Asian Journal of Andrology (2016) 18, 525–529

www.asiaandro.com; www.ajandrology.com

Open Access

INVITED ORIGINAL ARTICLE

Prostate Cancer

Race-specific genetic risk score is more accurate than nonrace-specific genetic risk score for predicting prostate cancer and high-grade diseases

Rong Na[1,2,3,4,*], Dingwei Ye[5,*], Jun Qi[6,*], Fang Liu[2], Xiaoling Lin[2], Brian T Helfand[7], Charles B Brendler[7], Carly Conran[3], Jian Gong[7], Yishuo Wu[1,2], Xu Gao[8], Yaqing Chen[9], S Lilly Zheng[3], Zengnan Mo[10], Qiang Ding[1,2], Yinghao Sun[8], Jianfeng Xu[2,3]

Genetic risk score (GRS) based on disease risk-associated single nucleotide polymorphisms (SNPs) is an informative tool that can be used to provide inherited information for specific diseases in addition to family history. However, it is still unknown whether only SNPs that are implicated in a specific racial group should be used when calculating GRSs. The objective of this study is to compare the performance of race-specific GRS and nonrace-specific GRS for predicting prostate cancer (PCa) among 1338 patients underwent prostate biopsy in Shanghai, China. A race-specific GRS was calculated with seven PCa risk-associated SNPs implicated in East Asians (GRS7), and a nonrace-specific GRS was calculated based on 76 PCa risk-associated SNPs implicated in at least one racial group (GRS76). The means of GRS7 and GRS76 were 1.19 and 1.85, respectively, in the study population. Higher GRS7 and GRS76 were independent predictors for PCa and high-grade PCa in univariate and multivariate analyses. GRS7 had a better area under the receiver-operating curve (AUC) than GRS76 for discriminating PCa (0.602 *vs* 0.573) and high-grade PCa (0.603 *vs* 0.575) but did not reach statistical significance. GRS7 had a better (up to 13% at different cutoffs) positive predictive value (PPV) than GRS76. In conclusion, a race-specific GRS is more robust and has a better performance when predicting PCa in East Asian men than a GRS calculated using SNPs that are not shown to be associated with East Asians.

Asian Journal of Andrology (2016) **18**, 525–529; doi: 10.4103/1008-682X.179857; published online: 3 May 2016

Keywords: Chinese; genetic risk score; genome-wide association study; prostate cancer; single nucleotide polymorphism

INTRODUCTION

Prostate cancer (PCa) is one of the most common cancers worldwide.[1] In China, the incidence of PCa is relatively low; however, it has increased rapidly in recent decades.[2,3] Inherited risk of developing the disease is one of the most important risk factors for determining the pathogenesis of PCa. Up to 42% of the disease risk could be explained by heritable factors.[4] Positive family history has been shown to be strongly associated with PCa[5] and is widely used in clinical practice for risk assessment of PCa. However, family history information may be influenced by family size, age, survival status of patient's relatives, family communication, recall abilities, etc. In addition, family history must be continually assessed as family history status may change.

Genetic risk scores (GRSs) as measures of inherited risk have been repeatedly shown to provide additional information to family history when assessing one's risk of developing PCa.[6–8] These GRSs are calculated based on the genotypes of PCa risk-associated single nucleotide polymorphisms (SNPs) implicated from genome-wide association studies (GWASs). To date, GWASs have found more than 100 SNPs associated with PCa; however, most of these studies were conducted primarily in Caucasian men.[9] As such, the vast majority of these PCa risk-associated SNPs were not found to be significantly associated with PCa in Chinese men.[10] A recent study demonstrated the predictive performance of GRS in several racial groups using all established PCa risk-associated SNPs.[11] Although these nonrace-specific GRSs could predict PCa risk, using nonrace-specific SNPs has the potential to lead to over- or under-estimates of disease risk. Whether or not a race-specific GRS (calculated only using SNPs that were significantly associated with the disease in a defined population) is more accurate for predicting disease risks remains unclear. In this study, we compared the performances of two GRSs for predicting PCa and high-grade PCa in a prostate biopsy cohort. The two GRSs were based on (a) East Asian population-specific (Chinese and Japanese), disease-associated SNPs and (b) disease-associated SNPs regardless of race information.

[1]Department of Urology, Huashan Hospital, Fudan University, Shanghai, China; [2]Fudan Institute of Urology, Huashan Hospital, Fudan University, Shanghai, China; [3]Program for Personalized Cancer Care, NorthShore University HealthSystem, Evanston, Illinois, USA; [4]Health Communication Institute, School of Public Health, Fudan University, Shanghai, China; [5]Department of Urology, Shanghai Cancer Center, Fudan University, Shanghai, China; [6]Department of Urology, Xinhua Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China; [7]Division of Urology, NorthShore University HealthSystem, Evanston, Illinois, USA; [8]Department of Urology, Shanghai Changhai Hospital, The Second Military Medical University, Shanghai, China; [9]Department of Medical Ultrasonic, Xinhua Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China; [10]Center for Genomic and Personalized Medicine, Guangxi Medical University, Nanning, China.
*These authors contributed equally to this study.
Correspondence: Dr. J Xu (jxu@northshore.org) or Dr. Y Sun (sunyhsmmu@126.com)

MATERIALS AND METHODS

Study population and study design

This was a multicenter study of a biopsy cohort from four tertiary medical centers in Shanghai (Huashan Hospital, Fudan University; Shanghai Cancer Center, Fudan University; Xinhua Hospital, Shanghai Jiao Tong University School of Medicine; Changhai Hospital, the Second Military Medical University), China. Consecutive patients (n = 1 617) undergoing initial prostate biopsy at these four centers were enrolled from August 2013 to December 2014. Written informed consent was obtained from each patient. The study was approved by the Institutional Review Board of each medical center.

Indications for prostate biopsy included (a) tPSA >10.0 ng ml^{-1}, (b) tPSA >4.0 ng ml^{-1} (with a confirmation after 2–3 months), (c) %f PSA <0.16 when patients had a suspicious total PSA level (>4.0 ng ml^{-1}), and (d) suspicious lesions detected by digital rectal examination (DRE) or ultrasound at any level of tPSA. Demographic and clinical information were documented before biopsy, including age, total prostate-specific antigen (tPSA) level, and free PSA (fPSA) level. Biopsy specimens were analyzed in the Department of Pathology at each hospital. Prostate cancer diagnosis and high-grade disease (defined as Gleason Score ≥7) were recorded. Patients were excluded from the study analyses if (a) records of pathological diagnosis were missing or (b) tPSA, fPSA, or p2PSA were unable to be tested due to poor blood sample quality. Thus, 79 patients were excluded, and 1538 patients were included for further analyses.

Genotyping

Blood samples were collected for extracting DNA. DNA samples were genotyped using Illumina BeadXpress platform with the Golden Gate SNP genotyping assay for 80 SNPs (**Supplementary Table 1**). All of the candidate SNPs were found to be significantly associated with PCa in Caucasian, of which seven SNPs were significantly associated (reached genome-wide significant level of $P < 5 \times 10^{-8}$) with PCa in East Asian (Chinese and Japanese) populations. Two-hundred samples (13.0%) and four SNPs failed to be genotyped because of DNA quality and assay design. The remaining 1338 samples reached the SNP call rate >95%.

Calculation of GRSs

Two GRSs based on 76 SNPs (GRS76; using all the SNPs regardless of race) and seven SNPs (GRS7; using race-specific SNPs) were calculated. Briefly, the allelic OR of each SNP was first obtained from external studies.[9] Second, a genotypic OR of each SNP was calculated from the allelic OR based on a multiplicative model (carrying two risk alleles at one locus, RR: OR^2; carrying one risk allele at the locus, RN: OR; and carrying two nonrisk alleles at the locus, NN: 1). Third, the risk relative to the average risk in the population was calculated based on genotypic OR and risk allele frequency (1000 Genome Project, CHB population) for each SNP in the Chinese populations; the final GRS was calculated by multiplying the risk of each SNP. Theoretically, individuals with GRS of 1.0 are considered to be at average risk of developing a disease compared to other members of their race at large. Individuals with a GRS lower or higher than 1.0 are considered to be at decreased or increased risk of developing the disease, respectively, compared to the general population (defined by race).

Statistical analysis

In univariate analyses, Student's t-tests and Mann–Whitney U-tests were used to compare different variables among groups for normal distribution variables and nonnormal distribution variables, respectively. GRSs were adjusted by age, tPSA, and fPSA using logistic regression in multivariate analyses. Areas under the receiver operating characteristic curve (AUC of ROC) analyses were used to evaluate the predictive values of GRS7 and GRS76. The AUCs were compared by Z-test. The differences between GRS7 and GRS76 at specific cutoff values were described using positive predictive values (PPVs). Net reclassification improvement analyses (NRIs) were used to evaluate the improvement of GRS7 from GRS76. All statistical analyses were performed using SPSS 22.0 (IBM, North Castle, NY, USA). Two-tailed $P < 0.05$ was considered statistically significant.

RESULTS

Characteristics of the study population and univariate analyses between each group are shown in **Table 1**. Age at diagnosis (mean: 67.59 *vs* 64.44, $P = 8.51 \times 10^{-6}$), tPSA (median: 25.61 *vs* 8.80, $P = 1.73 \times 10^{-73}$), GRS76 (median: 1.90 *vs* 1.50, $P = 7.76 \times 10^{-6}$), and GRS7 (median: 1.22 *vs* 0.99, $P = 3.47 \times 10^{-10}$) were significantly higher in PCa group than in non-PCa group while %fPSA was significantly lower in PCa group (median: 0.11 *vs* 0.16, $P = 2.50 \times 10^{-37}$). Similar results were found between high-grade PCa group and others (low-grade PCa and non-PCa). In multivariate analyses, both GRS76 (odds ratio, OR = 1.10 for PCa, $P = 4.02 \times 10^{-4}$; OR = 1.07 for high-grade PCa, $P = 0.014$) and GRS7 (OR = 1.45 for PCa, $P = 9.84 \times 10^{-6}$; OR = 1.34 for high-grade PCa, $P = 0.001$) were significant independent risk factors of PCa as well as high-grade PCa when adjusting for age, tPSA, and fPSA (**Table 2**).

To assess whether GRS of 1.0 represents average risk in the cohort, we calculated means of GRSs after excluding the extreme values (GRS above 75th percentile + 1.5 interquartile range, and GRS below 25th percentile - 1.5 interquartile range). For GRS7, the means were 1.19 in the entire population, 1.32 in PCa group, 1.11 in non-PCa group, 1.36 in high-grade PCa group, and 1.12 in nonhigh-grade PCa group. For GRS76, the means were 1.85 in the entire population, 1.99 in PCa group, 1.77 in non-PCa group, 2.04 in high-grade PCa group, and 1.77 in nonhigh-grade PCa group.

We then compared the performances of GRS76 and GRS7 for discriminating PCa and high-grade diseases (Gleason score ≥7). GRS7 (AUC = 0.602) had a better discriminative ability for PCa than GRS76 (AUC = 0.573) did; however, this difference did not reach a statistically significant level ($P = 0.20$). Similarly, GRS7 (AUC = 0.603) performed better for discriminating high-grade PCa than GRS76 (AUC = 0.575) did; however, no statistical significance was observed (**Table 3**). In NRI analyses, results showed that GRS7 slightly improved the predictive ability of PCa and high-grade diseases from GRS76 but did not reach statistical significance (**Table 4**). For instance, at the cutoff value of 1.5, the NRIs from GRS76 to GRS 7 were 0.033 for predicting PCa and 0.031 for high-grade PCa, indicating that using GRS7, there would be ~3% net improvement of predictive ability from GRS76.

Using PPV, we found that GRS7 had a better performance than GRS76 for predicting PCa and high-grade diseases at various cutoff values (**Figures 1** and **2**). For example, at a cutoff of 1.5, 29% of men were classified as higher risk using GRS7, with a PPV of 48.2% for predicting PCa. In comparison, 53% of men were classified as higher risk using GRS76, with a PPV of only 41.8% (**Figure 1**). The difference in PPV between the two GRSs was 6.4%. At a cutoff value of 2.0, 14% of men were classified as higher risk using GRS7, with a PPV of 56.4% for predicting PCa. In comparison, 40% of men were classified as higher risk using GRS76, with a PPV of only 43.9%. The difference in PPV between the two GRSs was even larger, at 12.5%, when using a cutoff of 2.0. Similar findings were observed for predicting high-grade PCa (**Figure 2**). At a cutoff of 1.5, the PPV of GRS7 was 40.5% for predicting high-grade PCa whereas the PPV of GRS76 was 34.5%. At a cutoff of 2.0, the PPV was 48.4% for GRS7 and 36.9% for GRS76.

Table 1: Characteristics of study population and the univariate analysis of each variable between PCa group and non-PCa group

	All	*PCa*	*Non-PCa*	*P*[a]	*High-grade PCa*	*Others*	*P*
Number of patients (%)	1338	499 (37.3)	839 (62.7)	-	403 (30.1)	935 (69.9)	-
Age (year, mean±s.d.)	65.61±12.55	67.59±14.08	64.44±11.38	8.51×10^{-6}	67.66±14.74	64.73±11.37	8.31×10^{-5}
tPSA (ng ml^{-1}, median, interquartile)	11.10 (6.95–23.67)	25.61 (11.31–98.47)	8.80 (5.79–13.63)	1.73×10^{-73}	36.55 (14.27–138.86)	9.02 (5.95–14.01)	1.01×10^{-82}
fPSA (ng ml^{-1}, median, interquartile)	1.66 (0.96–3.24)	2.69 (1.30–9.12)	1.42 (0.83–2.26)	5.57×10^{-34}	3.48 (1.50–13.99)	1.42 (0.83–2.26)	7.58×10^{-46}
%fPSA (median, interquartile)	0.13 (0.10–0.20)	0.11 (0.08–0.13)	0.16 (0.11–0.22)	2.50×10^{-37}	0.11 (0.08–0.13)	0.16 (0.11–0.22)	9.18×10^{-32}
GRS76[a] (median, interquartile)	1.63 (0.89–2.90)	1.90 (1.00–3.38)	1.50 (0.82–2.70)	7.76×10^{-6}	1.97 (1.02–3.46)	1.51 (0.84–2.72)	1.35×10^{-5}
GRS7[a] (median, interquartile)	1.07 (0.71–1.64)	1.22 (0.81–1.89)	0.99 (0.68–1.46)	3.47×10^{-10}	1.24 (0.81–1.94)	1.01 (0.68–1.50)	1.91×10^{-9}

[a]GRS76: using all genotyped 76 PCa risk-associated SNPs; GRS7: using 7 race-specific PCa risk-associated SNPs with confirmation $P<5\times10^{-8}$. PCa: prostate cancer; PSA: prostate-specific antigen; tPSA: total PSA; fPSA: free PSA; %fPSA: percentage fPSA; s.d.: standard deviation; SNPs: single nucleotide polymorphisms

Table 2: Multivariate logistic regression analyses of GRSs adjusting for different variables

	PCa		*High-grade PCa*	
	OR (95% CI)	*P*	*OR (95% CI)*	*P*
GRS76[a]	1.10 (1.04–1.16)	4.02×10^{-4}	1.07 (1.01–1.13)	0.014
Age	1.03 (1.02–1.05)	5.91×10^{-6}	1.03 (1.01–1.04)	1.05×10^{-4}
tPSA	1.08 (1.06–1.09)	5.91×10^{-24}	1.07 (1.06–1.08)	2.09×10^{-23}
fPSA	0.76 (0.69–0.84)	4.35×10^{-8}	0.80 (0.73–0.88)	3.82×10^{-6}
GRS7[a]	1.45 (1.23–1.70)	9.84×10^{-6}	1.34 (1.14–1.59)	0.001
Age	1.03 (1.02–1.05)	7.17×10^{-6}	1.03 (1.01–1.04)	1.08×10^{-4}
tPSA	1.08 (1.06–1.09)	1.58×10^{-23}	1.07 (1.06–1.08)	6.03×10^{-23}
fPSA	0.77 (0.70–0.84)	6.93×10^{-8}	0.81 (0.73–0.88)	5.52×10^{-6}

[a]GRS76: using all genotyped 76 PCa risk-associated SNPs; GRS7: using 7 race-specific PCa risk-associated SNPs with confirmation $P<5\times10^{-8}$. OR: odds ratio; PCa: prostate cancer; PSA: prostate-specific antigen; tPSA: total PSA; fPSA: free PSA; CI: confidence interval; SNPs: single nucleotide polymorphisms

Table 3: AUCs of receiver operating curve analyses of each GRS for predicting PCa and high-grade PCa

	PCa			*High-grade PCa*		
	AUC (95% CI)	*SE*	*P*	*AUC (95% CI)*	*SE*	*P*
GRS76[a]	0.573	0.016	0.20	0.575	0.017	0.24
GRS7[a]	0.602	0.016		0.603	0.017	

[a]GRS76: using all genotyped 76 PCa risk-associated SNPs; GRS7: using 7 race-specific PCa risk-associated SNPs with confirmation $P<5\times10^{-8}$. AUC: area under the curve; CI: confidence interval; PCa: prostate cancer; SE: standard error; SNPs: single nucleotide polymorphisms

Table 4: Net reclassification improvement from GRS76[a] to GRS7[a]

NRIs	*PCa*		*High-grade PCa*	
	NRIs	*P*[b]	*NRIs*	*P*[b]
Cutoff 1.0	0.051	0.09	0.053	0.09
Cutoff 1.5	0.033	0.31	0.031	0.35
Cutoff 2.0	0.0005	0.99	–0.007	1

[a]GRS76: using all genotyped 76 PCa risk-associated SNPs; GRS7: using 7 race-specific PCa risk-associated SNPs with confirmation $P<5\times10^{-8}$; [b]Comparing NRIs with zero. NRIs: net reclassification improvements; PCa: prostate cancer; SNPs: single nucleotide polymorphisms

DISCUSSION

In this study, we used genotyping data from a Chinese biopsy cohort to calculate two GRSs for each subject: one race-specific GRS (using seven SNPs previously shown to be significantly associated with East Asian men at $P < 5 \times 10^{-8}$) and one GRS that was based on all 76 SNPs previously reported to be significantly associated with PCa in any races. We found that while both GRS7 and GRS76 were significant predictors of PCa and high-grade PCa, GRS7 had slightly better, though nonstatistically significant, AUCs than GRS76 for discriminating PCa from non-PCa, and for discriminating high-grade PCa from nonhigh-grade PCa. More importantly, we found that GRS7 had considerably better PPVs than GRS76 for predicting PCa and high-grade PCa. Finally, we found that the mean of GRS7 (1.19) of the study population was closer to the expected value of 1.0 than GRS76 (1.85), a factor that is critical for risk assessment at an individual level to define higher or lower risk for PCa.

Multiple measurements can be used to assess whether a biomarker is a predictor of PCa risk at a population level, including a test for different means of a biomarker between cases and controls, AUC for discriminating cases from controls, and PPV for predicting probability of PCa among individuals classified as higher risk. These measurements in this study support that both race-specific GRS (GRS7) and nonrace-specific GRS (GRS76) can be used as predictors of PCa risk at a population level. However, based on the results of PPV, the race-specific GRS is a better choice. PPV is a more relevant measurement than AUC for the purpose of risk assessment for identifying high-risk men. Compared to GRS76, GRS7 identified fewer men at higher risk, but a higher proportion of them developed PCa and high-grade PCa, suggesting that GRS7 is more effective in identifying high-risk men for targeted intervention.

When assessing the performance of GRS for discriminating and predicting PCa, a mean GRS that closes to the expected value of 1.0 is critical for defining higher or lower risk for an individual. An advantage of the GRS used in this study is that it is population standardized, and the mean GRS is expected to be 1.0. This theoretical property is confirmed in a simulation study, in which up to 100 different risk-associated SNPs were simulated to be associated with disease risk (Yu *et al*. in the same issue). The study indicated when the OR and risk allele frequency of each SNP used in the calculation of GRS are correct (i.e., they are the same as simulated values), the mean GRS in the cohort was close to 1.0 regardless of the number of SNPs (top 30, 50, and 100 SNPs with highest ORs) used in the calculations of GRSs. The mean GRS, however, could deviate from 1.0 if ORs and risk allele frequencies used in the GRS calculation were different from the true simulated values. Therefore, observation of a mean GRS in the cohort close to 1.0 is important to ensure that the ORs and allele frequencies used in GRS calculation are appropriate. From this consideration, GRS7 is better than GRS76 because its mean is closer to 1.0.

It should be noted that although the mean of GRS7 in the cohort was close to 1.0, it was slightly higher than 1.0 (mean = 1.19). Two

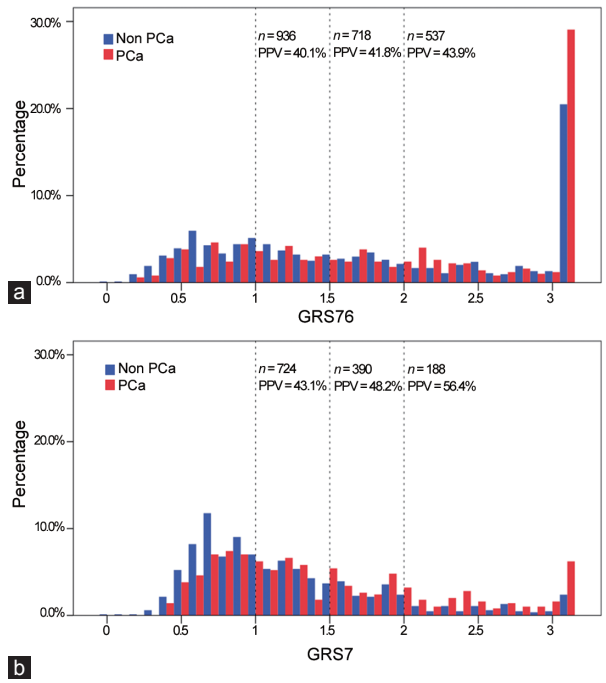

Figure 1: Distribution of PCa (red) and non-PCa (blue) at different GRS levels. Dash lines highlight the cutoff values of 1.0, 1.5, and 2.0, respectively. n: the number of study population with GRS over the cutoffs; PPV: positive predictive value. (**a**) GRS76; (**b**) GRS7.

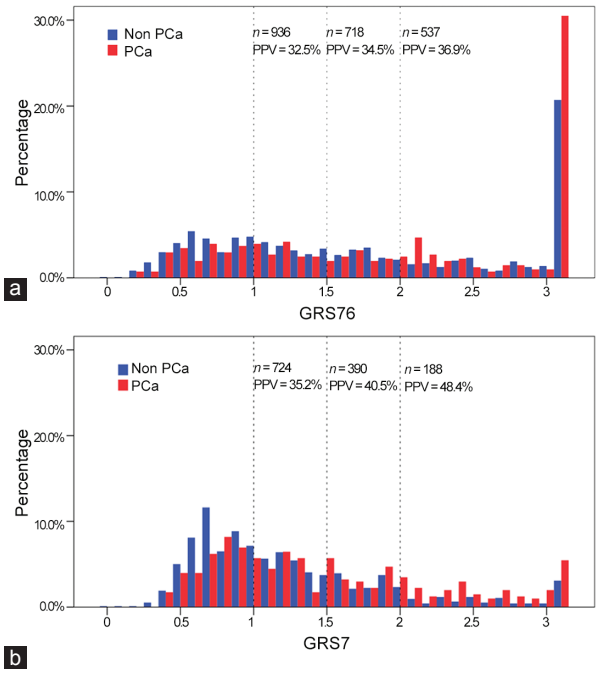

Figure 2: Distribution of high-grade PCa (red) and others (nonhigh-grade PCa, blue) at different GRS levels. Dash lines highlight the cutoff values of 1.0, 1.5, and 2.0, respectively. n: the number of study population with GRS over the cutoffs; PPV: positive predictive value. (**a**) GRS76; (**b**) GRS7.

possible reasons may account for this deviation. First, the estimates of ORs and risk allele frequencies of the seven SNPs used in GRS calculations may be over- or under-estimated. The estimated ORs of these SNPs are likely to be reliable because they were obtained from a large meta-analysis in East Asians.[9] The allele frequencies of these SNPs were obtained from a small sample (82 subjects) of the CHB (Chinese Han Beijing) population in the 1000 Genomes Project. These frequencies differed from the risk allele frequencies in this cohort, which ranged from 0.6% to 7.9% (**Supplementary Table 1**). The impact of over- and under-estimated OR and risk allele frequency is stronger on the GRS76 than on the GRS7 and may explain the larger deviation of its mean (1.85) from 1.0. The allele frequencies of these 76 SNPs were also from the CHB population of the 1000 Genomes Project. However, the differences in allele frequencies between the CHB population and our cohort ranged from 0.2% to 42% (**Supplementary Table 1**). Second, the higher mean GRS may indicate that men in this cohort have higher average genetic risk for PCa. This is plausible because a considerable proportion of men in this biopsy cohort have PCa. Furthermore, a subset of patients with negative biopsy results may also have PCa that could have been missed on initial needle biopsy.

A limitation of this study should be noticed. We only had genotypes of 80 SNPs rather than the ~100 PCa risk-associated SNPs that have now been identified. However, because the impact of any of these SNPs on GRS is small due to their modest ORs, missing a subset of these SNPs is unlikely to change the main conclusion of the study.

CONCLUSIONS

A race-specific GRS is more robust and accurate than a nonrace-specific GRS for assessing an individual's risk of developing PCa. Only risk-associated SNPs that have been previously confirmed in a specific racial group (*P* value reaching GWAS significant level) should be used for calculating GRS for a population of interest.

AUTHOR CONTRIBUTIONS

RN, DY, and JQ participated in statistical analyses and drafted the manuscript; FL, YW, XG, YC, and JG participated in acquisition of data and gave technical support; BTH, CBB, and CC participated in critical revision of the manuscript; SLZ and ZM participated in administrative, technical and material support; QD participated in study design; YS and JX participated in study design and supervision of the study.

All authors have read and approved the final version of the manuscript and agree with the order of presentation of the authors.

COMPETING INTERESTS

There is no conflict of interests in this paper.

ACKNOWLEDGMENTS

This work was in part supported by grants from the Key Project of the National Science Foundation of China to Jianfeng Xu (81130047), the National Key Basic Research Program Grant 973 of China to Jianfeng Xu (2012CB518301), the National Natural Science Foundation of China (Grant No. 81402339) to Rong Na, the intramural grants from Huashan Hospital Fudan University to Rong Na. This study is also partially supported by the Ellrodt-Schweighauser Family Chair of Cancer Genomic Research of NorthShore University HealthSystem to JX. Finally, We would like to thank all the subjects included in this study.

Supplementary information is linked to the online version of the paper on the *Asian Journal of Andrology* website.

REFERENCES

1 Jemal A, Bray F, Center MM, Ferlay J, Ward E, *et al*. Global cancer statistics. *CA Cancer J Clin* 2011; 61: 69–90.

2 Qi D, Wu C, Liu F, Gu K, Shi Z, *et al.* Trends of prostate cancer incidence and mortality in Shanghai, China from 1973 to 2009. *Prostate* 2015; 75: 1662–8.
3 Zhou M, Wang H, Zhu J, Chen W, Wang L, *et al.* Cause-specific mortality for 240 causes in China during 1990-2013: a systematic subnational analysis for the global burden of disease study 2013. *Lancet* 2016; 387: 251–72.
4 Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, *et al.* Environmental and heritable factors in the causation of cancer – Analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000; 343: 78–85.
5 Johns LE, Houlston RS. A systematic review and meta-analysis of familial prostate cancer risk. *BJU Int* 2003; 91: 789–94.
6 Aly M, Wiklund F, Xu J, Isaacs WB, Eklund M, *et al.* Polygenic risk score improves prostate cancer risk prediction: results from the Stockholm-1 cohort study. *Eur Urol* 2011; 60: 21–8.
7 Kader AK, Sun J, Reck BH, Newcombe PJ, Kim ST, *et al.* Potential impact of adding genetic markers to clinical parameters in predicting prostate biopsy outcomes in men following an initial negative biopsy: findings from the REDUCE trial. *Eur Urol* 2012; 62: 953–61.
8 Sun J, Na R, Hsu FC, Zheng SL, Wiklund F, *et al.* Genetic score is an objective and better measurement of inherited risk of prostate cancer than family history. *Eur Urol* 2013; 63: 585–7.
9 Wang M, Takahashi A, Liu F, Ye D, Ding Q, *et al.* Large-scale association analysis in Asians identifies new susceptibility loci for prostate cancer. *Nat Commun* 2015; 6: 8469.
10 Na R, Liu F, Zhang P, Ye D, Xu C, *et al.* Evaluation of reported prostate cancer risk-associated SNPs from genome-wide association studies of various racial populations in Chinese men. *Prostate* 2013; 73: 1623–35.
11 Hoffmann TJ, Van Den Eeden SK, Sakoda LC, Jorgenson E, Habel LA, *et al.* A large multiethnic genome-wide association study of prostate cancer identifies novel risk variants and substantial ethnic differences. *Cancer Discov* 2015; 5: 878–91.

Supplementary Table 1: List of genotyped SNPs in the present study

SNP ID	*Chromosome*	*Position (GRCh38)*	*Risk allele*	*Frequency in CHB*	*Frequency in the study population*	*Differences of frequency*[a]
rs10993994	10q11.23	51219502	T	0.418	0.483	0.066
rs12653946	5p15.33	1948829	T	0.350	0.356	0.006
rs1447295	8q24.21	127472793	A	0.121	0.154	0.033
rs1512268	8p21.2	23582408	T	0.359	0.302	–0.057
rs16901979	8q24.21	127112671	A	0.248	0.300	0.053
rs339331	6q22.1	117316745	T	0.641	0.671	0.030
rs6983267	8q24.21	127401060	G	0.388	0.467	0.079
rs10009409	4q13.3	74074117	T	0.577	0.505	–0.072
rs10086908	8q24.21	128081119	T	0.806	0.808	0.002
rs10187424	2p11.2	85647808	T	0.646	0.607	–0.039
rs103294	19q13.42	59489660	C	0.243	0.265	0.022
rs1041449	21q22.3	41823291	G	0.222	0.181	–0.041
rs10486567	7p15.2	27943088	A	0.876	0.853	–0.023
rs10875943	12q13.12	47962277	C	0.196	0.149	–0.047
rs10896449	11q13.3	68751243	G	0.077	0.080	0.003
rs10934853	3q21.3	129521063	A	0.407	0.428	0.021
rs10936632	3q26.2	171612796	A	0.206	0.291	0.085
rs11135910	8p21.2	25948059	T	0.031	0.021	–0.010
rs11568818	11q22.2	101906871	T	0.103	0.081	–0.022
rs11649743	17q12	33149092	G	0.660	0.673	0.013
rs11902236	2p25.1	10035319	T	0.124	0.133	0.009
rs12051443	16q22.2	71657426	A	0.778	0.500	–0.278
rs12155172	7p15.3	20961016	A	0.325	0.357	0.032
rs1218582	1q21.3	153100807	G	0.093	0.106	0.013
rs12418451	11q13.3	69167951	G	0.041	0.462	0.420
rs12480328	20q13.13	48961329	T	0.082	0.074	–0.008
rs12621278	2q31.1	173019799	A	0.258	0.275	0.017
rs1270884	12q24.21	113169954	A	0.289	0.286	–0.002
rs13385191	2p24.1	20751746	G	0.481	-	-
rs1465618	2p21	43407453	T	0.644	0.739	0.095
rs1571801	9q33.2	121665094	A	0.088	0.081	–0.006
rs16902094	8q24.21	128389528	G	0.201	0.259	0.058
rs17021918	4q22.3	94641726	C	0.366	0.353	–0.013
rs17599629	1q21.3	148924911	G	0.134	0.114	–0.020
rs17694493	9p21.3	22031998	G	0.031	0.037	0.006
rs1775148	1q32.1	205788696	A	0.459	-	-
rs1859962	17q24.3	66620348	G	0.418	0.404	–0.014
rs1894292	4q13.3	74568022	G	0.299	0.365	0.066
rs1933488	6q25.2	153482772	A	0.186	0.140	–0.046
rs1938781	11q12.1	58671686	G	0.286	0.332	0.045
rs1983891	6p21.1	41644405	T	0.384	0.350	–0.034
rs2055109	3p11.2	87550022	C	0.083	0.076	–0.006
rs2121875	5p12	44401302	C	0.500	0.513	0.013
rs2252004	10q26.12	122834699	C	0.767	0.722	–0.045
rs2273669	6q21	109391882	G	0.052	0.068	0.016
rs2292884	2q37.3	238107965	G	0.304	0.290	–0.014
rs2405942	Xp22.2	9774135	A	0.088	-	-
rs2427345	20q13.33	60449006	C	0.196	0.164	–0.032
rs2660753	3p12.1	87193364	T	0.320	0.301	–0.019
rs2735839	19q13.33	56056435	G	0.418	0.302	–0.115
rs2807031	Xp11.22	52913674	C	0.010	0.055	0.045
rs2928679	8p21.2	23494920	A	0.113	0.115	0.002
rs3771570	2q37.3	242031537	T	0.113	0.128	0.015
rs4245739	1q32.1	202785465	A	0.046	0.037	–0.009
rs4430796	17q21.2	37738049	A	0.743	0.741	–0.002
rs4713266	6p24.2	11327016	C	0.175	0.214	0.039
rs4962416	10q26.13	126686862	-	-	-	-

Contd...

Supplementary Table 1: Contd...

SNP ID	*Chromosome*	*Position (GRCh38)*	*Risk allele*	*Frequency in CHB*	*Frequency in the study population*	*Differences of frequency*[a]
rs5759167	22q13.2	41830156	G	0.675	0.697	0.022
rs5945619	Xp11.22	51498820	C	0.036	0.073	0.036
rs6062509	20q13.33	61833007	T	0.644	0.653	0.009
rs620861	8q24.21	127323428	G	0.519	0.564	0.045
rs6465657	7q21.3	97654263	C	0.874	0.854	–0.020
rs6625711	Xq13.1	70056575	A	0.485	0.498	0.013
rs6763931	3q23	142585523	A	0.340	0.358	0.018
rs684232	17p13.3	565715	C	0.552	0.515	–0.036
rs6869841	5q35.2	172872032	T	0.186	0.268	0.082
rs7141529	14q24.1	68196497	C	0.160	0.171	0.012
rs7153648	14q23	60655808	C	0.160	0.170	0.010
rs721048	2p15	62985235	A	0.041	0.051	0.009
rs7241993	18q23	74874961	C	0.402	0.393	–0.009
rs7611694	3q13.2	114758314	A	0.701	0.683	–0.018
rs7679673	4q24	106280983	C	0.160	0.241	0.081
rs8014671	14q24.2	70162009	G	0.268	0.301	0.033
rs8102476	19q13.2	43427453	C	0.701	0.614	–0.087
rs817826	9q31.2	109196121	C	0.078	0.102	0.024
rs887391	19q13.2	46677464	T	0.356	0.398	0.042
rs9287719	2p25.1	10628181	C	0.397	0.385	–0.012
rs9364554	6q25.3	160753654	T	0.376	0.335	–0.042
rs9600079	13q22.1	72626140	T	0.466	0.477	0.011
rs9623117	22q13.1	38782065	C	0.031	0.033	0.002

[a]Differences of frequency=frequency in study population – frequency in CHB (1000 Genome Project). SNP: single nucleotide polymorphism; Gray part: 7 SNPs used in GRS7; [a]GRS76: using all genotyped 76 PCa risk-associated SNPs; CHB: Chinese Han Beijing