Risk Analysis of Prostate Cancer in PRACTICAL, a Multinational Consortium, Using 25 Known Prostate Cancer Susceptibility Loci

Ali Amin Al Olama[1], Sara Benlloch[1], Antonis C. Antoniou[1], Graham G. Giles[2,3], Gianluca Severi[2], David E. Neal[4,5], Freddie C. Hamdy[6,7], Jenny L. Donovan[8], Kenneth Muir[9,10], Johanna Schleutker[11,12], Brian E. Henderson[13], Christopher A. Haiman[13], Fredrick R. Schumacher[13], Nora Pashayan[1,14], Paul D.P. Pharoah[1], Elaine A. Ostrander[15], Janet L. Stanford[16,17], Jyotsna Batra[18], Judith A. Clements[18], Suzanne K. Chambers[19,20,21], Maren Weischer[22], Børge G. Nordestgaard[22], Sue A. Ingles[13], Karina D. Sorensen[23], Torben F. Orntoft[23], Jong Y. Park[24], Cezary Cybulski[25], Christiane Maier[26], Thilo Doerk[27], Joanne L. Dickinson[28], Lisa Cannon-Albright[29,30], Hermann Brenner[31,32], Timothy R. Rebbeck[33], Charnita Zeigler-Johnson[34], Tomonori Habuchi[35], Stephen N. Thibodeau[36], Kathleen A. Cooney[37], Pierre O. Chappuis[38], Pierre Hutter[39], Radka P. Kaneva[40], William D. Foulkes[41], Maurice P. Zeegers[42], Yong-Jie Lu[43], Hong-Wei Zhang[44], Robert Stephenson[45], Angela Cox[46], Melissa C. Southey[47], Amanda B. Spurdle[48], Liesel FitzGerald[49], Daniel Leongamornlert[50], Edward Saunders[50], Malgorzata Tymrakiewicz[50], Michelle Guy[50], Tokhir Dadaev[50], Sarah J. Little[50], Koveela Govindasami[50], Emma Sawyer[50], Rosemary Wilkinson[50], Kathleen Herkommer[51], John L. Hopper[3], Aritaya Lophatonanon[10,11], Antje E. Rinckleb[26], Zsofia Kote-Jarai[50]; Rosalind A. Eeles[50,52], on behalf of The UK Genetic Prostate Cancer Study Collaborators/ British Association of Urological Surgeons' Section of Oncology[53], on behalf of The UK ProtecT Study Collaborators[54], and on behalf of The PRACTICAL Consortium[55]; and Douglas F. Easton[1]

Abstract

Background: Genome-wide association studies have identified multiple genetic variants associated with prostate cancer risk which explain a substantial proportion of familial relative risk. These variants can be used to stratify individuals by their risk of prostate cancer.

Methods: We genotyped 25 prostate cancer susceptibility loci in 40,414 individuals and derived a polygenic risk score (PRS). We estimated empirical odds ratios (OR) for prostate cancer associated with different risk strata defined by PRS and derived age-specific absolute risks of developing prostate cancer by PRS stratum and family history.

Results: The prostate cancer risk for men in the top 1% of the PRS distribution was 30.6 (95% CI, 16.4–57.3) fold compared with men in the bottom 1%, and 4.2 (95% CI, 3.2–5.5) fold compared with the median risk. The absolute risk of prostate cancer by age of 85 years was 65.8% for a man with family history in the top 1% of the PRS distribution, compared with 3.7% for a man in the bottom 1%. The PRS was only weakly correlated with serum PSA level (correlation = 0.09).

Conclusions: Risk profiling can identify men at substantially increased or reduced risk of prostate cancer. The effect size, measured by OR per unit PRS, was higher in men at younger ages and in men with family history of prostate cancer. Incorporating additional newly identified loci into a PRS should improve the predictive value of risk profiles.

Impact: We demonstrate that the risk profiling based on SNPs can identify men at substantially increased or reduced risk that could have useful implications for targeted prevention and screening programs. *Cancer Epidemiol Biomarkers Prev; 24(7); 1121–9. ©2015 AACR.*

[1]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge, United Kingdom. [2]Cancer Epidemiology Centre, the Cancer Council Victoria, Carlton, Victoria, Australia. [3]Centre for Molecular, Environmental, Genetic, and Analytic Epidemiology, The University of Melbourne, Melbourne, Victoria, Australia. [4]University of Cambridge, Department of Oncology, Addenbrooke's Hospital, Cambridge, United Kingdom. [5]Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge, United Kingdom. [6]Nuffield Department of Surgical Sciences, University of Oxford, Oxford, United Kingdom. [7]Faculty of Medical Science, University of Oxford, John Radcliffe Hospital, Oxford, United Kingdom. [8]School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom. [9]The University of Manchester, Centre for Epidemiology, Institute of Population Health, Manchester, United Kingdom. [10]University of Warwick, University House, Coventry, United Kingdom. [11]Institute of Biomedical Technology/BioMediTech, University of Tampere and FimLab Laboratories, Tampere, Finland. [12]Department of Medical Biochemistry and Genetics, University of Turku, Turku, Finland. [13]Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Centre, Los Angeles, California. [14]University College London,

AACR

Introduction

Genome-wide association studies (GWAS) have identified multiple common genetic variants associated with prostate cancer risk. The risks associated with such variants are generally modest, but in combination their effects may be substantial, and may provide the basis of targeted prevention (1). However, because the risks associated with these variants are modest, large studies are required to estimate their risks precisely. To facilitate this estimation, we genotyped 25 prostate cancer susceptibility SNPs in studies from the PRACTICAL consortium. PRACTICAL is an international prostate cancer consortium that includes more than 78 studies, including men of European, Asian, or African ancestry, and has a combined dataset of over 130,000 samples (http://practical.ccge.medschl.cam.ac.uk/). In the current analysis, we used data from 31,833 cases and controls from 24 studies in PRACTICAL and 8,581 samples from replication stage of a GWAS ("GWAS stage III"). Sixteen out of the 25 SNPs that we used in this study were identified through studies that included PRACTICAL (2–4) and nine SNPs were identified by other GWAS (5–10).

Materials and Methods

Samples

The current analysis was restricted to individuals of European ancestry, based on self-reported ethnicity, and thus we excluded samples with non-European ancestry.

Data were contributed from 25 studies in PRACTICAL and GWAS stage III. Twenty-five SNPs were genotyped specifically for this analysis in 31,833 cases and controls in PRACTICAL phase III, unless the genotype data were already available. We also included four studies from the GWAS stage III conducted in the United Kingdom and Australia, comprising a further 8,581 cases and controls (11). In this replication stage, 1,536 SNPs were genotyped, including the 25 susceptibility SNPs analyzed here. These two datasets were combined to give a total of 40,414 samples (20,288 cases and 20,126 controls). Three studies (MCCS, PFCS, and UKGPCS) that were included in the GWAS stage III also contributed genotyping of additional samples for PRACTICAL phase III (Table 1; Supplementary Table S1 and Supplementary Data). Studies provided a minimum core dataset that included disease status, age at diagnosis/observation, and ethnicity. Twenty-two studies provided data on family history and 18 studies provided data on Gleason score.

Where studied included more than one individual from the same family, only the index case was included, so that the analyses were based on unrelated men. For analyses of the polygenic risk score (PRS), we also excluded five studies (MAYO, PCFS, TASPRAC, ULM, and UTAH) that oversampled cases with family history of prostate cancer. This reduced the total number of samples to 34,986 (16,643 cases and 18,343 controls). All studies were approved by the relevant ethics committees.

Department of Applied Health Research, London, United Kingdom. [15]National Human Genome Research Institute, NIH, Bethesda, Maryland. [16]Division of Public Health Sciences, Fred Hutchinson Cancer Research Centre, Seattle, Washington. [17]Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington. [18]Australian Prostate Cancer Research Centre-Qld, Institute of Health and Biomedical Innovation and School of Biomedical Science, Queensland University of Technology, Brisbane, Queensland, Australia. [19]Griffith Health Institute, Griffith University, Gold Coast, Queensland, Australia. [20]Cancer Council Queensland, Brisbane, Queensland, Australia. [21]Prostate Cancer Foundation of Australia, Sydney, Australia. [22]Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark. [23]Department of Molecular Medicine (MOMA), Aarhus University Hospital, Aarhus N, Denmark. [24]Division of Cancer Prevention and Control, H. Lee Moffitt Cancer Centre, Tampa, Florida. [25]International Hereditary Cancer Centre, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland. [26]Department of Urology, University Hospital Ulm, Ulm, Germany. [27]Hannover Biomedical Research School, Hannover, Germany. [28]University of Tasmania, Menzies Research Institute Tasmania, Hobart, Tasmania, Australia. [29]Division of Genetic Epidemiology, Department of Medicine, University of Utah School of Medicine, Salt Lake City, Utah. [30]George E. Wahlen Department of Veterans Affairs Medical Centre, Salt Lake City, Utah. [31]Division of Clinical Epidemiology and Ageing Research, German Cancer Research Centre (DKFZ), Heidelberg, Germany. [32]German Cancer Consortium (DKTK), Heidelberg, Germany. [33]University of Pennsylvania, Philadelphia, Pennsylvania. [34]Division of Population Sciences, Department of Medical Oncology, Thomas Jefferson University, Philadelphia, Pennsylvania. [35]Department of Urology, Akita University School of Medicine, Akita, Japan. [36]Mayo Clinic, Rochester, Minnesota. [37]Division of Hematology/Oncology, University of Michigan Medical School, Ann Arbor, Michigan. [38]Divisions of Oncology and Genetic Medicine, Geneva University Hospitals, Geneva, Switzerland. [39]Hopital Cantonal Universitaire de Geneve, Geneva, Switzerland. [40]Department of Medical Chemistry and Biochemistry, Molecular Medicine Centre, Medical University–Sofia, Sofia, Bulgaria. [41]McGill University, Montreal, Quebec, Canada. [42]Department of Complex Genetics, Cluster of Genetics and Cell Biology, NUTRIM School for Nutrition, Toxicology, and Metabolism, Maastricht University Medical Centre+, Maastricht, the Netherlands. [43]Barts Cancer Institute, Queen Mary University of London, John Vane Science Centre, London, United Kingdom. [44]Second Military Medical University, Shanghai, P.R. China. [45]Division of Urology, Department of Surgery, University of Utah School of Medicine, Salt Lake City, Utah. [46]CR-UK/YCR Sheffield Cancer Research Centre, University of Sheffield, Sheffield, United Kingdom. [47]Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Parkville, Victoria, Australia. [48]Molecular Cancer Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Queensland, Australia. [49]Cancer Council Victoria, Cancer Epidemiology Centre, Melbourne, Victoria, Australia. [50]The Institute of Cancer Research, London, United Kingdom. [51]Department of Urology, Technical University Munich, Munich, Germany. [52]Royal Marsden NHS Foundation Trust, London and Surrey, United Kingdom. [53]UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons' Section of Oncology (Membership List Provided in Supplementary Data). [54]UK ProtecT Study Collaborators (Membership List Provided in Supplementary Data). [55]Full List of Participants of The PRACTICAL Consortium Is Provided in Supplementary Data.

Table 1. Total numbers of cases and controls used in the analyses

Study	Controls[a]	Cases[a]	Total[a]
GWAS stage III	4,076	4,505	8,581
PRACTICAL	16,050	15,783	31,833
Total	20,126	20,288	40,414
Total[b]	18,343	16,643	34,986

[a]Analyses were restricted to men of European ancestry (see the text).
[b]Total after excluding 5 studies that oversampled cases with family history.

Eighty-nine percent (31,150) of the samples had information on age at diagnosis (interview/blood draw for controls). The mean age at diagnosis for the cases was 64 years, slightly higher than the mean age at interview/blood draw for the controls (58 years; Supplementary Table S2A). Family history information was available for 21,209 (60.6%) samples and among samples with family history information, 10.7% of controls and 18.2% of cases had a family history of prostate cancer. Before excluding studies with oversampled familial cases, these percentages were 12.9% and 22.6%, respectively (Supplementary Table S2A and S2B).

Genotyping

Genotyping was performed in two experiments; these were subject to separate quality control procedures appropriate to the platforms used, before the data were combined for statistical analysis. In PRACTICAL phase III, genotyping of samples from two studies was performed by Sequenom, while 22 study sites performed the 5′-exonuclease assay (TaqMan) using the ABI Prism 7900HT sequence detection system according to the manufacturer's instructions. Primers and probes were supplied directly by Applied Biosystems as Assays-By-Design. Assays at all sites included at least four negative controls and 2% to 5% duplicates on each 384-well plate. Quality control guidelines were followed by all the participating groups as previously described (4). In addition, all sites also genotyped 16 CEPH samples. We excluded individuals that were not typed for at least 80% of the SNPs attempted. Data on a given SNP for a given site were also excluded if they failed any of the following quality control criteria: SNP call rate >95%, no deviation from the Hardy–Weinberg equilibrium in controls at $P < 0.00001$; <2% discordance between genotypes in duplicate samples and in the CEPH control samples. Cluster plots for SNPs that were close to failing any of the quality control criteria were re-examined centrally.

GWAS Stage III genotypes were generated using an Illumina Golden Gate Assay. All SNPs for this analysis passed the quality control filters used for this experiment: call rate >95%, a minor allele frequency in controls of >1%, or genotype frequency in controls consistent with the Hardy–Weinberg equilibrium at $P < 0.00001$. Duplicate concordance was 99.99% (11).

Statistical methods

We used combined data across all studies for the analysis. We assessed the association between each SNP and prostate cancer using a 1-degree-of-freedom Cochran–Armitage trend test, stratified by studies. Odds ratios (OR) and 95% confidence intervals (95% CI) associated with each genotype and cancer risk, and genotypes for pairs of SNPs, were estimated using unconditional logistic regression, stratified by study as a covariate. Both per-allele ORs, and genotype-specific ORs, were estimated. Heterogeneity in the OR estimates among studies was evaluated using a likelihood ratio test, by comparing with a model in which separate ORs were estimated for each study.

Modification of the ORs by disease aggressiveness and family history was assessed by using both family history (Yes vs. No) and Gleason score (<8 vs. ≥8) as binary variables. A test for association between SNP genotype at a locus and Gleason score as an ordinal variable was also performed, using polytomous regression. Modification of the ORs by age was assessed using a case-only analysis, assessing the association between age and SNP genotype in the cases using polytomous regression. The associations between SNP genotypes and PSA level were assessed using linear regression, after log-transformation of PSA level to correct for skewness.

Contribution to familial risk

The contribution of the known SNPs to the familial risk of prostate cancer, under a multiplicative model, was computed using the formula:

$$\frac{\sum_k (\log \lambda_k)}{(\log \lambda_0)}$$

where λ_0 is the observed familial risk to first-degree relatives of prostate cancer cases, assumed to be 2 (12), and λ_k is the familial relative risk due to locus k, given by:

$$\lambda_k = \frac{p_k r_k^2 + q_k}{(p_k r_k + q_k)^2}$$

where p_k is the frequency of the risk allele for locus k, $q_k = 1 - p_k$ and r_k is the estimated per-allele OR (13).

To evaluate evidence for interactions between pairs of SNPs, we used a likelihood ratio test and evaluated the evidence for departures from a multiplicative model, by comparing models with and a model without the interaction term for each pair of SNPs. The interaction term was the product of the allele doses for the two SNPs, hence leading to a 1-degree-of-freedom test for an interaction. On the basis of the assumption of a log-additive model, we constructed a PRS from the summed genotypes weighted by the estimated per-allele log-OR for each SNP, as estimated by logistic regression as above. Thus, for each individual j we derived:

$$Score_j = \sum_{i=1}^{N} \beta_i g_{ij}$$

Where:
N: number of SNPs (25)
g_{ij}: allele dose at SNP i (0, 1, 2) for individual j
β_i: per-allele log-OR of SNP i

The missing genotypes for an individual were replaced with the mean genotype of each SNP separately for cases and controls. A sensitivity analysis, in which analyses were based on samples with complete genotype data, gave very similar results (data not shown). We then standardized the PRS by dividing by the overall standard deviation of PRS in the controls.

The risk of prostate cancer was estimated for the percentiles of the distribution of the PRS; <1%, 1%–10%, 10%–25%, 25%–75% (defined here as "median risk"), 75%–90%, 90%–99%, >99%; and per standard deviation when fitted as a continuous covariate. We evaluated the fit of the combined risk score to a log-linear model by comparing the model with the PRS fit as a continuous covariate with a model in which separate parameters

were estimated for percentiles of risk adjusted for age at diagnosis and family history, using a likelihood ratio test.

We used a likelihood ratio test to evaluate the evidence for interaction between PRS and age at diagnosis/observation, PRS and family history, and also family history and age at diagnosis/observation by comparing models with and a model without an interaction term. Effect sizes by family history were compared using a case-only analysis. Analyses were performed using Stata 13.

The relative risk estimates were used to obtain estimates of the absolute risk of prostate cancer by PRS category and family history. Because we observed evidence for an interaction between PRS and age, we used both models with and without PRS × age interaction term. Absolute risks were constrained such that the age-specific incidences, averaged over all categories of PRS and family history, were consistent with the age-specific incidences of prostate cancer for the U.K. population for 2012 (http://ci5.iarc. fr/CI5plus; ref. 14). The model was adjusted for age at diagnosis (age <55, 55–59, 60–64, 65–70, and 70+ years). The procedure for deriving the age-specific incidences for each SNP profile category has been performed following the procedure explained by Antoniou and colleagues (15, 16), but adjusted to allow for competing causes of death.

For this purpose, we categorized PRS into seven risk groups (k = risk group 1 to 7), based on the percentile in the controls: <1%, 1%–10%, 10%–25%, 25%–75%, 75%–90%, 90%–99%, and >99%. We could not find any evidence for an interaction between PRS and family history of prostate cancer ($P = 0.49$) and assumed that family history and PRS are independently predictive of prostate cancer risk. Under this model, the prostate cancer incidence $\lambda_k^h(t)$ at age t for an individual in risk group k and family history group h (h = 1 with family history, h = 0 no family history) was assumed to follow a model of the form: $\lambda_k^h(t) = \lambda_0(t)\exp(\beta_k^h)$ where $\lambda_0(t)$ is the baseline prostate cancer incidence and $\exp(\beta_k^h)$ is the risk ratio in the risk group k and family history group h, relative to the baseline category (h = 0, k = 1), approximated by the OR estimates from the logistic regression analysis. To obtain the baseline incidence, $\lambda_0(t)$, we constrained the prostate cancer incidence averaged all risk groups to agree with the population age-specific prostate cancer incidences $\mu(t)$ (the incidence of prostate cancer at age t per 100,000 individuals in the United Kingdom; ref. 14). The baseline incidence can be obtained for each age by:

$$\lambda_0(t) = \mu(t)\frac{p_0\sum_1^k f_k S_k^0(t-1) + p_1\sum_1^k f_k S_k^1(t-1)}{p_0\sum_1^k f_k S_k^0(t-1)\exp(\beta_k^0) + p_1\sum_1^k f_k S_k^1(t-1)\exp(\beta_k^1)}$$

Here, p_0 is the probability of having no family history in the population (89.26% in the controls in this dataset) and $p_1 = 1 - p_0$ is the probability of having family history in the population (10.74% in the controls in this dataset). f_k is frequency of the SNP profile risk group k ($f_1 = 0.01$, $f_2 = 0.09$, $f_3 = 0.15$, $f_4 = 0.5$, $f_5 = 0.15$, $f_6 = 0.09$, and $f_7 = 0.01$) and $S_k^h(t)$ is the probability of surviving prostate cancer by age (t) in the risk group k for samples in the family history group h, which can be derived from incidence rates $\lambda_k^h(t)$ for ages $< t$ using the formula $S_k^h(t) = \exp(-\sum_0^t \lambda_k^h(t-1))$. Because definition $S_k^h(0) = 1$ for all k and h, it was possible to solve the above equation recursively, starting at age $t = 0$, to obtain the baseline incidences $\lambda_0(t)$ and hence the age-specific prostate cancer incidences at age (t), $\lambda_k^h(t)$,

for each group. We then computed the absolute risk by age t, adjusting for mortality from other causes, for each risk group, using the formula: $\sum_0^t S_k^h(t) \times \lambda_k^h(t) \times S_c(t)$

Where $S_c(t) = \exp(-\sum_0^t \mu_c(t-1))$ is the probability of not dying from another cause of death by age t, based on the age-specific mortality rates $\mu_c(t)$. The age-specific mortality rates, $\mu_c(t)$, was estimated by using all causes incidences of death per 100,000 individuals for England and Wales (http://www.ons. gov.uk/ons/index.html) and the prostate cancer death incidence per 100,000 individuals in the United Kingdom in year 2012 (14).

Results

All 25 SNPs showed evidence of association with prostate cancer ($P = 0.02$ to $P = 1.4 \times 10^{-46}$), with effect sizes that were consistent with previous reports. The largest per-allele OR estimate was 1.56 (95% CI, 1.44–1.68) for rs16901979 on 8q24 (Table 2). For each of the 24 autosomal SNPs, the effect size was larger for rare homozygotes than for heterozygotes, and the estimates were consistent with a multiplicative (log-additive) model. There was no evidence for heterogeneity among studies (Table 2).

Gleason score was available for 15,107 (74.5%) of the cases used in the analyses; of these, 2,139 had a score of 8+ and 12,968 had a score less than 8. One SNP, rs1447295, on chromosome 8, showed a larger effect size with increasing grade ($P = 0.001$), while four SNPs (rs17021918, rs1512268, rs7127900, and rs2735839) showed a larger effect sizes with decreasing grade ($P < 0.02$; Supplementary Table S3).

Thirteen of the SNPs (rs1465618, rs7679673, rs10486567, rs1447295, rs6983267, rs16901979, rs10993994, rs7931342, rs7127900, rs4430796, rs11649743, rs1859962, and rs5759167) showed a higher per-allele OR for cases with a prostate cancer family history than those without ($P < 0.05$), whereas no SNPs showed an effect in the opposite direction consistent with the predictions under a polygenic model (Supplementary Table S3; ref. 17).

Data on serum PSA level were available for 3,922 controls from six studies. Six SNPs (rs1447295, rs6983267, rs1512268, rs10993994, rs7127900, and rs2735839) showed association with PSA concentration levels significant at $P < 0.03$. rs1447295 showed an association with PSA in the opposite direction of the prostate cancer risk association but the rest of five SNPs showed an association with PSA in the same direction of the prostate cancer risk association (Supplementary Table S4).

Seven SNPs (rs1465618, rs12621278, rs10993994, rs7127900, rs1859962, rs2735839, and rs5945619) showed an evidence for a trend in the per-allele ORs with age; in each case, the effect size was larger for cases diagnosed at younger ages (Supplementary Table S5).

The combined effect of all pairs of SNPs was evaluated through a logistic regression model that included each pair of SNPs and an interaction term. The interaction term was significant at $P < 0.05$ level for 29 pairs (out of 300 possible pairs) compared with 15 expected by chance, and significant at the $P < 0.01$ level for 12 pairs compared with three expected by chance. However, no pair was significant at the $P < 0.05$ level after a Bonferroni correction for the number of tests (nominal significance $P = 1.6 \times 10^{-4}$; Supplementary Table S6).

Table 2. Summary results of 25 SNPs using PRACTICAL and GWAS stage III datasets in European

Marker[a] Chr/nearby gene	Alleles[b] Position[c]	MAF[d]	Per allele[e] OR (95%CI)	Het OR[e,f] (95%CI)	Hom OR[e,g] (95%CI)	P[h]	P[i]
rs721048 2/EHBP1	C/T 63131731	0.18	1.11 (1.07–1.16)	1.09 (1.04–1.15)	1.32 (1.17–1.48)	9.8×10^{-8}	0.13
rs1465618 2/THADA	G/A 43553949	0.2150	1.07 (1.03–1.11)	1.08 (1.03–1.13)	1.14 (1.03–1.26)	1.9×10^{-4}	0.39
rs12621278 2/ITGA6	A/G 173311553	0.06	0.75 (0.70–0.80)	0.76 (0.71–0.82)	0.38 (0.24–0.58)	4.9×10^{-17}	0.57
rs2660753 3/Unknown	G/A 87110674	0.10	1.12 (1.06–1.18)	1.12 (1.06–1.19)	1.32 (1.09–1.61)	1.2×10^{-5}	0.73
rs17021918 4/PDLIM5	G/A 95562877	0.35	0.88 (0.85–0.91)	0.86 (0.83–0.90)	0.80 (0.74–0.85)	6.7×10^{-15}	0.39
rs12500426 4/PDLIM5	G/T 95514609	0.46	1.10 (1.06–1.13)	1.11 (1.06–1.18)	1.20 (1.12–1.28)	4.8×10^{-8}	0.54
rs7679673 4/TET2	C/A 106061534	0.40	0.88 (0.85–0.90)	0.87 (0.83–0.91)	0.77 (0.72–0.82)	1.0×10^{-16}	0.08
rs9364554 6/SLC22A3	C/T 160833664	0.29	1.10 (1.06–1.14)	1.12 (1.07–1.18)	1.18 (1.09–1.27)	4.8×10^{-8}	0.85
rs10486567 7/JAZF1	G/A 27976563	0.23	0.85 (0.82–0.89)	0.86 (0.81–0.91)	0.72 (0.63–0.81)	4.5×10^{-12}	0.21
rs6465657 7/LMTK2	A/G 97816327	0.46	1.10 (1.06–1.13)	1.09 (1.04–1.15)	1.21 (1.13–1.28)	3.4×10^{-9}	0.32
rs1447295 8/Unknown	G/T 128485038	0.11	1.41 (1.35–1.48)	1.41 (1.34–1.49)	2.01 (1.69–2.41)	1.4×10^{-46}	0.50
rs6983267 8/Unknown	C/A 128413305	0.49	0.82 (0.79–0.85)	0.80 (0.76–0.85)	0.67 (0.63–0.72)	2.3×10^{-35}	0.61
rs16901979 8/Unknown	G/T 128124916	0.03	1.56 (1.44–1.68)	1.55 (1.43–1.69)	2.39 (1.47–3.86)	3.8×10^{-28}	0.29
rs2928679 8/SLC25A37	C/T 23438975	0.48	1.04 (1.01–1.07)	1.03 (.97–1.09)	1.08 (1.01–1.16)	0.02	0.10
rs1512268 8/NKX3.1	G/A 23526463	0.43	1.13 (1.10–1.17)	1.13 (1.08–1.19)	1.29 (1.21–1.37)	2.6×10^{-16}	0.19
rs4962416 10/CTBP2	A/G 126696872	0.28	1.04 (1.01–1.08)	1.03 (0.98–1.08)	1.11 (1.02–1.21)	0.02	0.68
rs10993994 10/MSMB	G/A 51549496	0.39	1.24 (1.20–1.28)	1.21 (1.15–1.27)	1.56 (1.46–1.66)	7.9×10^{-41}	0.36
rs7931342 11/Unknown	C/A 68994497	0.50	0.84 (0.81–0.86)	0.86 (0.82–0.91)	0.70 (0.65–0.74)	4.8×10^{-27}	0.86
rs7127900 11/Unknown	G/A 2233574	0.19	1.23 (1.18–1.28)	1.24 (1.18–1.30)	1.47 (1.32–1.65)	6.3×10^{-26}	0.63
rs4430796 17/HNF1B	A/G 36098040	0.48	0.81 (0.79–0.84)	0.81 (0.77–0.85)	0.66 (0.62–0.71)	2.7×10^{-38}	0.79
rs11649743 17/HNF1B	G/A 36074979	0.19	0.88 (0.85–0.92)	0.88 (0.83–0.92)	0.79 (0.70–0.90)	5.6×10^{-10}	0.25
rs1859962 17/Unknown	T/G 69108753	0.48	1.17 (1.14–1.21)	1.22 (1.15–1.28)	1.38 (1.30–1.47)	3.7×10^{-24}	0.19
rs2735839 19/KLK2/KLK3	G/A 51364623	0.15	0.81 (0.77–0.85)	0.82 (0.78–0.86)	0.62 (0.53–0.73)	1.1×10^{-19}	0.06
rs5759167 22/BIL/TTLL1	G/T 43500212	0.50	0.84 (0.82–0.87)	0.83 (0.79–0.87)	0.71 (0.67–0.76)	3.4×10^{-28}	0.87
rs5945619 X/NUDT11	T/C 51241672	0.36	1.13 (1.10–1.16)	—	1.28 (1.22–1.35)	1.9×10^{-20}	0.10

[a]dbSNP rs number.
[b]Major/minor allele, based on the frequencies in controls in PRACTICAL III data.
[c]Build 37 position.
[d]MAF in controls in combined European dataset.
[e]OR (minor allele) from a logistic regression using all European samples stratified by studies with no adjustment.
[f]OR in heterozygotes, relative to major allele homozygotes.
[g]OR in minor allele homozygotes, relative to major allele homozygotes.
[h]Cochran–Armitage test for trend.
[i]Heterogeneity P value among studies.

Under the assumption that these 25 SNPs combined approximately multiplicatively to alter the risk of prostate cancer, we constructed a PRS for 16,643 cases and 18,343 controls based on the estimated per-allele ORs of 25 SNPs, standardized by the standard deviation in controls. The standardized PRS had a mean = 0.651 (range, −3.81–5.36; SD = 0.98) in cases and mean = 0.104 (range, −4.05–4.15; SD = 1) in controls. The standardized PRS was strongly associated with disease risk (OR per unit PRS, 1.74; 95% CI, 1.70–1.78). The OR per unit increase of the standardized PRS declined with age from 1.76 (95% CI, 1.62–1.92) in cases diagnosed at age less than 55 years to 1.48 (95% CI, 1.37–1.60) in cases diagnosed at age 70+ ($P = 2.6 \times 10^{-4}$; Supplementary Table S5).

The OR per unit increase of PRS was larger for men with prostate cancer family history (1.79 vs. 1.70; $P = 1.8 \times 10^{-4}$; Supplementary Table S3). We found no evidence of an interaction between PRS and family history ($P = 0.49$) or between age at diagnosis and family history ($P = 0.11$), but there was some evidence for an interaction between PRS and age at diagnosis ($P = 0.003$).

There was no evidence of a difference in the OR per unit PRS according to Gleason score (OR, 1.75, GS<8 vs. OR = 1.65, GS 8+) after adjusting for age at diagnosis and family history ($P = 0.37$; Supplementary Table S3). The correlation between PSA and the PRS was weak, both in controls (correlation = 0.09) and in cases (correlation = 0.02).

When PRS was categorized by percentile, the top 1% of the population had an estimated OR of 30.6 (16.4–57.3) compared with the bottom 1% of the population, and an OR of 4.2 (95% CI, 3.2–5.5) compared with the median population risk (defined as the 25%–75% risk group). The bottom 1% of the population had an estimated OR of 0.14 (95% CI, 0.08–0.24) compared with the median risk (Table 3). After allowing for an interaction between PRS and age, the OR for the top 1% of the population, relative to the median risk group, decreased from 5.6, for men below age <55 years, to 3.8 for men aged 70+ years (Supplementary Tables S7 and S8).There was no difference between fit of the model with a continuous covariate for PRS and the model with separate parameters for percentiles of the PRS ($P = 0.24$). In particular, the predicted ORs for the top 1% and the bottom 1% of the population, based on a log-linear model, did not differ from that observed.

Table 3. ORs for prostate cancer by percentile of the PRS and family history

Percentiles	OR[a,b]	OR[a,c]
PRS group		
<1%	1 (baseline)	0.14 (0.08–0.24)
1%–10%	2.98 (1.66–5.35)	0.41 (0.36–0.47)
10%–25%	4.59 (2.58–8.17)	0.63 (0.57–0.70)
25%–75%	7.23 (4.08–12.80)	1 (baseline)
75%–90%	12.13 (6.83–21.54)	1.68 (1.54–1.83)
90%–99%	16.70 (9.38–29.72)	2.31 (2.09–2.56)
≥99%	30.63 (16.36–57.34)	4.24 (3.24–5.53)
Family history	2.52 (2.29–2.78)	2.52 (2.29–2.78)

[a]ORs obtained by fitting PRS group, family history, and age at diagnosis jointly.
[b]ORs compared with men in the 1st percentile as baseline.
[c]ORs compared with men in the 25th–75th percentile as baseline.

**Absolute risk (with family history of prostate cancer)
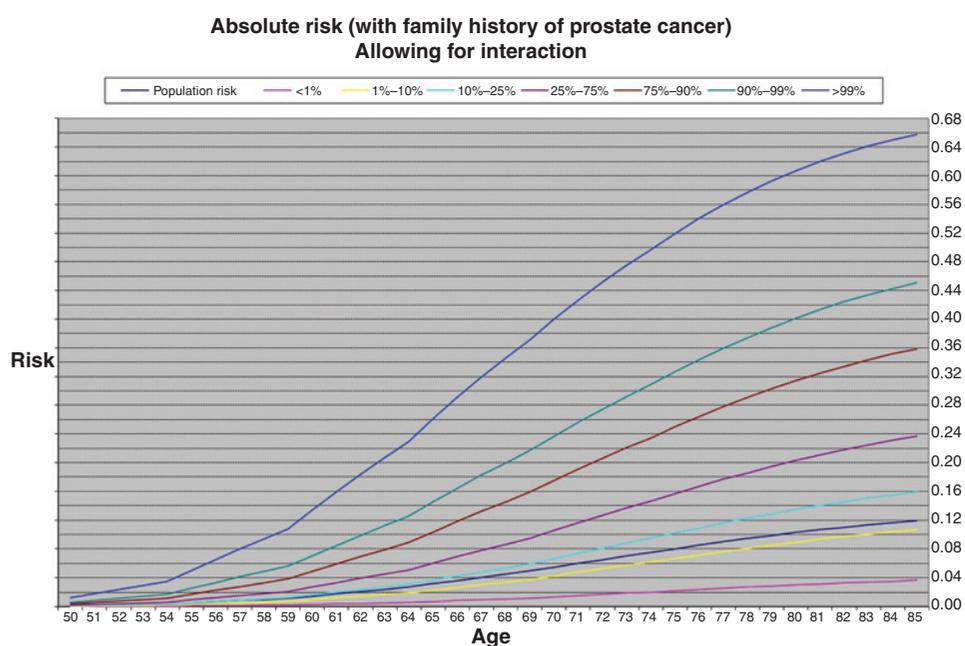Allowing for interaction**

Figure 1.
Absolute risk of prostate cancer by age in men with family history.

To estimate the absolute risk of prostate cancer for different risk groups defined by the combined genotypes at the 25 prostate cancer susceptibility loci, we fitted a logistic regression model that it included parameters for PRS (in seven categories) together with family history of prostate cancer once with (Supplementary Table S7) and once without a PRS × age at diagnosis interaction term (Table 3). We used both models (adjusted for age at diagnosis and family history) in order to estimate effect sizes for PRS. Then, we used the U.K. age-specific incidences of prostate cancer (0 to 85+ years; ref. 14) to estimate age-specific absolute risks of prostate cancer in the general population after considering competing causes of death for 14 risk groups defined by PRS and family history (seven PRS risk groups and two family history; see Materials and Methods). On the basis of this analysis, the absolute risk of prostate cancer by the age of 85 years for a man in the top 1% of the risk distribution with family history of prostate cancer was 65.8% (67.1% in a model not allowing for interaction) and for a man in the lowest 1% was 3.65% (3.67% in a model not allowing for interaction). The absolute risk for a man in the top 1% of the risk distribution with no family history of prostate cancer was 35.0% (36.1% in a model not allowing for interaction) and 1.46% (1.47% in a model not allowing for interaction) for someone in the lowest 1%. In comparison, the estimated absolute risk for a man in the 25% to 75% category was 10.2% in the absence of a family history of prostate cancer, and 23.7% for a man with family history (Figs. 1 and 2; Supplementary Figs. S1 and S2).

Discussion

These results demonstrate that risk profiling based on SNPs can identify men at substantially increased or reduced risk of prostate cancer. We derived a PRS based on a sum of SNP genotypes, weighted by their per-allele log ORs. The estimated ORs for the highest and lowest 1% of the population (4.2 and

0.14, respectively) were consistent with those predicted under a simple polygenic model in which the log OR increases linearly with the PRS. We have also shown that the effect size, measured by OR per unit PRS, was higher at younger ages. As expected, the majority of loci, and the PRS, showed a stronger effect for familial cases. In a logistic regression model, both PRS and family history were independently associated with prostate cancer risk. The OR due to family history was attenuated after adjustment for the PRS (from 2.63 to 2.50), as expected given that family history is, at least in part, a reflection of genetic susceptibility. However, the degree of attenuation (5% on a log-scale) was markedly less than 18%, the estimated contribution of these 25 loci to the familial risk of prostate cancer estimated on the basis of their ORs and allele frequencies in this study (see Materials and Methods). The reason for this difference is unclear but might reflect interactions between the known susceptibility loci summarized in the PRS and other factors influencing family history.

In order to investigate the added value of PRS, once we estimated the absolute risk for individuals with family history without fitting their PRS information and then repeated the same procedure after adding their PRS information. The absolute risk of prostate cancer for a man at the age of 85 years with family history was estimated to be 26.5% when PRS information was ignored. When we incorporated PRS information, a man at the age of 85 years, depending on his PRS risk group, could have an absolute risk ranging from 3.67% (if a man is in the bottom 1% of the risk distribution) to 67.1% (if a man is in the top 1% of the risk distribution; Supplementary Figs. S1 and S3). These observations indicate that family history and the PRS independently influence risk and can be combined to provide stronger discrimination.

Chatterjee and colleagues (18) derived theoretical estimates for the predictive performance of polygenic models for 10 complex traits or common diseases, including prostate cancer, using published estimates for individual SNPs. They estimated that

Absolute risk (with no family history of prostate cancer)
Allowing for interaction

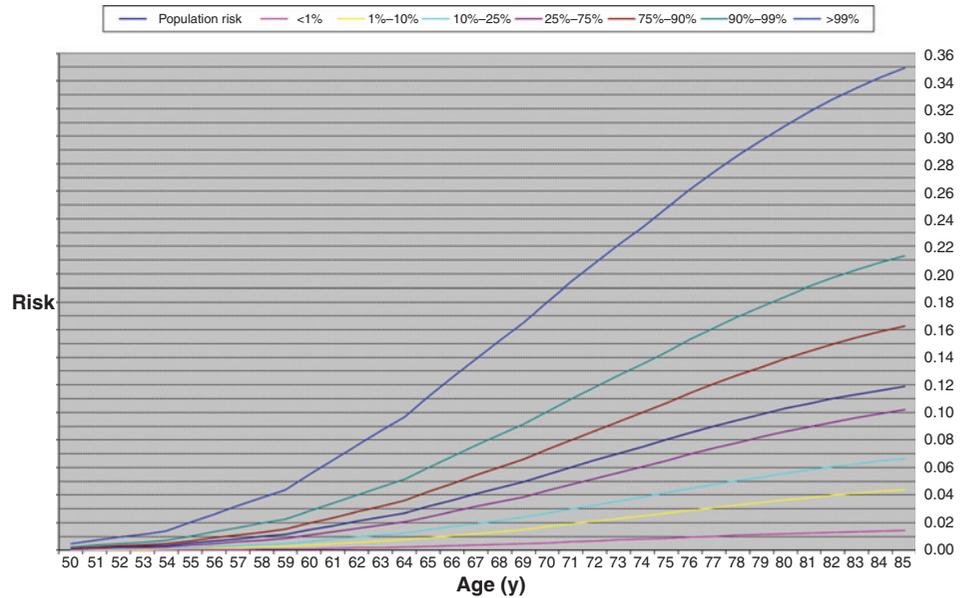Legend: Population risk — <1% — 1%–10% — 10%–25% — 25%–75% — 75%–90% — 90%–99% — >99%

Figure 2.
Absolute risk of prostate cancer by age in men with no family history.

approximately 7% of the population will be at 2-fold risk or greater for prostate cancer. We estimated, empirically, that the (average) risk to men in the 90% to 99% category of the PRS was 2.41-fold, relative to the population median, or approximately 2-fold relative to the population mean. However, this is an average risk over the 90% to 99% category, so that the percentile of the PRS at which the risk exceeds 2-fold will be >90%. On the basis of the estimated log(OR) per standardized PRS, approximately 6% of men will have a risk of greater than 2-fold, very close to the estimate of Chatterjee and colleagues (18).

These results show that genetic risk profiling using SNPs could be useful in defining men at high risk for the disease for targeted prevention and screening programs. The benefits of screening, relative to the costs, will be most favorable among men at higher risk. If, for example, the benefit:cost ratio is favorable for screening men at a greater than 2-fold risk, the PRS provides an effective method for identifying such men.

Although these analyses demonstrate the value of SNPs for risk prediction, a risk model could be improved in various ways. The analyses presented here are based on the 25 loci first identified to be associated with prostate cancer. Recently, however, additional loci have been identified (13, 19) and more than 100 common susceptibility loci are now known. In total, these loci increase the estimated proportion of the familial risk to 33% (19). Incorporating all known loci into a PRS should improve the predictive value of risk profiles.

In addition, the analyses presented here consider family history as a binary (yes/no) covariate. It is known that the risk of prostate cancer is dependent on both the number of affected relatives and their ages. MacInnis and colleagues (12, 20) have shown using segregation analysis that the familial aggregation of prostate cancer can be modeled as the combined effect of a recessive allele and a polygenic component, and that the polygenic component can be further partitioned into a component due to measured SNPs and an unmeasured component. This approach should provide more powerful prediction, particularly in families with

multiple cases of the disease. Finally, it is known that serum or urine PSA level is associated with prostate cancer risk, with the association persisting for several decades. Although some of the risks SNPs are also related to PSA level in the expected direction, the PSA level is only weakly correlated with PRS, indicating that incorporating PSA level and potentially other markers such as MSMB (21) into a risk algorithm should further improve the discrimination (22).

The absence of clear differences in the relative risk associated with SNPs by disease aggressiveness, even in this very large study, is striking. We did not find any convincing evidence for differences in the predictive values of the PRS by disease aggressiveness. The effect size was higher for less aggressive disease, but the difference was still small (1.75 vs. 1.65). This result is in contrast to the clear differences in SNP associations by disease pathology seen in other diseases, for example, in breast and ovarian cancer, and indicates that aggressive and nonaggressive disease, at least as measured by Gleason score, share these genetic risk factors as a common etiology.

Analysis of pairwise combinations of SNPs did not identify any clear examples of departure from a multiplicative model, after adjusting for multiple testing. We did, however, find an excess of interactions at the $P < 0.01$ level over the number that would be expected by chance. This suggests that interactions on this scale likely to exist, but their effect sizes are small and that very large sample sizes, exemplified by this collaborative study, will be required to identify and characterize them. If such interactions could be identified reliably, they may improve the predictive value of the risk profiling, and also provide insights into the biologic interactions between the underlying risk variants.

Disclosure of Potential Conflicts of Interest

Authors' Contributions

Conception and design: A. Amin Al Olama, D.F. Easton, G.G. Giles, G. Severi, F.R. Schumacher, J.Y. Park, H. Brenner, J.L. Hopper, Z. Kote-Jarai, R.A. Eeles
Development of methodology: A. Amin Al Olama, D.F. Easton, B.G. Nordestgaard, J.L. Hopper
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): S. Benlloch, G.G. Giles, G. Severi, D.E. Neal, F.C. Hamdy, J.L. Donovan, K. Muir, J. Schleutker, C.A. Haiman, F.R. Schumacher, N. Pashayan, P.D.P. Pharoah, E.A. Ostrander, J.L. Stanford, J. Batra, J.A. Clements, S.K. Chambers, M. Weischer, B.G. Nordestgaard, S.A. Ingles, K.D. Sorensen, T.F. Orntoft, J.Y. Park, C. Cybulski, C. Maier, J.L. Dickinson, L. Cannon-Albright, H. Brenner, T.R. Rebbeck, C. Zeigler-Johnson, T. Habuchi, S.N. Thibodeau, K.A. Cooney, P.O. Chappuis, P. Hutter, R.P. Kaneva, M.P. Zeegers, Y.-J. Lu, H.-W. Zhang, R. Stephenson, A. Cox, M.C. Southey, A.B. Spurdle, L. FitzGerald, D. Leongamornlert, E. Saunders, M. Guy, S.J. Little, K. Govindasami, R. Wilkinson, K. Herkommer, J.L. Hopper, A. Lophatonanon, A.E. Rinckleb, Z. Kote-Jarai, R.A. Eeles, D.F. Easton
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A. Amin Al Olama, A.C. Antoniou, N. Pashayan, K.D. Sorensen, H. Brenner, R. Stephenson, M.C. Southey, J.L. Hopper, D.F. Easton
Writing, review, and/or revision of the manuscript: A. Amin Al Olama, A.C. Antoniou, G.G. Giles, G. Severi, F.C. Hamdy, J.L. Donovan, K. Muir, J. Schleutker, N. Pashayan, P.D.P. Pharoah, E.A. Ostrander, J.L. Stanford, J. Batra, J.A. Clements, S.K. Chambers, M. Weischer, B.G. Nordestgaard, S.A. Ingles, K.D. Sorensen, T.F. Orntoft, J.Y. Park, C. Cybulski, C. Maier, T. Doerk, J.L. Dickinson, H. Brenner, T.R. Rebbeck, S.N. Thibodeau, P.O. Chappuis, R.P. Kaneva, W.D. Foulkes, M.P. Zeegers, Y.-J. Lu, R. Stephenson, A. Cox, M.C. Southey, A.B. Spurdle, L. FitzGerald, J.L. Hopper, Z. Kote-Jarai, R.A. Eeles, D.F. Easton
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S. Benlloch, G.G. Giles, M. Weischer, T.F. Orntoft, T. Doerk, C. Zeigler-Johnson, T. Habuchi, R.P. Kaneva, W.D. Foulkes, Y.-J. Lu, H.-W. Zhang, M.C. Southey, A.B. Spurdle, L. FitzGerald, D. Leongamornlert, M. Tymrakiewicz, M. Guy, T. Dadaev, S.J. Little, E. Sawyer, R. Wilkinson, R.A. Eeles
Study supervision: A. Amin Al Olama, D.F. Easton, B.E. Henderson, R.P. Kaneva, M.C. Southey, R.A. Eeles
Other (PI of a large study for which DNA samples and data are included as part of the PRACTICAL consortium): J.L. Stanford
Other (genotyping and phenotyping of participants): B.G. Nordestgaard
Other (generated laboratory data): J.L. Dickinson

Acknowledgments

The authors thank all the patients and control men who took part in this study. The authors also thank all the members from studies that they partic-

ipated in this phase of PRACTICAL consortium: Esther John, Amit Joshi, Ahva Shahabi, Joanne L. Dickinson, James R. Marthick, Mariana C. Stern, Roman Corral, David M.A. Wallace, Alan Doherty, R.I. Bhatt, K. Subramonian, John Arrand, Louise Flanagan, Sita Ann Bradley, The UK Genetic Prostate Cancer Study Collaborators (www.icr.ac.uk/ukgpcs); Prasad Bollina, Sue Bonnington, Lynne Bradshaw, James Catto, Debbie Cooper, Liz Down, Andrew Doble, Alan Doherty, Garrett Durkan, Emma Elliott, David Gillatt, Pippa Herbert, Peter Holding, Joanne Howson, Mandy Jones, Roger Kockelbergh, Rajeev Kumar, Peter Holding, Howard Kynaston, Athene Lane, Teresa Lennon, Norma Lyons, Hing Leung, Malcolm Mason, Hilary Moody, Philip Powell, Alan Paul, Stephen Prescott, Derek Rosario, Patricia O'Sullivan, Pauline Thompson, Lynne Bradshaw, Sarah Tidball, Paul M. Brown, Anne George, Gemma Marsden, Athene Lane, Michael Davis, Stephen Edwards, Cyril Fisher, Charles Jameson, Elizabeth Page, John Pedersen, Joanne Aitken, Robert A. Gardiner, Srilakshmi Srinivasan, Felicity Lose, Mary-Anne Kedda, Kimberly Alexander, Tracy O'Mara, Gail Risbridger, Wayne Tilley, Lisa Horvarth, Peter Heathcote, Glenn Wood, Greg Malone, Hema Samaratunga, Pamela Saunders, Allison Eckert, Trina Yeadon, Kris Kerr, Angus Collins, Megan Turner, Simon J. Foote, James R. Marthick, Andrea Polanowski, Rebekah M. McWhirter, Terrence Dwyer, Christopher L. Blizzard, Elenko Popov, Darina Kachakova, Atanaska Mitkova, Teodora Goranova, Gergana Stancheva, Olga Beltcheva, Rumyana Dodova, Aleksandrina Vlahova, Tihomir Dikov, Svetlana Christova, Michael Borre, Peter Klarskov, Sune F. Nielsen, Peter Iversen, Andreas Røder, Stig E. Bojesen, Aida Karina Dieffenbach, Manuel Luedeke, Mark Schrader, Josef Hoegel, Walther Vogel, Liisa Määttänen, Teuvo Tammela, Anssi Auvinen, Lori Tillmans, Shaun Riska, Liang Wang, Dan Stram, Kolonel Laurence N., Julio Pow-Sang, Hyun Y. Park, Selina Radlein, Maria Rincon, and Babu Zachariah (Supplementary Data).

The Editor-in-Chief of *Cancer Epidemiology, Biomarkers & Prevention* is an author of this article. In keeping with the AACR's Editorial Policy, the paper was peer reviewed and an AACR Journals' Editor not affiliated with *Cancer Epidemiology, Biomarkers & Prevention* rendered the decision concerning acceptability.

References

1. 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;104:1656–63.
2. Amin Al Olama A, Kote-Jarai Z, Giles GG, Guy M, Morrison J, Severi G, et al. Multiple loci on 8q24 associated with prostate cancer susceptibility. Nat Genet 2009;41:1058–60.
3. Eeles RA, Kote-Jarai Z, Amin Al Olama A, Giles GG, Guy M, Severi G, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet 2009;41:1116–21.
4. Eeles RA, Kote-Jarai Z, Giles GG, Amin Al Olama A, Guy M, Jugurnauth SK, et al. Multiple newly identified loci associated with prostate cancer susceptibility. Nat Genet 2008;40:316–21.
5. Amundadottir LT, Sulem P, Gudmundsson J, Helgason A, Baker A, Agnarsson BA, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet 2006;38:652–8.
6. Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet 2007;39:631–7.
7. Gudmundsson J, Sulem P, Rafnar T, Bergthorsson JT, Manolescu A, Gudbjartsson D, et al. Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. Nat Genet 2008;40:281–3.
8. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, et al. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet 2007;39:977–83.
9. Haiman CA, Patterson N, Freedman ML, Myers SR, Pike MC, Waliszewska A, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet 2007;39:638–44.
10. Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N, et al. Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet 2008;40:310–5.
11. Kote-Jarai Z, Amin Al Olama A, Giles GG, Severi G, Schleutker J, Weischer M, et al. Seven prostate cancer susceptibility loci identified by a multi-stage genome-wide association study. Nat Genet 2011;43:785–91.
12. MacInnis RJ, Antoniou AC, Eeles RA, Severi G, Guy M, McGuffog L, et al. Prostate cancer segregation analyses using 4390 families from UK and Australian population-based studies. Genet Epidemiol 2010;34:42–50.
13. Eeles RA, Amin Al Olama A, 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;45:385–2.

14. Ferlay J, Soerjomataram II, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–86.

15. Antoniou AC, Beesley J, McGuffog L, Sinilnikova OM, Healey S, Neuhausen SL, et al. Common breast cancer susceptibility alleles and the risk of breast cancer for BRCA1 and BRCA2 mutation carriers: implications for risk prediction. Cancer Res 2010;70:9742–54.

16. Antoniou AC, Pharoah PD, McMullan G, Day NE, Ponder BA, Easton D. Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. Genet Epidemiol 2001;21:1–18.

17. Antoniou AC, Easton DF. Polygenic inheritance of breast cancer: implications for design of association studies. Genet Epidemiol 2003;25:190–202.

18. Chatterjee N, Wheeler B, Sampson J, Hartge P, Chanock SJ, Park JH. Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. Nat Genet 2013;45:400–5.

19. Amin Al Olama A, Kote-Jarai Z, Berndt SI, Conti DV, Schumacher F, Han Y, et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nat Genet 2014;46:1103–9.

20. MacInnis RJ, Antoniou AC, Eeles RA, Severi G, Amin Al Olama A, McGuffog L, et al. A risk prediction algorithm based on family history and common genetic variants: application to prostate cancer with potential clinical impact. Genet Epidemiol 2011;35:549–56.

21. Whitaker HC, Kote-Jarai Z, Ross-Adams H, Warren AY, Burge J, George A, et al. The rs10993994 risk allele for prostate cancer results in clinically relevant changes in microseminoprotein-beta expression in tissue and urine. PLoS ONE 2010;5:e13363.

22. Aly M, Wiklund F, Xu J, Isaacs WB, Eklund M, D'Amato M, et al. Polygenic risk score improves prostate cancer risk prediction: results from the Stockholm-1 cohort study. Eur Urol 2011;60:21–8.