

available at www.sciencedirect.comjournal homepage: www.eu-openscience.europeanurology.com

European Association of Urology



Prostate Cancer

Reliability of Ancestry-specific Prostate Cancer Genetic Risk Score in Four Racial and Ethnic Populations

Zhuqing Shi^{a,†}, Jianan Zhan^{b,†}, Jun Wei^a, Skylar Ladson-Gary^a, Chi-Hsiung Wang^a, Peter J. Hulick^c, S. Lilly Zheng^a, Kathleen A. Cooney^d, William B. Isaacs^e, Brian T. Helfand^{a,f,g}, 23andMe Research Team,^b Bertram L. Koelsch^b, Jianfeng Xu^{a,f,g,*}

^a Program for Personalized Cancer Care, NorthShore University HealthSystem, Evanston, IL, USA; ^b 23andMe Inc, Sunnyvale, CA, USA; ^c Department of Medicine, NorthShore University HealthSystem, Evanston, IL, USA; ^d Department of Medicine, Duke University School of Medicine and Duke Cancer Institute, Durham, NC, USA; ^e Department of Urology and the James Buchanan Brady Urologic Institute, Johns Hopkins University School of Medicine, Baltimore, MD, USA; ^f Department of Surgery, NorthShore University HealthSystem, Evanston, IL, USA; ^g Department of Surgery, University of Chicago Medicine, Chicago, IL, USA

Article info

Article history:

Accepted September 2, 2022

Associate Editor:

Jochen Walz

Keywords:

Calibration
Ethnicity
Genetic risk score
Race

Abstract

Background: Reliability of prostate cancer (PCa) genetic risk score (GRS), that is, the concordance between its estimated risk and observed risk, is required for genetic testing at the individual level. Reliability data are lacking for non-European racial/ethnic populations, which hinders its clinical use and exacerbates racial disparity. **Objective:** To calibrate PCa ancestry-specific GRS in four racial/ethnic populations. **Design, setting, and participants:** PCa ancestry-specific GRSs, calculated from published risk-associated single-nucleotide polymorphisms in corresponding racial/ethnic populations, were evaluated in men who participated in 23andMe, Inc. genetic testing and consented for research, including 888 086 of European (EUR), 81 109 of Hispanic (HIS), 30 472 of African (AFR), and 13 985 of East Asian (EAS) ancestry, as classified by 23andMe's ancestry composition algorithm. **Outcome measurements and statistical analysis:** The concordance between the observed and estimated PCa risks at ten ancestry-specific GRS deciles was measured primarily by using the calibration slope (β), where 1 represents a perfect calibration. Platt scaling was used to correct the systematic bias of GRS. **Results and limitations:** A linear trend of an increased observed PCa prevalence in men with higher ancestry-specific GRS deciles was found in each racial population (all $p_{\text{-trend}} < 0.001$). A calibration analysis revealed a systematic bias of GRS; β was considerably lower than 1 (0.73, 0.64, 0.66, and 0.75 in EUR, HIS, AFR, and EAS ancestries, respectively). This bias was reduced after the Platt scaling correction: β for scaled GRS in the testing dataset (40% of individuals) approximated 1 for all groups (0.95, 1.05, 1.02, and 1.01 in EUR, HIS, AFR, and EAS populations, respectively). The generalizability of the Platt correction needs to be validated in independent cohorts.

[†] These authors contributed equally to this work.

* Corresponding author. 1001 University Place, Evanston, IL 60201, USA. Tel. +1 (224) 264-7501; Fax: +1 (224) 364-7675.

E-mail address: jxu@northshore.org (J. Xu).

<https://doi.org/10.1016/j.euros.2022.09.001>

2666-1683/© 2022 The Author(s). Published by Elsevier B.V. on behalf of European Association of Urology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Conclusions: A systematic bias of ancestry-specific GRS in the direction of an over-estimated risk for men in the highest decile was found in EUR and non-EUR populations. GRS is well calibrated after correction and is appropriate for genetic testing at the individual level for personalized PCa screening.

Patient summary: A corrected genetic risk score is more reliable (supported by the observed prostate cancer [PCa] risk) and appropriate for genetic testing for personalized PCa screening.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of European Association of Urology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Prostate cancer (PCa) is a major public health concern; in the USA, the lifetime risks of developing and dying of PCa are approximately 11% and 2.5%, respectively [1,2]. While prostate-specific antigen (PSA)-based screening programs reduce PCa mortality by ~20% [3], these also lead to over-diagnosis and overtreatment of PCa [4]. Accordingly, risk-based PSA screening is recommended by several guidelines [5,6]. In addition to the well-established risk factors of a family history of cancer and African (AFR) ancestry, the National Comprehensive Cancer Network guidelines also recommend germline testing of ten major genes. However, pathogenic germline mutations in these major genes are rare, detected in ~2% of the general population [7].

In contrast, >269 independent risk-associated single-nucleotide polymorphisms (SNPs) have been identified from multiple genome-wide association studies (GWASs) since 2007 [8]. The cumulative effect of these SNPs, measured by polygenic risk score methods, complements risk assessment of family history and major genes [8–15]. For example, in a study comparing three inherited risk measures in a PCa incidence cohort from a population-based UK Biobank (UKB), combination of family history and germline mutations in major genes (*BRCA2*, *HOXB13*, and *ATM*) identified 11% of men at a higher inherited PCa risk. Addition of a polygenic risk score identified an additional 15% of men at a higher PCa risk, and their observed PCa incidence and mortality rates were similar to men with a family history and germline mutations of major genes [12].

Despite the consistent results from large research populations, polygenic risk score testing has not been adopted by guidelines for risk stratification at the individual patient level [5,6]. Several outstanding challenges are generally cited [16], including (1) a lack of supporting data from prospective studies, (2) reliability of their risk estimates, and (3) generalizability in non-European racial/ethnic populations [17]. The first challenge was recently addressed in a large PCa prospective cohort from the UKB [12,18]. The reliability of a polygenic risk score, that is, the concordance between the estimated and observed risks, was also demonstrated in the UKB [12]. The top 10% of men in the UKB have a mean estimated risk of 2.29 for PCa, corroborated by their observed PCa risk of 2.43.

However, the reliability of a genetic risk score (GRS) in non-European racial/ethnic populations was not assessed in the above UKB study (because ~96% are of European [EUR] ancestry) [18] and in other studies. Several recent

transancestry studies demonstrated the performance of a cross-ancestry polygenic risk score in multiple racial/ethnic populations (EUR, AFR, East Asian [EAS], Hispanic [HIS], and South Asian) [8,19,20]. However, these studies demonstrated only the validity of polygenic risk score percentile at a population level (ie, higher percentiles and higher risks), not the validity of risk estimate (ie, concordance between the estimated and observed risks). The latter is required for calculating an individual's relative and lifetime risk. A lack of reliability data of polygenic risk scores in non-European populations hinders the broad implementation of polygenic risk scores in genetic testing and further exacerbates the racial/ethnic disparity in PCa screening [16,17]. This disparity is particularly prominent for men of AFR ancestry who are at the highest risk of developing and dying from PCa [21].

The objective of the current study is to assess the reliability of an ancestry-specific GRS in a large number of individuals from four racial/ethnic populations, including EUR and three non-European (AFR, HIS, and EAS) populations.

2. Patients and methods

2.1. Study population

Individuals included in this study are male consumers of 23andMe, Inc., a direct-to-consumer genetics company, who were genotyped as part of the 23andMe Personal Genome Service and provided informed consent to allow their aggregate and deidentified data to be used for research. The protocol was reviewed and approved by Ethical & Independent Review Services, a private institutional review board (<http://www.eandireview.com>). PCa case and control status was self-reported from an online survey. Owing to the late age of diagnosis of PCa, only men aged ≥ 50 yr were included in the analysis.

DNA samples were genotyped using several customized Illumina SNP arrays, and imputed using the Minimac3 software package (version 1.0.13) [22]. Individuals of EUR, HIS, AFR, and EAS ancestries, determined using 23andMe's ancestry composition algorithm [23], were included in this study. The numbers of EUR, HIS, AFR, and EAS men were 888 086, 81 109, 30 472, and 13 985, respectively (Table 1). A principal component analysis was performed independently for each ancestry, using ~65 000 high-quality genotyped variants, to measure genetic architecture of the participants.

2.2. Ancestry-specific GRS

The ancestry-specific PCa GRS was calculated based on PCa risk-associated SNPs in the four populations. These SNPs were identified from an evidence-based review of published GWASs that met the following criteria: (1) discovered from PCa GWASs and confirmed in additional stages

Table 1 – Key characteristics of study participants by racial/ethnic population

	EUR	HIS	AFR	EAS
No. of participants	888 086	81 109	30 472	13 985
Age (yr), median (IQR)	65 (57–73)	61 (55–69)	61 (55–70)	61 (54–69)
No. (%) of patients with PCa	53 220 (6.0)	3100 (3.8)	2267 (7.4)	305 (2.2)
Age at PCa diagnosis (yr), median (IQR)	63 (57–68)	62 (56–68)	60 (54–66)	65 (60–70)
GRS, mean (95% CI)	0.98 (0.98–0.98)	0.92 (0.91–0.92)	0.85 (0.84–0.86)	0.93 (0.91–0.95)

AFR = African; CI = confidence interval; EUR = European; GRS = genetic risk score; HIS = Hispanic; IQR = interquartile range; PCa = prostate cancer.

with combined $p < 5 \times 10^{-8}$ in at least one racial/ethnic population, (2) $p < 0.05$ in a racial population and with the same direction of association as GWASs, (3) linkage disequilibrium (LD) measurement ($r^2 < 0.2$) between any pair of SNPs within each population, and (4) available in our study (genotyped or imputed). The numbers of risk-associated SNPs and odds ratio (OR) used for calculating ancestry-specific GRSs were 232, 138, 128, and 67 for EUR, EAS, AFR, and HIS populations, respectively (Supplementary Tables 1–4).

A GRS was calculated by multiplying the per-allele OR for each SNP and normalizing the risk by the average risk expected in the population of specific races [24]. As the GRS is population standardized, its value can be considered as an individual's relative risk (RR) compared with the general population.

2.3. Statistical methods

The association of an ancestry-specific GRS with PCa risk was first tested by comparing the mean GRS value between cases and controls in each population using a logistic regression analysis with or without adjusting for age and the ten principal components. Furthermore, a dose-response association between higher GRS deciles and a higher observed PCa prevalence in each population was tested using a chi-square test for linear trend.

The reliability of an ancestry-specific GRS value (ie, estimated risk) was assessed by its concordance with the observed risk in individuals at each decile, measured by their RR to all individuals in each population. A calibration slope (β) was estimated from the regression line of the ten data points of deciles, with $\beta = 1.00$ representing perfect calibration [24]. In addition, a bias score (the mean absolute difference between the estimated and observed risks in ten deciles) was also estimated to measure the difference between the two risks. A bias score of 0 represents a perfect calibration [24].

Platt scaling was used to correct a systematic bias of a GRS [18,24–26]. Briefly, in each population, a logistic regression between a logarithm GRS and PCa status was performed in the training dataset (60% of randomly selected patients). The regression coefficient and intercept were then used to calculate scaled GRS values for patients in the testing dataset (the remaining 40% of participants from each cohort). The key characteristics were similar between the participants in the training and testing datasets (Supplementary Table 5).

2.4. Procedure

This project is a collaboration between NorthShore University HealthSystem (NorthShore) and 23andMe. Based on the considerations of (1) consent form (23andMe can share only aggregate and deidentified data) and (2) objectivity of an ancestry-specific GRS, we developed an unbiased approach for executing the project. The NorthShore team provided the method for calculating an ancestry-specific GRS to the 23andMe team, including GRS formula, ancestry-specific risk-associated

SNPs, and their OR and allele frequency in each population. The 23andMe team calculated the ancestry-specific GRS for each individual and provided the aggregated results to NorthShore.

3. Results

The prevalence of PCa, age at diagnosis, and mean GRS by race/ethnicity are presented in Table 1. The PCa prevalence was highest in AFR (7.44%) population, followed by EUR (5.99%), HIS (3.82%), and EAS (2.18%) populations. The mean age at PCa diagnosis (years) was lowest in AFR (60) population, followed by HIS (62), EUR (63), and EAS (65) populations. The mean ancestry-specific GRSs were 0.98, 0.93, 0.92, and 0.85 in EUR, EAS, HIS, and AFR populations, respectively.

The ancestry-specific GRS in each population was significantly associated with PCa risk; the mean GRS was significantly higher in men with than in those without PCa in each population, all $p < 0.001$ (Table 2). A higher GRS was significantly associated with an increased PCa risk in each population, with or without adjusting for age and top ten principal components in each population; OR ranged from 1.34 to 1.50 (all $p < 0.001$).

Furthermore, a dose-response association was found between higher ancestry-specific GRS deciles and the higher observed PCa prevalence; the linear trend was statistically significant in each population (all $p_{\text{trend}} < 0.001$; Fig. 1). PCa prevalence was highest for AFR among all racial/ethnic populations in each decile.

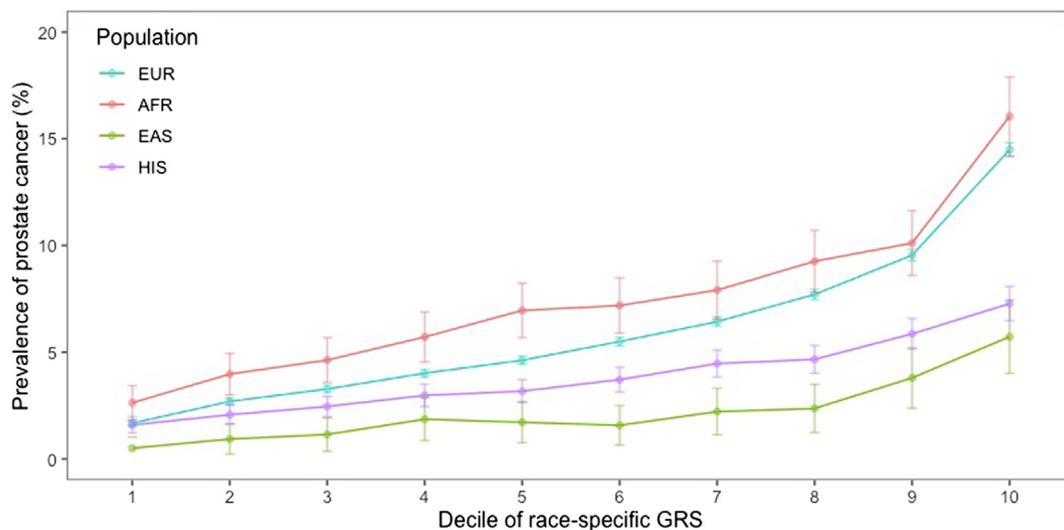
When assessing the concordance between the estimated PCa risk from the ancestry-specific GRS (mean GRS) and the observed RR for PCa in each decile using a calibration analysis, a systematic bias across the ten deciles was found in each population; β was considerably lower than 1 (0.73, 0.64, 0.66, and 0.75 in EUR, HIS, AFR, and EAS populations, respectively; Fig. 2A–D). The systematic bias was primarily driven by a considerably overestimated risk for men in the highest GRS decile and an underestimated risk for men in the lowest GRS decile. In AFR, for example, the mean GRS for men in the highest decile was 3.05, considerably higher than their observed RR of 2.42. In contrast, the mean GRS for men in the lowest decile was 0.14, considerably lower than their observed RR of 0.34. In addition, the bias scores were substantially higher than 0: 0.19, 0.23, 0.31, and 0.24 in EUR, HIS, AFR, and EAS populations, respectively.

The bias was reduced noticeably after the Platt scaling correction (Supplementary Table 6). In the testing dataset, the β values for the scaled GRS were 0.95, 1.05, 1.02, and 1.01 in EUR, HIS, AFR, and EAS populations, respectively

Table 2 – Mean GRS in cases and controls by racial/ethnic population

	No. of individuals		Test for different mean GRSs			Test for association of GRS and PCa			
	Cases	Controls	Mean GRS (95% CI)		T test	Logistic regression		Logistic regression ^a	
			Cases	Controls	p value	OR (95% CI)	p value	OR (95% CI)	p value
EUR	53 220	834 866	1.63 (1.61–1.64)	0.94 (0.93–0.94)	<0.001	1.39 (1.38–1.40)	<0.001	1.46 (1.45–1.46)	<0.001
HIS	3100	78 009	1.27 (1.23–1.30)	0.92 (0.91–0.92)	<0.001	1.44 (1.41–1.47)	<0.001	1.50 (1.46–1.53)	<0.001
AFR	2267	28 205	1.33 (1.26–1.39)	0.81 (0.80–0.82)	<0.001	1.37 (1.34–1.41)	<0.001	1.44 (1.40–1.47)	<0.001
EAS	305	13 680	1.57 (1.39–1.75)	0.91 (0.89–0.93)	<0.001	1.29 (1.23–1.36)	<0.001	1.34 (1.28–1.42)	<0.001

AFR = African; CI = confidence interval; EAS = East Asian; EUR = European; HIS = Hispanic; GRS = genetic risk score; OR = odds ratio.
^a Adjusting for age and top ten principal components.

**Fig. 1 – Observed prevalence of PCa in each ancestry-specific PCa GRS decile by racial/ethnic population. AFR = African; EAS = East Asian; EUR = European; GRS = genetic risk score; HIS = Hispanic; PCa = prostate cancer.**

(Fig. 3A–D). The bias scores were also greatly reduced to 0.03, 0.06, 0.05, and 0.11 in EUR, HIS, AFR, and EAS populations, respectively.

4. Discussion

While the reliability of risk estimates from polygenic risk scores has consistently been demonstrated in EUR ancestry, including a large population-based cohort from the UKB and a clinical trial population REDUCE (REDuction by DUTasteride of prostate Cancer Events) [18,26], it has not been assessed in non-European populations. This inequality poses one of the most important ethical and scientific challenges surrounding the routine implementation of polygenic risk scores into the clinic [16]. There is a major concern that current EUR-centric polygenic risk scores may exacerbate health disparities [17].

Using the large 23andMe cohort with >125 000 individuals from AFR, HIS, and EAS ancestries, we are able, for the first time, to assess the validity of an ancestry-specific PCa GRS in non-European racial/ethnic populations. Similar to the previous results from studies with individuals of EUR ancestry [18,26], a systematic bias in the direction of an

overestimated risk for men in the highest decile of an uncorrected GRS was also detected in non-European populations, but a scaled GRS was well calibrated. The estimated risk from the scaled GRS was corroborated by the observed PCa risk (β approximated to 1 and bias score approximated to 0 in each population). Results from this study provide the data required to support the clinical implementation of a GRS at an individual level for both non-European and EUR populations.

The necessity for establishing the reliability of risk estimates from polygenic risk scores is underappreciated. Most published studies report only the validity of percentiles from polygenic risk scores, such as a dose-response association between higher score percentiles and a higher PCa risk [8,17,27]. However, validity of percentile alone is insufficient to support its clinical use because percentile provides only disease risk ranking of an individual in a population but does not quantify the specific disease risk of an individual. Quantification of disease risk is needed in the clinical setting to estimate an individual's RR and lifetime risk (calculated based on an individual's RR and population-level incidence and mortality). Therefore, a clinically useful polygenic risk score should provide risk estimates for test individuals. Furthermore, it is important to ensure the

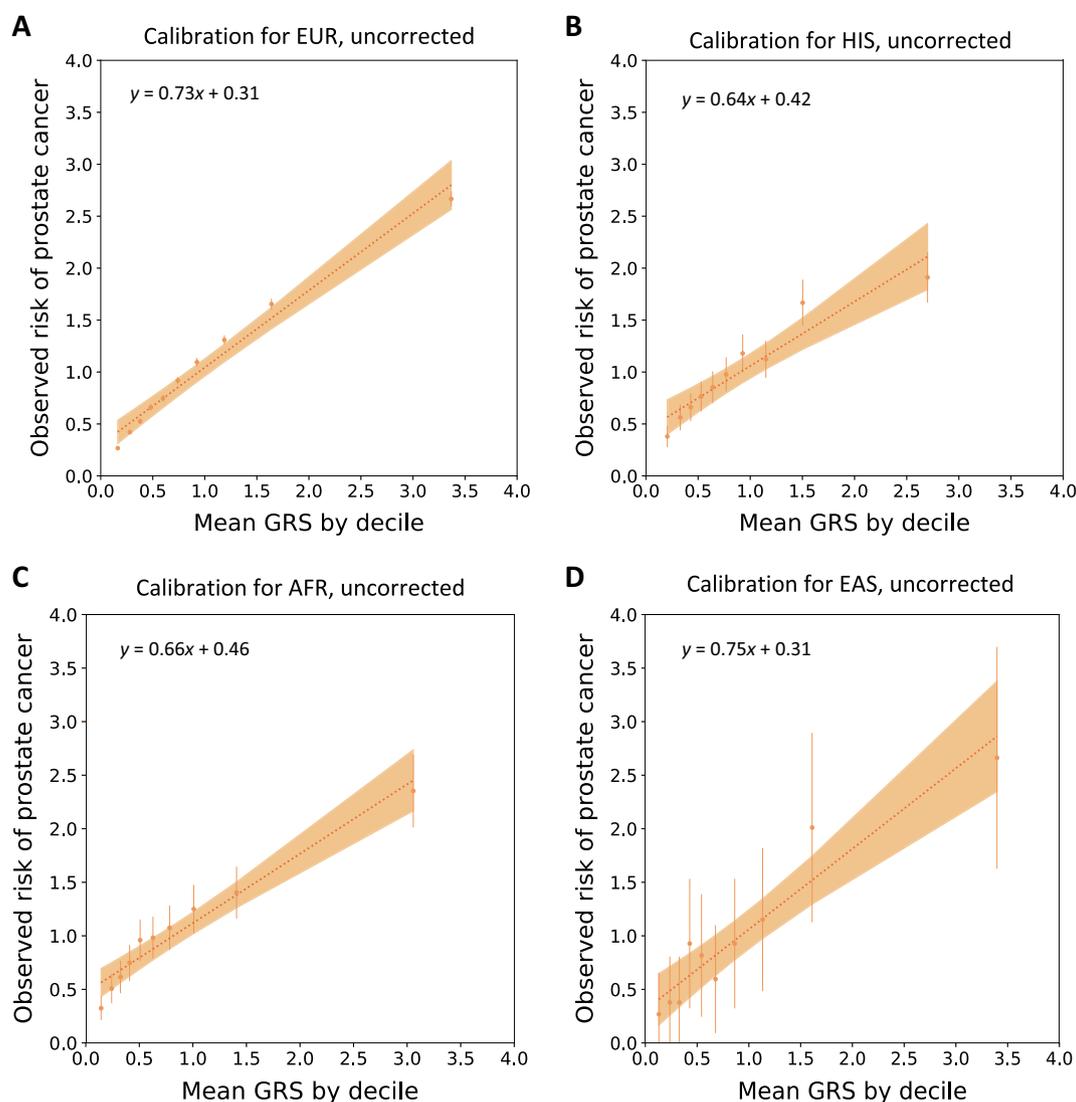


Fig. 2 – Calibration plots between observed relative risk for PCa and estimated risk (mean GRS) of each ancestry-specific PCa GRS decile among all study participants from (A) European, (B) African, (C) Hispanic, and (D) East Asian ancestries. AFR = African; EAS = East Asian; EUR = European; GRS = genetic risk score; HIS = Hispanic; PCa = prostate cancer.

accuracy of risk estimates reported to individuals from genetic testing. Results from our study suggest that an uncorrected GRS may be inaccurate, especially for those with the highest and lowest estimated risks. For example, an uncorrected GRS would considerably overestimate PCa for AFR men at top decile (mean RR of 3.05, with derived lifetime risk of 67% by age 85 yr). After the Platt scaling correction, their risk (mean RR of 2.37, with derived lifetime risk of 52% by age 85 yr) is more consistent with the observed risk of 2.42.

Multiple polygenic risk score methods have been reported, including those based on established GWAS-significant SNPs (eg, OR-weighted polygenic risk score, OR-weighted and population-standardized GRS, and polygenic hazard scores based on a survival analysis model to predict the time to age onset of PCa [28,29]), as well as millions of SNPs in the genome pruning and thresholding (P + T) and Bayesian genomic prediction methods (LDpred) [16,26,27,30,31]. The performance of these methods for dis-

criminating cases and controls, as measured by the area under the curve (AUC), is similar in PCa and other common diseases [26,32]. For example, the AUCs for differentiating PCa cases and controls in the REDUCE trial were 0.62, 0.62, and 0.60, respectively, for GRS (110 SNPs), P + T (397 SNPs), and LDpred (3 023 543 SNPs) [26]. The interpretation of score values, however, differs among various polygenic risk score methods. The score values of polygenic risk scores, P + T, and LDpred are difficult to interpret directly and can change with the evolving numbers of SNPs used in the calculation. In contrast, values of population-standardized GRSs can be interpreted as RR to the general population regardless of the number of SNPs used in the calculation [24]. For example, a GRS of 1.5 will always mean a 1.5-fold risk increase relative to the population. The simplicity in interpreting GRS values makes it easier for clinical implementation.

The unique RR property of a GRS value also makes it possible to directly compare the concordance between the

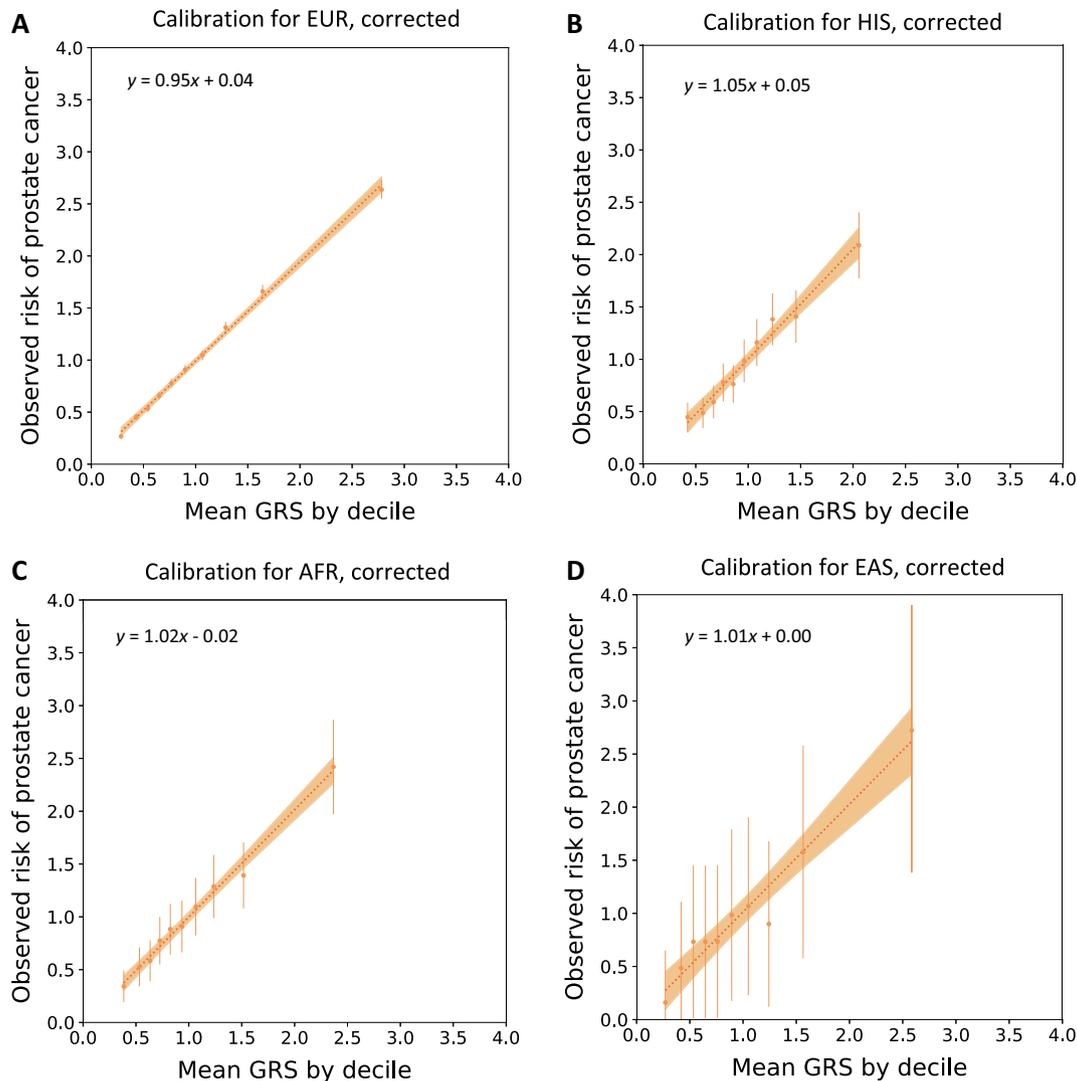


Fig. 3 – Calibration plots between observed relative risk for PCa and estimated risk (mean GRS) of each scaled ancestry-specific PCa GRS decile in the testing dataset (40% participants) for (A) European, (B) African, (C) Hispanic, and (D) East Asian ancestries. AFR = African; EAS = East Asian; EUR = European; GRS = genetic risk score; HIS = Hispanic; PCa = prostate cancer.

expected risk (ie, GRS value itself, without the need for fitting a regression model) and the observed risk using a calibration method. This differs from commonly used calibration approaches where the expected risk is first derived from fitting a logistic regression model and then its concordance with the observed risk of individuals in percentile groups (eg, deciles) is assessed using the Hosmer-Lemeshow goodness of fit test. This commonly used approach is susceptible to overfitting because the observed data are used two times (fitting regression model for estimating the expected risk and calibration between the expected and observed risks) [33].

While the calibration β of the scaled GRS in non-European racial/ethnic populations (1.01–1.05) was similar to that of EUR population (0.95), it is noted that the bias score in these populations (0.05–0.11) was slightly higher than that of EUR population (0.03). Furthermore, compared with the mean GRS in EUR population (0.98), the mean GRS in non-European racial/ethnic populations was substan-

tially lower than 1 (0.85–0.93), especially in AFR population (0.85). These results suggest that a systematic bias of the scaled GRS in non-European populations is modest and similar to that of EUR population, but the accuracy of risk estimate in these populations needs to be improved further. A combination of several factors may contribute to the relatively weaker performance in non-European populations, including (1) fewer ancestry-specific risk-associated SNPs (67–138 SNPs) than that of EUR population (232 SNPs), (2) smaller sample sizes (81 109, 30 472, and 13 985 in HIS, AFR, and EAS populations, respectively) than that of EUR population (888 086), and (3) a more heterogeneous genetic background. Additional studies in non-European populations are needed urgently.

Additional limitations are noted. First, PCa status of the participants of this study was self-reported from an online survey and was not validated. This may affect the accuracy of PCa diagnosis. However, the reliability of self-reported diagnoses in the 23andMe cohort has been evaluated sys-

tematically and demonstrated for multiple diseases, including PCa [34]. Second, the mean age of the participants in this study was relatively young and affects the penetrance and detection of PCa. This limitation may also contribute to an overestimated risk of an uncorrected GRS. Third, a lack of detailed clinicopathological variables in study participants prevents an in-depth analysis between a GRS and the aggressiveness of PCa. Such analyses may be performed in hospital-based studies. Fourth, the performance of the scaled GRS in the testing dataset of our study may be overfitted because individuals of the testing dataset were from the same cohort as the training dataset. In particular, the correction factor may be specific to the 23andMe cohort (prevalent cases, self-reported diagnoses, and relatively young age). To assess the generalizability of the Platt scaling estimates from the 23andMe cohort (training datasets), we applied the ancestry-specific scaling correction factors derived from the 23andMe cohort to the uncorrected GRS in the UKB, a population-based cohort in which the unaffected men were followed up for ~12 yr for incident PCa. The calibration β improved from the uncorrected GRS (0.81 and 0.64, respectively, in Whites and Blacks) to ~1 for the scaled GRS (1.04 and 0.97, respectively, in Whites and Blacks). Further evaluation of the generalizability for our ancestry-specific GRS in additional large non-European populations is needed. Finally, we recognized that many factors contribute to PCa risk in populations and therefore impact GRS calibration and generalizability. While genetic factors can be controlled partially using ancestry genetic structure derived from SNPs [28], other factors, such as environmental and lifestyle factors, cancer screening practice, and access of health care, are difficult to measure and control. Caution should be exercised when implementing the calibrated GRS.

5. Conclusions

In conclusion, the scaled ancestry-specific GRS in three non-European racial/ethnic populations was well calibrated (comparable with that in EUR population) and appropriate for genetic testing at the individual level for personalized PCa screening. However, more studies are needed to further improve the reliability of risk estimates in non-European populations.

Author contributions: Jianfeng Xu had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Xu, Koelsch.

Acquisition of data: Xu.

Analysis and interpretation of data: Shi, Zhan, Wei, Wang.

Drafting of the manuscript: Xu.

Critical revision of the manuscript for important intellectual content: Shi, Zhan, Wei, Ladson-Gary, Wang, Hulick, Zheng, Cooney, Isaacs, Helfand, 23andMe Research Team, Koelsch.

Statistical analysis: Shi, Zhan, Wei, Wang.

Obtaining funding: None.

Administrative, technical, or material support: None.

Supervision: Xu, Koelsch.

Other: None.

Financial disclosures: Jianfeng Xu certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: NorthShore University HealthSystem has an agreement with GoPath Laboratories for genetic tests of polygenic risk scores. Jianan Zhan and Bertram L. Koelsch are employed by and hold stock or stock options in 23andMe, Inc.

Funding/Support and role of the sponsor: None.

Acknowledgments: We are grateful to the Ellrodt-Schweighauser family for establishing Endowed Chair of Cancer Genomic Research (Dr. Xu), Chez and Melman families for establishing Endowed Chairs of Personalized Prostate Cancer Care at NorthShore University HealthSystem, and the Rob Brooks Fund for Personalized Prostate Cancer Care at NorthShore University HealthSystem. We would like to thank the research participants and employees of 23andMe for making this work possible. The following members of the 23andMe Research Team contributed to this study: Stella Aslibekyan, Adam Auton, Elizabeth Babalola, Robert K. Bell, Jessica Bielenberg, Katarzyna Bryc, Emily Bullis, Daniella Coker, Gabriel Cuellar Partida, Devika Dhamija, Sayantan Das, Sarah L. Elson, Nicholas Eriksson, Teresa Filshtein, Alison Fitch, Kipper Fletez-Brant, Pierre Fontanillas, Will Freyman, Julie M. Granka, Karl Heilbron, Alejandro Hernandez, Barry Hicks, David A. Hinds, Ethan M. Jewett, Yunxuan Jiang, Katelyn Kukar, Alan Kwong, Keng-Han Lin, Bianca A. Llamas, Maya Lowe, Jey C. McCreight, Matthew H. McIntyre, Steven J. Micheletti, Meghan E. Moreno, Priyanka Nandakumar, Dominique T. Nguyen, Elizabeth S. Noblin, Jared O'Connell, Aaron A. Petrakovitz, G. David Poznik, Alexandra Reynoso, Morgan Schumacher, Anjali J. Shastri, Janie F. Shelton, Jingchunzi Shi, Suyash Shringarpure, Qiaojuan Jane Su, Susana A. Tat, Christophe Toukam Tchakouté, Vinh Tran, Joyce Y. Tung, Xin Wang, Wei Wang, Catherine H. Weldon, Peter Wilton, and Corinna D. Wong.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.euro.2022.09.001>.

References

- [1] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA Cancer J Clin* 2021;71:7–33.
- [2] National Cancer Institute. Cancer stat facts: prostate cancer 2022. <https://seer.cancer.gov/statfacts/html/prost.html>.
- [3] Schroder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet* 2014;384:2027–35.
- [4] Kilpelainen TP, Tammela TL, Roobol M, et al. False-positive screening results in the European randomized study of screening for prostate cancer. *Eur J Cancer* 2011;47:2698–705.
- [5] U. S. Preventive Services Task Force, Grossman DC, Curry SJ, et al. Screening for prostate cancer: US Preventive Services Task Force recommendation statement. *JAMA* 2018;319:1901–13.
- [6] Carroll PR, Parsons JK, Andriole G, et al. NCCN guidelines insights: prostate cancer early detection, version 2.2016. *J Natl Compr Canc Netw* 2016;14:509–19.
- [7] Wei J, Yang W, Shi Z, et al. Observed evidence for guideline-recommended genes in predicting prostate cancer risk from a large population-based cohort. *Prostate* 2021;81:1002–8.
- [8] Conti DV, Darst BF, Moss LC, et al. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new

- susceptibility loci and informs genetic risk prediction. *Nat Genet* 2021;53:65–75.
- [9] Zheng SL, Sun J, Wiklund F, et al. Cumulative association of five genetic variants with prostate cancer. *N Engl J Med* 2008;358:910–9.
- [10] Kader AK, Sun J, Reck BH, 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.
- [11] Chen H, Liu X, Brendler CB, et al. Adding genetic risk score to family history identifies twice as many high-risk men for prostate cancer: results from the prostate cancer prevention trial. *Prostate* 2016;76:1120–9.
- [12] Shi Z, Platz EA, Wei J, et al. Performance of three inherited risk measures for predicting prostate cancer incidence and mortality: a population-based prospective analysis. *Eur Urol* 2021;79:419–26.
- [13] Darst BF, Sheng X, Eeles RA, Kote-Jarai Z, Conti DV, Haiman CA. Combined effect of a polygenic risk score and rare genetic variants on prostate cancer risk. *Eur Urol* 2021;80:134–8.
- [14] Seibert TM, Fan CC, Wang Y, et al. Polygenic hazard score to guide screening for aggressive prostate cancer: development and validation in large scale cohorts. *BMJ* 2018;360:j5757.
- [15] Na R, Labbate C, Yu H, et al. Single-nucleotide polymorphism-based genetic risk score and patient age at prostate cancer diagnosis. *JAMA Netw Open* 2019;2:e1918145.
- [16] Xu J, Resurreccion WK, Shi Z, et al. Inherited risk assessment and its clinical utility for predicting prostate cancer from diagnostic prostate biopsies. *Prostate Cancer Prostatic Dis* 2022;25:422–30.
- [17] Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* 2019;51:584–91.
- [18] Wei J, Shi Z, Na R, et al. Calibration of polygenic risk scores is required prior to clinical implementation: results of three common cancers in UKB. *J Med Genet* 2022;59:243–7.
- [19] Fritsche LG, Ma Y, Zhang D, et al. On cross-ancestry cancer polygenic risk scores. *PLoS Genet* 2021;17:e1009670.
- [20] Plym A, Penney KL, Kalia S, et al. Evaluation of a multiethnic polygenic risk score model for prostate cancer. *J Natl Cancer Inst* 2022;114:771–4.
- [21] Howlader N, Noone AM, Krapcho M, et al. SEER Cancer Statistics Review (CSR) 1975–2018. 2021. https://seer.cancer.gov/csr/1975_2018/.
- [22] Fuchsberger C, Abecasis GR, Hinds DA. minimac2: faster genotype imputation. *Bioinformatics* 2015;31:782–4.
- [23] Durand EY, Do CB, Mountain JL, Macpherson JM. Ancestry composition: a novel, efficient pipeline for ancestry deconvolution. *bioRxiv* 2014;010512.
- [24] Yu H, Shi Z, Wu Y, et al. Concept and benchmarks for assessing narrow-sense validity of genetic risk score values. *Prostate* 2019;79:1099–105.
- [25] Platt J. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Adv Large Margin Classifiers* 1999;10:61–74.
- [26] Yu H, Shi Z, Lin X, et al. Broad- and narrow-sense validity performance of three polygenic risk score methods for prostate cancer risk assessment. *Prostate* 2020;80:83–7.
- [27] Schumacher FR, Al Olama AA, Berndt SI, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet* 2018;50:928–36.
- [28] Karunamuni RA, Huynh-Le MP, Fan CC, et al. Performance of African-ancestry-specific polygenic hazard score varies according to local ancestry in 8q24. *Prostate Cancer Prostatic Dis* 2022;25:229–37.
- [29] Karunamuni RA, Huynh-Le MP, Fan CC, et al. Additional SNPs improve risk stratification of a polygenic hazard score for prostate cancer. *Prostate Cancer Prostatic Dis* 2021;24:532–41.
- [30] Vilhjalmsón BJ, Yang J, Finucane HK, et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am J Hum Genet* 2015;97:576–92.
- [31] Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun* 2019;10:1776.
- [32] Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;50:1219–24.
- [33] Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 1997;16:965–80.
- [34] Tung JY, Do CB, Hinds DA, et al. Efficient replication of over 180 genetic associations with self-reported medical data. *PLoS One* 2011;6:e23473.