

7-1-2015

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

Ali Amin Al Olama  
*University of Cambridge*

See next page for additional works at: <https://jdc.jefferson.edu/medoncfp>

 *University of Cambridge*  
Part of the [Oncology Commons](#)

[Let us know how access to this document benefits you](#)  
*University of Cambridge*

### Recommended Citation

Al Olama, Ali Amin; Benlloch, Sara; Antoniou, Antonis C; Giles, Graham G; Severi, Gianluca; Neal, David E; Hamdy, Freddie C; Donovan, Jenny L; Muir, Kenneth; Schleutker, Johanna; Henderson, Brian E; Haiman, Christopher; Schumacher, Fredrick R; Pashayan, Nora; Pharoah, Paul D P; Ostrander, Elaine A; Stanford, Janet L; Batra, Jyotsna; Clements, Judith A; Chambers, Suzanne K; Weischer, Maren; Nordestgaard, Børge G; Ingles, Sue A; Sorensen, Karina D; Orntoft, Torben F; Park, Jong Y; Cybulski, Cezary; Maier, Christiane; Doerk, Thilo; Dickinson, Joanne L; Cannon-Albright, Lisa; Brenner, Hermann; Rebbeck, Timothy R; Zeigler-Johnson, Charmita; Habuchi, Tomohiro; Thibodeau, Stephen N; Cooney, Kathleen A; Chappuis, Pierre O; Hutter, Pierre; Kaneva, Radka P; Foulkes, William D; Zeegers, Maurice P; Lu, Yong-Jie; Zhang, Hong-Wei; Stephenson, Robert; Cox, Angela; Southey, Melissa C; Spurdle, Amanda B; FitzGerald, Liesel; Leongamornlert, Daniel; Saunders, Edward; Tymrakiewicz, Malgorzata; Guy, Michelle; Dadaev, Tokhir; Little, Sarah J; Govindasami, Koveela; Sawyer, Emma; Wilkinson, Rosemary; Herkommer, Kathleen; Hopper, John L; Lophatonanon, Aritaya; Rinckleb, Antje E; Kote-Jarai, Zsofia; Eeles, Rosalind A; and Easton, Douglas F, "Risk Analysis of Prostate Cancer in PRACTICAL, a Multinational Consortium, Using 25 Known Prostate Cancer Susceptibility Loci." (2015). *Department of Medical Oncology Faculty Papers*. Paper 52.  
<https://jdc.jefferson.edu/medoncfp/52>

This Article is brought to you for free and open access by the Jefferson Digital Commons. The Jefferson Digital Commons is a service of Thomas Jefferson University's [Center for Teaching and Learning \(CTL\)](#). The Commons is a showcase for Jefferson books and journals, peer-reviewed scholarly publications, unique historical collections from the University archives, and teaching tools. The Jefferson Digital Commons allows researchers and interested readers anywhere in the world to learn about and keep up to date with Jefferson scholarship. This article has been accepted for inclusion in Department of Medical Oncology Faculty Papers by an authorized administrator of the Jefferson Digital Commons. For more information, please contact: [JeffersonDigitalCommons@jefferson.edu](mailto:JeffersonDigitalCommons@jefferson.edu).

---

## Authors

Ali Amin Al Olama, Sara Benlloch, Antonis C Antoniou, Graham G Giles, Gianluca Severi, David E Neal, Freddie C Hamdy, Jenny L Donovan, Kenneth Muir, Johanna Schleutker, Brian E Henderson, Christopher Haiman, Fredrick R Schumacher, Nora Pashayan, Paul D P Pharoah, Elaine A Ostrander, Janet L Stanford, Jyotsna Batra, Judith A Clements, Suzanne K Chambers, Maren Weischer, Børge G Nordestgaard, Sue A Ingles, Karina D Sorensen, Torben F Orntoft, Jong Y Park, Cezary Cybulski, Christiane Maier, Thilo Doerk, Joanne L Dickinson, Lisa Cannon-Albright, Hermann Brenner, Timothy R Rebbeck, Charnita Zeigler-Johnson, Tomonori Habuchi, Stephen N Thibodeau, Kathleen A Cooney, Pierre O Chappuis, Pierre Hutter, Radka P Kaneva, William D Foulkes, Maurice P Zeegers, Yong-Jie Lu, Hong-Wei Zhang, Robert Stephenson, Angela Cox, Melissa C Southey, Amanda B Spurdle, Liesel FitzGerald, Daniel Leongamornlert, Edward Saunders, Malgorzata Tymrakiewicz, Michelle Guy, Tokhir Dadaev, Sarah J Little, Koveela Govindasami, Emma Sawyer, Rosemary Wilkinson, Kathleen Herkommer, John L Hopper, Aritaya Lophatonanon, Antje E Rinckleb, Zsofia Kote-Jarai, Rosalind A Eeles, and Douglas F Easton



Published in final edited form as:

*Cancer Epidemiol Biomarkers Prev.* 2015 July ; 24(7): 1121–1129. doi:10.1158/1055-9965.EPI-14-0317.

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

*A full list of authors and affiliations appears at the end of the article.*

### Abstract

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

**Methods**—We genotyped 25 PrCa susceptibility loci in 40,414 individuals and derived a polygenic risk score (PRS). We estimated empirical Odds Ratios for PrCa associated with different risk strata defined by PRS and derived age-specific absolute risks of developing PrCa by PRS stratum and family history.

**Results**—The PrCa 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 PrCa by age 85 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 PrCa. The effect size, measured by OR per unit PRS, was higher in men at younger ages and in men with family history of PrCa. 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.

---

**Corresponding authors:** Prof Douglas F. Easton, Centre for Cancer Genetic Epidemiology, Strangeways Laboratory, Worts Causeway, Cambridge, UK, dfe20@medschl.cam.ac.uk, Tel: 0044 1223 748629, Fax:0044 1223 748628, Dr Ali Amin Al Olama, Centre for Cancer Genetic Epidemiology, Strangeways Laboratory, Worts Causeway, Cambridge, UK, aa461@medschl.cam.ac.uk, Tel: 0044 1223 748638, Fax:0044 1223 748628.

<sup>52</sup>UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons' Section of Oncology, membership list provided in Supplementary Note

<sup>53</sup>UK ProtecT Study Collaborators, membership list provided in Supplementary Note

<sup>54</sup>Full list of participants of The PRACTICAL Consortium is provided in the Supplementary Note

<sup>55</sup>These authors made a co-equal contribution in their respective positions in this paper

<sup>56</sup>These authors jointly directed this work

\*Current address: Gianluca Severi, Human Genetics Foundation, Via Nizza 52, 10126 Torino - Italy

**Conflict of Interest:** There is no conflict of interest.

## Keywords

Prostate Cancer risk; Genetic and Molecular Epidemiology; Genitourinary Cancers: Prostate; Risk Assessment; Methodology; Modelling and biostatistics; Methodology for SNP data analysis; Statistical methods in Genetics

---

## Introduction

Genome-wide association studies (GWAS) have identified multiple common genetic variants associated with prostate cancer (PrCa) 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, since the risks associated with these variants are modest, large studies are required to estimate their risks precisely. To facilitate this estimation, we genotyped 25 PrCa susceptibility SNPs in studies from the PRACTICAL consortium. PRACTICAL is an international PrCa 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 utilised data from 31,833 cases and controls from 24 studies in PRACTICAL and 8,581 samples from replication stage of a GWAS (“GWAS stage 3”). Sixteen out of the twenty five 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 3. 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 3 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 analysed 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 3 also contributed genotyping of additional samples for PRACTICAL phase III (Table 1, Supplementary Table 1 and Supplementary Notes). 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 eighteen 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 5 studies (MAYO, PCFS, TASPRAC, ULM and UTAH) that oversampled cases with family history of PrCa. 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.

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 2a). 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 PrCa. Before excluding studies with oversampled familial cases, these percentages were 12.9% and 22.6% respectively (Supplementary Table 2a and b).

## Genotyping

Genotyping was performed in two experiments; these were subject to separate QC procedures appropriate to the platforms used, before the data were combined for statistical analysis. In PRACTICAL phase III, genotyping of samples from 2 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 2–5% 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 QC criteria: SNP call rate >95%, no deviation from 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 QC criteria were re-examined centrally.

GWAS Stage 3 genotypes were generated using an Illumina Golden Gate Assay. All SNPs for this analysis passed the QC filters used for this experiment: call rate >95%, a minor allele frequency in controls of >1%, or genotype frequency in controls consistent with 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 PrCa 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 PrCa, under a multiplicative model, was computed using the formula:

$$\sum_k (\log \lambda_k) / (\log \lambda_0)$$

where  $\lambda_0$  is the observed familial risk to first degree relatives of PrCa cases, assumed to be 2 (12), and  $\lambda_k$  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 odds ratio (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. Based on the assumption of a log-additive model, we constructed a PRS from the summed genotypes weighted by the estimated per-allele log-odds ratios 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  $j$

$\beta_i$  : Per-allele log-odds ratio 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 standardised the PRS by dividing by the overall standard deviation of PRS in the controls.

The risk of PrCa 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 PrCa by PRS category and family history. Since we observed evidence for an interaction between PRS and age, we used both models with and without PRS  $\times$  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 PrCa for the UK population for 2012 (<http://ci5.iarc.fr/CI5plus>) (14). The model was adjusted for age at diagnosis (age <55, 55–59, 60–64, 65–70 and 70+). The procedure for deriving the age-specific incidences for each SNP profile category has been performed following the procedure explained by Antoniou et al. (15, 16), but adjusted to allow for competing causes of death.

For this purpose, we categorised 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 PrCa ( $P$ -value=0.49) and assumed that family history and PRS are independently predictive of PrCa risk. Under this model, the PrCa 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 PrCa 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 odds ratio estimates from the logistic regression analysis. To obtain the baseline incidence,  $\lambda_0(t)$ , we constrained the PrCa incidence averaged all risk groups to agree with the population age-specific PrCa incidences  $\mu(t)$  (the incidence of PrCa at age  $t$  per 100,000 individuals in the UK (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  $k$  ( $f_1=0.01$ , ( $f_2=0.09$ , ( $f_3=0.15$ , ( $f_4=0.5$ , ( $f_5=0.15$ , ( $f_6=0.09$ , ( $f_7=0.01$ ) and  $S_k^h(t)$  is the probability of surviving PrCa 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

$S_k^h(t) = \exp(-\sum_0^t \lambda_k^h(t-1))$ . Since 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 and hence the age-specific PrCa 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 PrCa death incidence per 100,000 individuals in UK in year 2012 (14).

## Results

All 25 SNPs showed evidence of association with PrCa ( $P=0.02$  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 3).

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 PrCa family history than those



without ( $P < 0.05$ ), while no SNPs showed an effect in the opposite direction consistent with the predictions under a polygenic model (17) (Supplementary Table 3).

Data on serum PSA level were available for 3,922 controls from 6 studies. Six SNPs (rs1447295, rs6983267, rs1512268, rs10993994, rs7127900 and rs2735839) showed association with PSA concentration levels significant at  $P$ -value  $< 0.03$ . rs1447295 showed an association with PSA in the opposite direction of the PrCa risk association but the rest of five SNPs showed an association with PSA in the same direction of the PrCa risk association (Supplementary Table 4).

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 5).

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  $P$ -value  $< 0.05$  level for 29 pairs (out of 300 possible pairs) compared with 15 expected by chance, and significant at the  $P$ -value  $< 0.01$  level for 12 pairs compared with 3 expected by chance. However, no pair was significant at the  $P$ -value  $< 0.05$  level after a bonferroni correction for the number of tests (nominal significance  $P$ -value =  $1.6 \times 10^{-4}$ , Supplementary Table 6).

Under the assumption that these 25 SNPs combined approximately multiplicatively to alter the risk of PrCa, we constructed a PRS for 16,643 cases and 18,343 controls based on the estimated per-allele ORs of 25 SNPs, standardised by the standard deviation in controls. The standardised 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 standardised 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 standardised PRS declined with age from 1.76 (95% CI 1.62–1.92) in cases diagnosed at age less than 55 to 1.48 (95% CI 1.37–1.60) in cases diagnosed at age 70+ ( $P$ -value =  $2.6 \times 10^{-4}$ , Supplementary Table 5).

The OR per unit increase of PRS was larger for men with PrCa family history (1.79 Vs 1.70;  $P$ -value =  $1.8 \times 10^{-4}$ , Supplementary Table 3). We found no evidence of an interaction between PRS and family history ( $P$ -value = 0.49) or between age at diagnosis and family history ( $P$ -value = 0.11) but there was some evidence for an interaction between PRS and age at diagnosis ( $P$ -value = 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 3). The correlation between PSA and the PRS was weak, both in controls (correlation = 0.09) and in cases (correlation = 0.02).

When PRS was categorised 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 Table 7 & 8). 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.

To estimate the absolute risk of PrCa for different risk groups defined by the combined genotypes at the 25 PrCa susceptibility loci, we fitted a logistic regression model that included parameters for PRS (in 7 categories) together with family history of PrCa once with (Supplementary Table 7) and once without a PRS  $\times$  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 UK age-specific incidences of PrCa (0 to 85+ years) (14) to estimate age-specific absolute risks of PrCa in the general population after considering competing causes of death for fourteen risk groups defined by PRS and family history (seven PRS risk groups and two family history, see methods). Based on this analysis, the absolute risk of PrCa by age 85 for a man in the top 1% of the risk distribution with family history of PrCa 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 PrCa 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%. By comparison, the estimated absolute risk for a man in the 25–75% category was 10.2% in the absence of a family history of PrCa, and 23.7% for a man with family history (Figure 1 & 2, Supplementary Figure 1 & 2).

## Discussion

These results demonstrate that risk profiling based on SNPs can identify men at substantially increased or reduced risk of PrCa. 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 also showed 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 PrCa 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 PrCa estimated based on their ORs and allele frequencies in this study (see methods). The reason for this difference is unclear but might reflect interactions between the known susceptibility loci summarised 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 PrCa for a man at age 85 with family history was estimated to be 26.5% when PRS information was ignored. When we incorporated PRS information, a man at age 85, depend 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 Figure 1 and 3). These observations indicate that family history and the PRS independently influence risk and can be combined to provide stronger discrimination.

Chatterjee et al. derived theoretical estimates for the predictive performance of polygenic models for ten complex traits or common diseases, including PrCa, using published estimates for individual SNPs (18). They estimated that ~7% of the population will be at two-fold risk or greater for PrCa. We estimated, empirically, that the (average) risk to men in the 90–99% 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 90–99% category, so that the percentile of the PRS at which the risk exceeds 2 fold will be >90%. Based on the estimated log(OR) per standardised PRS, approximately 6% of men will have a risk of greater than twofold, very close to the estimate of Chatterjee et al (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 favourable among men at higher risk. If, for example, the benefit-cost ratio is favourable for screening men at a greater than two-fold risk, the PRS provides an effective method for identifying such men.

While 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 PrCa. 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.

Additionally, the analyses presented here consider family history as a binary (yes/no) covariate. It is known that the risk of PrCa is dependent on both the number of affected relatives and their ages. MacInnis et al. (12, 20) have shown using segregation analysis that the familial aggregation of PrCa can be modelled 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 PrCa 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 non-aggressive disease, at least as measured by Gleason score, share these genetic risk factors as a common aetiology.

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 characterise them. If such interactions could be identified reliably, they may improve the predictive value of the risk profiling, and also provide insights into the biological interactions between the underlying risk variants.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Authors

Ali Amin Al Olama<sup>1,56</sup>, Sara Benlloch<sup>1,56</sup>, Antonis C. Antoniou<sup>1,56</sup>, Graham G. Giles<sup>2,3,55</sup>, Gianluca Severi<sup>2,55,\*</sup>, David E. Neal<sup>4,5,55</sup>, Freddie C. Hamdy<sup>6,55</sup>, Jenny L. Donovan<sup>7,55</sup>, Kenneth Muir<sup>8,9,55</sup>, Johanna Schleutker<sup>10,11,55</sup>, Brian E. Henderson<sup>12,55</sup>, Christopher Haiman<sup>12,55</sup>, Fredrick R. Schumacher<sup>12,55</sup>, Nora Pashayan<sup>1,13,55</sup>, Paul D.P. Pharoah<sup>1,55</sup>, Elaine A. Ostrander<sup>14,55</sup>, Janet L. Stanford<sup>15,16,55</sup>, Jyotsna Batra<sup>17,55</sup>, Judith A. Clements<sup>17,55</sup>, Suzanne K. Chambers<sup>18,19,20,55</sup>, Maren Weischer<sup>21,55</sup>, Børge G. Nordestgaard<sup>21,55</sup>, Sue A. Ingles<sup>12,55</sup>, Karina D. Sorensen<sup>22,55</sup>, Torben F. Orntoft<sup>22,55</sup>, Jong Y. Park<sup>23,55</sup>, Cezary Cybulski<sup>24,55</sup>, Christiane Maier<sup>25,55</sup>, Thilo Doerk<sup>26,55</sup>, Joanne L. Dickinson<sup>27,55</sup>, Lisa Cannon-Albright<sup>28,29,55</sup>, Hermann Brenner<sup>30,31,55</sup>, Timothy R. Rebbeck<sup>32,55</sup>, Charnita Zeigler-Johnson<sup>33</sup>, Tomonori Habuchi<sup>34,55</sup>, Stephen N. Thibodeau<sup>35,55</sup>, Kathleen A. Cooney<sup>36,55</sup>, Pierre O. Chappuis<sup>37,55</sup>, Pierre Hutter<sup>38,55</sup>, Radka P. Kaneva<sup>39,55</sup>, William D Foulkes<sup>40,55</sup>, Maurice P. Zeegers<sup>41,55</sup>, Yong-Jie Lu<sup>42,55</sup>, Hong-Wei Zhang<sup>43,55</sup>, Robert Stephenson<sup>44</sup>, Angela Cox<sup>45</sup>, Melissa C. Southey<sup>46</sup>, Amanda B. Spurdle<sup>47</sup>, Liesel FitzGerald<sup>48</sup>, Daniel Leongamornlert<sup>49</sup>, Edward Saunders<sup>49</sup>, Malgorzata Tymrakiewicz<sup>49</sup>, Michelle Guy<sup>49</sup>, Tokhir Dadaev<sup>49</sup>, Sarah J. Little<sup>49</sup>, Koveela Govindasami<sup>49</sup>, Emma Sawyer<sup>49</sup>, Rosemary Wilkinson<sup>49</sup>, Kathleen Herkommer<sup>50</sup>, John L. Hopper<sup>3</sup>, Aritaya Lophatonanon<sup>9,10</sup>, Antje E. Rinckleb<sup>25</sup>, Zsofia Kote-Jarai<sup>49,56</sup>, Rosalind A. Eeles<sup>49,51,52,53,54,56</sup> **on behalf of The UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons' Section of**

**Oncology on behalf of The UK ProtecT Study Collaborators on behalf of The PRACTICAL Consortium, and Douglas F. Easton<sup>1,56</sup>**

**Affiliations**

<sup>1</sup>Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Strangeways Laboratory, Worts Causeway, Cambridge, UK <sup>2</sup>Cancer Epidemiology Centre, the Cancer Council Victoria, 1 Rathdowne Street, Carlton, Victoria, Australia <sup>3</sup>Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, The University of Melbourne, Victoria, Australia <sup>4</sup>University of Cambridge, Department of Oncology, Box 279, Addenbrooke's Hospital, Hills Road Cambridge CB2 0QQ <sup>5</sup>Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge, UK <sup>6</sup>Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK, Faculty of Medical Science, University of Oxford, John Radcliffe Hospital, Oxford, UK <sup>7</sup>School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS, UK <sup>8</sup>The University of Manchester, Centre for Epidemiology, Institute of Population Health, Oxford Road, Manchester, M13 9PL, United Kingdom <sup>9</sup>University of Warwick, University House, Coventry, CV4 8UW, United Kingdom <sup>10</sup>Institute of Biomedical Technology/BioMediTech, University of Tampere and FimLab Laboratories, Tampere, Finland <sup>11</sup>Department of Medical Biochemistry and Genetics, University of Turku, Turku, Finland <sup>12</sup>Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Centre, Los Angeles, California, US <sup>13</sup>University College London, Department of Applied Health Research, 1–19 Torrington Place, London, WC1E 7HB <sup>14</sup>National Human Genome Research Institute, National Institutes of Health, 50 South Drive, Rm. 5351, Bethesda, Maryland, USA <sup>15</sup>Division of Public Health Sciences, Fred Hutchinson Cancer Research Centre, Seattle, Washington, 98109-1024, USA <sup>16</sup>Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington, US <sup>17</sup>Australian Prostate Cancer Research Centre-Qld, Institute of Health and Biomedical Innovation and School of Biomedical Science, Queensland University of Technology, Brisbane, Australia <sup>18</sup>Griffith Health Institute, Griffith University, Gold Coast, QLD, Australia <sup>19</sup>Cancer Council Queensland, Brisbane, Australia <sup>20</sup>Prostate Cancer Foundation of Australia, Sydney, Australia <sup>21</sup>Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, Herlev Ringvej 75, DK-2730 Herlev, Denmark <sup>22</sup>Department of Molecular Medicine (MOMA), Aarhus University Hospital, Brendstrupgaardsvej 100, 8200 Aarhus N, Denmark <sup>23</sup>Division of Cancer Prevention and Control, H. Lee Moffitt Cancer Centre, 12902 Magnolia Drive, Tampa, FL 33612, USA <sup>24</sup>International Hereditary Cancer Centre, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland <sup>25</sup>Department of Urology, University Hospital Ulm, Germany <sup>26</sup>Hannover Biomedical Research School, Carl-Neuberg Str. 1, D- 30625 Hannover, Germany <sup>27</sup>University of Tasmania, Menzies Research Institute Tasmania, Medical Sciences Building 2, 17 Liverpool Street, Private Bag 23, Hobart, Tasmania, Australia 7001 <sup>28</sup>Division of Genetic Epidemiology, Department of Medicine, University of Utah School of

Medicine, Salt Lake City, Utah, USA <sup>29</sup>George E. Wahlen Department of Veterans Affairs Medical Centre, Salt Lake City, Utah, US <sup>30</sup>Division of Clinical Epidemiology and Ageing Research, German Cancer Research Centre (DKFZ), Heidelberg, Germany <sup>31</sup>German Cancer Consortium (DKTK), Heidelberg, Germany <sup>32</sup>University of Pennsylvania. 217 Blockley Hall. 423 Guardian Drive. Philadelphia, PA. 19104., USA <sup>33</sup>Division of Population Sciences, Department of Medical Oncology, Thomas Jefferson University, 834 Chestnut Street, Suite 314, Philadelphia, PA 19107, USA <sup>34</sup>Department of Urology, Akita University School of Medicine, 1-1-1 Hondo, Akita 010-8543, Japan <sup>35</sup>Mayo Clinic, Rochester, Minnesota, USA <sup>36</sup>Division of Haematology/Oncology University of Michigan Medical School, Associate Director for Faculty Affairs, UMCCC, Michigan, US <sup>37</sup>Divisions of Oncology and Genetic Medicine, Geneva University Hospitals, 30 Bd de la Cluse, 1211 Geneva 14, Switzerland <sup>38</sup>Hopital Cantonal Universitaire de Geneve, 24 rue Micheli-du-Crest, 1211, Geneve 14, Switzerland <sup>39</sup>Department of Medical Chemistry and Biochemistry, Molecular Medicine Centre, Medical University – Sofia, 2 Zdrave Str., 1431 Sofia, Bulgaria <sup>40</sup>McGill University, Montreal, Quebec, Canada <sup>41</sup>Department of Complex Genetics, Cluster of Genetics and Cell Biology, NUTRIM School for Nutrition, Toxicology and Metabolism, Maastricht University Medical Centre+, Maastricht, The Netherlands <sup>42</sup>Barts Cancer Institute, Queen Mary University of London, John Vane Science Centre, Charterhouse Square, London EC1M 6BQ <sup>43</sup>Second Military Medical University, 800 Xiangyin Rd., Shanghai 200433, P. R. China <sup>44</sup>Division of Urology, Department of Surgery, University of Utah School of Medicine, Salt Lake City, Utah, USA <sup>45</sup>CR-UK/YCR Sheffield Cancer Research Centre, University of Sheffield, Sheffield, UK <sup>46</sup>Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Grattan Street, Parkville, Victoria 3010, Australia <sup>47</sup>Molecular Cancer Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia <sup>48</sup>Cancer Council Victoria, Cancer Epidemiology Centre, 615 St Kilda Rd, Melbourne, Australia <sup>49</sup>The Institute of Cancer Research, Cotswold Road, Sutton, Surrey, SM2 5NG, UK <sup>50</sup>Department of Urology, Technical University Munich, Germany <sup>51</sup>Royal Marsden NHS Foundation Trust, Fulham and Sutton, London and Surrey, UK

## Acknowledgments

We thank all the patients and control men who took part in this study. We also thank all the members from studies that they participated in this phase of PRACTICAL consortium: Esther John, Amit Joshi and 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/ukgps](http://www.icr.ac.uk/ukgps)), 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, Romyana 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, Colonel Laurence N., Julio Pow-Sang, Hyun Y. Park, Selina Radlein, Maria Rincon, Babu Zachariah (Supplementary Notes).

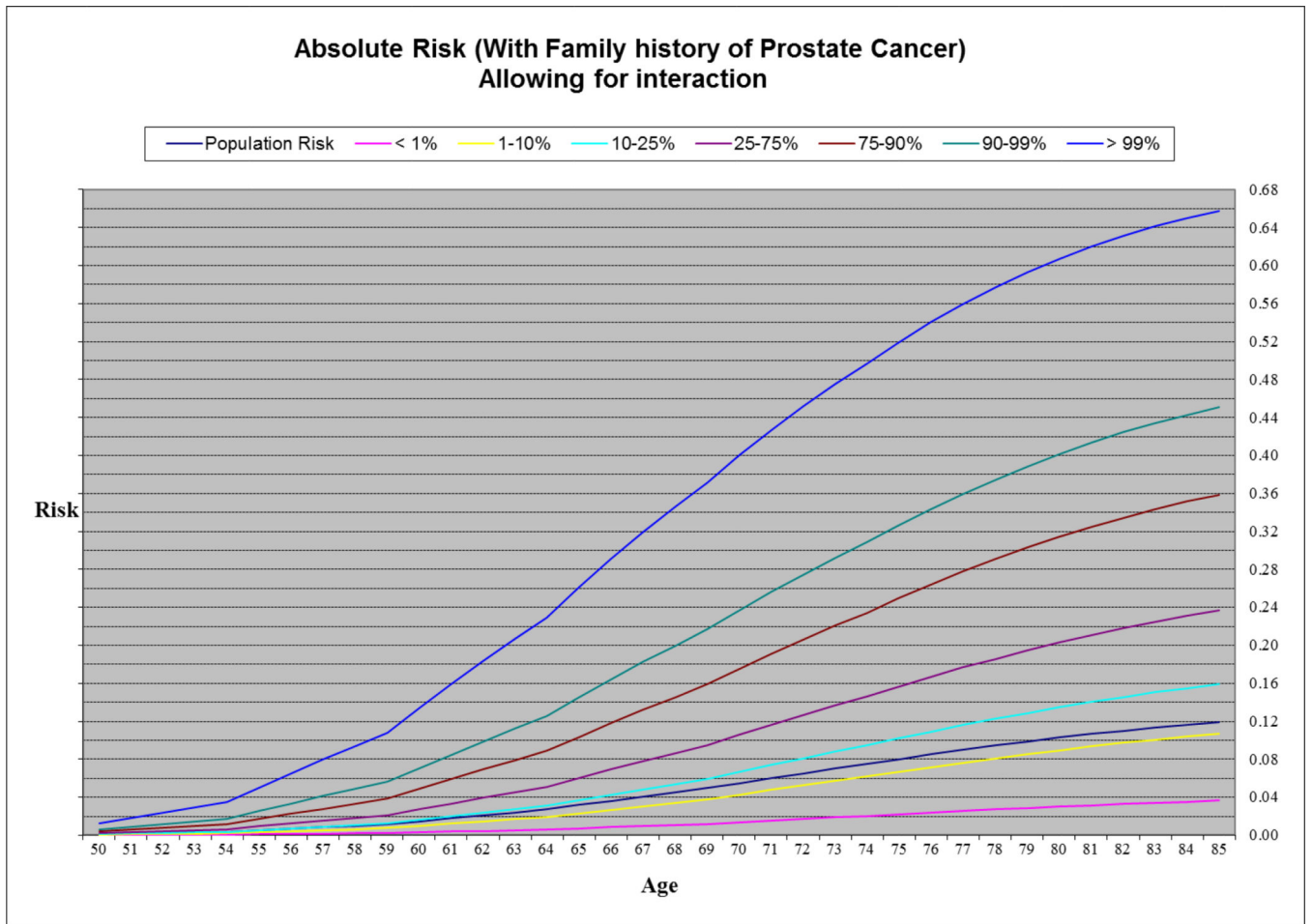
**Financial Support:** D F. Easton was recipient of the CR-UK grant C1287/A10118. R A. Eeles was recipient of the CR-UK grant C5047/A10692 and B E. Henderson was recipient of the NIH grant 1U19CA148537-01.

## Reference List

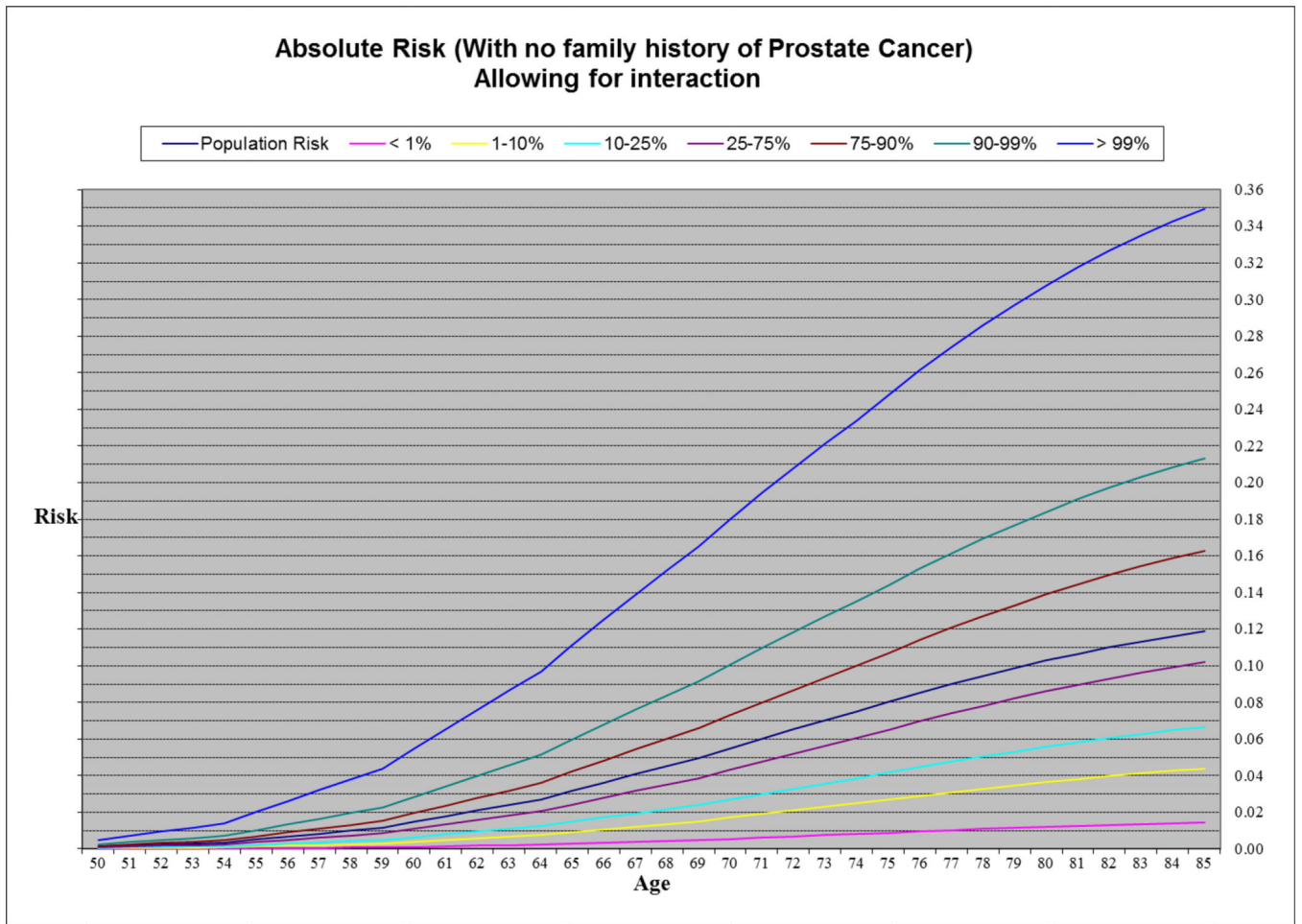
1. Pashayan N, Duffy SW, Chowdhury S, Dent T, Burton H, Neal DE, et al. Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. *Br J Cancer*. 2011; 104:1656–1663. [PubMed: 21468051]
2. Amin Al Olama A, Kote-Jarai Z, Giles GG, Guy M, Morrison J, Severi G, et al. Multiple loci on 8q24 associated with prostate cancer susceptibility. *Nat Genet*. 2009; 41:1058–1060. [PubMed: 19767752]
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; 41:1116–1121. [PubMed: 19767753]
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; 40:316–321. [PubMed: 18264097]
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; 38:652–658. [PubMed: 16682969]
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; 39:631–637. [PubMed: 17401366]
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; 40:281–283. [PubMed: 18264098]
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; 39:977–983. [PubMed: 17603485]
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; 39:638–644. [PubMed: 17401364]
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; 40:310–315. [PubMed: 18264096]
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; 43:785–791. [PubMed: 21743467]
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. [PubMed: 19492347]
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; 45:385–382. [PubMed: 23535732]
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*. 2014 Epub 2014/09/16.

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; 70:9742–9754. [PubMed: 21118973]
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):1–18. [PubMed: 11443730]
17. Antoniou AC, Easton DF. Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet Epidemiol.* 2003; 25(3):190–202. [PubMed: 14557987]
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(4):400–405. [PubMed: 23455638]
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. 2014 PubMed PMID: 25217961.
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
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. [PubMed: 20967219]
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; 60:21–28. [PubMed: 21295399]





**Figure 1.**  
Absolute risk of PrCa by age in men with family history.



**Figure 2.**  
Absolute risk of PrCa by age in men with no family history.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 1**

Total numbers of cases and controls used in the analyses

<b>Study</b>	<b>Controls<sup>a</sup></b>	<b>Cases<sup>a</sup></b>	<b>Total<sup>a</sup></b>
GWAS Stage 3	4,076	4,505	8,581
PRACTICAL	16,050	15,783	31,833
Total	20,126	20,288	40,414
Total <sup>b</sup>	18,343	16,643	34,986

<sup>a</sup> Analyses were restricted to men of European ancestry (see text).

<sup>b</sup> Total after excluding 5 studies that oversampled cases with family history.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**  
 Summary results of 25 SNPs using PRACTICAL and GWAS Stage 3 datasets in European.

Marker <sup>d</sup> Chr/Nearby Gene	Alleles <sup>b</sup> Position <sup>c</sup>	MAF <sup>d</sup>	Per allele <sup>e</sup> OR (95%CI)	Het OR <sup>e,f</sup> 95%CI)	Hom OR <sup>e,g</sup> (95%CI)	P-value <sup>h</sup>	P-value <sup>i</sup>
rs721048 2 / <i>EHHBP1</i>	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 <sup>-8</sup>	0.13
rs1465618 2 / <i>THADA</i>	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 <sup>-4</sup>	0.39
rs12621278 2 / <i>ITGA6</i>	A/G 173311553	0.06	.75 (.70–.80)	.76 (.71–.82)	.38 (.24–.58)	4.9×10 <sup>-17</sup>	0.57
rs260753 3 / <i>Unknown</i>	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 <sup>-5</sup>	0.73
rs17021918 4 / <i>PDLIM5</i>	G/A 95562877	0.35	.88 (.85–.91)	.86 (.83–.90)	.80 (.74–.85)	6.7×10 <sup>-15</sup>	0.39
rs12500426 4 / <i>PDLIM5</i>	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 <sup>-8</sup>	0.54
rs7679673 4 / <i>TET2</i>	C/A 106061534	0.40	.88 (.85–.90)	.87 (.83–.91)	.77 (.72–.82)	1.0×10 <sup>-16</sup>	0.08
rs9364554 6 / <i>SLC22A3</i>	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 <sup>-8</sup>	0.85
rs10486567 7 / <i>JAZF1</i>	G/A 27976563	0.23	.85 (.82–.89)	.86 (.81–.91)	.72 (.63–.81)	4.5×10 <sup>-12</sup>	0.21
rs6465657 7 / <i>LMTK2</i>	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 <sup>-9</sup>	0.32
rs1447295 8 / <i>Unknown</i>	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 <sup>-6</sup>	0.50
rs6983267 8 / <i>Unknown</i>	C/A 128413305	0.49	.82 (.79–.85)	.80 (.76–.85)	.67 (.63–.72)	2.3×10 <sup>-35</sup>	0.61
rs16901979 8 / <i>Unknown</i>	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 <sup>-28</sup>	0.29
rs2928679 8 / <i>SLC25A37</i>	C/T 23438975	0.48	1.04 (1.01–1.07)	1.03 (.97–1.09)	1.08 (1.01–1.16)	.02	0.10
rs1512268 8 / <i>NKX3.1</i>	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 <sup>-16</sup>	0.19
rs4962416 10 / <i>CTBP2</i>	A/G 126696872	0.28	1.04 (1.01–1.08)	1.03 (.98–1.08)	1.11 (1.02–1.21)	.02	0.68
rs10993994 10 / <i>MSMB</i>	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 <sup>-41</sup>	0.36

Marker <sup>a</sup> Chr/Nearby Gene	Alleles <sup>b</sup> Position <sup>c</sup>	MAF <sup>d</sup>	Per allele <sup>e</sup> OR (95%CI)	Het OR <sup>e,f</sup> 95%CI)	Hom OR <sup>e,g</sup> (95%CI)	P-value <sup>h</sup>	P-value <sup>i</sup>
rs7931342 11 / <i>Unknown</i>	C/A 68994497	0.50	.84 (.81–86)	.86 (.82–91)	.70 (.65–74)	4.8×10 <sup>-27</sup>	0.86
rs7127900 11 / <i>Unknown</i>	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 <sup>-26</sup>	0.63
rs4430796 17 / <i>HNF1B</i>	A/G 36098040	0.48	.81 (.79–84)	.81 (.77–85)	.66 (.62–71)	2.7×10 <sup>-38</sup>	0.79
rs11649743 17 / <i>HNF1B</i>	G/A 36074979	0.19	.88 (.85–92)	.88 (.83–92)	.79 (.70–90)	5.6×10 <sup>-10</sup>	0.25
rs1859962 17 / <i>Unknown</i>	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 <sup>-24</sup>	0.19
rs2735839 19 / <i>KLK2/KLK3</i>	G/A 51364623	0.15	.81 (.77–85)	.82 (.78–86)	.62 (.53–73)	1.1×10 <sup>-19</sup>	0.06
rs5759167 22 / <i>BIL/TLL1</i>	G/T 43500212	0.50	.84 (.82–87)	.83 (.79–87)	.71 (.67–76)	3.4×10 <sup>-28</sup>	0.87
rs5945619 X / <i>NUDT11</i>	T/C 51241672	0.36	1.13 (1.10–1.16)	-	1.28 (1.22–1.35)	1.9×10 <sup>-20</sup>	0.10

<sup>a</sup> dbSNP rs number.

<sup>b</sup> Major/minor allele, based on the frequencies in controls in PRACTICAL III data.

<sup>c</sup> Build 37 position.

<sup>d</sup> MAF in controls in combined European dataset.

<sup>e</sup> OR = odds ratio (minor allele) from a logistic regression using all European samples stratified by studies with no adjustment.

<sup>f</sup> OR in heterozygotes, relative to major allele homozygotes.

<sup>g</sup> OR in minor allele homozygotes, relative to major allele homozygotes.

<sup>h</sup> Cochran-Armitage test for trend.

<sup>i</sup> Heterogeneity *P*-value among studies

**Table 3**

Odds ratios for PrCa by percentile of the PRS and family history.

Percentiles		OR <sup>a,b</sup>	OR <sup>a,c</sup>
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)

<sup>a</sup> ORs obtained by fitting PRS group, family history and age at diagnosis jointly.

<sup>b</sup> ORs compared to men in the 1st percentile as baseline.

<sup>c</sup> ORs compared to men in the 25<sup>th</sup>–75<sup>th</sup> percentile as baseline.