Additional SNPs improve risk stratification of a polygenic hazard score for prostate cancer

UKGPCS collaborators (2021). Additional SNPs improve risk stratification of a polygenic hazard score for prostate cancer. *Prostate Cancer and Prostatic Diseases*. Advance online publication. https://doi.org/10.1038/s41391-020-00311-2

Published in:
Prostate Cancer and Prostatic Diseases

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright 2021 Springer. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access
This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: http://go.qub.ac.uk/oa-feedback

Download date:18. Apr. 2024

Additional SNPs improve risk stratification of a polygenic hazard score for prostate cancer

Roshan A. Karunamuni[1], Minh-Phuong Huynh-Le[2], Chun C. Fan[3], Wesley Thompson[4], Rosalind A. Eeles[5], Zsofia Kote-Jarai[5], Kenneth Muir[6], Artitaya Lophatananon[6], UKGPCS collaborators[7], Johanna Schleutker[8], Nora Pashayan[9], Jyotsna Batra[10], APCB BioResource (Australian Prostate Cancer BioResource)[11], Henrik Grönberg[12], Eleanor I. Walsh[13], Emma L. Turner[13], Athene Lane[13,14], Richard M. Martin[13,14,15], David E. Neal[16], Jenny L. Donovan[17], Freddie C. Hamdy[16,18], Børge G. Nordestgaard[19], Catherine M. Tangen[20], Robert J. MacInnis[21,22], Alicja Wolk[23], Demetrius Albanes[24], Christopher A. Haiman[25], Ruth C. Travis[26], Janet L. Stanford[27], Lorelei A. Mucci[28], Catharine M. L. West[29], Sune F. Nielsen[30], Adam S. Kibel[31], Fredrik Wiklund[12], Olivier Cussenot[32], Sonja I. Berndt[24], Stella Koutros[24], Karina Dalsgaard Sørensen[33], Cezary Cybulski[34], Eli Marie Grindedal[35], Jong Y. Park[36], Sue A. Ingles[37], Christiane Maier[38], Robert J. Hamilton[39], Barry S. Rosenstein[40], Ana Vega[41], The IMPACT Study Steering Committee and Collaborators[42], Manolis Kogevinas[43], Kathryn L. Penney[44], Manuel R. Teixeira[45], Hermann Brenner[46], Esther M. John[47], Radka Kaneva[48], Christopher J. Logothetis[49], Susan L. Neuhausen[50], Azad Razack[51], Lisa F. Newcomb[27], Canary PASS Investigators[27], Marija Gamulin[52], Nawaid Usmani[53], Frank Claessens[54], Manuela Gago-Dominguez[55], Paul A. Townsend[56], Monique J. Roobol[57], Wei Zheng[58], The Profile Study Steering Committee[59], Ian G. Mills[60], Ole A. Andreassen[61], Anders M. Dale[62], Tyler M. Seibert[1,62,63], The PRACTICAL Consortium[64]

[1]Department of Radiation Medicine and Applied Sciences, University of California San Diego, La Jolla, CA, USA
[2]Radiation Oncology, George Washington University, Washington, DC
[3]Center for Human Development, University of California San Diego, La Jolla, CA, USA
[4]Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, CA, USA
[5]The Institute of Cancer Research, London, SM2 5NG, UK
[6]Division of Population Health, Health Services Research and Primary Care, University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[7]http://www.icr.ac.uk/our-research/research-divisions/division-of-genetics-and-epidemiology/oncogenetics/research-projects/ukgpcs/ukgpcs-collaborators
[8]Institute of Biomedicine, University of Turku, Finland
[9]Department of Applied Health Research, University College London, London, WC1E 7HB, UK
[10]Australian Prostate Cancer Research Centre-Qld, Institute of Health and Biomedical Innovation and School of Biomedical Sciences, Queensland University of Technology, Brisbane

QLD 4059, Australia
[11]Australian Prostate Cancer Research Centre-Qld, Queensland University of Technology, Brisbane; Prostate Cancer Research Program, Monash University, Melbourne; Dame Roma Mitchell Cancer Centre, University of Adelaide, Adelaide; Chris O'Brien Lifehouse and
[12]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, SE-171 77 Stockholm, Sweden
[13]Bristol Medical School, Department of Population Health Sciences, University of Bristol, Bristol, United Kingdom
[14]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom
[15]National Institute for Health Research (NIHR) Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and the University of Bristol, Bristol, United Kingdom
[16]Nuffield Department of Surgical Sciences, University of Oxford, Room 6603, Level 6, John Radcliffe Hospital, Headley Way, Headington, Oxford, OX3 9DU, UK
[17]School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom
[18]Faculty of Medical Science, University of Oxford, John Radcliffe Hospital, Oxford, United Kingdom
[19]Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark
[20]SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[21]Cancer Epidemiology Division, Cancer Council Victoria, 615 St Kilda Road, Melbourne, VIC 3004, Australia
[22]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Grattan Street, Parkville, VIC 3010, Australia
[23]Unit of Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, SE-171 77 Stockholm, Sweden
[24]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, Maryland, 20892, USA
[25]Center for Genetic Epidemiology, Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, CA 90015, USA
[26]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, OX3 7LF, UK
[27]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, 98109-1024, USA
[28]Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA 02115, USA
[29]Division of Cancer Sciences, University of Manchester, Manchester Academic Health Science Centre, Radiotherapy Related Research, The Christie Hospital NHS Foundation Trust, Manchester, M13 9PL UK
[30]Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, 2200 Copenhagen, Denmark
[31]Division of Urologic Surgery, Brigham and Womens Hospital, 75 Francis Street, Boston, MA 02115, USA
[32]Sorbonne Universite, GRC n°5 , AP-HP, Tenon Hospital, 4 rue de la Chine, F-75020 Paris,

France
[33]Department of Molecular Medicine, Aarhus University Hospital, Palle Juul-Jensen Boulevard 99, 8200 Aarhus N, Denmark
[34]International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, 70-115 Szczecin, Poland
[35]Department of Medical Genetics, Oslo University Hospital, 0424 Oslo, Norway
[36]Department of Cancer Epidemiology, Moffitt Cancer Center, 12902 Magnolia Drive, Tampa, FL 33612, USA
[37]Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, CA 90015, USA
[38]Humangenetik Tuebingen, Paul-Ehrlich-Str 23, D-72076 Tuebingen, Germany
[39]Dept. of Surgical Oncology, Princess Margaret Cancer Centre, Toronto ON M5G 2M9, Canada
[40]Department of Radiation Oncology and Department of Genetics and Genomic Sciences, Box 1236, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, NY 10029, USA
[41]Fundación Pública Galega Medicina Xenómica, Santiago de Compostela, 15706, Spain.
[42]http://impact.icr.ac.uk
[43]ISGlobal, Barcelona, Spain
[44]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital/Harvard Medical School, Boston, MA 02184, USA
[45]Department of Genetics, Portuguese Oncology Institute of Porto (IPO-Porto), 4200-072 Porto, Portugal
[46]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), D-69120, Heidelberg, Germany
[47]Departments of Epidemiology & Population Health and of Medicine, Division of Oncology, Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA 94304 USA
[48]Molecular Medicine Center, Department of Medical Chemistry and Biochemistry, Medical University of Sofia, Sofia, 2 Zdrave Str., 1431 Sofia, Bulgaria
[49]The University of Texas M. D. Anderson Cancer Center, Department of Genitourinary Medical Oncology, 1515 Holcombe Blvd., Houston, TX 77030, USA
[50]Department of Population Sciences, Beckman Research Institute of the City of Hope, 1500 East Duarte Road, Duarte, CA 91010, 626-256-HOPE (4673)
[51]Department of Surgery, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
[52]Division of Medical Oncology, Urogenital Unit, Department of Oncology, University Hospital Centre Zagreb, University of Zagreb, School of Medicine, 10 000 Zagreb, Croatia
[53]Department of Oncology, Cross Cancer Institute, University of Alberta, 11560 University Avenue, Edmonton, Alberta, Canada T6G 1Z2
[54]Molecular Endocrinology Laboratory, Department of Cellular and Molecular Medicine, KU Leuven, BE-3000, Belgium
[55]Genomic Medicine Group, Galician Foundation of Genomic Medicine, Instituto de Investigacion Sanitaria de Santiago de Compostela (IDIS), Complejo Hospitalario Universitario de Santiago, Servicio Galego de Saúde, SERGAS, 15706, Santiago de Compostela, Spai

[56]Division of Cancer Sciences, Manchester Cancer Research Centre, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, NIHR Manchester Biomedical Research Centre, Health Innovation Manchester, Univeristy of Manchester, M13 9WL
[57]Department of Urology, Erasmus University Medical Center, 3015 CE Rotterdam, The Netherlands
[58]Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 800, Nashville, TN 37232 USA.
[59]http://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/a-study-find-out-looking-gene-changes-would-be-useful-in-screening-for-prostate-cancer-profile-pilot
[60]Center for Cancer Research and Cell Biology, Queen's University of Belfast, Belfast, UK
[61]NORMENT, KG Jebsen Centre, Oslo University Hospital and University of Oslo, Oslo, Norway
[62]Department of Radiology, University of California San Diego, La Jolla, CA, USA
[63]Department of Bioengineering, University of California San Diego, La Jolla, CA, USA
[64]Institute of Cancer Research, Sutton, SW7 3RPm UK

* Additional members from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium (PRACTICAL,http://practical.icr.ac.uk/) and associated funding are provided in the Supplementary Data Appendix 1 and 2.

Corresponding Authors:

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

Tyler M. Seibert, MD, PhD
University of California, San Diego
Department of Radiation Medicine and Applied Sciences
9500 Gilman Dr, Mail Code 0861
La Jolla, CA 92093
tseibert@ucsd.edu

Running Title

Adding SNPs improves prostate cancer polygenic score

Funding

This study was funded in part by a grant from the United States National Institute of Health/National Institute of Biomedical Imaging and Bioengineering (#K08EB026503), the University of California Cancer Research Coordinating Committee (#C21CR2060), the Research Council of Norway (#223273), KG Jebsen Stiftelsen, and South East Norway Health Authority.

Abstract

Background: Polygenic hazard scores (PHS) can identify individuals with increased risk of prostate cancer. We estimated the benefit of additional SNPs on performance of a previously validated PHS (PHS46).

Materials and Method: 180 SNPs, shown to be previously associated with prostate cancer, were used to develop a PHS model in men with European ancestry. A machine-learning approach, LASSO-regularized Cox regression, was used to select SNPs and to estimate their coefficients in the training set (75,596 men). Performance of the resulting model was evaluated in the testing/validation set (6,411 men) with two metrics: (1) hazard ratios (HRs) and (2) positive predictive value (PPV) of prostate-specific antigen (PSA) testing. HRs were estimated between individuals with PHS in the top 5% to those in the middle 40% (HR95/50), top 20% to bottom 20% (HR80/20), and bottom 20% to middle 40% (HR20/50). PPV was calculated for the top 20% (PPV80) and top 5% (PPV95) of PHS as the fraction of individuals with elevated PSA that were diagnosed with clinically significant prostate cancer on biopsy.

Results: 166 SNPs had non-zero coefficients in the Cox model (PHS166). All HR metrics showed significant improvements for PHS166 compared to PHS46: HR95/50 increased from 3.72 to 5.09, HR80/20 increased from 6.12 to 9.45, and HR20/50 decreased from 0.41 to 0.34. By contrast, no significant differences were observed in PPV of PSA testing for clinically significant prostate cancer.

Conclusion: Incorporating 120 additional SNPs (PHS166 vs PHS46) significantly improved HRs for prostate cancer, while PPV of PSA testing remained the same.

Introduction

Optimal prostate cancer screening strategies seek to strike a balance between identifying clinically significant and potentially lethal cases that require treatment, while minimizing overdiagnosis of indolent, lower-risk cases that do not need radical treatment[1–3]. Genetic risk models have emerged as potentially useful tools that identify individuals with greater risk for being diagnosied with prostate cancer[4,5], and so help inform if and when to initiate screening for an individual. A subset of these models called polygenic hazard scores (PHS) seeks to directly identify associations between common genetic variants and the age of diagnosis of prostate cancer by utilizing the framework of time-to-event analyses[1,6].

We have previously reported on a PHS model for prostate cancer, PHS46, that demonstrated excellent performance in an independent test set of men from varied genetic ancestries[6]. The model incorporates genetic data of 46 unique single nucleotide polymorphisms (SNPs), and was identified through a systematic search of European men genotyped on the iCOGS chipset (Illumina, San Diego, CA). With an ever-increasing list of loci associated with prostate cancer in the literature[7–9], we sought to determine what effect, if any, the incorporation of additional SNPs would have on the performance of PHS46.

To this end, we employed a machine-learning approach, LASSO-regularized Cox regression,[10,11] to select SNPs from a list that included the 46 used in PHS46, as well as over 100 SNPs identified in previous analyses as having genome-wide significance for association with prostate cancer[7]. LASSO-regularized regression is an established variable selection technique in datasets with a large number of predictors and has been previously implemented as a SNP selection tool for a breast cancer polygenic risk score[12]. Performance metrics describing statistical model goodness-of-fit and clinically actionable screening utility of the LASSO-regularized PHS model for prostate cancer were compared with those achieved with PHS46 to determine the potential benefit of incorporating additional SNPs in polygenic hazard models.

Material and Methods

Study dataset

We obtained genotype and phenotype data from the PRACTICAL[13] consortium for this analysis. Genotyping was performed previously on either OncoArray[13] or iCOGS[9] chips, and these data were previously imputed using the 1000 Genomes reference panel[14]. Missing SNP calls were replaced with the mean of the genotyped data for that SNP in the training set[1,15]. In total, data from 82,007 men with European genetic ancestry (Supplementary Table 1)[13,16] were available for this analysis. A testing set consisting of 6,411 men (4,828 controls and 1,583 cases) enrolled in the ProtecT clinical trial was set aside for estimating the performance of the final PHS models. The data from ProtecT were chosen as the testing set because they are well characterized and were previously used for validation of PHS46[1], allowing us to directly benchmark the performance of the updated model against previous iterations. The ProtecT trial also included biopsies of participants with elevated prostate-specific antigen (PSA) level, which permits analysis of the positive predictive value of the current clinical standard for screening, PSA testing. The remaining 75,596 individuals (25,127 controls and 50,469 cases) were used for training of the model. This first analysis was limited to men of European descent because of much greater data availability in that population, but our previous work has shown that development in Europeans can inform careful future work to assess and improve performance in other ancestries[17].

Model development using LASSO regularization

A list of published SNPs previously identified[1,7] to be associated with prostate cancer was compiled. In total, 180 unique SNPs were considered for estimation within the PHS model framework. An initial screening was conducted to identify pairs of SNPs that were highly correlated ($R^2 > 0.95$). For each pair of highly correlated SNPs, a univariable Cox proportional hazards model using age of diagnosis of prostate cancer as the time to event was calculated for

each SNP in the pair, and the one with the larger p-value was discarded. The remaining SNPs were included as candidates for the new PHS model. The R (v.4.0.1) package 'glmnet' was used to estimate a LASSO-regularized Cox-proportional hazards model[10,11] using age of diagnosis of prostate cancer as the time to event. The genetic data of candidate SNPs and first four European ancestry principal components were included as predictors. Controls were censored at age of last follow-up. The hyper-parameter of the LASSO-regularized model, lambda, was selected using 10-fold cross-validation[10,11]. The final form of the LASSO model was estimated at the value of lambda that minimized the mean cross-validated error.

Characterization of LASSO-regularized PHS model

The PHS score for each of the individuals in the training and testing set was estimated as the weighted sum of the genetic counts of each of the SNPs in the PHS model, using the LASSO model coefficients as weights. Distributions of the new PHS score were compared qualitatively between training and testing groups to confirm that the model was appropriately calibrated for use in the testing set.

We also sought to assess how the LASSO-regularized PHS score compared to family history in explaining the variation in age at diagnosis of prostate cancer. A multivariable Cox proportional hazards model was estimated using the age at diagnosis of any prostate cancer as the time to event, and the PHS score and family history as predictors in both training and testing sets, separately. The family history variable was coded as a binary variable: "None" or "One or more affected first-degree relatives". Observations with missing family history values were removed from the analysis. The explained relative risk[18] (ERR) of each of the covariables as well as the full model were estimated using the "clinfun" software package in R, and provided a quantifiable measure for the importance of each variable in the model. Empirical confidence intervals for ERR were estimated using 1000 bootstrapped iterations.

Performance comparison between PHS46 and LASSO-regularized PHS

Performance in the testing set was assessed using hazard ratios (HRs) and positive predictive value (PPV), as described below. In each case, performance metrics were generated for the newly developed LASSO PHS model and for PHS46. Model coefficients for PHS46 were obtained from the literature[17]. For each performance metric, one thousand bootstrap samples of the testing set were used to generate empirical 95% confidence intervals for LASSO PHS and for PHS46. In addition, bootstrapped 95% confidence intervals were generated for the percentage change of each performance metric between the two models, using PHS46 as the reference. Percent changes were deemed statistically significant if the bootstrapped 95% confidence interval did not include 0.

HR performance

Calibration Cox proportional hazards models were fit to the bootstrapped testing data using the PHS score as the sole predictor and the age-of-diagnosis of prostate cancer as the dependent variable. The model coefficient of this Cox regression model is referred to as the calibration factor. Next, the hazard ratio between two PHS groups, such as those in the top 5% to the middle 40% (HR95/50), is estimated as the exponential of the product of the calibration factor and the difference in mean PHS scores of each group. Hazard ratios between the top 20% to the bottom 20% (HR80/20) and the bottom 20% to the middle 40% (HR20/50) were similarly calculated. The PHS cutoffs used to define these groups were determined from the distribution of PHS in the training set controls under 70 years of age[1,15].

A similar strategy was used to estimate the HR performance for clinically significant prostate cancer. The criteria for clinical significance were any of: Gleason score >=7, stage T3-T4, PSA concentration >= 10ng/mL, pelvic lymph nodal metastasis, or distant metastasis[19]. In this analysis, controls and low-risk (i.e., not clinically significant) cancers were censored at age of last follow-up and age of diagnosis, respectively. HRs are reported after sample-weight

correction[1,17,20] using the total number of cases and controls in the ProtecT trial to generate weighting factors.

Sample-weight corrected HR values were also generated using the age at diagnosis of non-clinically significant prostate cancer. Individuals with clinically significant prostate cancer were removed from this secondary analysis.

PPV performance

One indicator of clinical utility of a risk-stratification approach like PHS is whether it can be used to improve the PPV of the standard clinical screening test, prostate-specific antigen (PSA). As a population-based screening study, ProtecT provides biopsy results of both cases and controls with a positive PSA result (i.e., ≥3 ng/mL). PPV performance of each model was estimated by randomly sampling individuals within the testing set with positive PSA results, while maintaining the case to control ratio of the ProtecT study (1:2). PPV is calculated as the fraction of positive PSA individuals in the top 20% (PPV80) or top 5% (PPV95) of PHS scores that had clinically significant prostate cancer.

Cumulative incidence curves for LASSO-PHS in United Kingdom

To illustrate the utility of the LASSO PHS model in informing prostate cancer screening, cumulative incidence curves for various PHS risk groups were estimated, as described previously[21]. The age-specific general cumulative incidence curve for prostate cancer was estimated for the United Kingdom population, aged 40 to 70, using data from Cancer Research UK 2015-2017[22]. The proportion of clinically significant and non-clinically significant prostate cancer at each age was estimated using data from the Cluster Randomized Trial of PSA Testing for Prostate Cancer (CAP) trial[23]. Disease-specific cumulative incidence curves for clinically significant and non-clinically-significant prostate cancer were estimated by multiplying the general cumulative incidence curve by their respective proportions. The risk-adjusted incidence

curves for individuals in the upper 5th percentile and upper 20th percentile were estimated by multiplying the disease-specific cumulative incidence curves by the mean value of HR95/50 and HR80/50 in the testing set, respectively. Hazard ratios were obtained using the age of diagnosis of clinically significant prostate cancer as the time-to-event and after sample-weight correction.

Results

SNP screening and PHS model training

Of the 180 SNPs originally considered for this study, 6 SNPs were discarded in the initial screening process of removing highly correlated SNPs. Of the 174 remaining candidate SNPs (Supplementary Table 2), 166 had non-zero LASSO model coefficients and were selected for the final PHS model (PHS166).

The majority of the 166 variants (ninety-seven, 53%) used in PHS166 were classified as intron variants (Supplementary Table 3). Of the genes associated with variants from PHS166, HNF1B on chromosome 17 was associated with the greatest number of variants (4). Additional genes that were associated with multiple variants included ITGA6(x2), LINC00506(x2), PDLIM5(x2), TERT(x2), CTD-2194D22.4(x2), RGS17(x2), LOC105375751(x2), and CASC8(x3). Two of the SNPs used in PHS166 (rs721048 and rs10993994) were designated as 'pathogenic' by ClinVar[24] and associated with hereditary prostate cancer.

PHS166 model characterization

Distributions of PHS166 score were visually consistent between training and testing sets (Supplementary Figure 1). The 20th, 30th, 70th, 80th, and 98th percentiles of the reference PHS risk scores (controls in training set) were estimated as -0.411, -0.307, 0.048, 0.154, and 0.557, respectively.

PHS166 contributed roughly 80 to 90 percent of the total explained relative risk (Supplementary Table 4) of a Cox proportional hazards model containing both family history and PHS166. Family history was not found to be statistically significantly associated with age at diagnosis of prostate cancer in the testing set[1].

Performance comparison – PHS46 vs. PHS166

All PHS166 HR-based performance metrics showed statistically significant improvements compared to PHS46 (Table 1), for both any and clinically significant prostate cancer. The mean HR95/50 and HR80/20 values for PHS166 were roughly 36 to 55% greater than those for PHS46. For example, HR80/20 for clinically significant prostate cancer increased from 6.12 to 9.45. Similarly, HR20/50 for PHS166 was, on average, 18% lower than that for PHS46. Similar trends were observed for non-clinically significant prostate cancer (Supplementary Table 5). No significant differences between models were observed in either of the PPV-based performance metrics (Table 2). Among individuals in the top 20% of risk scores with a positive PSA test, the estimated mean PPV for clinically significant prostate cancer was roughly 0.19 irrespective of the model used – indicating approximately 19% of positive PSA tests in this risk group yielded a diagnosis of clinically significant prostate cancer. By comparison, approximately 13% of all positive PSA tests resulted in a diagnosis of clinically significant prostate cancer.

Cumulative incidence curves for PHS166 in United Kingdom

Cumulative incidence curves for clinically significant and non-clinically significant prostate cancer for the upper 5th percentile (>95th percentile) and upper 20th percentile (>80th percentile) of PHS166 scores in the United Kingdom demonstrated expected stratification of prostate cancer risk (Figure 1).

Discussion

Using a machine-learning, LASSO-regularized Cox framework, we identified 166 SNPs to be included in a polygenic hazard model (PHS166) for association with age of diagnosis of prostate cancer in men of European genetic ancestry. Variants used in PHS166 were associated with several genes, including those encoding for hepatocyte nuclear factor-1 beta (HNF1B), cancer susceptibility 8 (CASC8), and telomerase (TERT). PHS166 also explained a much larger percentage of the total explained relative risk compared to family history, suggesting that the former is important for stratifying patients' risk. When compared to the original PHS, consisting of 46 SNPs, PHS166 demonstrated substantially improved HR performance. For example, the HR for clinically significant prostate cancer comparing the upper and lower quintiles of genetic risk increased by 56% when using PHS166. No significant improvements were found in the PPV of PSA testing when using PHS to stratify risk.

Increased separation in hazard rates between PHS risk groups may allow for more nuance in clinical decision making in certain scenarios. Accurate identification of low, intermediate, and high PHS risk groups in prostate cancer may help in decisions of when (or if) to initiate screening as well as possibly improving the interpretation of the disease screens[25]. Targeting screening to men in the upper percentiles of polygenic risk as opposed to those in the lowest risk group may reduce the proportion of overdiagnosed indolent cancers from 43% to 19%[26,27]. Risk stratification achieved here by PHS166 is similar or better than commonly used clinical tools for diseases such as breast cancer, diabetes, and cardiovascular disease[25,28–30]. Clinically meaningful risk stratification is illustrated by the estimated cumulative incidence curves in Figure 1. This effect is particularly pronounced for clinically significant disease because of the increased proportion of clinically significant cases observed at older ages[2,21,23].

The lack of improvement in PPV in this study may suggest a "performance plateau" when using PHS to define broad risk categories for certain clinical applications. A similar effect has been previously described for prostate cancer polygenic models, in the context of using risk

scores to discriminate prostate biopsy outcomes[31]. Some of the precision in a score may also be diluted in broad clinical applications. The PPV analysis here is applied to participants in the ProtecT trial, which enrolled men aged 50 to 69 years, and screening in the trial was offered irrespective of underlying genetic risk[2]. Further investigation is needed to learn whether timing screening according to genetic risk might better leverage the superior HR performance of PHS166 risk score to improve the PPV of PSA testing.

LASSO frameworks have been used to identify SNPs for polygenic risk scores of several phenotypes, including fracture risk[32], type 2 diabetes[33], and breast cancer[12]. In this work, we have extended the application of LASSO to select SNPs in a polygenic hazard model of prostate cancer from a list of candidates previously identified through logistic and time-to-event analysis. Simulation studies[11] have suggested that LASSO provides more robust estimates than stepwise selection in cases with both a few large effects, as well as many small effects. As new prostate cancer associated variants are discovered, this framework can be easily implemented to develop updated polygenic hazard models.

One limitation of PHS166 is that it was entirely developed and tested in European men. However, a well-vetted, well-tested PHS model for men of European genetic ancestry can be used as a starting block for developing models for other genetic ancestries, where large-scale databases are often more scarce, as has been shown for PHS46[17,34]. Furthermore, some of the SNPs selected for incorporation into PHS166 were originally discovered in analyses that included men from the ProtecT testing set. Therefore, the improvements in HRs observed for PHS166 may be somewhat overestimated. However, this bias is likely small, given that the testing set was only a small fraction (less than 5%) of the data used in prior discovery analyses, and the ProtecT data were not used to calculate SNP weights in PHS166. The LASSO-regularized Cox framework was also used to minimize any potential for over-fitting[35] by introducing penalties for large effect sizes. In addition, this study uses age of diagnosis as the time-to-event variable, and any preceding period of undiagnosed disease is unknown.

Hypothetical perfect measurement of age of onset would likely further improve performance of the PHS model.

In conclusion, we applied a machine-learning, LASSO-regularized Cox regression framework to develop a larger PHS that includes 166 previously discovered SNPs. When comparing the performance of PHS166 to the original model, PHS46, we found that incorporating 120 more SNPs significantly improved HRs for clinically significant prostate cancer. However, incorporating more SNPs did not improve on the ability of PHS46 to inform the PPV of PSA testing in the ProtecT dataset, perhaps illustrating a plateau effect and/or dilution of risk stratification in a broad clinical application.

Ethics Statement

All contributing studies were approved by the relevant ethics committees and performed in accordance with the Declaration of Helsinki; written informed consent was obtained from the study participants. The present analyses used de-identified 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, Sep 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): https://www.ncbi.nlm.nih.gov/gap/?term=Icogs+prostatehttps://www.ncbi.nlm.nih.gov/gap/?term=Icogs+prostate.

Anyone can apply to join the consortium. The eligibility requirements are listed here: http://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 contact information below:

Rosalind Eeles
Principal Investigator for PRACTICAL
Professor of Oncogenetics
Institute of Cancer Research (ICR)
Sutton, UK
Email: PRACTICAL@icr.ac.uk
URL: http://practical.icr.ac.uk
Tel: ++44 (0)20 8722 4094

References

1 Seibert TM, Fan CC, Wang Y, Zuber V, Karunamuni R, Parsons JK *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 Huynh-Le MP, Myklebust TÅ, Feng CH, Karunamuni R, Johannesen TB, Dale AM *et al.* Age dependence of modern clinical risk groups for localized prostate cancer—A population-based study. *Cancer* 2020; **126**: 1691–1699.

3 Pashayan N, Duffy SW, Chowdhury S, Dent T, Burton H, Neal DE *et al.* Polygenic susceptibility to prostate and breast cancer: Implications for personalised screening. *Br J Cancer* 2011; **104**: 1656–1663.

4 Witte JS. Personalized prostate cancer screening: Improving PSA tests with genomic information. *Sci Transl Med* 2010; **2**: 1–5.

5 Chen H, Liu X, Brendler CB, Ankerst DP, Leach RJ, Goodman PJ *et al.* Adding genetic risk score to family history identifies twice as many high-risk men for prostate cancer: Results from the prostate cancer prevention trial. *Prostate* 2016; **76**: 1120–1129.

6 Huynh-Le M-P, Fan CC, Karunamuni R, Thompson WK, Martinez ME, Eeles RA *et al.* Polygenic hazard score is associated with prostate cancer in multi-ethnic populations. *medRxiv* 2020; : 1–34.

7 Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ *et al.* Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet* 2018; **50**: 928–936.

8 Kote-Jarai Z, Easton DF, Stanford JL, Ostrander EA, Schleutker J, Ingles SA *et al.* Multiple novel prostate cancer predisposition loci confirmed by an international study: The PRACTICAL consortium. *Cancer Epidemiol Biomarkers Prev* 2008; **17**: 2052–2061.

9 Eeles RA, Olama AA Al, Benlloch S, Saunders EJ, Leongamornlert DA, Tymrakiewicz M *et al.* Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom

genotyping array. *Nat Genet* 2013; **45**: 385–391.

10 Tibshiranit BR. Regression Shrinkage and Selection via the Lasso. *J R Stat Soc B* 1996; : 267–288.

11 Tibshirani R. The lasso method for variable selection in the cox model. *Stat Med* 1997; **16**: 385–395.

12 Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A *et al.* Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet* 2019; **104**: 21–34.

13 Amos CI, Dennis J, Wang Z, Byun J, Schumacher FR, Gayther SA *et al.* The oncoarray consortium: A network for understanding the genetic architecture of common cancers. *Cancer Epidemiol Biomarkers Prev* 2017; **26**: 126–135.

14 Eeles R. Prostate cancer genome-wide association study from 89,000 men using the OncoArray chip to identify novel prostate cancer susceptibility loci. *J Clin Oncol* 2016; **34**: 1525.

15 Karunamuni RA, Huynh-Le MP, Fan CC, Eeles RA, Easton DF, Kote-Jarai ZsS *et al.* The effect of sample size on polygenic hazard models for prostate cancer. *Eur J Hum Genet* 2020. doi:10.1038/s41431-020-0664-2.

16 Li Y, Byun J, Cai G, Xiao X, Han Y, Cornelis O *et al.* FastPop: A rapid principal component derived method to infer intercontinental ancestry using genetic data. *BMC Bioinformatics* 2016; **17**: 1–8.

17 Huynh-Le M-P, Chieh Fan C, Karunamuni R, Martinez ME, Eeles RA, Kote-Jarai Z *et al.* Polygenic hazard score is associated with prostate cancer in multi-ethnic populations. *medRxiv* 2019. doi:https://doi.org/10.1101/19012237.

18 Heller G. A measure of explained risk in the proportional hazards model. *Biostatistics* 2012; **13**: 315–325.

19 NCCN Clinical Practice Guidelines in Oncology. Prostate Cancer. Version 1.2019. .

20 Therneau TM, Li H. Computing the Cox Model for Case Cohort Designs. *Lifetime Data Anal* 1999; **5**: 99–112.

21 Huynh-Le M-P, Fan CC, Karunamuni R, Walsh EI, Turner EL, Lane JA *et al.* A genetic risk score to personalize prostate cancer screening, applied to population data. *Cancer Epidemiol Biomarkers Prev* 2020; : cebp.1527.2019.

22 Prostate cancer incidence statistics | Cancer Research UK. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/incidence#heading-One (accessed 23 Jul2020).

23 Martin RM, Donovan JL, Turner EL, Metcalfe C, Young GJ, Walsh EI *et al.* Effect of a low-intensity PSA-based screening intervention on prostate cancer mortality: The CAP randomized clinical trial. *JAMA - J Am Med Assoc* 2018; **319**: 883–895.

24 Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM *et al.* ClinVar: Public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* 2014; **42**: 980–985.

25 Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. *Nat Rev Genet* 2018; **19**: 581–590.

26 Pashayan N, Duffy SW, Neal DE, Hamdy FC, Donovan JL, Martin RM *et al.* Implications of polygenic risk-stratified screening for prostate cancer on overdiagnosis. *Genet Med* 2015; **17**: 789–795.

27 Pashayan N, Pharoah PDP, Schleutker J, Talala K, Tammela TLJ, Määttänen L *et al.* Reducing overdiagnosis by polygenic risk-stratified screening: Findings from the Finnish section of the ERSPC. *Br J Cancer* 2015; **113**: 1086–1093.

28 Wang TJ, Larson MG, Levy D, Benjamin EJ, Leip EP, Omland T *et al.* Plasma Natriuretic Peptide Levels and the Risk of Cardiovascular Events and Death. *N Engl J Med* 2004; **350**: 655–663.

29 Yang X, Leslie G, Gentry-Maharaj A, Ryan A, Intermaggio M, Lee A *et al.* Evaluation of

polygenic risk scores for ovarian cancer risk prediction in a prospective cohort study. *J Med Genet* 2018; **55**: 546–554.

30 Yeh HC, Duncan BB, Schmidt MI, Wang NY, Brancati FL. Smoking, smoking cessation, and risk for type 2 diabetes mellitus: A cohort study. *Ann Intern Med* 2010; **152**: 10–17.

31 Ren S, Xu J, Zhou T, Jiang H, Chen H, Liu F *et al.* Plateau effect of prostate cancer risk-associated SNPs in discriminating prostate biopsy outcomes. *Prostate* 2013; **73**: 1824–1835.

32 Forgetta V, Keller-baruch J, Forest M, Durand A, Bhatnagar S, Kemp JP *et al.* Development of a polygenic risk score to improve screening for fracture risk : A genetic risk prediction study. *PLoS Med* 2020; : 1–19.

33 Chen T-H, Chatterjee N, Landi MT, Shi J. A penalized regression framework for building polygenic risk models based on summary statistics from genome-wide association studies and incorporating external information. *J Am Stat Assoc* 2020; **1459**: 1–19.

34 Karunamuni R, Huynh-Le M-P, Fan CC, Thompson W, Eeles RA, Kote-Jarai Z *et al.* African-specific improvement of a polygenic hazard score for age at diagnosis of prostate cancer. *medRxiv* 2020; : 1–32.

35 McNeish DM. Using Lasso for Predictor Selection and to Assuage Overfitting: A Method Long Overlooked in Behavioral Sciences. *Multivariate Behav Res* 2015; **50**: 471–484.

Figure Legends

Figure 1. Cumulative incidence curves for PHS166. Risk-adjusted cumulative incidence curves for the upper 5th percentile (>95th percentile) and upper 20th percentile (>80th percentile) of PHS166 scores for clinically significant and non-clinically-significant prostate cancer. Reference curves representing the population average cumulative incidence (i.e., unadjusted for genetic risk).

Table 1. HR performance in testing set. Sample-weight-corrected hazard ratios are estimated for PHS166 and PHS46 in the testing set, using age-of-onset of any or clinically significant prostate cancer. The percent change for each metric is calculated using the value of PHS46 as the reference. Mean values and 95% confidence intervals are reported.

Type of cancer	HR	PHS46	PHS166	Change (%)
Any	HR95/50	3.29 [2.73,3.77]	4.45 [3.68,5.06]	36 [18,53]
	HR80/20	5.15 [3.92,6.18]	7.85 [6.04,9.33]	53 [25,78]
	HR20/50	0.44 [0.40,0.49]	0.37 [0.33,0.40]	-18 [-25,-10]
Clinically Significant	HR95/50	3.72 [2.89,4.43]	5.09 [3.84,6.05]	37 [13,59]
	HR80/20	6.12 [4.18,7.67]	9.45 [6.17,11.79]	55 [17,88]
	HR20/50	0.41 [0.35,0.47]	0.34 [0.29,0.39]	-18 [-28,-9]

Table 2. PPV performance in testing set. Positive predictive value (PPV) of PSA testing for clinically significant prostate cancer using top 5% (PPV95) and top 20% (PPV80) cutoffs of PHS166 and PHS46 risk scores. The percent change for each metric is calculated using the value of PHS46 as the reference.

PPV	PHS46	PHS166	Change (%)
PPV95	0.227 [0.159,0.292]	0.239 [0.171,0.305]	6.3 [-25.5,32.1]
PPV80	0.192 [0.155,0.231]	0.187 [0.150,0.222]	-2.8 [-16.3,9.9]

5

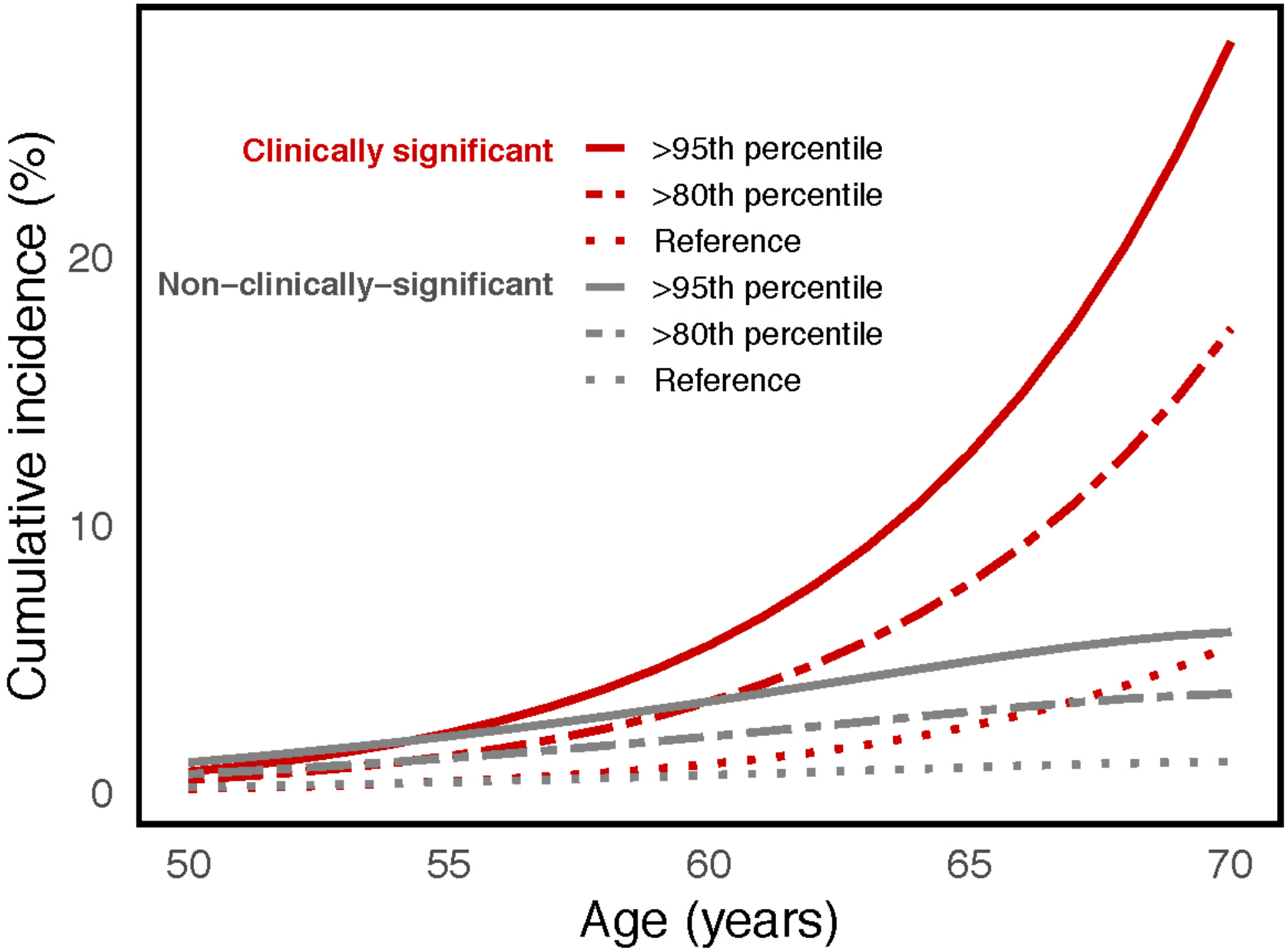

Clinically significant
>95th percentile
>80th percentile
Reference
Non−clinically−significant
>95th percentile
>80th percentile
Reference
Cumulative incidence (%)
0
10
20
50
55
60
65
70
Age (years)

Table of Contents

Supplementary Table 1. Contributing studies	2
Supplementary Table 2. SNP characteristics	6
Supplementary Table 3. Annotations for PHS166 variants	11
Supplementary Figure 1. Distribution of PHS166 score.	15
Supplementary Table 4. Explained relative risk (ERR) comparison	16
Supplementary Table 5. Testing performance for non-clinically significant prostate cancer	17
Appendix 1. Additional members of the PRACTICAL Consortium	16
Appendix 2. Funding sources and Acknowledgements for the PRACTICAL consortium	21

Supplementary Table 1.Contributing Studies. Descriptions of contributing studies to training and testing sets.

Study Group Acronym	Study Group Name	cases	controls	Average age of cases
	Training Set			
AHS	The Agricultural Health Study	491	1159	67.6
ATBC	Alpha-Tocopherol Beta-Carotene (BPC3)	1281	1913	72.2
Aarhus	Aarhus Prostate Cancer Study	1076	545	64.0
CCI	Cross Cancer Institute Prostate Brachytherapy Cohort	266	0	63.7
COH	City Of Hope	257	259	60.4
COSM	The Cohort Of Swedish Men	2298	1117	70.6
CPCS1	Copenhagen Prostate Cancer Study 1	536	258	68.2
CPCS2	Copenhagen Prostate Cancer Study 2	444	228	63.8
Canary PASS	Canary Prostate Active Surveillance Study (PASS)	364	0	62.1
CeRePP	French Prostate Case-Control Study	923	644	65.8
EPIC	European Prospective Investigation Into Cancer and Nutrition (BPC3)	635	693	66.7
ERSPC	European Randomised study of Screening for Prostate Cancer	71	65	71.2
ESTHER	Epidemiological investigations of the chances of preventing, recognizing early and optimally treating chronic diseases in an elderly population	324	315	64.8
FHCRC	Fred Hutchinson Prostate Cancer Studies	407	388	60.4
Gene-PARE	Genetic Predictors of Adverse Radiotherapy Response	242	0	66.2

HPFS	Health Professionals Follow-up Study (BPC3)	1168	1044	69.8
Hamburg-Zagreb		146	149	68.1
IMPACT	Identification of Men with a genetic predisposition to ProstAte Cancer: Targeted screening in men at a higher genetic risk and controls	49	867	63.8
IPO-Porto	Portuguese Oncology Institute of Porto	374	180	56.3
KULEUVEN	Katholieke Universiteit Leuven	166	103	65.8
LAAPC	University of Southern California – Los Angeles Prostate Cancer Study	440	280	67.3
MCC-Spain	Multi Case Control Study-Spain	520	397	66.8
MCCS	Melbourne Collaborative Cohort Study	715	315	69.6
MDACC_AS	MD Anderson Cancer Center, Active surveillance trial	501	0	64.7
MEC	Multiethnic Cohort Study (BPC3)	598	642	69.9
MOFFITT	The Moffitt Group	403	203	64.7
Malaysia	Prostate cancer study in Malaysia	1	0	78.4
Oslo	COhort of NORway (CONOR)	1443	0	72.2
PCMUS	Prostate Cancer study Medical University Sofia	192	60	68.2
PHS	Physicians Health Study (BPC3)	622	257	68.7
PLCO	Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (BPC3)	678	980	72.4
PRAGGA	PRostate cAncer Genetics in Galicia	129	100	68.1
PROCAP	PROgression in CAncer of the Prostate	659	236	64.3
PROFILE	Germline genetic profiling: correlation with targeted prostate cancer screening	13	21	59.4

	and treatment – The Pilot Profile Study			
PROGReSS	Prostate Cancer Group, Santiago, Spain	673	322	69.7
Poland	The Poland Group	484	317	69.1
ProMPT	Prostate cancer : Mechanisms of progression and Treatment	839	12	64.5
QLD	Prostate Cancer Supportive Care and Patient Outcomes Project (ProsCan), The QldMen and the Red Cross study	3282	1232	62.5
RAPPER	Radiogenomics: Assessment of Polymorphisms for Predicting the Effects of Radiotherapy	1914	0	70.2
SEARCH	Study of Epidemiology and Risk factors in Cancer Heredity	2511	1440	63.1
SFPCS	San Francisco Bay Area Prostate Cancer Study (former NC_CCPC)	279	205	64.8
SNP Prostate Ghent		316	135	65.3
SPAG	Serum Proteomic analysis for biomarkers of Aggressive prostate disease in the Guernsey population	39	169	65.3
STHM2	Stockholm 2	3019	1481	65.3
SWOG-PCPT	Prostate Cancer Prevention Trial	1072	1084	69.9
SWOG-SELECT	Selenium and Vitamin E Cancer Prevention Trial	1479	2070	67.4
TAMPERE	Finnish Genetic Predisposition to Prostate Cancer Study	2421	1183	67.2
TORONTO	Princess Margaret Biopsy Database	644	449	64.6
UKGPCS	U.K. Genetic Prostate Cancer Study and The Prostate Cancer Research Foundation Study	11939	1432	61.1

ULM	Familial Prostate Cancer Study Germany	457	178	64.3
WUGS	Washington University Genetics Study	669	0	61.2
	Testing Set			
ProtecT	Prostate Testing for Cancer and Treatment	1553	1464	62.8
UKGPCS	U.K. Genetic Prostate Cancer Study and The Prostate Cancer Research Foundation Study	30	3364	61.0

Supplementary Table 2. SNP characteristics. RS-ID, Chromosome (Chr), Effect allele, Reference Allele (Ref), Base pair position (version 37), and LASSO-derived PHS coefficient for each of the 174 SNPs considered for this study. SNPs highlighted in gray (n=8) were not included in the final model – PHS166.

RS-ID	Chr	Effect	Ref	Position	PHS coefficient
rs56391074	1	AT	A	88210715	0.019
rs17599629	1	G	A	150658287	0.028
rs34579442	1	C	CT	153899900	0.033
rs1218582	1	G	A	154834183	0.023
rs4245739	1	C	A	204518842	-0.039
rs62106670	2	T	C	8597123	0.023
rs9287719	2	C	T	10710730	0.037
rs9306895	2	C	T	20878153	0.013
rs1465618	2	T	C	43553949	0.030
rs721048	2	A	G	63131731	0.002
rs6545977	2	A	G	63301164	-0.049
rs74702681	2	T	C	66652885	0.054
rs10187424	2	C	T	85794297	-0.035
rs11691517	2	G	T	111893096	-0.032
rs12621278	2	G	A	173311553	-0.116
rs16860513	2	A	T	173342367	-0.004
rs34925593	2	C	T	174234547	0.018
rs59308963	2	TATTCTGTC	T	202123479	-0.017
rs2292884	2	G	A	238443226	0.036
rs3771570	2	T	C	242382864	0.044
rs2660753	3	T	C	87110674	0.000
rs75219487	3	A	T	87147922	0.057
rs6788616	3	G	A	87205079	0.050
rs7611694	3	C	A	113275624	-0.041
rs4857841	3	A	G	128046643	0.048
rs6763931	3	A	G	141102833	0.016
rs182314334	3	C	T	152004202	-0.041
rs142436749	3	G	A	169093100	0.083
rs78416326	3	C	G	170074517	-0.091
rs10936632	3	C	A	170130102	-0.009
rs10009409	4	T	C	73855253	0.008
rs1894292	4	A	G	74349158	-0.026
rs12500426	4	A	C	95514609	0.000
rs6853490	4	G	A	95544718	0.020
rs17021918	4	T	C	95562877	-0.036
rs7679673	4	A	C	106061534	-0.073
rs2242652	5	A	G	1280028	-0.030

rs7725218	5	A	G	1282414	-0.040
rs2736108	5	T	C	1297488	0.039
rs10866527	5	T	C	1891800	0.026
rs12653946	5	T	C	1895829	0.003
rs2121875	5	C	A	44365545	0.009
rs10793821	5	C	T	133836209	-0.019
rs76551843	5	G	A	169172133	-0.138
rs4976790	5	T	G	177968915	0.027
rs4713266	6	T	C	11219030	-0.033
rs7767188	6	A	G	30073776	0.021
rs12665339	6	G	A	30601232	0.008
rs3096702	6	A	G	32192331	0.016
rs9296068	6	G	T	32988695	-0.018
rs9469899	6	A	G	34793124	0.034
rs1983891	6	T	C	41536427	0.034
rs4711748	6	T	C	43694598	0.019
rs9443189	6	G	A	76495882	-0.012
rs2273669	6	G	A	109285189	0.032
rs339331	6	C	T	117210052	-0.041
rs3910736	6	T	C	153412476	-0.020
rs1933488	6	G	A	153441079	-0.022
rs9364554	6	T	C	160833664	0.058
rs527510716	7	C	G	1944537	0.018
rs11452686	7	TA	T	20414110	0.000
rs12155172	7	A	G	20994491	0.048
rs10486567	7	A	G	27976563	-0.060
rs17621345	7	C	A	40875192	-0.029
rs56232506	7	A	G	47437244	0.023
rs6965016	7	C	A	97807882	0.051
rs2928679	8	A	G	23438975	0.028
rs11782388	8	C	T	23525358	0.047
rs11135910	8	T	C	25892142	0.043
rs9297746	8	C	T	127909361	-0.040
rs12543663	8	C	A	127924659	0.000
rs10086908	8	C	T	128011937	0.000
rs28556804	8	G	A	128014315	-0.061
rs77541621	8	A	G	128077146	0.152
rs1016343	8	T	C	128093297	0.081
rs183373024	8	G	A	128104117	0.361
rs16901979	8	A	C	128124916	0.022
rs60163266	8	A	G	128323157	0.025
rs620861	8	A	G	128335673	-0.051
rs6983267	8	T	G	128413305	-0.081

rs1447295	8	A	C	128485038	0.000
rs7812894	8	A	T	128520479	0.135
rs12549761	8	G	C	128540776	-0.083
rs1048169	9	C	T	19055965	0.021
rs17694493	9	G	C	22041998	0.022
rs10122495	9	T	A	34049779	0.004
rs1182	9	A	C	132576060	0.034
rs141536087	10	GCGCA	G	854691	0.058
rs76934034	10	C	T	46082985	-0.002
rs10993994	10	T	C	51549496	0.117
rs1935581	10	T	C	90195149	-0.029
rs3850699	10	G	A	104414221	-0.024
rs7094871	10	C	G	114712154	-0.008
rs4962416	10	C	T	126696872	0.037
rs1881502	11	T	C	1507512	0.008
rs72853963	11	A	G	2224664	0.022
rs7127900	11	A	G	2233574	0.069
rs61890184	11	A	G	7547587	0.032
rs547171081	11	CGG	C	47421962	0.014
rs2277283	11	C	T	61908440	0.033
rs12785905	11	C	G	66951965	0.033
rs12275055	11	G	A	68981359	0.076
rs7929962	11	C	T	68985583	-0.047
rs11290954	11	A	AC	76260543	-0.025
rs11568818	11	C	T	102401661	-0.031
rs1800057	11	G	C	108143456	0.052
rs11214775	11	A	G	113807181	-0.040
rs138466039	11	T	C	125054793	0.086
rs878987	11	G	A	134266372	0.025
rs2066827	12	G	T	12871099	-0.034
rs10845938	12	A	G	14416918	-0.034
rs80130819	12	C	A	48419618	-0.051
rs10875943	12	C	T	49676010	0.038
rs902774	12	A	G	53273904	0.008
rs55914512	12	T	G	53282274	0.067
rs7968403	12	C	T	65012824	-0.030
rs5799921	12	G	GA	90160530	-0.032
rs1270884	12	A	G	114685571	0.022
rs7295014	12	G	A	133067989	0.026
rs1004030	14	C	T	23305649	-0.014
rs11629412	14	G	C	37138294	-0.032
rs8008270	14	T	C	53372330	-0.042
rs4643253	14	C	T	69106108	0.000

rs7141529	14	T	C	69126744	-0.022
rs8014671	14	A	G	71092256	-0.021
rs4924487	15	G	C	40922915	-0.032
rs33984059	15	G	A	56385868	-0.074
rs112293876	15	C	CA	66764641	0.034
rs11863709	16	T	C	57654576	-0.061
rs201158093	16	TAA	TA	82178893	0.034
rs684232	17	C	T	618965	0.056
rs28441558	17	C	T	7803118	0.061
rs142444269	17	T	C	30098749	-0.027
rs11649743	17	A	G	36074979	-0.050
rs718961	17	A	G	36077099	-0.016
rs4430796	17	G	A	36098040	-0.011
rs11651052	17	A	G	36102381	-0.085
rs117576373	17	T	C	46820676	0.098
rs11650494	17	A	G	47345186	0.047
rs2680708	17	A	G	56456120	-0.024
rs1859962	17	G	T	69108753	0.083
rs8093601	18	C	G	51772473	0.018
rs28607662	18	C	T	53230859	0.026
rs12956892	18	T	G	56746315	0.013
rs533722308	18	CT	C	60961193	0.026
rs10460109	18	T	C	73036165	0.022
rs7241993	18	T	C	76773973	-0.048
rs11666569	19	T	C	17214073	-0.032
rs118005503	19	C	G	32167803	-0.024
rs8102476	19	T	C	38735613	-0.047
rs11672691	19	A	G	41985587	-0.049
rs61088131	19	C	T	42700947	-0.009
rs17632542	19	C	T	51361757	-0.197
rs2735839	19	A	G	51364623	-0.020
rs11480453	20	CA	C	31347512	-0.025
rs12480328	20	C	T	49527922	-0.039
rs6091758	20	G	A	52455205	0.045
rs2427345	20	T	C	61015611	-0.032
rs35897249	20	G	A	62233638	-0.017
rs6062509	20	G	T	62362563	-0.017
rs1041449	21	G	A	42901421	0.028
rs9625483	22	A	G	28888939	0.059
rs58133635	22	T	C	40471188	0.025
rs5759167	22	T	G	43500212	-0.064
rs73179053	22	C	T	43501620	-0.085
rs747745	22	C	T	43503547	-0.007

rs2405942	23	G	A	9814135	-0.021
rs17321482	23	T	C	11482634	-0.024
rs4907775	23	G	A	51263200	0.053
rs2807031	23	C	T	52896949	0.000
rs7888856	23	G	A	66751555	-0.034
rs5919432	23	C	T	67021550	-0.005
rs11795627	23	T	C	69957441	-0.020
rs6625711	23	A	T	70139850	0.004

Supplementary Table 3. Annotations for PHS166 variants. Annotations and nearby genes for 166 variants used in PHS166 were tabulated using publicly available data from dbSNP.

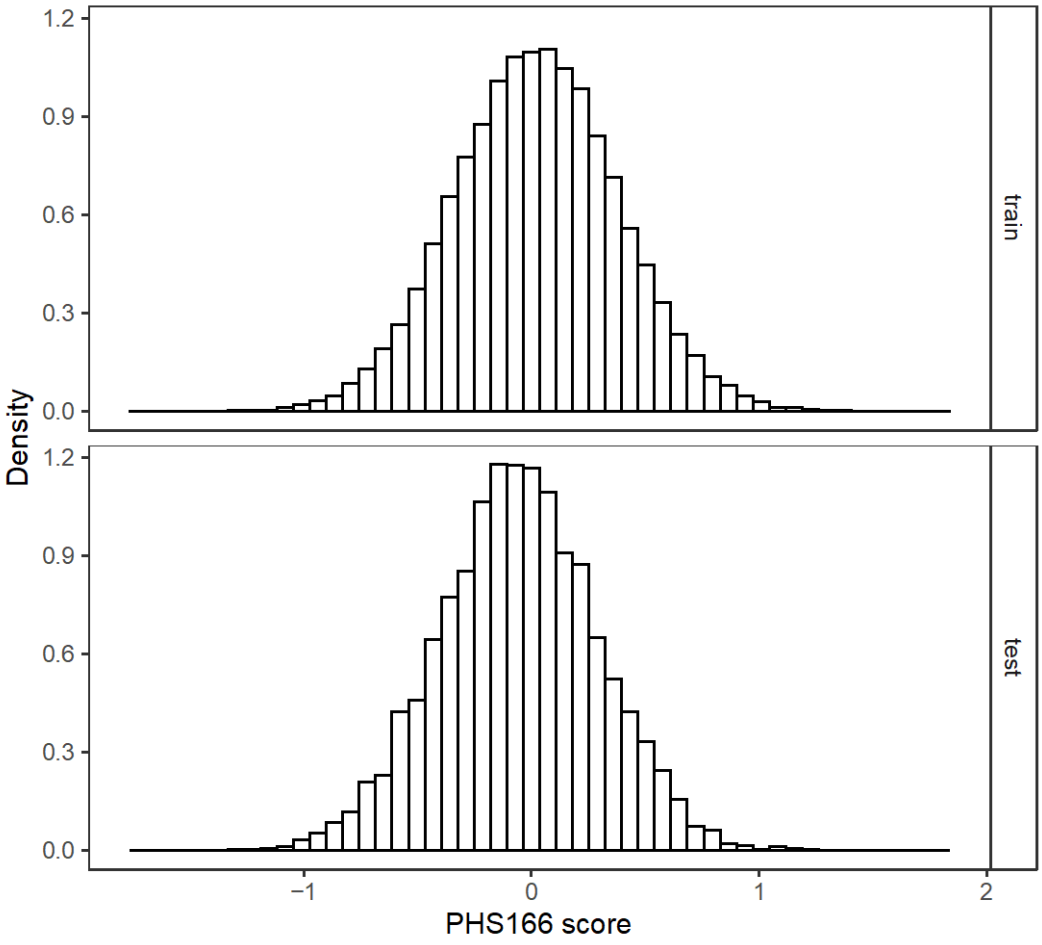
RS-ID	Gene annotations
rs56391074	None
rs17599629	GOLPH3L / Intron Variant
rs34579442	LOC101928059 / Intron Variant
rs1218582	KCNN3 / Intron Variant
rs4245739	MDM4 / Non Coding Transcript Variant
rs62106670	None
rs9287719	NOL10 / 500B Downstream Variant
rs9306895	GDF7 / 3 Prime UTR Variant
rs1465618	THADA / Intron Variant
rs721048	EHBP1 / Intron Variant
rs6545977	None
rs74702681	MEIS1-AS3 / Non Coding Transcript Variant
rs10187424	None
rs11691517	BCL2L11 / Intron Variant
rs12621278	ITGA6 / Intron Variant
rs16860513	ITGA6 / Intron Variant
rs34925593	None
rs59308963	CASP8 / Intron Variant
rs2292884	MLPH / Missense Variant
rs3771570	FARP2 / Intron Variant
rs75219487	LINC00506 / Intron Variant
rs6788616	LINC00506 / Intron Variant
rs7611694	SIDT1 / Intron Variant
rs4857841	EEFSEC / Intron Variant
rs6763931	ZBTB38 / Intron Variant
rs182314334	MBNL1 / Intron Variant
rs142436749	MECOM / Intron Variant
rs78416326	SKIL / 2KB Upstream Variant
rs10936632	None
rs10009409	LOC105377273 / Intron Variant
rs1894292	AFM / Intron Variant
rs6853490	PDLIM5 / Intron Variant
rs17021918	PDLIM5 / Intron Variant
rs7679673	None
rs2242652	TERT / Intron Variant
rs7725218	TERT / Intron Variant
rs2736108	None
rs10866527	CTD-2194D22.4 / Intron Variant
rs12653946	CTD-2194D22.4 / Intron Variant

rs2121875	FGF10 / Intron Variant
rs10793821	None
rs76551843	DOCK2 / Intron Variant
rs4976790	COL23A1 / Intron Variant
rs4713266	NEDD9 / Intron Variant
rs7767188	TRIM31 / Intron Variant
rs12665339	ATAT1 / Intron Variant
rs3096702	NOTCH4 / 2KB Upstream Variant
rs9296068	None
rs9469899	UHRF1BP1 / Intron Variant
rs1983891	FOXP4 / Intron Variant
rs4711748	None
rs9443189	MYO6 / Intron Variant
rs2273669	ARMC2 / Intron Variant
rs339331	RFX6 / Intron Variant
rs3910736	RGS17 / Intron Variant
rs1933488	RGS17 / Intron Variant
rs9364554	SLC22A3 / Intron Variant
rs527510716	MAD1L1 / Intron Variant
rs12155172	LINC01162 / Intron Variant
rs10486567	JAZF1 / Intron Variant
rs17621345	SUGCT / Intron Variant
rs56232506	TNS3 / Intron Variant
rs6965016	LMTK2 / Intron Variant
rs2928679	None
rs11782388	LOC107986930 / Intron Variant
rs11135910	EBF2 / Intron Variant
rs9297746	LOC105375751 / Intron Variant
rs28556804	LOC105375751 / Intron Variant
rs77541621	None
rs1016343	PCAT2 / Intron Variant
rs183373024	PRNCR1 / Non Coding Transcript Variant
rs16901979	None
rs60163266	CASC8 / Intron Variant
rs620861	CASC8 / Intron Variant
rs6983267	CASC8 / Intron Variant
rs7812894	None
rs12549761	None
rs1048169	HAUS6 / 3 Prime UTR Variant
rs17694493	CDKN2B-AS1 / Intron Variant
rs10122495	UBAP2 / 2KB Upstream Variant
rs1182	TOR1A / Non Coding Transcript Variant
rs141536087	LARP4B / 3 Prime UTR Variant

rs76934034	MARCHF8 / Intron Variant
rs10993994	MSMB / 2KB Upstream Variant
rs1935581	RNLS / Intron Variant
rs3850699	TRIM8 / Intron Variant
rs7094871	TCF7L2 / Intron Variant
rs4962416	CTBP2 / Intron Variant
rs1881502	MOB2 / Intron Variant
rs72853963	None
rs7127900	None
rs61890184	PPFIBP2 / Intron Variant
rs547171081	MIR4487 / 2KB Upstream Variant
rs2277283	INCENP / Missense Variant
rs12785905	KDM2A / Intron Variant
rs12275055	None
rs7929962	None
rs11290954	EMSY / Intron Variant
rs11568818	MMP7 / 2KB Upstream Variant
rs1800057	ATM / Missense Variant
rs11214775	HTR3B / Intron Variant
rs138466039	None
rs878987	B3GAT1 / Intron Variant
rs2066827	CDKN1B / Missense Variant
rs10845938	None
rs80130819	LOC105369750 / Intron Variant
rs10875943	None
rs902774	None
rs55914512	None
rs7968403	RASSF3 / Intron Variant
rs5799921	LOC107984543 / Intron Variant
rs1270884	None
rs7295014	FBRSL1 / Intron Variant
rs1004030	MMP14 / 2KB Upstream Variant
rs11629412	PAX9 / Intron Variant
rs8008270	FERMT2 / Intron Variant
rs7141529	RAD51B / Intron Variant
rs8014671	LOC101928075 / Intron Variant
rs4924487	KNL1 / Intron Variant
rs33984059	RFX7 / Missense Variant
rs112293876	None
rs11863709	ADGRG1 / Intron Variant
rs201158093	None

rs684232	VPS53 / 2KB Upstream Variant
rs28441558	CHD3 / Intron Variant
rs142444269	None
rs11649743	HNF1B / Intron Variant
rs718961	HNF1B / Intron Variant
rs4430796	HNF1B / Intron Variant
rs11651052	HNF1B / Intron Variant
rs117576373	LOC105371811 / Non Coding Transcript Variant
rs11650494	None
rs2680708	RNF43 / Intron Variant
rs1859962	CASC17 / Intron Variant
rs8093601	None
rs28607662	TCF4 / Intron Variant
rs12956892	None
rs533722308	BCL2 / Intron Variant
rs10460109	None
rs7241993	LOC105372225 / Intron Variant
rs11666569	MYO9B / Intron Variant
rs118005503	None
rs8102476	None
rs11672691	PCAT19 / Intron Variant
rs61088131	POU2F2 / Intron Variant
rs17632542	KLK3 / Missense Variant
rs2735839	None
rs11480453	None
rs12480328	ADNP / Intron Variant
rs6091758	None
rs2427345	LOC105372710 / Intron Variant
rs35897249	GMEB2 / Intron Variant
rs6062509	ZGPAT / Intron Variant
rs1041449	None
rs9625483	TTC28 / Intron Variant
rs58133635	TNRC6B / Intron Variant
rs5759167	None
rs73179053	None
rs747745	None
rs2405942	SHROOM2 / Intron Variant
rs17321482	ARHGAP6 / Intron Variant
rs4907775	None
rs7888856	None
rs5919432	None
rs11795627	TEX11 / Intron Variant
rs6625711	None

Supplementary Figure 1. Distribution of PHS166 score. Histograms of PHS166 scores for training and testing sets show similar patterns in distribution.



Supplementary Table 4. Explained relative risk (ERR) comparison. Values of ERR are tabulated for PHS166, Family History (None or >=1 affected relative), and the full model (PHS166 + Family History). PHS166 contributed 82% of the overall ERR in the training set, and 98% of the overall ERR in the testing set. z

dataset	PHS166	Family History	PHS166 + Family History
training	0.116 [0.109,1.122]	0.016 [0.013, 0.018]	0.139 [0.132, 0.1460
testing	0.147 [0.116,0.182]	0.001 [0, 0.006]	0.15 [0.12,0.185]

Supplementary Table 5. Testing performance for non-clinically significant prostate cancer. Sample-weight hazard ratios are estimated for PHS166 and PHS46 in the testing set using the age-of-onset of non-clinically significant prostate cancer.

HR	PHS46	PHS166	Change (%)
HR95/50	3.20 [2.59,3.78]	4.28 [3.49,4.96]	34 [16,51]
HR80/20	4.96 [3.61,6.14]	7.47 [5.49,9.07]	51 [23,77]
HR20/50	0.45 [0.39,0.50]	0.37 [0.33,0.42]	-17 [-25,-10]

Appendix 1. Additional members of the PRACTICAL Consortium

Teuvo L. J. Tammela[1,2], Csilla Sipeky[3], Anssi Auvinen[4], Alison M. Dunning[5], Maya Ghoussaini[6], Suzanne Chambers[7,8], Judith A. Clements[9,10], Lisa Horvath[11,12], Leire Moya[9,10], Gail P. Risbridger[13,14], Wayne Tilley[15], Markus Aly[16,17,18], Tobias Nordström[19,20], Stig E. Bojesen[21,22], Peter Iversen[23], Phyllis J. Goodman[24], Ian M. Thompson Jr.[25], Melissa C. Southey[26,27,28], Graham G. Giles[27], Roger L. Milne[27,29,26], Niclas Håkansson[30], Stephanie Weinstein[31], Fredrick R. Schumacher[32,33], Loic Le Marchand[34], Xin Sheng[35], Tim J. Key[36], Elaine A. Ostrander[37], Edward Giovannucci[38], Neil Burnet[39], Gill Barnett[40], Martin Andreas Røder[23], Bettina F. Drake[41], Géraldine Cancel-Tassin[42,43], Stephen Chanock[31], Gerald L. Andriole[41], Robert N. Hoover[31], Mitchell J. Machiela[31], Laura E. Beane Freeman[31], Michael Borre[44,45], Dominika Wokolorczyk[46], Jan Lubinski[46], Lovise Maehle[47], Hui-Yi Lin[48], Mariana C Stern[49], Manuel Luedeke[50], Thomas Schnoeller[51], Neil E. Fleshner[52], Antonio Finelli[53], Sarah L. Kerns[54], Harry Ostrer[55], Antonio Gómez-Caamaño[56], Laura Fachal[57,58,59,60], Javier Llorca[61,62], Gemma Castaño-Vinyals[63,64,65,66], Antonio Alcaraz[67], Robert Szulkin[68,69], Martin Eklund[70], Peter Kraft[71], Sara Lindström[72], Meir Stampfer[73], Paula Paulo[74], Andreia Brandão[75,74], Marta Cardoso[74], Maria P. Silva[75,74], Xin Gao[76], Bernd Holleczek[77], Ben Schöttker[76], Chavdar Slavov[78], Vanio Mitev[79], Jeri Kim[80], Yuan Chun Ding[81], Linda Steele[81], Jasmine Lim[82], Soo-Hwang Teo[83], Daniel W. Lin[84,85], Davor Lessel[86], Zeljko Kastelan[87], Matthew Parliament[88,89], Aswin Abraham[88], Steven Joniau[90], Thomas Van den Broeck[91,90], Jose Esteban Castelao[92], Maria Elena Martinez[93], Ron H.N. van Schaik[94], Guido Jenster[95], Maria Elena Martinez[98], Ron H.N. van Schaik[99], Guido Jenster[100]

[1]Department of Urology, Tampere University Hospital, FI-33521 Tampere, Finland
[2]Faculty of Medicine and Health Technology, Tampere University, FI-33100 Tampere, Finland
[3]Institute of Biomedicine, University of Turku, Finland
[4]Unit of Health Sciences, Faculty of Social Sciences, Tampere University, Tampere, Finland
[5]Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Strangeways Laboratory, Worts Causeway, Cambridge, CB1 8RN, UK
[6]Open Targets, Wellcome Sanger Institute, Hinxton, Saffron Walden, CB10 1SA, UK
[7]University of Technology, Sydney
[8]Cancer Council Queensland, Fortitude Valley, QLD 4006, Australia
[9]Australian Prostate Cancer Research Centre-Qld, Institute of Health and Biomedical Innovation and School of Biomedical Sciences, Queensland University of Technology, Brisbane QLD 4059, Australia
[10]Translational Research Institute, Brisbane, Queensland 4102, Australia
[11]Chris O'Brien Lifehouse (COBLH), Camperdown, Sydney, NSW 2010, Australia
[12]Garvan Institute of Medical Research, Sydney NSW 2010, Australia
[13]Department of Anatomy and Developmental Biology, Biomedicine Discovery Institute, Monash University, Melbourne,Victoria 3800, Australia
[14]Prostate Cancer Translational Research Program, Cancer Research Division, Peter MacCallum Cancer Centre, Melbourne, VIC 3000, Australia
[15]Dame Roma Mitchell Cancer Research Laboratories, University of Adelaide, Adelaide, South Australia, Australia
[16]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, SE-171 77 Stockholm, Sweden

[17]Department of Molecular Medicine and Surgery, Karolinska Institutet
[18]Department of Urology, Karolinska University Hospital, Solna, 171 76 Stockholm
[19]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden
[20]Department of Clinical Sciences at Danderyds Hospital, Karolinska Institutet, Stockholm, Sweden
[21]Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark
[22]Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, 2200 Copenhagen, Denmark
[23]Copenhagen Prostate Cancer Center, Department of Urology, Rigshospitalet, Copenhagen University Hospital, DK-2730 Herlev, Copenhagen, Denmark
[24]SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[25]CHRISTUS Santa Rosa Hospital – Medical Center, San Antonio, TX, USA
[26]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria 3168, Australia
[27]Cancer Epidemiology Division, Cancer Council Victoria, 615 St Kilda Road, Melbourne, VIC 3004, Australia
[28]Department of Clinical Pathology, The Melbourne Medical School, The University of Melbourne, Melbourne, VIC 3010, Australia.
[29]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Grattan Street, Parkville, VIC 3010, Australia
[30]Unit of Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, SE-171 77 Stockholm, Sweden
[31]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, Maryland, 20892, USA
[32]Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH 44106-7219, USA
[33]Seidman Cancer Center, University Hospitals, Cleveland, OH 44106, USA.
[34]Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI 96813, USA
[35]Center for Genetic Epidemiology, Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, CA 90015, USA
[36]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, OX3 7LF, UK
[37]National Human Genome Research Institute, National Institutes of Health, 50 South Drive, Rm. 5351, Bethesda, MD 20892, USA
[38]Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA
[39]Division of Cancer Sciences, University of Manchester, Manchester Cancer Research Centre, Manchester Academic Health Science Centre, and The Christie NHS Foundation Trust, Manchester, UK
[40]University of Cambridge, Department of Oncology, Oncology Centre, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge, CB2 0QQ, UK
[41]Washington University School of Medicine, 660 S. Euclid Avenue, Campus Box 8242, St.

Louis, MO 63110 , USA
[42]CeRePP, Tenon Hospital, F-75020 Paris, France.
[43]Sorbonne Universite, GRC n°5 , AP-HP, Tenon Hospital, 4 rue de la Chine, F-75020 Paris, France
[44]Department of Urology, Aarhus University Hospital, Palle Juul-Jensen Boulevard 99, 8200 Aarhus N, Denmark
[45]Department of Clinical Medicine, Aarhus University, DK-8200 Aarhus N
[46]International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, 70-115 Szczecin, Poland
[47]Department of Medical Genetics, Oslo University Hospital, 0424 Oslo, Norway
[48]School of Public Health, Louisiana State University Health Sciences Center, New Orleans, LA 70112, USA
[49]Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, CA 90015, USA
[50]Humangenetik Tuebingen, Paul-Ehrlich-Str 23, D-72076 Tuebingen, Germany
[51]Department of Urology, University Hospital Ulm, Germany
[52]Dept. of Surgical Oncology, Princess Margaret Cancer Centre, Toronto ON M5G 2M9, Canada
[53]Division of Urology, Princess Margaret Cancer Centre, Toronto ON M5G 2M9, Canada
[54]Department of Radiation Oncology, University of Rochester Medical Center, Saunders Research Building, Rm 2.202, 265 Crittenden Boulevard, Rochester, NY 14620, USA
[55]Professor of Pathology and Pediatrics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Ullman 817, Bronx, NY 10461, USA
[56]Department of Radiation Oncology, Complexo Hospitalario Universitario de Santiago, SERGAS, Santiago de Compostela 15706, Spain
[57]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge, CB2 0SR, UK.
[58]Fundación Pública Galega Medicina Xenómica, Santiago De Compostela, 15706, Spain.
[59]Instituto de Investigación Sanitaria de Santiago de Compostela, Santiago De Compostela, 15706, Spain.
[60]Centro de Investigación en Red de Enfermedades Raras (CIBERER), Spain
[61]University of Cantabria-IDIVAL, 39005 Santander, Spain
[62]CIBER Epidemiología y Salud Pública (CIBERESP), 28029 Madrid, Spain
[63]ISGlobal, Barcelona, Spain
[64]IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain
[65]Universitat Pompeu Fabra (UPF), Barcelona, Spain
[66]CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[67]Department and Laboratory of Urology. Hospital Clínic. Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona. Spain. C/Villarroel 170; 08036 Barcelona, Spain
[68]Division of Family Medicine, Department of Neurobiology, Care Science and Society, Karolinska Institutet, Huddinge, Sweden
[69]Scandinavian Development Services, Danderyd, Sweden
[70]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, 171 77 Stockholm,

Sweden
[71]Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA
[72]Department of Epidemiology, Health Sciences Building, University of Washington
[73]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital/Harvard Medical School, Boston, MA 02184, USA
[74]Cancer Genetics Group, IPO-Porto Research Center (CI-IPOP), Portuguese Oncology Institute of Porto (IPO-Porto), Porto, Portugal
[75]Department of Genetics, Portuguese Oncology Institute of Porto (IPO-Porto), Porto, Portugal
[76]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), D-69120, Heidelberg, Germany
[77]Saarland Cancer Registry, 66119 Saarbrücken, Germany
[78]Department of Urology and Alexandrovska University Hospital, Medical University of Sofia, 1431 Sofia, Bulgaria
[79]Molecular Medicine Center, Department of Medical Chemistry and Biochemistry, Medical University of Sofia, Sofia, 2 Zdrave Str., 1431 Sofia, Bulgaria
[80]The University of Texas M. D. Anderson Cancer Center, Department of Genitourinary Medical Oncology, 1515 Holcombe Blvd., Houston, TX 77030, USA
[81]Department of Population Sciences, Beckman Research Institute of the City of Hope, 1500 East Duarte Road, Duarte, CA 91010, 626-256-HOPE (4673)
[82]Department of Surgery, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
[83]Cancer Research Malaysia (CRM), Outpatient Centre, Subang Jaya Medical Centre, Subang Jaya, Selangor, Malaysia
[84]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, 98109-1024, USA
[85]Department of Urology, University of Washington, 1959 NE Pacific Street, Box 356510, Seattle, WA 98195, USA
[86]Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, D-20246 Hamburg, Germany
[87]Department of Urology, University Hospital Center Zagreb, University of Zagreb School of Medicine , Zagreb, Croatia
[88]Department of Oncology, Cross Cancer Institute, University of Alberta, 11560 University Avenue, Edmonton, Alberta, Canada T6G 1Z2
[89]Division of Radiation Oncology, Cross Cancer Institute, 11560 University Avenue, Edmonton, Alberta, Canada T6G 1Z2
[90]Department of Urology, University Hospitals Leuven, Herestraat 49, Box 7003 41, BE-3000 Leuven, Belgium
[91]Molecular Endocrinology Laboratory, Department of Cellular and Molecular Medicine, KU Leuven, BE-3000, Belgium
[92]Genetic Oncology Unit, CHUVI Hospital, Complexo Hospitalario Universitario de Vigo, Instituto de Investigación Biomédica Galicia Sur (IISGS), 36204, Vigo (Pontevedra), Spain
[93]University of California San Diego, Moores Cancer Center, Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA 92093-0012, USA

[94]Department of Clinical Chemistry, Erasmus University Medical Center, 3015 CE Rotterdam, The Netherlands
[95]Department of Urology, Erasmus University Medical Center, 3015 CE Rotterdam, The Netherlands
[96]Molecular Endocrinology Laboratory, Department of Cellular and Molecular Medicine, KU Leuven, BE-3000, Belgium
[97]Genetic Oncology Unit, CHUVI Hospital, Complexo Hospitalario Universitario de Vigo, Instituto de Investigación Biomédica Galicia Sur (IISGS), 36204, Vigo (Pontevedra), Spain
[98]University of California San Diego, Moores Cancer Center, Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA 92093-0012, USA
[99]Department of Clinical Chemistry, Erasmus University Medical Center, 3015 CE Rotterdam, The Netherlands
[100]Department of Urology, Erasmus University Medical Center, 3015 CE Rotterdam, The Netherlands

Appendix 2. Funding sources and Acknowledgements for the PRACTICAL consortium

CRUK and PRACTICAL consortium
This work was supported by the Canadian Institutes of Health Research, European Commission's Seventh Framework Programme grant agreement n° 223175 (HEALTH-F2-2009-223175), Cancer Research UK Grants C5047/A7357, C1287/A10118, C1287/A16563, C5047/A3354, C5047/A10692, C16913/A6135, and The National Institute of Health (NIH) Cancer Post-Cancer GWAS initiative grant: No. 1 U19 CA 148537-01 (the GAME-ON initiative). We 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.
The Prostate Cancer Program of Cancer Council Victoria also acknowledge grant support from The National Health and Medical Research Council, Australia (126402, 209057, 251533, , 396414, 450104, 504700, 504702, 504715, 623204, 940394, 614296,), VicHealth, Cancer Council Victoria, The Prostate Cancer Foundation of Australia, The Whitten Foundation, PricewaterhouseCoopers, and Tattersall's. EAO, DMK, and EMK acknowledge the Intramural Program of the National Human Genome Research Institute for their support.
Genotyping of the OncoArray was funded by the US National Institutes of Health (NIH) [U19 CA 148537 for ELucidating Loci Involved in Prostate cancer SuscEptibility (ELLIPSE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract number HHSN268201200008I].
This study would not have been possible without the contributions of the following: Coordination team, bioinformatician and genotyping centers: Genotyping at CCGE, Cambridge: Caroline Baines and Don Conroy

Funding for the iCOGS infrastructure came from: the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund.

This study would not have been possible without the contributions of the following: Per Hall (COGS); Douglas F. Easton, Paul Pharoah, Kyriaki Michailidou, Manjeet K. Bolla, Qin Wang (BCAC), Andrew Berchuck (OCAC), Rosalind A. Eeles, Douglas F. Easton, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL), Georgia Chenevix-Trench, Antonis Antoniou, Lesley McGuffog, Fergus Couch and Ken Offit (CIMBA), Joe Dennis, Alison M. Dunning, Andrew Lee, and Ed Dicks, Craig Luccarini and the staff of the Centre for Genetic Epidemiology

Laboratory, Javier Benitez, Anna Gonzalez-Neira and the staff of the CNIO genotyping unit, Jacques Simard and Daniel C. Tessier, Francois Bacot, Daniel Vincent, Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, Stig E. Bojesen, Sune F. Nielsen, Borge G. Nordestgaard, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility

Additional funding and acknowledgments from studies in PRACTICAL:

Information of the PRACTICAL consortium can be found at http://practical.icr.ac.uk/

Aarhus
This study was supported by Innovation Fund Denmark, the Danish Cancer Society and The Velux Foundation (Veluxfonden).
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
The ATBC Study is supported by the Intramural Research Program of the U.S. National Cancer Institute, National Institutes of Health, and by U.S. Public Health Service contract HHSN261201500005C from the National Cancer Institute, Department of Health and Human Services.

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

CAP
The CAP trial is funded by Cancer Research UK and the UK Department of Health (C11043/A4286, C18281/A8145, C18281/A11326, C18281/A15064, and C18281/A24432). RMM is supported by a Cancer Research UK (C18281/A19169) programme grant (the Integrative Cancer Epidemiology Programme) and by the NIHR Biomedical Research Centre at University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

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.

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

COSM
COSM is funded by The Swedish Research Council (grant for the Swedish Infrastructure for Medical Population-based Life-course Environmental Research – SIMPLER), the Swedish Cancer Foundation.

CPCS1 / CPCS2
Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev Ringvej 75, DK-2730 Herlev, Denmark
We thank participants and staff of the Copenhagen General Population Study for their important contributions.

EPIC
The coordination of EPIC was financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The national cohorts (that recruited male participants) are supported by Danish Cancer Society (Denmark); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy); 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), Statistics Netherlands (The Netherlands); Health Research Fund (FIS), PI13/00061 to Granada; , PI13/01162 to EPIC-Murcia), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020) (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford) (United Kingdom). For information on how to submit an application for gaining access to EPIC data and/or biospecimens, please follow the instructions at http://epic.iarc.fr/access/index.php

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 unconditional 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-CA080122, R01-CA056678, R01-CA082664, and R01-CA092579, and K05-CA175147 from the US National Cancer Institute, National Institutes of Health, with additional support from the Fred Hutchinson Cancer Research Center (P30-CA015704).
We thank all the individuals 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. S.L.K. is supported by 1K07CA187546 from the U.S. National Cancer Institute.

HPFS
The Health Professionals Follow-up Study was supported by grants UM1CA167552, CA133891, CA141298, and P01CA055075.
We 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.
We 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 FEDER (POCI-01-0145-FEDER-028245). MC (SFRH/BD/116557/2016) and MPS (SFRH/BD/132441/2017) are research fellows from FCT.
We would like to express our gratitude to all patients and families who have participated in this study.

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 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 to A.R).
We thank all associates in the Urology Unit, University of Malaya, Cancer Research Malaysia (CRM) and the Malaysian Men's Health Initiative (MMHI).

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, PI15/00069), 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). ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya.
We acknowledge 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.

MCCS
Melbourne Collaborative Cohort Study (MCCS) cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further augmented by Australian National Health and Medical Research Council grants 209057, 396414 and 1074383 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database.

MEC
The MEC was supported by NIH grants CA063464, CA054281, CA098758, and CA164973.

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

Oslo
CONOR was supported by grants from the Nordic Cancer Union, the Swedish Cancer Society (2012/823) and the Swedish Research Council (2014/2269).
The authors wish to acknowledge the services of CONOR, the contributing research centres delivering data to CONOR, and all the study participants.

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.
We 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, NIH
The authors thank 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.

PRAGGA
PRAGGA was supported by Programa Grupos Emergentes, Cancer Genetics Unit, CHUVI Vigo Hospital, Instituto de Salud Carlos III, Spain.
PRAGGA wishes to thank Victor Muñoz Garzón, Manuel Enguix Castelo, Sara Miranda Ponte, Carmen M Redondo, Manuel Calaza, Francisco Gude Sampedro, Joaquín González-Carreró and the staff of the Department of Pathology and Biobank of University Hospital Complex of Vigo, Instituto de Investigación Sanitaria Galicia Sur (IISGS), SERGAS, Vigo, Spain; Máximo Fraga, José Antúnez and the Biobank of University Hospital Complex of Santiago, Santiago de Compostela, Spain; and Maria Torres, Angel Carracedo and the Galician Foundation of Genomic Medicine.

PROCAP
PROCAP was supported by the Swedish Cancer Foundation (08-708, 09-0677).

We thank and acknowledge 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.

PROFILE

We would like to acknowledge the support of the Ronald and Rita McAulay Foundation and Cancer Research UK. We also 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

We acknowledge the Profile study steering committee and participants.

PROGReSS

This research was supported by Spanish Instituto de Salud Carlos III (ISCIII) funding, an initiative of the Spanish Ministry of Economy and Innovation partially supported by European Regional Development FEDER Funds (INT15/00070, INT16/00154, INT17/00133; PI19/01424; PI16/00046; PI13/02030; PI10/00164), and through the Autonomous Government of Galicia (Consolidation and structuring program: IN607B) given to A.Vega.

We would like to thank the patients for their contribution to the study

ProtecT/ProMPT

The ProtecT team acknowledges the support of The Universities of Oxford, Bristol and Cambridge, Cancer Research UK and NIHR. CR-UK grants (C8197/A10123; C8197/A10865) supported genotyping. We acknowledge the support of the National Institute for Health Research Cambridge Bio-medical Research Centre, Cambridge, UK, and Oxford Biomedical Research Centre, Oxford, UK. The national ProMPT (Prostate cancer Mechanisms of Progression and Treatment) collaborative (grant G0500966/75466) supported sample collections in Cambridge and Oxford. The University of Oxford sponsors ProMPT and ProtecT. We are grateful to staff at the Bristol Medical School, Bristol, UK, the Nuffield Department of Surgical Sciences at the University of Oxford, Oxford, UK, the Wellcome Trust Clinical Research Facility, Addenbrooke's Clinical Research Centre, Cambridge, UK, and all research groups in the nine ProtecT participating Clinical centres in the UK (Birmingham, Bristol, Cambridge, Cardiff, Edinburgh, Leeds, Leicester, Newcastle, Sheffield), for conducting the ProtecT study. The NIHR Health Technology Assessment Programme (HTA) funded the ProtecT study (projects 96/20/06, 96/20/99), and supported its linked ProMPT and CAP (Cluster Randomized Trial of PSA Testing for Prostate Cancer) studies, with Cancer Research UK (grants C522/A8649, C11043/A4286, C18281/A8145, C18281/A11326, and C18281/A15064), the Medical Research Council (grant G0500966, ID 75466) and NCRI, UK. The ProtecT epidemiological data were generated through funding from the Southwest National Health Service Research and Development. ProtecT DNA extraction was supported by the USA Dept of Defense award W81XWH-04-1-0280, Yorkshire Cancer Research and CR-UK. We thank the

National Institute for Health Research, Hutchison Whampoa Limited, the Human Research Tissue Bank (Addenbrooke's Hospital), the Oxford Human Tissue Bank, and Cancer Research UK for supporting the biorepository. The authors thank men with prostate cancer and all the participants who donated their time and samples to the Oxford and Cambridge Biorepositories used in this research. We acknowledge the support of research staff who so carefully curated the samples and follow-up data (Jo Burge, Marie Corcoran, Anne George, Sara Stearn, Rajeev Kumar, Michael Davis, Peter Holding). The views and opinions expressed herein are those of the authors and do not necessarily reflect those of the Department of Health and Social Care in England.

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, Cancer Australia PdCCRS and Cancer Council Queensland funding to J Batra
The QLD team would like to acknowledge and sincerely thank the urologists, pathologists, data managers and patient participants who have generously and altruistically supported the QLD cohort.

RAPPER
RAPPER is supported by Cancer Research UK [C1094/A11728; C1094/A18504] and Experimental Cancer Medicine Centre funding [C1467/A7286], and the NIHR Manchester Biomedical Research Centre.
The RAPPER group thank Dr. Holly Summersgill for project management.

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. The University of Cambridge has received salary support in respect of PP from the NHS in the East of England through the Clinical Academic Reserve.

SFPCS
SFPCS was funded by California Cancer Research Fund grant 99-00527V-10182

SPAG
Manchester Cancer Research Centre and MRC – Medical Research Council for PhD studentship work, MUG - Male Uprising in Guernsey, Hope for Guernsey and Wessex Medical Research for research support

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 2012-0073
The authors acknowledge 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 / SWOG-SELECT
PCPT is funded by Public Health Service grants U10CA37429 and 5UM1CA182883 from the National Cancer Institute.
The authors thank the site investigators and staff and, most importantly, the participants from PCPT / SELECT 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.

Toronto
Prostate Cancer Canada Movember Discovery Grant (D2013-17) to RJH; Canadian Cancer Society Research Institute Career Development Award in Cancer Prevention (2013-702108) to RJH

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. UKGPCS would like to thank all urologists and other persons involved in the planning, coordination, and data collection of the study. KM and AL were in part supported from the NIHR Manchester Biomedical Research Centre

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

WUGS
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.