Validation of a Prostate Cancer Polygenic Risk Score for Clinical Use

September 2018

Mary Helen Black[1], Shuwei Li[1], Holly LaDuca[1], Jefferey Chen[1], Robert Hoiness[1], Stephanie Gutierrez[1], Brigette Tippin-Davis[1], Hsiao-Mei Lu[1], Marta Gielzak[2], Kathleen Wiley[2], Zhuqing Shi[3], Hongjie Yu[3], Brian T. Helfand[3], William Isaacs[2], Jianfeng Xu[3]

[1]Ambry Genetics, Aliso Viejo, CA

[2]Johns Hopkins University, Baltimore, MD

[3]NorthShore University Health System, Evanston, IL

Abstract

BACKGROUND: Genome-wide association studies (GWAs) have identified over 100 common single nucleotide polymorphisms (SNPs) associated with prostate cancer (PrCa) risk. While the risk associated with each SNP is modest, a polygenic risk score (PRS) based on their combined genotypes may have substantial predictive value for risk stratification. The overarching goal of this study is to provide evidence for the analytical and clinical validity of a PrCa PRS in the patient care setting.

METHODS: We genotyped 72 PrCa-associated SNPs using next-generation sequencing in order to examine whether a risk score based on these SNPs was predictive of PrCa in 3,140 Caucasian men (1,725 cases who previously underwent radical prostatectomy and 1,415 controls unaffected with PrCa) from a large, multi-site collaborative study. We constructed a population-standardized PRS, with each SNP weighted by per-allele relative risks in Caucasians from large GWAs and population-specific allele frequencies, and tested PRS association with PrCa using multivariate logistic regression models adjusted for age alone, as well as age and family history of PrCa.

RESULTS: Mean±SD age of PrCa cases at diagnosis and age of controls at the time of testing or last clinic visit was 59.0±6.7 and 58.9±12.0 years, respectively. Among cases, 58.5% had Gleason score ≤6, while 41.5% had Gleason ≥8. In addition, 38.6% of cases and 15.7% of controls had ≥1 first- or second-degree relative with PrCa. The PRS was significantly higher in cases than controls (mean±SD 1.37±0.87 vs. 0.99±0.61, p<0.0001). Compared to men in the 1st quartile of age-adjusted PRS, those in the 2nd, 3rd and 4th quartile were 1.88 (95% CI: 1.52-2.32), 2.35 (95% CI: 1.90-2.90) and 4.37 (95% CI: 3.50-5.47) times as likely to have PrCa (all p<0.0001), respectively. Additional adjustment for family history among the subset of patients with this information yielded similar results. PRS predictive performance was consistent with prior literature (AUROC=0.65, 95% CI: 0.63-0.67).

CONCLUSIONS: These data suggest that a 72-SNP PRS is predictive of PrCa. Our ongoing analyses will include PRS association with clinical features of PrCa, validation in non-Caucasian populations, and interaction with high- and moderate-penetrance PrCa predisposition genes, as well as clinical utility for personalized PrCa screening.

Introduction

Genome-wide association studies (GWAs) have identified approximately 100 single nucleotide polymorphisms (SNPs) associated with increased susceptibility to prostate cancer (PrCa). While each individual SNP contributes only modestly to this susceptibility, their combined effects are estimated to account for ~33% of the familial risk of PrCa[1]. It has further been shown that a polygenic risk score (PRS) based on a combination of SNP genotypes can be consistently estimated across a variety of populations and geographic regions, that it may have substantial predictive value for PrCa risk stratification, and that the predictive performance of such a PRS is superior to that based on family history information alone[2-6].

In a comparison of self-reported family history and PRS as two measures of inherited risk for PrCa in several study populations, Sun *et al.* observed substantial variation in the proportion of positive family history but similar mean PRS among participating sites of the REDUCE trial [3]. Despite utilization of the same protocol to obtain family history information, sites in Eastern and Western Europe reported 4.2% and 10.9% of men with PrCa-affected first- or second-degree family members, respectively, while the proportion in North America was 22.8%; similar variation was also observed among sites within the same country. The family history-associated relative risks (RRs) for PrCa among these 3 regions ranged from 1.20 to 1.91. Recent studies in men of European ancestry from the PRACTICAL consortium and others have reported RR for PrCa associated with family history to be

as high as 2.5[7, 8]. In contrast, RRs associated with a 33-SNP PRS were relatively homogenous among REDUCE study sites, ranging 1.69 to 1.89. PRS-associated RRs recently reported by PRACTICAL and others have ranged from 1.74 to 1.86, despite differences in the number of SNPs included in each PRS model[7, 8]. Moreover, Sun *et al.* reported superior predictive performance of the PRS compared to family history alone, and demonstrated that the addition of family history information to the PRS model did not further improve prediction of PrCa[3]. Taken together, the consistent risk estimates associated with PRSs across multiple studies, and its ability to more accurately identify men who develop PrCa, provide evidence supporting its potential use in clinical risk assessment.

In this report we examined the extent to which a PRS, based on the combined effects of 72 SNPs previously reported to be associated with PrCa in multiple large GWAs, was predictive of PrCa in Caucasian men who tested negative for pathogenic or likely pathogenic variants in known PrCa-susceptibility genes. We also assessed whether the contribution of overall genetic risk represented by the PRS varied by age and/or family history, and estimated absolute lifetime risk for PrCa to age 85 accordingly. Ongoing analyses will include evaluating PRS association with clinical features of PrCa, validation in non-Caucasian populations, and interaction with high- and moderate-penetrance PrCa predisposition genes, as well as clinical utility for personalized PrCa screening.

Methods

Patient population

Samples from a total of 3,140 patients were ascertained from three study sites: Johns Hopkins University Hospital (JHH: 1,725 cases and 541 controls), Ambry Genetics (AG: 490 controls), and NorthShore University HealthSystem's Genomic Health Initiative (NSGHI: 384 controls), and were eligible for study inclusion if they were male, self-reported Caucasian race, and ≥18 years of age at the time of PrCa diagnosis (JHH cases), genetic testing (AG controls), or last clinic visit (JHH and NSGHI controls). Those who tested positive in the present study for a pathogenic or likely pathogenic variant in a PrCa-susceptibility gene (*ATM, BRCA1, BRCA2, CHEK2, EPCAM, HOXB13, NBN, PALB2, MLH1, MSH2, MSH6, PMS2, RAD51D, TP53*) were excluded (n=203). Cases were patients undergoing radical prostatectomy for treatment of clinically localized PrCa at JHH, and were included if disease was organ-confined and Gleason score ≤6 or ≥8, as determined upon pathological evaluation of the prostatectomy specimen. All controls were unaffected with PrCa; 59% of AG controls had a personal history of at least one non-PrCa primary. NSGHI controls had a minimum of 1 year of clinical history available in the electronic health record (EHR), and were excluded if any ICD-9/10 diagnosis of cancer was present at any time in the EHR. History of additional non-PrCa primaries was not available for JHH controls. PrCa-specific family history information was available for the majority of cases (n=1,484) and controls from AG and JHH (n=828).

Molecular Analysis

Sequencing quality for Illumina NextSeq 500 are monitored during the sequencing run, and include visualization of Intensity-vs-Cycle (IVC) plots, and cluster intensity over the duration of the run. Other quality metrics that are evaluated for the entire sequencing run upon completion of sequencing and demultiplexing of the samples include metrics for the % Perfect Index Reads, % of ≥Q30 Bases, and overall Mean Quality Score. Samples passing the sequencing quality metrics were fed into proprietary NGS data processing pipeline in a parallelized fashion, starting with alignment of sequencing reads to human reference genome build (GRCh37/hg19), followed by variant and genotype calling on the panel genes and the 72 PrCa-associated SNPs. Additionally, NGS coverage is evaluated for all 72 associated SNPs for every sample, and any SNPs with no or low coverage (<20X) were excluded from genotype calling, and were not included in downstream statistical analysis.

Statistical Analysis

SNPs that met the following criteria were selected for inclusion in the analysis: 1) reported with genome-wide significance in >1 GWA analysis, based on a sample size of >500 cases and >500 controls in any population; 2) are not strongly correlated; 3) effects are race/ethnicity-specific (*i.e.* only SNPs that meet 1) and 2) in a specific population were included). For this study, we selected SNPs reported in studies of individuals of Caucasian ancestry[1, 9-20]. We identified 72 SNPs meeting these criteria, of which 3 could not be directly genotyped and were substituted by tag SNPs, with pairwise r²=1.0 and distance ranging from 519 to 8,504 bp between each proxy and originally reported SNP.

NGS data were examined to assess missing rates for each sample, and each SNP. Samples were excluded if >7 SNPs were missing due to bioinformatics quality control thresholds (n=13; 0.4% of samples passing bioinformatics QC). SNP calls were checked for consistency with publically available databases (GRCh37/hg19; Ensembl release 91[21]) and literature-reported reference and risk alleles. We compared SNP allele frequencies among control subjects to those available in the gnomAD non-Finnish European population to ensure consistency with the reference population. Hardy Weinberg Equilibrium (HWE) was assessed for all SNPs among controls using R package HardyWeinberg[22]. To assess the assumption of SNP effects consistent with a multiplicative model, we examined all possible pair-wise SNP*SNP interactions associated with PrCa using logistic regression, with a 1-df likelihood ratio test for the interaction; p-values were FDR-corrected for multiple testing and p<0.05 was considered significant. We additionally tested for higher-order SNP interactions using logic regression[23].

Using an approach consistent with prior literature[5], we computed a SNP-based population-standardized PRS for each individual. Using previously published estimates of the per-allele ORs and risk allele frequencies (p) for each SNP, and assuming independent and additive risks on the log OR scale, we computed the unscaled population average risk for autosomal SNPs as:

$$\mu = (1 - p)^2 + 2p(1 - p)OR + p^2 OR^2$$

and for sex chromosome SNPs as:

$$\mu = pOR + (1 - p)$$

Adjusted risk values were then calculated as:

$$\frac{1}{\mu}, \frac{OR}{\mu}, \frac{OR^2}{\mu}$$

for the three genotypes defined by the number of risk alleles: 0, 1 or 2, respectively. Missing genotypes were imputed with a population average risk of 1.0. Adjusted risk values for each SNP were multiplied

Table 1. Odds ratios (OR) and 95% confidence intervals (CI) for prostate cancer per quartile of PRS.

PRS Quartile or covariate	PRS Threshold	Case n	Control n	Unadjusted			Adjusted for Age			Adjusted for Age and FH[1]		
				OR	95% CI	p-value	OR	95% CI	p-value	OR	95% CI	p-value
<=25%	<0.68	268	447	1.00	-	-	-	1.00	-	-	-	1.00	-	-	-
25-50%	0.68-1.03	379	334	1.88	1.52	2.33	<0.001	1.88	1.52	2.32	<0.001	1.70	1.33	2.18	<0.001
50-75%	1.03-1.58	416	298	2.33	1.88	2.88	<0.001	2.35	1.90	2.90	<0.001	2.37	1.84	3.05	<0.001
>75%	>1.58	516	199	4.33	3.47	5.41	<0.001	4.37	3.50	5.47	<0.001	4.01	3.07	5.27	<0.001
Age								1.01	1.00	1.01	0.16	1.05	1.04	1.06	<0.001
FH												3.47	2.79	4.35	<0.001

[1]Family History (FH) is defined as presence/absence of at least one first- or second-degree relative with prostate cancer. Adjustment for family history resulted in reduced sample size and therefore altered PRS quartile thresholds, because once samples were removed, the distribution is altered. Analyses including family history included 2,312 (81%) of 2,859 subjects meeting all inclusion/exclusion criteria and missing ≤7 SNPs; a total of 1,484 (94%) cases and 828 (65%) controls were included.

Table 2. Odds ratios (OR) and 95% confidence intervals (CI) per Standard Deviation of PRS by presence or absence of family history of prostate cancer, compared to Al Olama *et al.* 2015[7] and Schumacher et al. 2018[8].

Family History			Present Study[1]			Al Olama *et al.* 2015[2]			Schumacher *et al.* 2018[3]		
	Case n	Control n	OR	95% CI		OR	95% CI		OR	95% CI	
Positive	574	130	1.96	1.50	2.63	1.79	1.63	1.96	1.95	1.85	2.05
Negative	910	698	1.77	1.56	2.02	1.70	1.64	1.76	1.84	1.80	1.88

[1]ORs for the present study are adjusted for age
[2]ORs shown for Al Olama *et al.* 2015[7] were adjusted for age
[3]ORs shown for Schumacher *et al.* 2018[8] were not adjusted

to compute the overall PRS-associated risk for each individual based on his observed genotypes. Logistic regression models were used to estimate the ORs for PrCa by quartile of the PRS, with the 1st quartile category (<25th percentile) as the reference. ORs for PrCa were also estimated for the following percentiles of the distribution of the PRS, assuming 25th-75th percentile as the reference: <1st, 1st-10th, 10th-25th, 25th-75th, 75th-90th, 90th-99th, >99th percentile. A continuous PRS was tested for association with family history, defined as a binary indicator for presence of ≥1 first- or second-degree family member with a history of prostate cancer, using logistic regression. We used a 1-df likelihood ratio test to assess evidence for interaction between continuous PRS and age at diagnosis, genetic testing or last clinic visit, and continuous PRS and family history, by comparing models with and without an interaction term. Area under the receiver-operating curve (AUROC) was computed using R package pROC. R (v.3.3.3) was used for all statistical analyses; all statistical tests were two sided, and p-values <0.05 were considered nominally statistically significant.

Lastly, absolute lifetime risk to age 85 years was estimated based on PRS-specific ORs, population-based age-specific prostate cancer incidences rates, and both all-cause and PrCa-specific mortality rates, to account for competing causes of death. Population rates were based on non-Hispanic white males in the United States (SEER, www.seer.cancer.gov, [24, 25]). Absolute risk was estimated with ABSRISK v.1.0 [26].

Results

A total of 3,140 patient samples (1,725 PrCa cases and 1,415 controls) underwent NGS for the present study. After assessment of bioinformatics QC and inclusion/exclusion criteria, data from 2,859 patients (1,579 PrCa cases and 1,280 controls) were available for analysis. The mean±SD age at diagnosis for cases and at testing or last clinic visit for controls was 59.0±6.7 and 58.9±12.0 years, respectively. Among cases, 58.5% had Gleason score ≤6, while 41.5% had Gleason ≥8. Of those with available family history information, 38.6% of cases and 15.7% of controls had ≥1 first- or second-degree relative with PrCa.

The mean±SD SNP call rate, or the proportion of individuals for whom a genotype was successfully determined for a given SNP, was 99.7%±0.8% (range 93.8% to 100.0%). SNP risk allele frequencies (RAF) among controls ranged from 3.6% to 94.8%, and were consistent with the gnomAD non-Finnish European population (range: 3.1% to 94.3%; mean±SD absolute difference among SNPs: 1.7%±1.6%, p>0.05). Among the 69 SNPs located on autosomal chromosomes, none deviated significantly from HWE (all p ≥0.01); the 3 SNPs located on the X chromosome are not expected to meet classic HWE in an all-male population[27]. Consistent with the findings of previous large studies, we did not detect any significant pairwise or high-order interactions among the SNPs after FDR correction for multiple testing[7].

The sum of the risk alleles across the 72 SNPs was approximately normally distributed among cases and controls, and ranged from 49 to 83 and 41 to 79, respectively (mean±SD risk allele count: 64.8±5.5 vs. 62.1±5.3, p<0.0001; Figure 1). The mean±SD PRS was significantly higher for cases compared to controls (1.37±0.87 vs. 0.99±0.61, p<0.0001). The continuous PRS was significantly associated with PrCa risk, with an OR per standard deviation of the PRS of 1.80 (95% CI: 1.64-1.98). Compared to men in the 1st quartile of age-adjusted PRS, those in the 2nd, 3rd and 4th quartile were 1.88 (95% CI: 1.52-2.32), 2.35 (95% CI: 1.90-2.90) and 4.37 (95% CI: 3.50-5.47) times as likely to have PrCa (all p<0.0001). Maximum AUROC for PRS discrimination of cases and controls was reached at a threshold of 1.06, corresponding to a PPV=0.67 and NPV=0.55 (AUROC=0.65, 95% CI: 0.63-0.67; Figure 2). Adjustment for age resulted in nearly identical PRS ORs per quartile and model discriminatory performance (AUROC=0.64) compared to the unadjusted model (Table 1). Further adjustment for presence/absence of a first- or second-degree relative with PrCa resulted in slight attenuation of PRS ORs per quartile (Table 1), but similar discriminatory performance (AUROC=0.64), which was superior to the model of family history alone (AUROC=0.61, 95% CI: 0.59-0.63). In the fully adjusted model, both family history and PRS quartile were significantly associated with PrCa (all p<0.001), implying that each factor independently contributes to disease. The small reduction in PRS effect estimates was driven primarily by the strong relationship between family history and PrCa; the PRS was not associated with family history among cases and controls (p=0.47), or within the subset of cases (p=0.45) or controls (p=0.97). Moreover, we observed no significant interactions between the PRS and age (p=0.92), or family history (p=0.68). Consistent with the findings of others, PRS-associated risk for PrCa did not vary by increasing age or presence of PrCa in the family history (Table 2).

Examining age- and family history-adjusted effects stratified by PRS percentiles on a finer scale, we observed monotonic increasing ORs with increasing PRS compared to the 25th-75th percentile range (median risk modeled as the reference) (Table 3). Importantly, the ORs above and below the median risk were comparable to and had 95% CIs overlapping with two other large studies: a 25-SNP PRS and 147-SNP PRS assessed in 40,414 and 72,729 PrCa cases and controls of European ancestry, respectively[7, 8].

The mean±SD absolute lifetime risk to age 85 estimated for all cases vs. controls based on PRS-associated risk was 15.0%±8.5% vs. 10.5%±6.5%. By PRS percentile, cumulative lifetime risk to age 85 for men in the <1st percentile of the PRS was 2.4%, and increased with increasing PRS percentile to 42.0% for men in the >99th percentile (Figure 3). Cumulative risk from age 25 to 45 was close to 0% for all PRS percentiles, as PrCa incidence is extremely low in this age range. Risk by PRS percentile began to diverge at ~50 years of age, the age that most men are first screened for PrCa in the clinical setting.

Table 3. Odds ratios (OR) and 95% confidence intervals (CI) for prostate cancer per PRS percentile, compared to Al Olama *et al.* 2015[7] and Schumacher *et al.* 2018[8].

	PRS Threshold	Case n	Control n	p-value	Present Study[1] (72 SNPs)			Al Olama *et al.* 2015[2] (25 SNPs)			Schumacher *et al.* 2018[3] (147 SNPs)		
					OR	95% CI		OR	95% CI		OR	95% CI	
<1%	<0.28	9	15	0.003	0.26	0.10	0.63	0.14	0.08	0.24	0.15	0.11	0.20
1-10%	0.28-0.48	88	120	<0.0001	0.39	0.28	0.53	0.41	0.36	0.47	0.35	0.32	0.37
10-25%	0.48-0.68	184	162	0.0002	0.61	0.47	0.79	0.63	0.57	0.70	0.54	0.51	0.57
25-75%	0.68-1.58	752	404	-	1.00	-	-	1.00	-	-	1.00	-	-
75-90%	1.58-2.17	256	90	0.001	1.60	1.22	2.13	1.68	1.54	1.83	1.74	1.67	1.82
90-99%	2.17-4.19	172	36	<0.0001	2.72	1.86	4.08	2.31	2.09	2.56	2.69	2.55	2.82
≥99%	>4.19	23	1	0.02	12.16	2.49	219.53	4.24	3.24	5.53	5.71	5.04	6.08

[1]ORs for the present study were estimated for percentiles of a population-standardized PRS, adjusted for age and family history. Family history was defined as presence/absence of ≥1 first- or second-degree family member with a personal history of PrCa. ORs for some percentiles of the PRS may be unstable due to very small sample size (eg. ≥99%).

[2]ORs shown for Al Olama *et al.* 2015 [7] were estimated for percentiles of a weighted PRS, adjusted for age and family history. Family history was defined as presence/absence of ≥1 first- or second-degree family member with a personal history of PrCa. Despite model differences, the present study ORs for PrCa risk are similar to those estimated by Al Olama *et al.*, except for ≥99% category.

[3]ORs shown for Schumacher *et al.* 2018[8] were estimated for percentiles of a weighted PRS, and are not adjusted for age or family history. Despite model differences, the present study ORs for PrCa risk are similar to those estimated by Schumacher *et al.*, except for ≥99% category.

Figure 1. Distribution of the sum of risk alleles across 72 SNPs, for cases (navy blue) compared to controls (blue). Probability density on the y-axis represents the proportion of cases and controls, respectively, with a given risk allele count (x-axis). The mean±SD risk allele count in cases vs. controls: 64.8±5.5 vs. 62.1±5.3, p<0.0001.
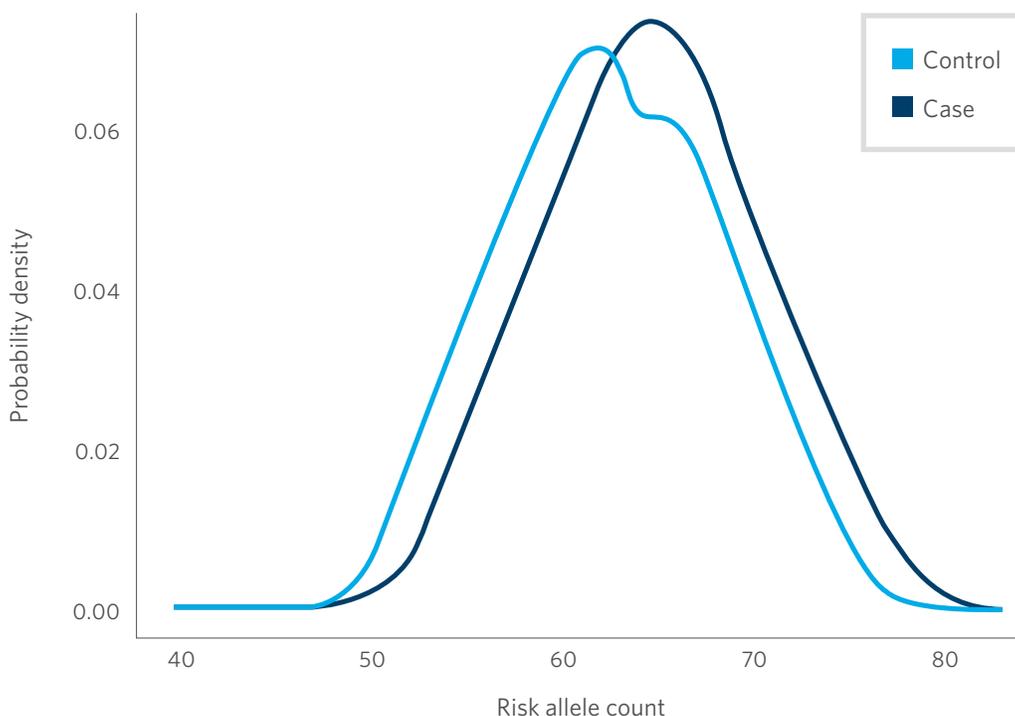
Figure 2. Area under the receiver operating curve (AUROC) for the accuracy of the PRS in distinguishing between prostate cancer cases and controls (AUROC=0.65, 95% CI: 0.63-0.67).
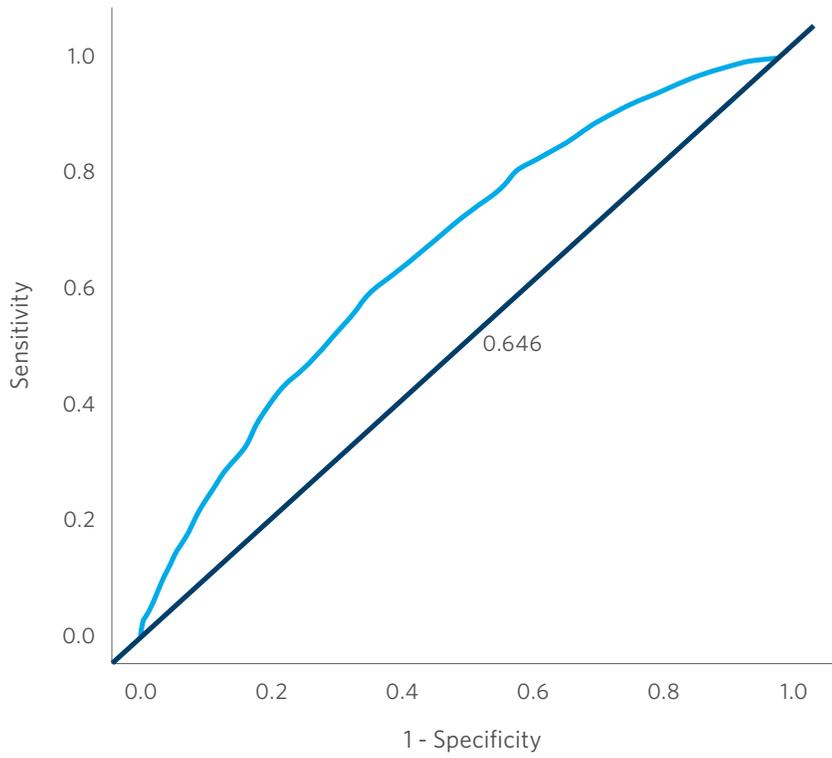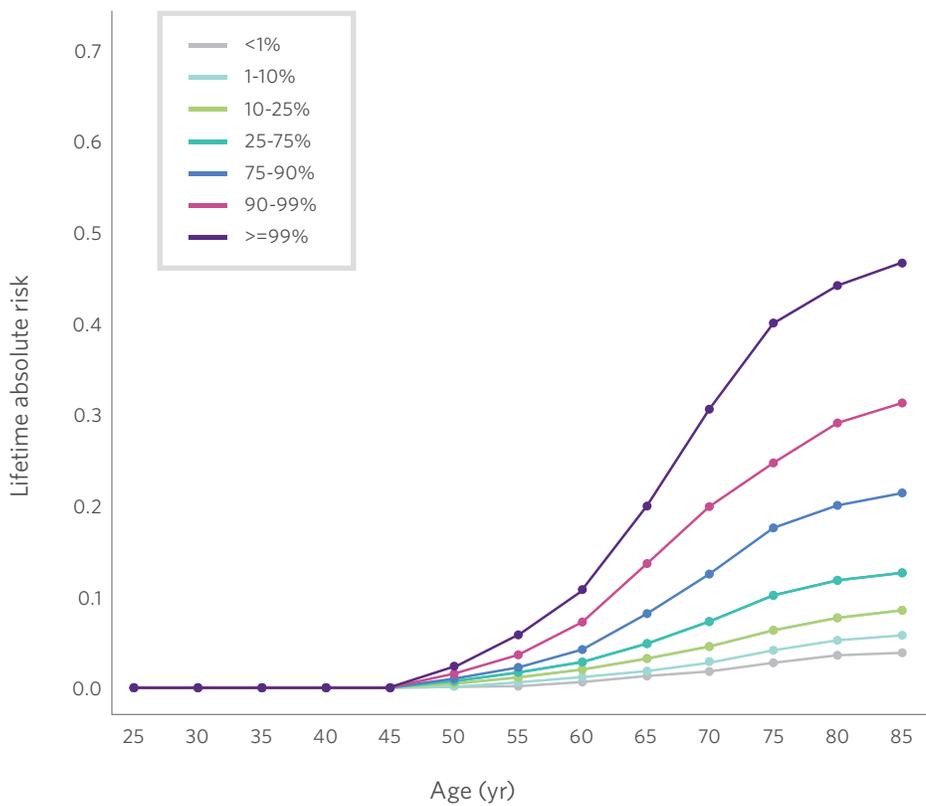

Figure 3. Cumulative lifetime (absolute) risk for prostate cancer by percentiles of PRS.

Discussion

We examined the performance of a 72-SNP PRS in a population of men who were negative for pathogenic/likely pathogenic variants in reported PrCa-susceptibility genes. Our results demonstrate that the relative and absolute lifetime risks of PrCa associated with the PRS for these men are similar to those reported previously in larger and/or broader populations. We also found that the predictive performance of the 72 SNP PRS, while modest, outperformed that of family history alone, as is consistent with prior literature.

Amin Al Olama and colleagues from PRACTICAL, an international collaboration of nearly 80 PrCa studies conducted across multi-ethnic populations, estimated risks associated with and evaluated the predictive performance of a 25-SNP PRS in 40,414 Caucasian men [7]. Recently, Shumacher *et al.* reported results from the prostate cancer OncoArray project, which aimed to identify additional loci associated with PrCa and assessed a 147-SNP PRS in 72,729 men of European ancestry[8]. Investigators from both studies reported a strong association between PRS and PrCa risk, with an OR per standard deviation of PRS of 1.74 (95% CI: 1.70-1.78) and 1.86 (95% CI: 1.83-1.89), respectively. Similarly, we estimated an OR of 1.80 (95% CI: 1.64-1.98) for association with PrCa per PRS standard deviation. After adjusting for age and presence of any PrCa family history, Al Olama *et al.* reported that men with a PRS in the highest 10% of the distribution were 2.31 (95% CI: 2.09-2.56) times as likely to develop PrCa as those with a median-risk PRS (25th-75th percentile of PRS distribution). Shumacher *et al.* reported an OR of 2.69 (95% CI: 2.55-2.82) for this same percentile category of PRS relative to the median-risk category. Consistent with these findings, our age- and family history-adjusted PRS-associated risk for men in the 90-99th PRS percentile compared to the median range was 2.72 (95% CI: 1.86-4.08). As observed in both studies, the PRS and family history were independently associated with PrCa, and the PrCa risk associated with the PRS did not vary by presence/absence of family history.

Our AUROC for models of family history alone and PRS alone were 0.61 (95% CI: 0.59-0.63) and 0.65 (95% CI: 0.63-0.67), respectively. Further adjustment of the PRS model for age and/or family history did not improve its discriminatory performance (AUROC=0.64 for both adjusted models, respectively), which implies that age and/or family history do not add discriminatory power for distinguishing PrCa cases and controls above that already accounted for by the PRS. Several other large studies have reported similar observations, with AUROCs for the PRS alone ranging from 0.59-0.62[3, 5, 28], and AUROCs for the joint model of PRS and family history ranging from 0.60-0.64[3].

Absolute risks to age 85 reported by Amin Al Olama *et al.* ranged from 1.5% to 35.0% across the percentiles of PRS distribution for men without a family history of PrCa, and from 3.7% to 65.8% for men with a positive family history[7]. Among all men included in our study, irrespective of family history, the observed absolute risks by PRS percentile ranged from an average of 2.4% (<1st percentile) to 42.0% (>99th percentile), in line with prior studies. We also observed a divergence of cumulative risk by PRS at approximately age 50, with increasing separation of lifetime risks by PRS percentile with increasing age. Given that 50 years is the age that most men are initially screened for prostate cancer, an approach combining clinical assessment with PRS profiling for estimation of lifetime risk may be useful for clinical risk assessment.

In conclusion, these data suggest that a 72-SNP PRS is predictive of PrCa. Our ongoing analyses will include PRS association with clinical features of PrCa, validation in non-Caucasian populations, and interaction with high- and moderate-penetrance PrCa predisposition genes, as well as clinical utility for personalized PrCa screening.

References

1. Al Olama, A.A., *et al.*, A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nat Genet, 2014. **46**(10): p. 1103-9.
2. Helfand, B.T., *et al.*, Clinical validity and utility of genetic risk scores in prostate cancer. Asian J Androl, 2016. **18**(4): p. 509-14.
3. Sun, J., *et al.*, Genetic score is an objective and better measurement of inherited risk of prostate cancer than family history. Eur Urol, 2013. **63**(3): p. 585-7.
4. Liss, M.A., *et al.*, Prostate genetic score (PGS-33) is independently associated with risk of prostate cancer in the PLCO trial. Prostate, 2015. **75**(12): p. 1322-8.
5. Conran, C.A., *et al.*, Population-standardized genetic risk score: the SNP-based method of choice for inherited risk assessment of prostate cancer. Asian J Androl, 2016. **18**(4): p. 520-4.
6. Chen, H., *et al.*, Adding genetic risk score to family history identifies twice as many high-risk men for prostate cancer: Results from the prostate cancer prevention trial. Prostate, 2016. **76**(12): p. 1120-9.
7. Amin Al Olama, A., *et al.*, Risk Analysis of Prostate Cancer in PRACTICAL, a Multinational Consortium, Using 25 Known Prostate Cancer Susceptibility Loci. Cancer Epidemiol Biomarkers Prev, 2015. **24**(7): p. 1121-9.
8. Schumacher, F.R., *et al.*, Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nature Genetics, 2018. **50**(7): p. 928-936.
9. Al Olama, A.A., *et al.*, Multiple loci on 8q24 associated with prostate cancer susceptibility. Nat Genet, 2009. **41**(10): p. 1058-60.
10. Eeles, R.A., *et al.*, Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet, 2009. **41**(10): p. 1116-21.
11. Eeles, R.A., *et al.*, Multiple newly identified loci associated with prostate cancer susceptibility. Nat Genet, 2008. **40**(3): p. 316-21.
12. Eeles, R.A., *et al.*, Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. Nat Genet, 2013. **45**(4): p. 385-91, 391e1-2.
13. Gudmundsson, J., *et al.*, Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. Nat Genet, 2009. **41**(10): p. 1122-6.
14. Gudmundsson, J., *et al.*, Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet, 2007. **39**(5): p. 631-7.
15. Gudmundsson, J., *et al.*, Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. Nat Genet, 2008. **40**(3): p. 281-3.
16. Hoffmann, T.J., *et al.*, A large multiethnic genome-wide association study of prostate cancer identifies novel risk variants and substantial ethnic differences. Cancer Discov, 2015. **5**(8): p. 878-91.
17. Kote-Jarai, Z., *et al.*, Seven prostate cancer susceptibility loci identified by a multi-stage genome-wide association study. Nat Genet, 2011. **43**(8): p. 785-91.
18. Schumacher, F.R., *et al.*, Genome-wide association study identifies new prostate cancer susceptibility loci. Hum Mol Genet, 2011. **20**(19): p. 3867-75.
19. Thomas, G., *et al.*, Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet, 2008. **40**(3): p. 310-5.
20. Yeager, M., *et al.*, Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet, 2007. **39**(5): p. 645-9.
21. Zerbino, D.R., *et al.*, Ensembl 2018. Nucleic Acids Res, 2018. **46**(D1): p. D754-d761.
22. Graffelman, J. and J.M. Camarena, Graphical tests for Hardy-Weinberg equilibrium based on the ternary plot. Hum Hered, 2008. **65**(2): p. 77-84.
23. Schwender, H. and K. Ickstadt, Identification of SNP interactions using logic regression. Biostatistics, 2008. **9**(1): p. 187-98.
24. Fay, M.P., *et al.*, Age-conditional probabilities of developing cancer. Stat Med, 2003. **22**(11): p. 1837-48.
25. Fay, M.P., Estimating age conditional probability of developing disease from surveillance data. Popul Health Metr, 2004. **2**(1): p. 6.
26. Dupont, W.D. and W.D. Plummer, Jr., Understanding the relationship between relative and absolute risk. Cancer, 1996. **77**(11): p. 2193-9.
27. Graffelman, J. and B.S. Weir, Testing for Hardy-Weinberg equilibrium at biallelic genetic markers on the X chromosome. Heredity (Edinb), 2016. **116**(6): p. 558-68.
28. Aly, M., *et al.*, Polygenic risk score improves prostate cancer risk prediction: results from the Stockholm-1 cohort study. Eur Urol, 2011. **60**(1): p. 21-8.