinson Cancer Research Center. FHCRC would like to thank all the men who participated in these studies.

Gene-PARE: The Gene-PARE study was supported by grants 1R01CA134444 from the U.S. National Institutes of Health , PC074201 and W81XWH-15-1-0680 from the Prostate Cancer Research Program of the Department of Defense and RSGT-05-200-01-CCE from the American Cancer Society

Hamburg-Zagreb: None reported

HPFS: The Health Professionals Follow-up Study was supported by grants UM1CA167552, CA133891, CA141298, and P01CA055075.HPFS are grateful to the participants and staff of the Physicians' Health Study and Health Professionals Follow-Up Study for their valuable contributions, as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY.

IMPACT: The IMPACT study was funded by The Ronald and Rita McAulay Foundation, CR-UK Project grant (C5047/A1232), Cancer Australia, AICR Netherlands A10-0227, Cancer Australia and Cancer Council Tasmania, NIHR, EU Framework 6, Cancer Councils of Victorial and South Australia, Philanthropic donation to Northshore University Health System. We acknowledge support from the National Institute for Health Research (NIHR) to the Biomedical Research Centre at The Institute of Cancer Research and Royal Marsden Foundation NHS Trust. IMPACT acknowledge the IMPACT study steering committee, collaborating centres and participants.

IPO-Porto: The IPO-Porto study was funded by Fundação para a Ciência e a Tecnologia (FCT; UID/DTP/00776/2013 and PTDC/DTP-PIC/1308/2014) and by IPO-Porto Research Center (CI-IPOP-16-2012 and CI-IPOP-24-2015). MC and MPS are research fellows from Liga Portuguesa Contra o Cancro, Núcleo Regional do Norte. SM is a research fellow from FCT (SFRH/BD/71397/2010).IPO-Porto would like to express our gratitude to all patients and families who have participated in this study.

Karuprostate: The Karuprostate study was supported by the the Frech National Health Directorate and by the Association pour la Recherche sur les Tumeurs de la ProstateKarusprostate thanks Séverine Ferdinand

KULEUVEN: F.C. and S.J. are holders of grants from FWO Vlaanderen (G.0684.12N and G.0830.13N), the Belgian federal government (National Cancer Plan KPC_29_023), and a Concerted Research Action of the KU Leuven (GOA/15/017). TVDB is holder of of a doctoral fellowship of the FWO.

LAAPC: This study was funded by grant R01CA84979 (to S.A. Ingles) from the National Cancer Institute, NIH.

Malaysia: The study was funded by the University Malaya High Impact Research Grant (HIR/MOHE/MED/35).Malaysia thanks all associates in the Urology Unit, University of Malaya, Cancer Research Initiatives Foundation (CARIF) and the Malaysian Men's Health Initiative (MMHI).

MCCS: MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index and the Australian Cancer Database.

MCC-Spain: The study was partially funded by the Accion Transversal del Cancer, approved on the Spanish Ministry Council on the 11th October 2007, by the Instituto de Salud Carlos III-FEDER (PI08/1770, PI09/00773-Cantabria, PI11/01889-FEDER, PI12/00265, PI12/01270, PI12/00715), by the Fundación Marqués de Valdecilla (API 10/09), by the Spanish Association Against Cancer (AECC) Scientific Foundation and by the Catalan Government DURSI grant 2009SGR1489. Samples: Biological samples were stored at the Parc de Salut MAR Biobank (MARBiobanc; Barcelona) which is supported by Instituto de Salud Carlos III FEDER (RD09/0076/00036). Also sample collection was supported by the Xarxa de Bancs de Tumors de Catalunya sponsored by Pla Director d'Oncologia de Catalunya (XBTC).MCC-Spain acknowledges the contribution from Esther Gracia-Lavedan in preparing the data. We thank all the subjects who participated in the study and all MCC-Spain collaborators.

MD Anderson: Prostate Cancer Case-Control Studies at MD Anderson (MDA) supported by grants CA68578, ES007784, DAMD W81XWH-07-1-0645 and CA140388.

MDACC_AS: None reported

MEC: Funding provided by NIH grant U19CA148537 and grant U01 CA 164973.

MIAMI (WFPCS): ACS

MOFFITT: The Moffitt group was supported by the US National Cancer Institute (R01CA128813, PI: J.Y. Park).

NMHS: Funding for the Nashville Men's Health Study (NMHS) was provided by the National Institutes of Health Grant numbers: RO1CA121060

PCaP only data: The North Carolina - Louisiana Prostate Cancer Project (PCaP) is carried out as a collaborative study supported by the Department of Defense contract DAMD 17-03-2-0052. For HCaP-NC follow-up data: The Health Care Access and Prostate Cancer Treatment in North Carolina (HCaP-NC) study is carried out as a collaborative study supported by the American Cancer Society award RSGT-08-008-01-CPHPS. For studies using both PCaP and HCaP-NC follow-up data please use: The North Carolina - Louisiana Prostate Cancer Project (PCaP) and the Health Care Access and Prostate Cancer Treatment in North Carolina (HCaP-NC) study are carried out as collaborative studies supported by the Department of Defense contract DAMD 17-03-2-0052 and the American Cancer Society award RSGT-08-008-01-CPHPS, respectively. For any PCaP data please include: The authors thank the staff, advisory committees and research subjects participating in the PCaP study for their important contributions. For studies using PCaP DNA/genotyping data please include: We would like to acknowledge the UNC BioSpecimen Facility and the LSUHSC Pathology Lab for our DNA extractions, blood processing, storage and sample disbursement (https://genome.unc.edu/bsp). For studies using PCaP tissue please include: We would like to acknowledge the RPCI Department of Urology Tissue Microarray and Immunoanalysis Core for our tissue processing, storage and sample disbursement. For studies using HCaP-NC follow-up data please use: The Health Care Access and Prostate Cancer Treatment in North Carolina (HCaP-NC) study is carried out as a collaborative study supported by the American Cancer Society award RSGT-08-008-01-CPHPS. The authors thank the staff, advisory committees and research subjects participating in the HCaP-NC study for their important contributions. For studies that use both PCaP and HCaP-NC please use: The authors thank the staff, advisory committees and research subjects participating in the PCaP and HCaP-NC studies for their important contributions.

PCMUS: The PCMUS study was supported by the Bulgarian National Science Fund, Ministry of Education and Science (contract DOO-119/2009; DUNK01/2-2009; DFNI-B01/28/2012) with additional support from the Science Fund of Medical University - Sofia (contract 51/2009; 8I/2009; 28/2010;).

PHS: The Physicians' Health Study was supported by grants CA34944, CA40360, CA097193, HL26490 and HL34595.PHS members are grateful to the participants and staff of the Physicians' Health Study and Health Professionals Follow-Up Study for their valuable contributions, as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY.

PLCO: This PLCO study was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIHPLCO thanks Drs. Christine Berg and Philip Prorok, Division of Cancer Prevention at the National Cancer Institute, the screening center investigators and staff of the PLCO Cancer Screening Trial for their contributions to the PLCO Cancer Screening Trial. We thank Mr. Thomas Riley, Mr. Craig Williams, Mr. Matthew Moore, and Ms. Shannon Merkle at Information Management Services, Inc., for their management of the data and Ms. Barbara O'Brien and staff at Westat, Inc. for their contributions to the PLCO Cancer Screening Trial. We also thank the PLCO study participants for their contributions to making this study possible.

Poland: None reported

PROCAP: PROCAP was supported by the Swedish Cancer Foundation (08-708, 09-0677).PROCAP thanks and acknowledges all of the participants in the PROCAP study. We thank Carin Cavalli-Björkman and Ami Rönnberg Karlsson for their dedicated work in the collection of data. Michael Broms is acknowledged for his skilful work with the databases. KI Biobank is acknowledged for handling the samples and for DNA extraction. We acknowledge The NPCR steering group: Pär Stattin (chair), Anders Widmark, Stefan Karlsson, Magnus Törnblom, Jan Adolfsson, Anna Bill-Axelson, Ove Andrén, David

Robinson, Bill Pettersson, Jonas Hugosson, Jan-Erik Damber, Ola Bratt, Göran Ahlgren, Lars Egevad, and Roy Ehrnström.

PROGReSS: The PROGReSS study was suported by grants from the Instituto de Salud Carlos III (FIS PI10/00164 and FIS PI13/02030) and Fondo Europeo de Desarrollo Regional (FEDER 2007-2013).

ProMPT : Founded by CRUK, NIHR, MRC, Cambride Biomedical Research Centre

ProtecT: Founded by NIHRProtecT and ProMPT would like to acknowledge the support of The University of Cambridge, Cancer Research UK. Cancer Research UK grants [C8197/A10123] and [C8197/A10865] supported the genotyping team. We would also like to acknowledge the support of the National Institute for Health Research which funds the Cambridge Bio-medical Research Centre, Cambridge, UK. We would also like to acknowledge the support of the National Cancer Research Prostate Cancer: Mechanisms of Progression and Treatment (PROMPT) collaborative (grant code G0500966/75466) which has funded tissue and urine collections in Cambridge. We are grateful to staff at the Welcome Trust Clinical Research Facility, Addenbrooke's Clinical Research Centre, Cambridge, UK for their help in conducting the ProtecT study. We also acknowledge the support of the NIHR Cambridge Biomedical Research Centre, the DOH HTA (ProtecT grant) and the NCRI / MRC (ProMPT grant) for help with the bio-repository. The UK Department of Health funded the ProtecT study through the NIHR Health Technology Assessment Programme (projects 96/20/06, 96/20/99). The ProtecT trial and its linked ProMPT and CAP (Comparison Arm for ProtecT) studies are supported by Department of Health, England; Cancer Research UK grant number C522/A8649, Medical Research Council of England grant number G0500966, ID 75466 and The NCRI, UK. The epidemiological data for ProtecT were generated though funding from the Southwest National Health Service Research and Development. DNA extraction in ProtecT was supported by USA Dept of Defense award W81XWH-04-1-0280, Yorkshire Cancer Research and Cancer Research UK. The authors would like to acknowledge the contribution of all members of the ProtecT study research group. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Department of Health of England. The bio-repository from ProtecT is supported by the NCRI (ProMPT) Prostate Cancer Collaborative and the Cambridge BMRC grant from NIHR.

PROtEuS: PROtEuS was supported financially through grants from the Canadian Cancer Society [13149, 19500, 19864, 19865] and the Cancer Research Society, in partnership with the Ministère de l'enseignement supérieur, de la recherche, de la science et de la technologie du Québec, and the Fonds de la recherche du Québec - Santé.PROtEuS would like to thank its collaborators and research personnel, and the urologists involved in subjects recruitment. We also wish to acknowledge the special contribution made by Ann Hsing and Anand Chokkalingam to the conception of the genetic component of PROtEuS.

QLD: The QLD research is supported by The National Health and Medical Research Council (NHMRC) Australia Project Grants [390130, 1009458] and NHMRC Career Development Fellowship and Cancer Australia PdCCRS funding to J Batra.The QLD research is supported by The National Health and Medical Research Council (NHMRC) Australia Project Grants [390130, 1009458] and NHMRC Career Development Fellowship and Cancer Australia PdCCRS funding to J Batra.

RAPPER: RAPPER is funded by Cancer Research UK [C1094/A11728; C1094/A18504] and Experimental Cancer Medicine Centre funding [C1467/A7286]The RAPPER group thank Rebecca Elliott for project management.

SABOR: The SABOR research is supported by NIH/NCI Early Detection Research Network, grant U01 CA0866402-12. Also supported by the Cancer Center Support Grant to the Cancer Therapy and Research Center from the National Cancer Institute (US) P30 CA054174

SCCS: SCCS is funded by NIH grant R01 CA092447, and SCCS sample preparation was conducted at the Epidemiology Biospecimen Core Lab that is supported in part by the Vanderbilt-Ingram Cancer Center (P30 CA68485). Data on SCCS cancer cases used in this publication were provided by the Alabama Statewide Cancer Registry; Kentucky Cancer Registry, Lexington, KY; Tennessee Department of Health, Office of Cancer Surveillance; Florida Cancer Data System; North Carolina Central Cancer Registry, North Carolina Division of Public Health; Georgia Comprehensive Cancer Registry; Louisiana Tumor Registry; Mississippi Cancer Registry; South Carolina Central Cancer Registry; Virginia Department of Health, Virginia Cancer Registry; Arkansas Department of Health, Cancer Registry, 4815 W. Markham, Little Rock, AR 72205. The Arkansas Central Cancer Registry is fully funded by a grant from National Program of Cancer Registries, Centers for Disease Control and Prevention (CDC). Data on SCCS cancer cases from Mississippi were collected by the Mississippi Cancer Registry which participates in the National Program of Cancer Registries (NPCR) of the Centers for Disease Control and Prevention (CDC). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the CDC or the Mississippi Cancer Registry.

SCPCS: SCPCS is funded by CDC grant S1135-19/19, and SCPCS sample preparation was conducted at the Epidemiology Biospecimen Core Lab that is supported in part by the Vanderbilt-Ingram Cancer Center (P30 CA68485).

SEARCH: SEARCH is funded by a programme grant from Cancer Research UK [C490/A10124] and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge.

SNP_Prostate_Ghent: The study was supported by the National Cancer Plan, financed by the Federal Office of Health and Social Affairs, Belgium.

SPAG: Wessex Medical ResearchHope for Guernsey, MUG, HSSD, MSG, Roger Allsopp

STHM2: STHM2 was supported by grants from The Strategic Research Programme on Cancer (StratCan), Karolinska Institutet; the Linné Centre for Breast and Prostate Cancer (CRISP, number 70867901), Karolinska Institutet; The Swedish Research Council (number K2010-70X-20430-04-3) and The Swedish Cancer Society (numbers 11-0287 and 11-0624); Stiftelsen Johanna Hagstrand och Sigfrid Linnérs minne; Swedish Council for Working Life and Social Research (FAS), number 2012-0073STHM2 acknowledges the Karolinska University Laboratory, Aleris Medilab, Unilabs and the Regional Prostate Cancer Registry for performing analyses and help to retrieve data. Carin Cavalli–Björkman and Britt-Marie Hune for their enthusiastic work as research nurses. Astrid Björklund for skilful data management. We wish to thank the BBMRI.se biobank facility at Karolinska Institutet for biobank services.

SWOG-PCPT & SELECT: PCPT & SELECT are funded by Public Health Service grants U10CA37429 and 5UM1CA182883 from the National Cancer Institute. SWOG and SELECT thank the site investigators and staff and, most importantly, the participants who donated their time to this trial.

TAMPERE: The Tampere (Finland) study was supported by the Academy of Finland (251074), The Finnish Cancer Organisations, Sigrid Juselius Foundation, and the Competitive Research Funding of the Tampere University Hospital (X51003). The PSA screening samples were collected by the Finnish part of ERSPC (European Study of Screening for Prostate Cancer).TAMPERE would like to thank Riina Liikanen, Liisa Maeaettaenen and Kirsi Talala for their work on samples and databases.

UGANDA: None reported

UKGPCS: UKGPCS would also like to thank the following for funding support: The Institute of Cancer Research and The Everyman Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign UK (now Prostate Action), The Orchid Cancer Appeal, The National Cancer Research Network UK, The National Cancer Research Institute (NCRI) UK. We are grateful for support of NIHR funding to the NIHR Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust. UKGPCS should also like to acknowledge the NCRN nurses, data managers and Consultants for their work in the UKGPCS study.UKGPCS would like to thank all urologists and other persons involved in the planning, coordination, and data collection of the CAPS study.

ULM: The Ulm group received funds from the German Cancer Aid (Deutsche Krebshilfe).

WUGS/WUPCS: WUGS would like to thank the following for funding support: The Anthony DeNovi Fund, the Donald C. McGraw Foundation, and the St. Louis Men's Group Against Cancer.

References

1. Khera, A.V., Chaffin, M., Aragam, K.G., Haas, M.E., Roselli, C., Choi, S.H., Natarajan, P., Lander, E.S., Lubitz, S.A., Ellinor, P.T., et al. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 50, 1219-1224.
2. Yeager, M., Orr, N., Hayes, R.B., Jacobs, K.B., Kraft, P., Wacholder, S., Minichiello, M.J., Fearnhead, P., Yu, K., Chatterjee, N., et al. (2007). Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet 39, 645-649.
3. Kolonel, L.N., Henderson, B.E., Hankin, J.H., Nomura, A.M., Wilkens, L.R., Pike, M.C., Stram, D.O., Monroe, K.R., Earle, M.E., and Nagamine, F.S. (2000). A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol 151, 346-357.
4. Haiman, C.A., Patterson, N., Freedman, M.L., Myers, S.R., Pike, M.C., Waliszewska, A., Neubauer, J., Tandon, A., Schirmer, C., McDonald, G.J., et al. (2007). Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet 39, 638-644.
5. Siddiq, A., Couch, F.J., Chen, G.K., Lindstrom, S., Eccles, D., Millikan, R.C., Michailidou, K., Stram, D.O., Beckmann, L., Rhie, S.K., et al. (2012). A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. Hum Mol Genet 21, 5373-5384.
6. Schumacher, F.R., Berndt, S.I., Siddiq, A., Jacobs, K.B., Wang, Z., Lindstrom, S., Stevens, V.L., Chen, C., Mondul, A.M., Travis, R.C., et al. (2011). Genome-wide association study identifies new prostate cancer susceptibility loci. Hum Mol Genet 20, 3867-3875.
7. Garcia-Closas, M., Brinton, L.A., Lissowska, J., Chatterjee, N., Peplonska, B., Anderson, W.F., Szeszenia-Dabrowska, N., Bardin-Mikolajczak, A., Zatonski, W., Blair, A., et al. (2006). Established breast cancer risk factors by clinically important tumour characteristics. Br J Cancer 95, 123-129.
8. Cook, M.B., Wang, Z., Yeboah, E.D., Tettey, Y., Biritwum, R.B., Adjei, A.A., Tay, E., Truelove, A., Niwa, S., Chung, C.C., et al. (2014). A genome-wide association study of prostate cancer in West African men. Hum Genet 133, 509-521.
9. Al Olama, A.A., Kote-Jarai, Z., Berndt, S.I., Conti, D.V., Schumacher, F., Han, Y., Benlloch, S., Hazelett, D.J., Wang, Z., Saunders, E., et al. (2014). A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nat Genet 46, 1103-1109.
10. Wang, Z., Zhu, B., Zhang, M., Parikh, H., Jia, J., Chung, C.C., Sampson, J.N., Hoskins, J.W., Hutchinson, A., Burdette, L., et al. (2014). Imputation and subset-based association analysis across different cancer types identifies multiple independent risk loci in the TERT-CLPTM1L region on chromosome 5p15.33. Hum Mol Genet 23, 6616-6633.
11. Hsing, A.W., Yeboah, E., Biritwum, R., Tettey, Y., De Marzo, A.M., Adjei, A., Netto, G.J., Yu, K., Li, Y., Chokkalingam, A.P., et al. (2014). High prevalence of screen detected prostate cancer in West Africans: implications for racial disparity of prostate cancer. J Urol 192, 730-735.
12. Chokkalingam, A.P., Yeboah, E.D., Demarzo, A., Netto, G., Yu, K., Biritwum, R.B., Tettey, Y., Adjei, A., Jadallah, S., Li, Y., et al. (2012). Prevalence of BPH and lower urinary tract symptoms in West Africans. Prostate Cancer Prostatic Dis 15, 170-176.

13. Chu, L.W., Ritchey, J., Devesa, S.S., Quraishi, S.M., Zhang, H., and Hsing, A.W. (2011). Prostate cancer incidence rates in Africa. Prostate Cancer 2011, 947870.
14. Chung, C.C., Hsing, A.W., Edward, Y., Biritwum, R., Tettey, Y., Adjei, A., Cook, M.B., De Marzo, A., Netto, G., Tay, E., et al. (2014). A comprehensive resequence-analysis of 250 kb region of 8q24.21 in men of African ancestry. Prostate 74, 579-589.
15. Berndt, S.I., Wang, Z., Yeager, M., Alavanja, M.C., Albanes, D., Amundadottir, L., Andriole, G., Beane Freeman, L., Campa, D., Cancel-Tassin, G., et al. (2015). Two susceptibility loci identified for prostate cancer aggressiveness. Nat Commun 6, 6889.
16. Bailey-Wilson, J.E., Childs, E.J., Cropp, C.D., Schaid, D.J., Xu, J., Camp, N.J., Cannon-Albright, L.A., Farnham, J.M., George, A., Powell, I., et al. (2012). Analysis of Xq27-28 linkage in the international consortium for prostate cancer genetics (ICPCG) families. BMC Med Genet 13, 46.
17. Lu, L., Cancel-Tassin, G., Valeri, A., Cussenot, O., Lange, E.M., Cooney, K.A., Farnham, J.M., Camp, N.J., Cannon-Albright, L.A., Tammela, T.L., et al. (2012). Chromosomes 4 and 8 implicated in a genome wide SNP linkage scan of 762 prostate cancer families collected by the ICPCG. Prostate 72, 410-426.
18. Xu, J., Lange, E.M., Lu, L., Zheng, S.L., Wang, Z., Thibodeau, S.N., Cannon-Albright, L.A., Teerlink, C.C., Camp, N.J., Johnson, A.M., et al. (2013). HOXB13 is a susceptibility gene for prostate cancer: results from the International Consortium for Prostate Cancer Genetics (ICPCG). Hum Genet 132, 5-14.
19. Teerlink, C.C., Leongamornlert, D., Dadaev, T., Thomas, A., Farnham, J., Stephenson, R.A., Riska, S., McDonnell, S.K., Schaid, D.J., Catalona, W.J., et al. (2016). Genome-wide association of familial prostate cancer cases identifies evidence for a rare segregating haplotype at 8q24.21. Hum Genet 135, 923-938.
20. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19, 1655-1664.
21. Bergstrom, A., McCarthy, S.A., Hui, R.Y., Almarri, M.A., Ayub, Q., Danecek, P., Chen, Y., Felkel, S., Hallast, P., Kamm, J., et al. (2020). Insights into human genetic variation and population history from 929 diverse genomes. Science 367, 1339-+.