Title

African-specific improvement of a polygenic hazard score for age at diagnosis of prostate cancer

Journal

International Journal of Cancer, 148(1)

ISSN

0020-7136

Authors

Karunamuni, Roshan A

Huynh-Le, Minh-Phuong

Fan, Chun C

et al.

Publication Date

2021

DOI

10.1002/ijc.33282

Peer reviewed

Received: 13 May 2020 | Revised: 7 August 2020 | Accepted: 12 August 2020

DOI: 10.1002/ijc.33282

CANCER EPIDEMIOLOGY

IJC International Journal of Cancer

African-specific improvement of a polygenic hazard score for age at diagnosis of prostate cancer

Roshan A. Karunamuni[1†] | Minh-Phuong Huynh-Le[1] | Chun C. Fan[2] | Wesley Thompson[3] | Rosalind A. Eeles[4,5] | Zsofia Kote-Jarai[4] | Kenneth Muir[6,7] | UKGPCS Collaborators[‡] | Artitaya Lophatananon[6] | Catherine M. Tangen[8] | Phyllis J. Goodman[8] | Ian M. Thompson Jr[9] | William J. Blot[10,11] | Wei Zheng[10] | Adam S. Kibel[12] | Bettina F. Drake[13] | Olivier Cussenot[14,15] | Géraldine Cancel-Tassin[14,15] | Florence Menegaux[16] | Thérèse Truong[16] | Jong Y. Park[17] | Hui-Yi Lin[18] | Jeannette T. Bensen[19,20] | Elizabeth T. H. Fontham[18] | James L. Mohler[20,21] | Jack A. Taylor[22,23] | Luc Multigner[24] | Pascal Blanchet[25] | Laurent Brureau[25] | Marc Romana[26] | Robin J. Leach[27] | Esther M. John[28] | Jay Fowke[29,30] | William S. Bush[31] | Melinda Aldrich[32] | Dana C. Crawford[31] | Shiv Srivastava[33,34] | Jennifer C. Cullen[33,34] | Gyorgy Petrovics[33,34] | Marie-Élise Parent[35,36] | Jennifer J. Hu[37] | Maureen Sanderson[38] | Ian G. Mills[39] | Ole A. Andreassen[40] | Anders M. Dale[41] | Tyler M. Seibert[1,41] | The PRACTICAL Consortium[42]

[1]Department of Radiation Medicine and Applied Sciences, University of California San Diego, La Jolla, California

[2]Healthlytix, San Diego, California

[3]Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, California

[4]The Institute of Cancer Research, London, UK

[5]Royal Marsden NHS Foundation Trust, London, UK

[6]Division of Population Health, Health Services Research and Primary Care, University of Manchester, Oxford Road, Manchester, M13 9PL, UK

[7]Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK

[8]SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, WA

[9]CHRISTUS Santa Rosa Hospital—Medical Center, San Antonio, Texas

[10]Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee

[11]International Epidemiology Institute, Rockville, Maryland

[12]Division of Urologic Surgery, Brigham and Womens Hospital, Boston, Massachusetts

[13]Washington University School of Medicine, St. Louis, Missouri

[14]Sorbonne Universite, GRC n°5, AP-HP, Tenon Hospital, Paris, France

[15]CeRePP, Tenon Hospital, Paris, France

[16]Université Paris-Saclay, UVSQ, Inserm, CESP, Villejuif, France

Abbreviations: HR, hazard ratio; PHS, polygenic hazard score; PHS46+African, updated polygenic hazard score model for men with African genetic ancestry; PHS46, previously published polygenic hazard score model using 46 SNPs; PRACTICAL, Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium; SNP, single nucleotide polymorphism.

[†]Additional members from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium (PRACTICAL, http://www.practical.icr.ac.uk/) and associated funding are provided in Appendices A and B (Supporting Information).

[‡]The collaborators included are in mentioned in http://www.icr.ac.uk/our-research/research-divisions/division-of-genetics-and-epidemiology/oncogenetics/research-projects/ukgpcs/ukgpcs-collaborators.

[17]Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, Florida

[18]School of Public Health, Louisiana State University Health Sciences Center, New Orleans, Louisiana

[19]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

[20]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

[21]Department of Urology, Roswell Park Comprehensive Cancer Center, Buffalo, New York

[22]Epidemiology Branch, National Institute of Environmental Health Sciences, North Carolina

[23]Epigenetic and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, North Carolina

[24]Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail)—UMR_S 1085, Rennes, France

[25]CHU de Pointe-à-Pitre, Univ Antilles, Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail)—UMR_S 1085, Pointe-à-Pitre, France

[26]UMR Inserm 1134 Biologie Intégrée du Globule Rouge, INSERM/Université Paris Diderot—Université Sorbonne Paris Cité/INTS/Université des Antilles, Paris, France

[27]Department of Cell System and Anatomy and Mays Cancer Center, University of Texas Health San Antonio, San Antonio, Texas

[28]Department of Medicine, Division of Oncology, Stanford Cancer Institute, Stanford University School of Medicine, Stanford, California

[29]Department of Medicine and Urologic Surgery, Vanderbilt University Medical Center, 1211 Medical Center Drive, Nashville, Tennessee

[30]Division of Epidemiology, Department of Preventive Medicine, The University of Tennessee Health Science Center, Tennessee

[31]Case Western Reserve University, Department of Population and Quantitative Health Sciences, Cleveland Institute for Computational Biology, Cleveland, Ohio

[32]Department of Thoracic Surgery, Vanderbilt University Medical Center, Nashville, Tennessee

[33]Uniformed Services University, Bethesda, Maryland

[34]Center for Prostate Disease Research, Bethesda, Maryland

[35]Epidemiology and Biostatistics Unit, Centre Armand-Frappier Santé Biotechnologie, Institut national de la recherche scientifique, Quebec, Canada

[36]Department of Social and Preventive Medicine, School of Public Health, University of Montreal, Montreal, Quebec, Canada

[37]Sylvester Comprehensive Cancer Center, The University of Miami School of Medicine, Miami, Florida

[38]Department of Family and Community Medicine, Meharry Medical College, Nashville, Tennessee

[39]Center for Cancer Research and Cell Biology, Queen's University of Belfast, Belfast, UK

[40]NORMENT, KG Jebsen Centre, Oslo University Hospital and University of Oslo, Oslo, Norway

[41]Department of Radiology, University of California San Diego, La Jolla, California

[42]Institute of Cancer Research, Sutton, UK

Correspondence
Roshan A. Karunamuni, Department of Radiation Medicine and Applied Sciences, University of California, San Diego, 3960 Health Sciences Dr, Mail Code 0865, La Jolla, CA 92093.
Email: rakarunamuni@health.ucsd.edu

Tyler M. Seibert, University of California, San Diego, Department of Radiation Medicine and Applied Sciences, 3960 Health Sciences Dr, Mail Code 0865, La Jolla, CA 92093.
Email: tseibert@health.ucsd.edu

Funding information
KG Jebsen Stiftelsen; National Institute of Biomedical Imaging and Bioengineering, Grant/Award Number: K08EB026503; Research Council of Norway, Grant/Award Number: 223273; South East Norway Health Authority; U.S. Department of Defense, Grant/Award Number: W81XWH-13-1-0391

Abstract

Polygenic hazard score (PHS) models are associated with age at diagnosis of prostate cancer. Our model developed in Europeans (PHS46) showed reduced performance in men with African genetic ancestry. We used a cross-validated search to identify single nucleotide polymorphisms (SNPs) that might improve performance in this population. Anonymized genotypic data were obtained from the PRACTICAL consortium for 6253 men with African genetic ancestry. Ten iterations of a 10-fold cross-validation search were conducted to select SNPs that would be included in the final PHS46 +African model. The coefficients of PHS46+African were estimated in a Cox proportional hazards framework using age at diagnosis as the dependent variable and PHS46, and selected SNPs as predictors. The performance of PHS46 and PHS46 +African was compared using the same cross-validated approach. Three SNPs (rs76229939, rs74421890 and rs5013678) were selected for inclusion in PHS46 +African. All three SNPs are located on chromosome 8q24. PHS46+African showed substantial improvements in all performance metrics measured, including a 75% increase in the relative hazard of those in the upper 20% compared to the bottom 20% (2.47-4.34) and a 20% reduction in the relative hazard of those in the bottom 20% compared to the middle 40% (0.65-0.53). In conclusion, we identified three

SNPs that substantially improved the association of PHS46 with age at diagnosis of prostate cancer in men with African genetic ancestry to levels comparable to Europeans.

KEYWORDS

African, genome wide association study, genomics, genotypic ancestry, health disparities, polygenic risk, prostate cancer

1 | INTRODUCTION

Polygenic models can provide personalized estimates of the risk of developing prostate cancer. In the context of survival analysis, these models can provide insight into age at diagnosis of prostate cancer, and thus could be used to guide decisions on whether and when to offer screening.[1] Studies of polygenic models have often included only individuals of European genetic ancestry, owing to greater availability of data from that population.[2,3] As a consequence, these models have been tailored to identify and estimate coefficients of genetic common variants for that particular population, while potentially missing variants that may hold value in other populations.[2] There is concern that using these European-focused models could actually exacerbate health disparities.[2-4]

As an example, our group recently published on the performance of a polygenic hazard score (PHS) originally developed using a European dataset, in a multiethnic dataset consisting of individuals of European, African and Asian genetic ancestry.[5] The model (called here PHS46, referred to in the referenced manuscript as PHS_2), includes 46 single nucleotide polymorphisms (SNPs) in its calculation and was strongly associated with age at diagnosis in all three genetic populations ($P < 10^{-16}$). However, the hazard ratio (HR) for prostate cancer between individuals in the upper 20th percentile and those in the lower 20th percentile of PHS46 was approximately half as large for those with African genetic ancestry (2.6) as it was for those with European (5.6) or Asian (4.6) ancestry. A similar pattern was observed for clinically significant prostate cancer and for death from prostate cancer.

In the current study, we attempt to bridge the apparent gap in model performance of PHS46 for individuals with African genetic ancestry. To this end, we used a machine learning approach to systematically search for SNPs that add statistical value to a base model of PHS46 among African men (PHS46+African). By including PHS46 as a covariate in our SNP search, we sought to identify those SNPs that may hold particular value for individuals with African genetic ancestry.

What's new?

Polygenic models can provide personalized estimates of cancer risk. However, Africans are a population generally underserved in genome-wide studies. In this report, the authors used machine learning to identify three SNPs that significantly improved the performance of an established polygenic-hazard model of prostate-cancer risk vs age at diagnosis, specifically tailored to men of African genetic ancestry. Inclusion of these SNPs improved the performance of this novel 'PHS46+African' score by as much as 79% compared to the original model, to levels comparable with Europeans.

2 | MATERIAL AND METHODS

2.1 | Study dataset

We obtained genotype and phenotype data from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL)[6] consortium for this study. Genotyping was performed using the OncoArray platform[6,7] and had undergone quality assurance steps, as described previously.[8] The study dataset contains no overlap with that used to estimate model coefficients of PHS46. It is a subset of another dataset wherein the performance discrepancy of PHS46 between different genotypic ancestries was first observed.[5] All 46 SNPs of PHS46 were directly genotyped on the OncoArray platform.

The genotypic ancestry of each individual was also determined previously.[6,9] In total, the African dataset consisted of data from 6253 men with African genotypic ancestry. Missing SNP calls were replaced with the mean of the genotyped data for that SNP in the African dataset. The percentage of individuals with missing SNP calls ranged from 0% to 5.9% across the SNPs, while the percentage of SNPs with missing calls ranged from 0.5% to 5.1% across the individuals. Individuals without prostate cancer were censored at age at last follow-up in the Cox proportional hazard models. A description of the PRACTICAL study groups that contributed data toward this analysis are described in Supplementary Table 1. PHS46 risk score for each individual in the African dataset was estimated as the sum of SNP allele counts (X) multiplied by their respective coefficients (β)[5]:

$$PHS46 = \sum_{i=1}^{46} X_i \beta_i.$$

2.2 | SNP can

A multistep approach was used to select SNPs, from those directly genotyped on the OncoArray platform, that would improve the performance of PHS46 in the African dataset. Training and testing sets

were generated using 10 iterations of a 10-fold cross-validation structure resulting in 100 total permutations. For each permutation, a multivariable logistic regression model using case/control status as the dependent variable was estimated using each genotyped SNP in turn, adjusting for PHS46 and four principal components based on genetic ancestry, determined previously.[9] SNPs with *P* values less than 1×10^{-6} were considered for further analysis. In order of increasing *P* value, each SNP was tested in a multiple Cox proportional hazards model, after adjusting for PHS46, four ancestral principal components and previously selected SNPs. The Cox model in the SNP scan used age at diagnosis of prostate cancer as the dependent variable. If the *P* value of the coefficient of the tested SNP was less than 1×10^{-6}, it was considered for the final model in that permutation. SNPs that reached this *P* value threshold in more than 50% of the permutations were selected to construct the PHS46+African model, consisting of PHS46 and the newly identified SNPs.

2.3 | Comparing performance between PHS46 and PHS46+African—hazard ratio

For each permutation of the previously described cross-validation structure, an PHS46+African Cox proportional hazards model was estimated in the training set using PHS46 and the selected SNPs as independent predictors. The PHS46+African risk score for each individual is then estimated using the corresponding PHS46 score, selected SNP allele counts (*Y*) and their respective coefficients (α):

$$PHS46+African = PHS46 + \sum_{j=1}^{SNPs} Y_j \alpha_j.$$

The performance of the PHS46+African and PHS46 models was then determined in the cross-validation testing set, and the resulting HR were obtained, as previously described.[1] For each model, the PHS risk scores within the cross-validation testing set are assigned to quantile groups identified using the corresponding training set control values. The HR between two quantile groups, such as those in the top 20% to those in the bottom 20%, is estimated as the exponential of the difference in mean PHS values for each group. In this calculation, the PHS values are linearly scaled by a sample-weight correction factor to account for case-control sampling.[1,5,10] Three HR values were calculated: HR80/20 (top 20% to bottom 20%), HR98/50 (top 2% to middle 40%) and HR20/50 (bottom 20% to middle 40%). The average HR across permutations for both PHS46+African and PHS46 are reported.

To allow for comparisons with previously published results, the performance metrics for PHS46 and PHS46+African were also estimated for age at diagnosis of clinically significant prostate cancer. When estimating performance for clinically significant prostate cancer, controls and nonclinically significant cancers were censored at age of last follow-up and age of diagnosis, respectively. The previously used criteria for clinically significant cancer were any of: Gleason score ≥ 7, stage T3-T4, PSA concentration ≥ 10 ng/mL, pelvic lymph nodal metastasis or distant metastasis.[1] Paired *t*-tests were used to test for statistically significant differences (α = 0.05) in HR between PHS46 +African and PHS46.

Additionally, in each permutation, the performance of a Cox model consisting of PHS46 and SNPs that were considered in that permutation was also estimated. These results are provided within Supplementary Table 2 and performance estimates are provided that are not prone to information leakage from training to testing set.

2.4 | Comparing performance between PHS46 and PHS46+African - C-index

In addition to the HR, the performances of PHS46 and PHS46 +African were compared using Harrell's c-index.[11] For each permutation of the aforementioned cross-validation structure, the c-index of PHS46 and PHS46+African scores were estimated in the testing fold using the "coxph" function in the R "survival" package. Paired *t*-tests were used to test for statistically significant differences (α = 0.05) between the two models.

2.5 | Characterization of PHS46+African

Coefficients of the PHS46+African model, consisting of PHS46 and the SNPs selected in the SNP-scan, were estimated using 1000 bootstrapped samples of the African dataset.

2.6 | Clinical utility of PHS46+African

As an example of the clinical utility of the PHS46+African risk score, the risk-equivalent age was estimated for those individuals in the upper 2 percentile of the distribution of PHS46+African risk scores. The risk-equivalent age, as defined previously,[12] is when an individual from a given PHS percentile has prostate cancer risk equivalent to the average 60-year-old man. The age-specific general cumulative incidence curve was generated using data from SEER*Explorer incidence rates by age at diagnosis, 2003-2017 for Black Americans.[13] The corresponding risk-adjusted incidence curve was estimated by multiplying the general cumulative incidence curve by the mean value of HR98/50 for PHS46+African obtained from the analysis of the age-of-diagnosis of prostate cancer. The risk-equivalent age was then calculated as the age at which the risk-adjusted cumulative incidence curve had the same value as the general cumulative incidence curve at age 60.

3 | RESULTS

3.1 | Individual and OncoArray characteristics

In total, there were 3013 men with (cases) and 3240 men without (controls) prostate cancer in the African dataset. The mean [95% CI]

ages of cases and controls were 62.4 [62.1, 62.7] and 61.8 [61.4, 62.1] years, respectively. The OncoArray genotypic data, after the quality assurance process, included 444 323 SNPs.

3.2 | SNP scan

Across the 100 permutations of the cross-validation iterations, a total of 12 SNPs were considered for final selection (Supplementary Table 3). Three SNPs were selected in more than 50% of the permutations and included in the final PHS46+African model. By cross-referencing the chromosomal positions against dbSNP,[14] these variants were identified as rs76229939,[15] rs74421890[16] and rs5013678.[17] All three SNPs (Table 1) are located on chromosome 8q24, a region of the chromosome previously identified as containing common variants associated with prostate cancer.[18,19] An examination of R^2 (Supplementary Table 4) showed little association, ranging from 0.0027 to 0.0057, among genotype data from the three SNPs in the African dataset.

Reference threshold (Supplementary Table 5) and mean (Supplementary Table 6) values for PHS46+African in the African dataset are presented in the Supplemental Data.

3.3 | HR performance of PHS46+African

Figure 1 shows the difference in HRs between PHS46+African and PHS46 within the African dataset using age at diagnosis of any prostate cancer (Supplementary Table 7). Overall, we observed an improvement in all the metrics calculated: a 75% increase in HR98/50 from 2.10 to 3.67; a 79% increase in HR80/20 from 2.47 to 4.42 and a 23% decrease in HR20/50 from 0.65 to 0.51. We also observed improvements in all performance metrics when using age at diagnosis of clinically significant prostate cancer: 103% increase in HR98/50 from 1.91 to 3.88, 113% improvement in HR80/20 from 2.21 to 4.71, and 29% improvement in HR20/50 from 0.70 to 0.50. All observed changes in HR were statistically significant ($P < 1x10^{-16}$).

3.4 | C-index of PHS46+African

The mean c-indices of PHS46 and PHS46+African across the cross-validation folds were estimated as 0.55 and 0.58 ($P < 1x10^{-16}$), respectively.

3.5 | Risk-equivalent age for PHS46+African

The risk-equivalent age for those individuals in the top 2 percentiles of the distribution of PHS46+African scores was estimated as 50 years old, suggesting that a man with a PHS46+African score in the top 2 percentiles reached a prostate cancer detection risk equivalent to that of a standard 60-year-old roughly 10 years earlier, at an age of 50 years. The corresponding risk-equivalent age when using PHS46 scores was 54 years.

4 | DISCUSSION

Using a cross-validated search of a dataset made up entirely of men with African genetic ancestry, we were able to identify three SNPs that substantially improved the performance of PHS46 in this population to levels that are comparable to those observed in Europeans and Asians. Performance improvements were observed in HRs tracking risk between PHS groups, concordance indices tracking the overall utility of PHS as a continuous variable and risk-equivalent age tracking the potential clinical utility of PHS. The three SNPs, rs76229939, rs74421890 and rs5013678, are all located on chromosome 8q24—a region of the genome where variants have been associated with prostate cancer in both the general population and specifically in men with African genetic ancestry.[19,20] Despite the relative proximity of the three SNPs on chromosome 8, their genetic data were not strongly associated in our dataset, suggesting that each SNP provides non-redundant information for an individual's genetic score.

Each of the three SNPs have been previously identified in the literature to be associated with prostate cancer: rs76229939 is an intron variant of the prostate cancer-associated transcript 2 (*PCAT2*) gene, while rs74421890 and rs5013678 are both noncoding transcript variants of the prostate-cancer-associated noncoding RNA 1 (*PRNCR1*) gene. The minor allele frequencies of rs76229939 and rs74421890 in Europeans, as reported by dbSNP,[14] are approximately zero to three decimal places, which may explain why they were not selected in the original formulation of PHS46.

This study is not meant to be an exhaustive search for all possible SNPs that are associated with the age of diagnosis of prostate cancer in individuals with African genetic ancestry. Our study is also limited by the small number of available observations relative to those often found in many genome-wide association studies, which can have tens or hundreds of thousands of individuals. However, we were able to

TABLE 1 Characteristics of PHS46 +African SNPs

RS number	Chromosome	Position	Effect	Ref	Beta	Frequency (%)
rs76229939	8	128085394	G	A	0.441	4.8
rs74421890	8	128096183	A	G	0.415	4.1
rs5013678	8	128103979	G	A	−0.260	8.1

Note: RS-ID, chromosome and base-pair position (based on version 37), effect and reference alleles, bootstrap-estimated beta and effect allele frequencies in aggregated Africans from 1000Genomes (referenced from dbSNP) of the three SNPs selected for addition to PHS46.

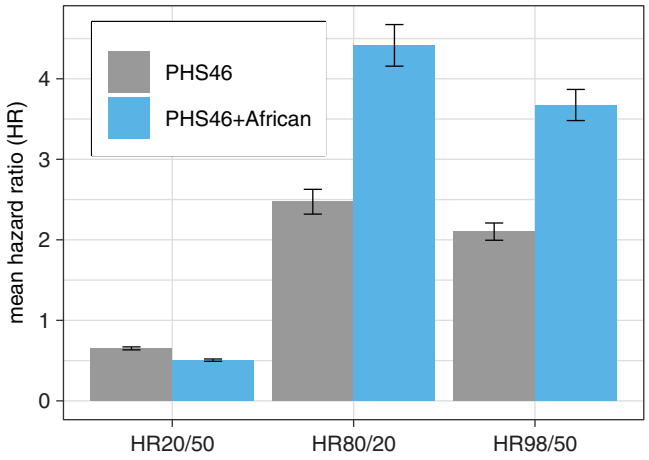

FIGURE 1 Comparison between PHS46 and PHS46+African. Mean hazard ratio metrics plotted for PHS46 and PHS46+African models in the African data set. Improvements were observed in all performance metrics investigated. Error bars represent 95% confidence interval [Color figure can be viewed at wileyonlinelibrary.com]

extract information that is likely robust by employing a cross-validated search for those SNPs that specifically add value to the performance of PHS46, and not simply independently associated with prostate cancer. Future analysis will include a more detailed analysis of the 8q24 region, including SNPs that are imputed using TOPMed reference panels. We also note that no SNP score, including PHS46 and PHS46 +African, has been shown to discriminate men at risk of aggressive prostate cancer from those at risk of indolent prostate cancer. Finally, the performance metrics reported in this study may be biased by the leakage of information across cross-validated folds of the data when identifying those SNPs to include in the final African-PHS model. This bias is expected to be similar for all SNPs and should not have influenced selection of the three SNPs included in the final model over those not selected.

In conclusion, we identified three SNPs (rs76229939, rs74421890 and rs5013678) on 8q24 that substantially improved the performance of PHS46 in a dataset of men with African genetic ancestry. The addition of these SNPs to the polygenic risk score substantially improved its association with age at diagnosis of prostate cancer in Africans, to levels comparable with those found in Europeans.

ETHICS STATEMENT

All contributing studies were approved by the relevant ethics committees; written informed consent was obtained from the study participants.[21] The present analyses used deidentified data from the PRACTICAL consortium and have been approved by the review board at the corresponding authors' institution.

CONFLICT OF INTEREST

All authors declare no personal or financial conflicts of interest for the submitted work except as follows. CCF is a scientific consultant for CorTechs Labs, Inc. RE reports honorarium as a speaker for GU-ASCO meeting in San Francisco Jan 2016, support from Janssen, and honorarium as speaker for RMH-FR meeting Nov 2017. She reports honorarium as a speaker at the University of Chicago invited talk May 2018, and an educational honorarium by Bayer & Ipsen to attend GU Connect "Treatment sequencing for mCRPC patients within the changing landscape of mHSPC" at ESMO Barcelona, September 2019. She reports member of external Expert Committee on the Prostate Dx Advisory Panel. OAA received speaker's honorarium from Lundbeck, and is a consultant for Healthlytix. AMD reports that he was a founder and holds equity in CorTechs Labs Inc., and serves on its Scientific Advisory Board. He is a member of the Scientific Advisory Board of Human Longevity, Inc., and the Mohn Medical Imaging and Visualization Centre. He received funding through research grants from GE Healthcare to UCSD. The terms of these arrangements have been reviewed by and approved by UCSD in accordance with its conflict of interest policies. TMS reports honoraria, outside of the present work, from: University of Rochester, Varian Medical Systems, Multimodal Imaging Servcies Corporation; and WebMD. He reports research funding from NIH/NBIB, U.S. Department of Defense, Radiological Society of North America, American Society for Radiation Oncology, and Varian Medical Systems.

DATA AVAILABILITY STATEMENT

The data used in this work were obtained from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) consortium. Readers who are interested in accessing the data must first submit a proposal to the Data Access Committee. If the reader is not a member of the consortium, their concept form must be sponsored by a principal investigator (PI) of one of the PRACTICAL consortium member studies. If approved by the Data Access Committee, PIs within the consortium, each of whom retains ownership of their data submitted to the consortium, can then choose to participate in the specific proposal. In addition, portions of the data are available for request from dbGaP (database of Genotypes and Phenotypes), which is maintained by the National Center for Biotechnology Information (NCBI): www.ncbi.nlm.nih.gov/gap/?term=Icogs+prostate https://www.ncbi.nlm.nih.gov/gap/?term=Icogs+prostate https://www.ncbi.nlm.nih.gov/gap/?term=Icogs+prostate. Anyone can apply to join the consortium. The eligibility requirements are listed here: http://www.practical.icr.ac.uk/blog/?page_id=9. Joining the consortium would not guarantee access, as a proposal for access would still be submitted to the Data Access Committee, but there would be no need for a separate member sponsor. Readers may find information about application by using the following contact information: Rosalind Eeles, Principal Investigator for PRACTICAL, Professor of Oncogenetics, Institute of Cancer Research (ICR), Sutton, UK, Email: practical@icr.ac.uk URL: http://www.practical.icr.ac.uk.

ORCID

Roshan A. Karunamuni https://orcid.org/0000-0001-8723-8123
Minh-Phuong Huynh-Le https://orcid.org/0000-0002-5025-4709

Wei Zheng https://orcid.org/0000-0003-1226-070X
Florence Menegaux https://orcid.org/0000-0003-3363-6411
Luc Multigner https://orcid.org/0000-0003-3205-8568
Esther M. John https://orcid.org/0000-0003-3259-8003
Gyorgy Petrovics https://orcid.org/0000-0003-3732-284X
Marie-Élise Parent https://orcid.org/0000-0002-4196-3773
Tyler M. Seibert https://orcid.org/0000-0002-4089-7399

REFERENCES

1. 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;360:1-7.
2. Duncan L, Shen H, Gelaye B, et al. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat Commun*. 2019;10(3328):1-9.https://doi.org/10.1038/s41467-019-11112-0.
3. Popejoy A, Fullerton S. Genomics is falling on diversity. *Nature*. 2016; 538:161-164.
4. 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;51:584-591. https://doi.org/10.1038/s41588-019-0379-x.
5. Huynh-Le M-P, Chieh Fan C, Karunamuni R, et al. Polygenic hazard score is associated with prostate cancer in multi-ethnic populations. *medRxiv*. 2019. https://doi.org/10.1101/19012237.
6. Amos CI, Dennis J, Wang Z, et al. The oncoarray consortium: a network for understanding the genetic architecture of common cancers. *Cancer Epidemiol Biomarkers Prev*. 2017;26:126-135.
7. Infinium OncoArray-500K BeadChip Research cancer predisposition and risk https://www.illumina.com/products/by-type/microarray-kits/infinium-oncoarray-500k.html. Accessed July 21, 2020.
8. Leitzmann MF, Rohrmann S. Risk factors for the onset of prostatic cancer: age, location, and behavioral correlates. *Clin Epidemiol*. 2012; 4:1-11.
9. Li Y, Byun J, Cai G, et al. FastPop: a rapid principal component derived method to infer intercontinental ancestry using genetic data. *BMC Bioinformat*. 2016;17:1-8. https://doi.org/10.1186/s12859-016-0965-1.
10. Therneau TM, Li H. Computing the Cox model for case cohort designs. *Lifetime Data Anal*. 1999;5:99-112.
11. Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *J Am Med Assoc*. 1982;247:2543-2546.
12. Huynh-Le M-P, Fan CC, Karunamuni R, et al. A genetic risk score to personalize prostate cancer screening, applied to population data. *Cancer Epidemiol Biomarkers Prev*. 2020;29(9):1731-1738.https://doi.org/10.1158/1055-9965.EPI-19-1527.
13. SEER*Explorer: An interactive website for SEER cancer statistics Surveillance Research Program, National Cancer Institute [Internet]. 2020https://seer.cancer.gov/explorer/. Accessed July 21, 2020.
14. Sherry ST, Ward M, Sirotkin K. dbSNP—database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res*. 1999;9:677-679.
15. rs76229939 RefSNP Report—dbSNP—NCBI [Internet]. https://www.ncbi.nlm.nih.gov/snp/rs76229939. Accessed July 14, 2020.
16. rs74421890 RefSNP Report—dbSNP—NCBI [Internet]. https://www.ncbi.nlm.nih.gov/snp/rs74421890. Accessed Jul 14, 2020.
17. rs5013678 RefSNP Report—dbSNP—NCBI [Internet]. https://www.ncbi.nlm.nih.gov/snp/rs5013678
18. Yeager M, Orr N, Hayes RB, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet*. 2007;39:645-649.
19. Haiman CA, Patterson N, Freedman ML, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. *Nat Genet*. 2007; 39:638-644.
20. Han Y, Rand KA, Hazelett DJ, et al. Prostate cancer susceptibility in men of African ancestry at 8q24. *J Natl Cancer Inst*. 2016;108:1-5.
21. Kote-Jarai Z, Easton DF, Stanford JL, et al. Multiple novel prostate cancer predisposition loci confirmed by an international study: the PRACTICAL consortium. *Cancer Epidemiol Biomarkers Prev*. 2008;17: 2052-2061.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Karunamuni RA, Huynh-Le M-P, Fan CC, et al. African-specific improvement of a polygenic hazard score for age at diagnosis of prostate cancer. *Int. J. Cancer*. 2021;148:99–105. https://doi.org/10.1002/ijc.33282