Genetics in Medicine (2022) 24, 576–585

ELSEVIER

Genetics in Medicine

www.journals.elsevier.com/genetics-in-medicine

ARTICLE

Breast and prostate cancer risk: The interplay of polygenic risk, rare pathogenic germline variants, and family history

Emadeldin Hassanin[1], Patrick May[2], Rana Aldisi[1], Isabel Spier[3,4], Andreas J. Forstner[3,5,6], Markus M. Nöthen[3], Stefan Aretz[3,4], Peter Krawitz[1], Dheeraj Reddy Bobbili[2], Carlo Maj[1,*]

[1]*Institute for Genomic Statistics and Bioinformatics, University of Bonn, Bonn, Germany;* [2]*Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg;* [3]*Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany;* [4]*National Center for Hereditary Tumor Syndromes, University Hospital Bonn, Bonn, Germany;* [5]*Centre for Human Genetics, Philipps-University Marburg, Marburg, Germany;* [6]*Institute of Neuroscience and Medicine (INM-1), Research Center Jülich, Jülich, Germany*

ARTICLE INFO

Article history:
Received 6 June 2021
Received in revised form
12 September 2021
Accepted 12 November 2021
Available online 18 November 2021

Keywords:
Breast cancer
Family history
Polygenic risk score
Prostate cancer
Rare pathogenic variants

ABSTRACT

Purpose: We aimed to investigate to what extent polygenic risk scores (PRS), rare pathogenic germline variants (PVs), and family history jointly influence breast cancer and prostate cancer risk.

Methods: A total of 200,643 individuals from the UK Biobank were categorized as follows: (1) heterozygotes or nonheterozygotes for PVs in moderate to high-risk cancer genes, (2) PRS strata, and (3) with or without a family history of cancer. Multivariable logistic regression and Cox proportional hazards models were used to compute the odds ratio across groups and the cumulative incidence through life.

Results: Cumulative incidence by age 70 years among the nonheterozygotes across PRS strata ranged from 9% to 32% and from 9% to 35% for breast cancer and prostate cancer, respectively. Among the PV heterozygotes it ranged from 20% to 48% in moderate-risk genes and from 51% to 74% in high-risk genes for breast cancer, and it ranged from 30% to 59% in prostate cancer risk genes. Family history was always associated with an increased cancer odds ratio.

Conclusion: PRS alone provides a meaningful risk gradient leading to a cancer risk stratification comparable to PVs in moderate risk genes, whereas acts as a risk modifier when considering high-risk genes. Including family history along with PV and PRS further improves cancer risk stratification.

Dheeraj Reddy Bobbili and Carlo Maj contributed equally.

*Correspondence and requests for materials should be addressed to Carlo Maj, Institute for Genomic Statistics and Bioinformatics, University of Bonn, Venusberg-Campus 1, 53127 Bonn, Germany. *E-mail address:* cmaj@uni-bonn.de

doi: https://doi.org/10.1016/j.gim.2021.11.009

Introduction

Breast cancer and prostate cancer represent 2 of the most common cancers in women and men, respectively. Within the UK Biobank (UKB) cohort, breast cancer is the most prevalent cancer diagnosis in females, and prostate cancer is the most prevalent cancer diagnosis in males (https://biobank.ctsu.ox.ac.uk/~bbdatan/CancerSummaryReport.html). Along with several other factors, predisposing genetic variants (constitutional/germline variants) play a crucial role in the risk of developing breast cancer and prostate cancer.

Both breast cancer and prostate cancer are characterized by a high heritability, estimated to be around 31% for breast cancer[1] and 58% for prostate cancer.[2] Within breast cancer cases, approximately 5% to 10% are monogenic forms caused by moderate to high penetrant pathogenic germline variants.[3] Similarly, in prostate cancer familial subtypes following a Mendelian inheritance have been identified.[4] It is noteworthy that in 17% of the patients with family history for prostate cancer, who were referred for genetic testing, a pathogenic germline variant could be identified.[5] Breast cancer and prostate cancer share some susceptibility genes suggesting a potential shared genetic predisposition between the 2 cancer types.[6] It has also been observed that family history in first-degree relatives for prostate cancer increases women's risk of developing breast cancer by 14%.[7] Similarly, having a first-degree relative with breast cancer increases the chance of developing prostate cancer by 18%,[8] which further underpins the hypothesis of shared genetic risk factors.

Several studies have shown the crucial role of predisposing germline variants in the etiology of breast cancer: rare high-risk variants in *BRCA1* and *BRCA2*[9]; rare intermediate-/moderate-risk variants in *PALB2, CHEK2*, and *ATM*[10]; and various common low risk variants.[11] In particular, *BRCA1/2* pathogenic variants are most commonly linked to monogenic breast cancer, usually designated as hereditary breast cancer and ovarian cancer.[3]

In addition to the risk conferred by rare pathogenic variants in the strongly associated genes, different genome-wide association studies (GWAS) have identified hundreds of single-nucleotide variations associated with breast cancer risk. Although each single-nucleotide variation has a negligible effect size, their cumulative effect calculated as polygenic risk score (PRS) contributes significantly to the cancer risk, and it can improve disease risk stratification in the general population.[12] Although it is well-established that both rare and common constitutive variants are associated with breast cancer, only few studies have explored their combined effect and specifically to what extent the polygenic background acts as a risk modifier of monogenic variants of breast cancer.

For instance, the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model is a comprehensive breast cancer prediction tool incorporating *BRCA1, BRCA2, PALB2, ATM,* and *CHEK2* variants, along with other risk factors such as family and medical history, lifestyle, and, recently, also PRS.[13] In a recent study, the impact of PRS on the penetrance of the breast cancer risk variants was assessed for NM_024675.3:c.1592del (rs180177102) in *PALB2* and NM_007194.3:c.1100del (rs555607708) in *CHEK2* in Finnish population[14] and for *BRCA1/2* cancer-associated variants in a previous release of UKB including a smaller cohort of 49,960 individuals with exome-sequencing data.[15]

Similarly, different genes are associated with the etiology of prostate cancer, in particular *BRCA1/2, ATM, CHEK2,* and *HOXB13*.[16-18] Moreover, several studies have shown that for prostate cancer also the cumulative risk driven by the presence of common variants as summarized by PRS models is strongly associated with the cancer risk.[19] Few studies showed the effect of PRS stratification among heterozygotes for p.G84E in *HOXB13*,[20] and heterozygotes for *BRCA1/2* pathogenic variant.[21] However, those studies focused only on specific variants or genes.

In this work, we compared the prevalence and the lifetime risk of breast cancer and prostate cancer among 200,643 individuals from the UKB. Individuals were categorized into heterozygotes and nonheterozygotes of rare pathogenic or likely pathogenic (P/LP) variants (hereafter defined as PV) in moderate or high susceptibility genes; low, intermediate, and high PRS; and with or without a family history for the respective cancer.

Material and Methods

Data source

This study was performed using genetic and phenotypic data from UKB (application number 52446). UKB is a long-term prospective population-based study, and the volunteers are being recruited mainly from England, Scotland, and Wales; it involves more than 500,000 participants aged between 40 and 69 years at recruitment. An abundant diversity of phenotypic and health-related information is available on each participant; for 487,410 samples, genotyping data are available, and for 200,643 individuals, exome sequencing (ES) data are also available. The data set is accessible for research purposes, and all participants provided documented consent.[22]

Study participants

Breast cancer cases were defined on the basis of self-reported code 1002 (in data field 20001), International Classification of Diseases (ICD)-10 code C50.X, or ICD-9 code 174.X in hospitalization records. For prostate cancer, cases with self-reported code 1044 (in data field 20001), ICD 10 code C61 and D075, or ICD-9 code 185 in

hospitalization records were included. The remaining samples with no other cancer diagnosis were considered as controls. Individuals of all ancestries were included in the analysis. Only individuals with both genotyping and ES data were included ($N = 200,643$). On the basis of the available genotype data, we excluded outliers for heterozygosity or genotype missing rates, putative sex chromosome aneuploidy, and discordant reported sex vs genotypic sex. In the analysis, we included only females for breast cancer and only males for prostate cancer. We excluded 1 from each pair of related individuals if the genetic relationship was closer than the second degree, defined as kinship coefficient > 0.0884 as calculated by the UKB (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/ukbgene_instruct.html).

Variant selection

Annovar[23] was used to annotate the variant call format files per chromosome from the 200,643 ES data. Variant frequencies were retrieved from the Genome Aggregation Database (gnomAD),[24] whereas ClinVar[25] annotations were considered to interpret the pathogenicity of germline variants.

The following inclusion criteria were applied to select rare PV in the UKB data: (1) only variants in protein-coding regions of the *BRCA1/2, CHEK2, ATM,* and *PALB2* genes for breast cancer and *BRCA1/2, CHEK2, ATM,* and *HOXB13* genes for prostate cancer; (2) allele frequency < 0.005 in at least 1 ethnic subpopulation of gnomAD and also allele frequency < 0.005 in gnomAD overall; (3) not annotated as synonymous, nonframeshift deletion, and nonframeshift insertion; and (4) annotated as P/LP on the basis of ClinVar, ie, if the variant is consistently classified as such or, in case of a conflicting interpretation, if at least 3 P/LP annotations were available without any benign/likely benign classification. A similar variant filtering approach has been applied in a recent analysis aimed at identifying disease causing monogenic variants.[15] Individuals carrying any of the identified variants in the moderate to high penetrant genes in heterozygous or homozygous state were classified as PV heterozygotes. We use the term nonheterozygote to refer to individuals who are not heterozygous for a PV variant.

PRS

To generate the PRS, we used a previously validated PRS for breast cancer and prostate cancer containing 313 and 103 variants, respectively.[21,26] The PRS was calculated from the UKB genotype data using the PLINK 2.0[27] scoring function. We applied a previous approach to minimize variance in PRS distributions across genetic ancestries.[28] Specifically, we fit linear regression model using the first 4 ancestry principal components (PCs) in the controls (PC_PRS = PC1 + PC2 + PC3 + PC4). The derived model was applied to predict the PC_PRS over the entire data set. The PC adjusted PRS was calculated by subtracting PC_PRS from the raw PRS (ie, the residual PRSs were computed) and used for the subsequent analyses.

Statistical analysis

Individuals were stratified on the basis of the PRS percentile, presence or absence of PV (ie, heterozygous or nonheterozygous), and family history. We considered the corresponding family history of cancer in parents and siblings as reported by participants (UKB Data-fields: 20110, 20107, 20111). We assigned individuals to low (<10%), intermediate (10%-90%), and high (>90%) PRS where the definition of a high PRS (above the 90th percentile) followed a previous study.[18] The rationale to stratify PRS into 3 risk classes was in line with the hypothesis that PRS is associated with a nonlinear decrease of risk for extremely low PRS and nonlinear increase of risk for extremely high PRS as observed in other studies.[12]

Intermediate PRS, nonheterozygote, and an absent family history corresponded to the large majority of individuals (69.9% and 72.1% for breast cancer and prostate cancer, respectively); therefore, this group was used as a reference to assess cancer prevalence in the population (ie, to compute the odds ratios [ORs]). We performed the analysis considering all genes (ie, heterozygotes of variants in any of the susceptibility genes) and also performed gene-specific analysis. For breast cancer, we stratified between PV heterozygotes in genes characterized by moderate/intermediate penetrance (ie, *ATM, CHEK2, PALB2*, in the following summarized as moderate) and heterozygotes in highly penetrant genes (ie, *BRCA1/2*) to assess the effect of PRS in the 2 risk groups. In contrast, for prostate cancer, we defined only a single group because there is no clear difference in the penetrance of the included genes. For each group, we computed the OR using a logistic regression model adjusted for age at recruitment and the first 4 PCs. We then predicted the cancer ORs across PRS percentiles from a logistic regression model by considering nonheterozygotes without family history with intermediate PRS as reference and conditioning on the mean of covariates (age and the first 4 PCs).

We estimated the lifetime risk by age 70 years resulting from PV status and the PRS. We fit a Cox proportional hazards model using the R package *survival*. We used age as the time scale representing the time-to-event, considering age at diagnosis in cases and age of last assessment in controls. The model included PV heterozygote status, PRS strata (ie, low, intermediate, high), age, and the first 4 ancestry PCs, whereas adjusted survival curves were plotted with the R package *survminer*. For all statistical analyses, we used R 3.6.3.

Results

Stratification of UKB cohort individuals for cancer prevalence, family history, and genetic risk factors

Within the 200,643 UKB individuals with available genotyping and exome data, we identified 6288 breast cancer cases (3838 prevalent cases and 2450 incident cases) with a mean age at diagnosis of 55.6 years. The remaining 85,903 women with no other cancer diagnosis were considered as controls, and the mean age at last visit was 56.8 years (Supplemental Table 1).

For prostate cancer, a total of 4021 cases (1331 prevalent cases and 2690 incident cases) were identified with a mean age at diagnosis of 64.4 years. The remaining 73,053 men with no other cancer diagnosis were considered as controls, and the mean age at last visit was 57.0 years (Supplemental Table 2).

It is noteworthy that both in breast cancer and prostate cancer, there was a significantly higher proportion of individuals with a family history for cancers not only among heterozygotes of PV in the selected cancer susceptibility genes (OR = 2.09 and 1.62, $P < .01$) but also among individuals with high-PRS (OR = 1.38 and 1.37, $P < .01$) (Tables 1 and 2).

Distribution of PV heterozygotes within the UKB cohort

We identified 1622 heterozygotes of 309 PV in the 5 analyzed breast cancer susceptibility genes ie, *BRCA1/2*, *PALB2*, *CHEK2*, and *ATM*.

In addition, 1492 heterozygotes of 259 PV were found in the 5 considered prostate cancer susceptibility genes, ie, *BRCA1/2*, *ATM*, *CHEK2*, and *HOXB13*. The list of the considered variants, annotations, and number of heterozygotes are available in the Supplementary File 2.

Among the study participants, homozygous PVs were not identified either in breast cancer or in prostate cancer.

PRS distribution within the UKB cohort

The breast cancer and prostate cancer PRSs followed a normal distribution (raw and PC-adjusted PRS are shown in Supplemental Figure 1) and were significantly higher in cases than in controls ($P < .01$) (Supplemental Figure 2).

We observed a nonlinear increase of cancer risk for individuals in the extreme right tail of the PRS distribution and a less evident nonlinear decrease in the left tail (Supplemental Figure 3–disease prevalence by PRS percentile for both breast and prostate cancer). This corroborates the hypothesis that PRS can be used to stratify individuals into risk classes according to a liability threshold model[29] (ie, low, intermediate, and high risk).

Interplay between PV heterozygosity and PRS

None of the selected PV was included in the PRS, and thus, they represent an independent genetic signal. We observed that the mean and median of PRS was significantly higher in affected heterozygotes than in unaffected heterozygotes (Supplemental Figure 4).

For breast cancer, we performed a separate analysis for the high-risk genes *BRCA1/2* and the moderate-risk genes *PALB2, CHEK2*, and *ATM*. We estimated how breast cancer risk is influenced by PRS and the heterozygous status for PV in cancer susceptibility genes by computing the ORs for cancer across groups with respect to nonheterozygotes with intermediate PRS because they represent the major group in the population. Heterozygotes with intermediate PRS represent the heterozygotes population, and therefore, they are designated as heterozygotes for simplicity. The high-risk genes PV heterozygotes had a higher OR than individuals with only a high PRS (5.9 vs 2.0, Figure 1A). Instead, PV heterozygotes in the moderate risk genes had an OR comparable with the OR in case of nonheterozygotes with high PRS (OR = 2.2 vs 2.0), but the number of nonheterozygote women with high PRS was considerably larger than the number of heterozygotes (Figure 1A). Notably, women heterozygous for PV in moderate risk genes (ie, *ATM*, *CHEK2* and *PALB2*) with low PRS had a lower risk than nonheterozygote women with only high PRS (OR 1.2 vs 2.0).

In general, PRS modifies the penetrance of PVs in both moderate- and high-risk genes. Of note, PV heterozygote women with low PRS in case of both high-risk and moderate-risk genes had lower ORs (ie, 2.9 and 1.2, respectively), whereas heterozygote women with high PRS had the largest absolute ORs (OR = 8.6 and 3.3, respectively; Figure 1A).

For prostate cancer, PV heterozygotes with intermediate PRS had OR comparable with that of nonheterozygotes with high-PRS (OR = 2.3 vs 2.2) and even lower in case of low PRS (OR = 1.6). Notably, similar to the number observed in women for breast cancer, the number of nonheterozygote men with high PRS was considerably larger than the number of heterozygotes (Figure 1C). As expected, among PV heterozygotes, men with low PRS had the lowest ORs and the men with high PRS had the highest ORs (1.6 and 6.1, respectively, Figure 1C).

Similarly, analysis of the lifetime cancer risk showed a joint effect of PV and PRS. The cumulative incidence by age 70 years in heterozygotes was the lowest in case of low PRSs and the highest in the case of high PRS. In breast cancer, values ranged from 51% to 74% for high-risk genes and from 20% to 48% for moderate-risk genes (Figure 1B), whereas for prostate cancer the incidence ranged from 30% to 59% (Figure 1D). Notably, for nonheterozygotes the cumulative incidence ranged between 9% and 32% for breast cancer and between 9% and 35% for prostate cancer.

Table 1 Characteristics of the participants categorized by PV heterozygosity status and PRS strata in prostate cancer

	Heterozygote and High PRS	Heterozygote and Intermediate PRS	Heterozygote and Low PRS	Nonheterozygote and High PRS	Nonheterozygote and Intermediate PRS	Nonheterozygote and Low PRS
Participants, *n*	187	1185	120	7520	60,474	7588
Cases, *n* (%)	42 (22.46)	118 (9.96)	8 (6.67)	728 (9.68)	2971 (4.91)	154 (2.03)
Controls, *n*	145 (77.54)	1067 (90.04)	112 (93.33)	6792 (90.32)	57,503 (95.09)	7434 (97.97)
Age,[a] mean (SD)	57.89 (8.92)	57.19 (8.68)	56.31 (8.41)	57.31 (8.73)	57.43 (8.7)	57.31 (8.75)
Family history of prostate cancer, *n* (%)	33 (17.65)	135 (11.39)	19 (15.83)	798 (10.61)	4880 (8.07)	455 (6)

PRS, polygenic risk score; *PV,* pathogenic variant.
[a]Age at diagnosis for cases and age at last visit for controls.

Inclusion of family history on the cancer risk stratification

A family history of the corresponding cancer was present in 19% and 16% of cases and 10.7% and 7.8% of controls (OR = 2.0 and 2.3, $P < .01$) for breast cancer and prostate cancer, respectively (Supplemental Tables 1 and 2). Considering individuals with no family history and intermediate PRS as reference, we found that both family history and PRS were associated with higher risk (see ORs in Supplemental Figures 5 and 6). The risk was lowest for low PRS and no family history (ORs of 0.45 and 0.42 for breast cancer and prostate cancer, respectively) and the highest in the presence of both family history and high PRS (ORs of 3.5 in breast cancer and 4.6 in prostate cancer).

The full models considering the underlying continuous distribution of PRS by computing the predicted ORs across PRS percentiles in individuals stratified for family history and PV status in moderate-risk and high-risk genes showed that the cancer risk is strongly influenced by PRS in all groups (Figure 2). Considering the nonheterozygotes with no family history and median PRS percentile group as reference, the predicted breast cancer ORs in the lower tail of PRS was 0.36 for nonheterozygotes with no family history, whereas in the upper tail of PRS, for PV heterozygotes with family history, the OR reached 6.6 and 10.3 in moderate-risk and high-risk genes, respectively. A similar trend was observed for prostate cancer in which the lowest predicted OR of 0.3 was reached for PV nonheterozygotes without family history and OR of 13.1 for heterozygotes with family history and high PRS.

The effect of PRS in single gene heterozygotes

We estimated how PRS influences breast cancer prevalence among PV heterozygote women in each of the analyzed susceptibility genes.

The gene-specific analysis revealed a strong variability in risk conferred by rare PV in different genes. In particular, for breast cancer, the largest effect sizes were attributable to *BRCA1/2*, a comparably lower effect size was present for *PALB2* and *ATM*, and the lowest effect size was observed for *CHEK2* (Supplemental Figure 7). Gene-specific analysis in prostate cancer also showed heterogeneity across gene effect sizes with the largest effect observed for *HOXB13* and the smallest effect observed for *BRCA1* (Supplemental Figure 8). Despite having single genes, both breast and prostate cancers were characterized by different effect sizes, and the PRS modifies the relative risk across all genes.

Similar to the overall analysis, the gene-specific analysis showed that family history, PV, and PRS are associated with increased cancer risk. Despite the genes characterized by

Table 2 Characteristics of the participants by PV heterozygosity status and PRS strata in breast cancer

	Heterozygote and High PRS	Heterozygote and Intermediate PRS	Heterozygote and Low PRS	Nonheterozygote and High PRS	Nonheterozygote and Intermediate PRS	Nonheterozygote and Low PRS
Participants, *n*	222	1279	121	8997	72,473	9099
Cases, *n* (%)	56 (25.23)	241 (18.84)	16 (13.22)	1068 (11.87)	4634 (6.39)	273 (3)
Controls, *n*	166 (74.77)	1038 (81.16)	105 (86.78)	7929 (88.13)	67,839 (93.61)	8826 (97)
Age,[a] mean (SD)	55.51 (8.83)	56.03 (8.7)	54.74 (9.52)	56.52 (8.35)	56.75 (8.4)	57.08 (8.31)
Family history of breast cancer, *n* (%)	56 (25.23)	260 (20.33)	20 (16.53)	1317 (14.64)	8023 (11.07)	710 (7.8)

PRS, polygenic risk score; *PV,* pathogenic variant.
[a]Age at diagnosis for cases and age at last visit for controls.

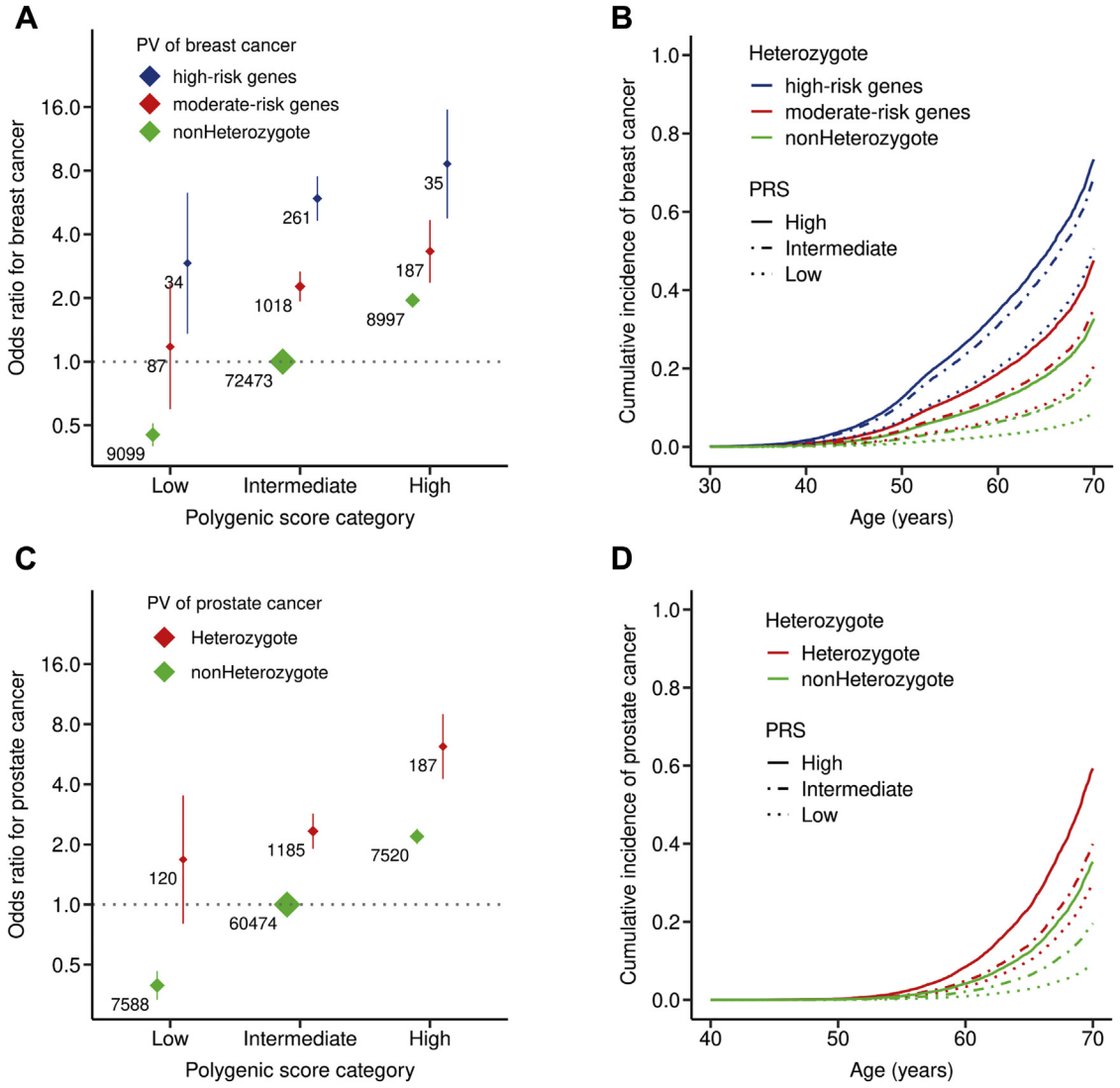

Figure 1 Cancer odds ratio and cumulative incidence among individuals categorized according to the presence of PV heterozygotes and PRS. Heterozygotes and nonheterozygotes were categorized into 3 strata on the basis of their PRS: low (<10 percentile), intermediate (10-90 percentile), or high (>90 percentile) PRS. The odds ratio was calculated from a logistic regression model with age, and the first 4 principal components of ancestry as covariates for breast cancer (A), and prostate cancer (C). The reference group was nonheterozygotes with intermediate PRS. The adjusted odds ratio is indicated by the colored boxes. The numbers next to the odds ratios indicate the sample size of the corresponding group. The 95% CI are indicated by the vertical lines around the boxes. Cumulative incidence was estimated from a Cox proportional hazards model using age, and the first 4 ancestry principal components for breast cancer (B), and prostate cancer (D). PRS, polygenic risk score; PV, pathogenic variant.

different risk levels, family history lead to larger ORs, and this trend was observed across different PRS strata (Figure 3A and B for breast and prostate cancer, respectively).

Discussion

In this study, we analyzed how breast and prostate cancer prevalence and cumulative incidence within the UKB cohort is affected by genetic susceptibility and family history. We considered both the genetic component driven by rare PV in genes associated with hereditary forms of cancer and the polygenic background present in all individuals.

Our results support the hypothesis of cumulative genetic risks caused by both rare PV and the polygenic background. We observed a higher prevalence of cancer in PV heterozygotes with high PRS (ie, individuals with suspected hereditary forms of breast cancer and prostate cancer). This result corroborates the role of the polygenic background as a modifier of the breast cancer and prostate cancer risk among PV heterozygotes unselected for specific clinical criteria (as the UKB cohort), and this is in line with that observed in other studies focused on specific genes or variants.[14,15] Lifetime risk analysis of breast cancer and prostate cancer indicated that the cumulative disease incidence can be jointly influenced by the presence of PV and the polygenic contribution over the course of life.

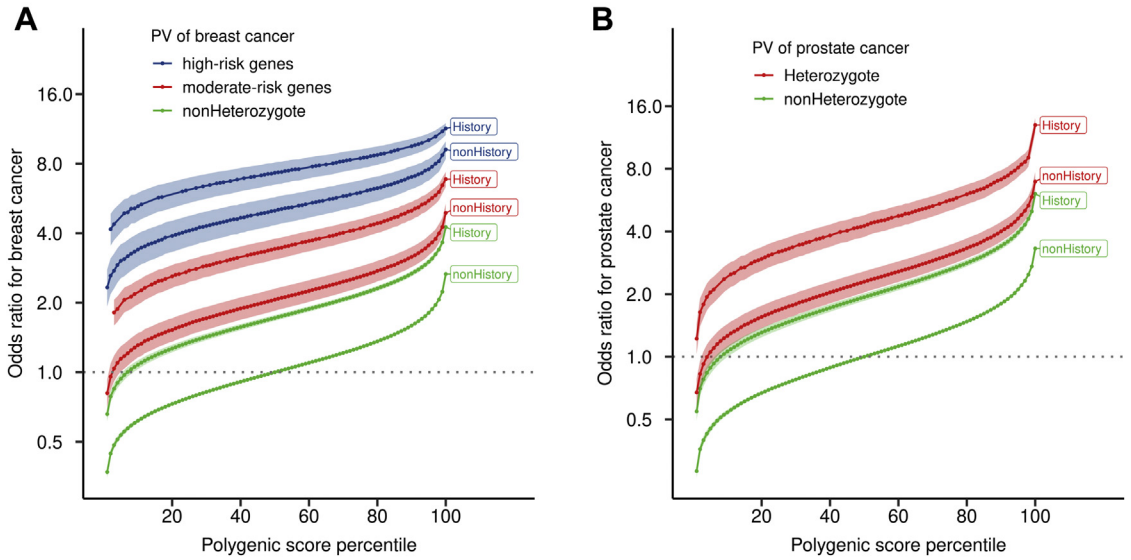

Figure 2 Interplay of PV, family history, and polygenic risk score (PRS). Predicted odds ratios for cancer were estimated from logistic models adjusted for age and first 4 ancestry principal components for breast cancer (A), and for prostate cancer (B). Nonheterozygotes with median PRS and no family history served as the reference group. PRS, polygenic risk score; PV, pathogenic variant.

Single-gene analysis revealed heterogeneous effects across genes, and therefore, the modifier role exerted by PRS should be framed within the absolute risk attributable to individual genes. This is in line with a recent study suggesting that PRS inclusion in risk stratification may prevent excess of surveillance for breast cancer in PV heterozygotes in moderate-risk genes such as *CHEK2* and *ATM*, whereas the cancer risk for PV heterozygotes in high-risk genes such as *BRCA1/2* is clinically relevant irrespective of the PRS.[30] Another recent work showed that there is a wide-range of absolute risks for breast cancer and prostate cancer in PV heterozygotes in terms of different genes and across PRS stratification.[31]

Our results showed that the PRS acts as a risk modifier for breast cancer and prostate cancer among both the general population and PV heterozygotes in all the well-known cancer susceptibility risk genes. PRS can define a significant proportion of the general population that is at a risk comparable with PV heterozygotes for moderate-risk genes or even more when considering family history. According to these findings, there should be a potential benefit including PRS in health care prevention policies for both the general population and at-risk individuals carrying PVs because risk-stratified surveillance might improve early disease detection and prevention.[32,33]

In particular, we observed that women with PVs in moderate-risk genes *ATM, CHEK2*, or *PALB2* with a high PRS had a cumulative incidence comparable with women with PV in high-risk genes *BRCA1/2* with a low PRS. On the contrary, women heterozygous for PV in moderate-risk genes with a low PRS had a cumulative incidence comparable to the general population. These results suggest that for women with PV in moderate-risk genes, the addition of PRS can optimize the risk stratification, which is often based on the life-time risk. Therefore, especially in the presence of PV in moderate-risk genes for breast cancer, intense surveillance programs and potential preventive measures can be better assessed when including the modifier role of PRS.

Moreover, with increasing population-based cohort sizes, PRS can better define a small group of very high-risk nonheterozygote individuals in the extreme tail of the PRS distribution characterized by even larger ORs and cumulative incidences than the ones observed in the current analysis.

In addition, our results showed that the inclusion of family history can further and independently improve the risk stratification along with genetic factors. Previous studies have discussed that family history is mainly associated with monogenic variants and minimally with PRS.[34,35] However, the PRS predictions are affected by estimation errors in variant effect sizes from the reference GWAS; thus, it can be expected that more accurate PRS models will be developed with the increased availability of population-based data.[36] Moreover, the additional effect of family history can be caused by unconsidered variants in the genetic risk models (eg, copy number variations), but it can also capture nongenetic contributors such as environmental/lifestyle factors.

Our study has different limitations. First, there is evidence of a healthy volunteers selection bias of the UKB cohort, and thus, the results might not be generalizable in terms of effect sizes.[37] Second, our risk assessment was based solely on genetic variants and family history and did not include other risk factors. Previous studies with UKB showed that lifestyle modifiable risk factors play a pivotal role in cancer prevalence,[38] and a shared lifestyle within families could influence family history with the disease.[39] This might explain the additional effect of family history

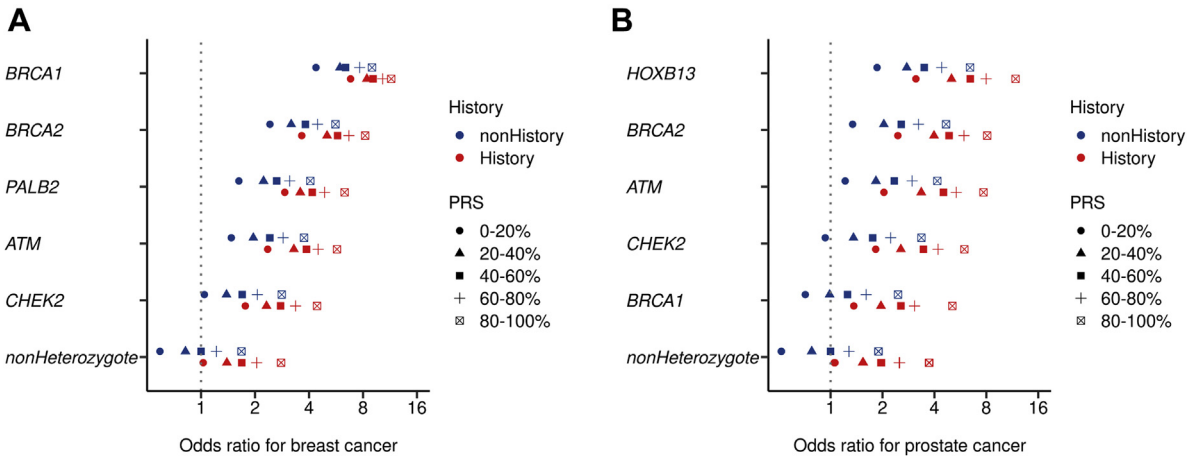

Figure 3 Interplay of pathogenic variant, family history, and PRS in single genes. Odds ratios for cancer were estimated from logistic models adjusted for age and first 4 ancestry principal components for breast cancer (A), and prostate cancer (B). Nonheterozygotes with 40% to 60% PRS and no family history served as the reference group. PRS, polygenic risk score.

of cancer with respect to the genetic risk. Finally, although we performed the analysis on the whole UKB cohort, we could not test the risk stratification generalizability across different populations because of the limited sample size. PRS could be biased toward the European population because PRS was constructed on the basis of European reference GWAS. Thus, PRS might be a worse predictor in non-European or admixed individuals, as previously discussed in different studies.[40]

In conclusion, we showed the significant role of PRS in both general population and heterozygotes of rare pathogenic germline variants in moderate to high-risk cancer genes. PRS strongly alters the penetrance of moderate-risk and high-risk variants and influences the lifetime disease risk. The data suggest that stratification of individuals based solely on the PRS can reach ORs comparable with those associated with heterozygotes of PV in moderate-risk genes that are currently subject to risk-adapted tailored surveillance programs. Consequently, PRS can identify a relatively large group of individuals within the general population for whom intense surveillance measures such as those offered to heterozygotes of moderate-risk genes should be considered. These findings highlight the potential usefulness of PRS in the context of cancer risk stratification. Our analysis shows that family history along with rare PV and PRS represents an additional stratification level to the cancer risk.

Data Availability

Genome-wide genotyping data, exome-sequencing data, and phenotypic data from the UK Biobank are available upon successful project application (http://www.ukbiobank.ac.uk/about-biobank-uk/).

Restrictions apply to the availability of these data, which were used under license for the current study (Project ID: 52446). Summary statistics are available from the Polygenic Score Catalog (pgs-info@ebi.ac.uk): for breast cancer at https://www.pgscatalog.org/score/PGS000007/ and for prostate cancer at https://www.pgscatalog.org/score/PGS000049/.

Acknowledgments

This research was conducted using the UK Biobank Resource under application number 52446. C.M. and E.H. were supported by the BONFOR-program of the Medical Faculty, University of Bonn (O-147.0002). P.M. was supported by the Luxembourg National Research Fund as part of the National Centre of Excellence in Research on Parkinson's disease (NCER-PD, FNR11264123) and the Deutsche Forschungsgemeinschaft Research Units FOR2715 (INTER/DFG/17/11583046) and FOR2488 (INTER/DFG/19/14429377). S.A. and I.S. were members of the European Reference Network on Genetic Tumor Risk Syndromes (ERN GENTURIS) Project ID No 739547. The authors acknowledge the use of de.NBI cloud and the support by the High Performance and Cloud Computing Group at the Zentrum für Datenverarbeitung of the University of Tübingen and the Federal Ministry of Education and Research (BMBF) through grant no 031A535A.

Author Information

Conceptualization: E.H., D.R.B., C.M.; Analysis: E.H.; Supervision: P.K., P.M., D.R.B., C.M.; Results Interpretation: A.J.F., I.S., S.A., M.M.N., P.K.; Writing-review and editing: E.H., P.M., R.A., I.S., S.A., M.M.N., P.K., D.R.B., C.M.

Ethics Declaration

Ethics approval for the UK Biobank (UKB) study was obtained from the North West Multicentre for Research Ethics Committee (MREC). The UKB ethics statement is available at https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/ethics. All UKB participants provided informed consent at recruitment.

Conflict of Interest

The authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2021.11.009) contains supplementary material, which is available to authorized users.

References

1. Möller S, Mucci LA, Harris JR, et al. The heritability of breast cancer among women in the Nordic twin study of cancer. *Cancer Epidemiol Biomarkers Prev*. 2016;25(1):145–150. http://doi.org/10.1158/1055-9965.EPI-15-0913.
2. Hjelmborg JB, Scheike T, Holst K, et al. The heritability of prostate cancer in the Nordic twin study of cancer. *Cancer Epidemiol Biomarkers Prev*. 2014;23(11):2303–2310. http://doi.org/10.1158/1055-9965.EPI-13-0568.
3. Economopoulou P, Dimitriadis G, Psyrri A. Beyond BRCA: new hereditary breast cancer susceptibility genes. *Cancer Treat Rev*. 2015;41(1):1–8. http://doi.org/10.1016/j.ctrv.2014.10.008.
4. Potter SR, Partin AW. Hereditary and familial prostate cancer: biologic aggressiveness and recurrence. *Rev Urol*. 2000;2(1):35–36.
5. Nicolosi P, Ledet E, Yang S, et al. Prevalence of germline variants in prostate cancer and implications for current genetic testing guidelines. *JAMA Oncol*. 2019;5(4):523–528. http://doi.org/10.1001/jamaoncol.2018.6760.
6. Sakoda LC, Jorgenson E, Witte JS. Turning of COGS moves forward findings for hormonally mediated cancers. *Nat Genet*. 2013;45(4):345–348. http://doi.org/10.1038/ng.2587.
7. Beebe-Dimmer JL, Yee C, Cote ML, et al. Familial clustering of breast and prostate cancer and risk of postmenopausal breast cancer in the Women's Health Initiative Study. *Cancer*. 2015;121(8):1265–1272. http://doi.org/10.1002/cncr.29075.
8. Ren ZJ, Cao DH, Zhang Q, et al. First-degree family history of breast cancer is associated with prostate cancer risk: a systematic review and meta-analysis. *BMC Cancer*. 2019;19(1):871. http://doi.org/10.1186/s12885-019-6055-9.
9. Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *JAMA*. 2017;317(23):2402–2416. http://doi.org/10.1001/jama.2017.7112.
10. Easton DF, Pharoah PD, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med*. 2015;372(23):2243–2257. http://doi.org/10.1056/NEJMsr1501341.
11. Mavaddat N, Pharoah PD, Michailidou K, et al. Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst*. 2015;107(5):1–15. http://doi.org/10.1093/jnci/djv036.
12. 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(9):1219–1224. http://doi.org/10.1038/s41588-018-0183-z.
13. Lee A, Mavaddat N, Wilcox AN, et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med*. 2019;21(8):1708–1718. Published correction appears in *Genet Med*. 2019;21(6):1462. https://doi.org/10.1038/s41436-018-0406-9.
14. Mars N, Widén E, Kerminen S, et al. The role of polygenic risk and susceptibility genes in breast cancer over the course of life. *Nat Commun*. 2020;11(1):6383. http://doi.org/10.1038/s41467-020-19966-5.
15. Fahed AC, Wang M, Homburger JR, et al. Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. *Nat Commun*. 2020;11(1):3635. http://doi.org/10.1038/s41467-020-17374-3.
16. Giri VN, Hegarty SE, Hyatt C, et al. Germline genetic testing for inherited prostate cancer in practice: implications for genetic testing, precision therapy, and cascade testing. *Prostate*. 2019;79(4):333–339. http://doi.org/10.1002/pros.23739.
17. Pritchard CC, Mateo J, Walsh MF, et al. Inherited DNA-repair gene mutations in men with metastatic prostate cancer. *N Engl J Med*. 2016;375(5):443–453. http://doi.org/10.1056/NEJMoa1603144.
18. Ewing CM, Ray AM, Lange EM, et al. Germline mutations in HOXB13 and prostate-cancer risk. *N Engl J Med*. 2012;366(2):141–149. http://doi.org/10.1056/NEJMoa1110000.
19. Sipeky C, Talala KM, Tammela TLJ, Taari K, Auvinen A, Schleutker J. Prostate cancer risk prediction using a polygenic risk score. *Sci Rep*. 2020;10(1):17075. http://doi.org/10.1038/s41598-020-74172-z.
20. Karlsson R, Aly M, Clements M, et al. A population-based assessment of germline HOXB13 G84E mutation and prostate cancer risk. *Eur Urol*. 2014;65(1):169–176. http://doi.org/10.1016/j.eururo.2012.07.027.
21. Lecarpentier J, Silvestri V, Kuchenbaecker KB, et al. Prediction of breast and prostate cancer risks in male BRCA1 and BRCA2 mutation carriers using polygenic risk scores. *J Clin Oncol*. 2017;35(20):2240–2250. http://doi.org/10.1200/JCO.2016.69.4935.
22. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203–209. http://doi.org/10.1038/s41586-018-0579-z.
23. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat Protoc*. 2015;10(10):1556–1566. http://doi.org/10.1038/nprot.2015.105.
24. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434–443. Published correction appears in *Nature*. 2021;590(7846):E53. https://doi.org/10.1038/s41586-020-2308-7.
25. Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*. 2014;42(Database issue):D980–D985. http://doi.org/10.1093/nar/gkt1113.
26. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet*. 2019;104(1):21–34. http://doi.org/10.1016/j.ajhg.2018.11.002.
27. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7. http://doi.org/10.1186/s13742-015-0047-8.
28. Khera AV, Chaffin M, Zekavat SM, et al. Whole-genome sequencing to characterize monogenic and polygenic contributions in patients hospitalized with early-onset myocardial infarction. *Circulation*. 2019;139(13):1593–1602. http://doi.org/10.1161/CIRCULATIONAHA.118.035658.
29. Wray NR, Maier R. Genetic basis of complex genetic disease: the contribution of disease heterogeneity to missing heritability. *Curr Epidemiol Rep*. 2014;1(4):220–227. http://doi.org/10.1007/s40471-014-0023-3.

30. Gao C, Polley EC, Hart SN, et al. Risk of breast cancer among carriers of pathogenic variants in breast cancer predisposition genes varies by polygenic risk score. *J Clin Oncol*. 2021;39(23):2564–2573. http://doi.org/10.1200/JCO.20.01992.
31. Barnes DR, Silvestri V, Leslie G, et al. Breast and prostate cancer risks for male BRCA1 and BRCA2 pathogenic variant carriers using polygenic risk scores. *J Natl Cancer Inst*. Published online July 28, 2021. https://doi.org/10.1093/jnci/djab147.
32. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009;360(13):1320–1328. http://doi.org/10.1056/NEJMoa0810084.
33. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer*. 2013;108(11):2205–2240. http://doi.org/10.1038/bjc.2013.177.
34. Shiovitz S, Korde LA. Genetics of breast cancer: a topic in evolution. *Ann Oncol*. 2015;26(7):1291–1299. http://doi.org/10.1093/annonc/mdv022.
35. Brandão A, Paulo P, Teixeira MR. Hereditary predisposition to prostate cancer: from genetics to clinical implications. *Int J Mol Sci*. 2020;21(14):5036. http://doi.org/10.3390/ijms21145036.
36. Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc*. 2020;15(9):2759–2772. http://doi.org/10.1038/s41596-020-0353-1.
37. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of socio-demographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol*. 2017;186(9):1026–1034. http://doi.org/10.1093/aje/kwx246.
38. Kachuri L, Graff RE, Smith-Byrne K, et al. Pan-cancer analysis demonstrates that integrating polygenic risk scores with modifiable risk factors improves risk prediction. *Nat Commun*. 2020;11(1):6084. http://doi.org/10.1038/s41467-020-19600-4.
39. Al Ajmi K, Lophatananon A, Mekli K, Ollier W, Muir KR. Association of nongenetic factors with breast cancer risk in genetically predisposed groups of women in the UK Biobank cohort. *JAMA Netw Open*. 2020;3(4):e203760. http://doi.org/10.1001/jamanetworkopen.2020.3760.
40. Kim MS, Patel KP, Teng AK, Berens AJ, Lachance J. Genetic disease risks can be misestimated across global populations. *Genome Biol*. 2018;19(1):179. http://doi.org/10.1186/s13059-018-1561-7.