Am J Cancer Res 2024;14(12):5935-5951
www.ajcr.us /ISSN:2156-6976/ajcr0160629

Original Article

Interaction of breast cancer-relevant DNA repair genes and air pollution in relation to breast cancer risk in UK biobank

Carmen Smotherman[1], Brian Sprague[2], Dejana Braithwaite[1], Lusine Yaghjyan[1]

[1]*Department of Epidemiology, University of Florida, College of Public Health and Health Professions and College of Medicine, Gainesville, FL, USA;* [2]*Department of Surgery, University of Vermont, Burlington, VT, USA*

Received September 19, 2024; Accepted November 25, 2024; Epub December 15, 2024; Published December 30, 2024

Abstract: We investigated if selected polymorphisms in DNA repair genes modify the association between exposure to particulate matter ≤ 10 micron in diameter (PM_{10}) and breast cancer (BCa) risk. We included 150,929 postmenopausal women (5,969 with BCa) from UK Biobank, a population-based prospective cohort. Cancer diagnoses were ascertained through the linkage to the UK National Health Service Central Registers. Information on BCa risk factors was collected at baseline. Blood samples were collected from participants at enrollment and genotyped using the Applied Biosystems UK BiLEVE Axiom Array or the Applied Biosystems UK Biobank Axiom Array. Cox proportional hazards regression was used to examine interactions of exposure (2007 PM_{10} and cumulative average PM_{10}) with 14 SNPs, adjusting for BCa risk factors. The positive associations of 2007 PM_{10} and cumulative average PM_{10} with BCa risk were stronger in women with one or two copies of XRCC2 rs3218536 C allele vs. none (2007 PM_{10} Hazard Ratio [HR] per 10 μg/m^3 = 1.54, 95% Confidence Interval [CI] 1.22, 1.95 or HR = 1.14, 95% CI 1.03, 1.30 vs. HR = 0.52, 95% CI 0.16, 1.75, p-interaction = 0.02; cumulative average PM_{10} HR per 10 μg/m^3 = 2.80, 95% CI 1.99, 3.96 or HR = 1.89, 95% CI 1.64, 2.18 vs. HR = 0.45, 95% CI 0.08, 2.37, p-interaction = 0.05). We observed no interactions of PM_{10} with other SNPs. Our results suggest stronger associations of 2007 PM_{10} and cumulative average PM_{10} with postmenopausal BCa risk in carriers of XRCC2 rs3218536 C allele.

Keywords: Air pollution, breast cancer, UK Biobank, DNA repair genes, prospective cohort

Introduction

Accumulation of DNA damage and impaired DNA repair capacity are important events in breast cancer carcinogenesis [1]. Because of the extensive breast tissue remodeling that occurs repeatedly during a woman's lifetime (in utero development, puberty, monthly cycles, pregnancy, lactation and involution), breast tissue is particularly prone to DNA damage. Continuous DNA assaults on the breast tissue from a variety of endogenous and exogenous factors result in accumulation of cytogenetic alterations, such as deletions, amplifications and/or mutations in critical oncogenes and tumor suppressor genes, which can subsequently lead to malignant transformation [1]. Specific DNA repair mechanisms have evolved to repair different types of DNA damage and to maintain genomic integrity by repairing DNA damage or inducing apoptosis and cell cycle arrest if repair is not possible [2-4]. For example, the importance of the DNA damage repair in breast cancer is further demonstrated by the studies of mutations in p53, a tumor suppressor gene which controls cellular growth and division and protects cells against genomic alterations resulting from DNA damage by suppressing proliferation or activating apoptosis [5]; p53 gene has been linked to breast cancer risk, especially for early onset breast cancer [5, 6]. The cellular response to DNA damage and ability to maintain genomic stability by DNA repair are crucial in preventing cancer initiation and progression [7]. Higher levels of DNA damage and impaired DNA repair may increase an individual's susceptibility to breast cancer [4]. Genomic instability resulting from the failure to repair DNA damage and subsequent accumula-

https://doi.org/10.62347/WNIY6250

tion of various alterations has been linked to an increased breast cancer risk [8, 9].

Several polymorphisms have been described in DNA damage signaling and repair genes and some of the polymorphisms have been implicated in genetic susceptibility to breast cancer, including breast cancer susceptibility gene 1 (BRCA1) and breast cancer susceptibility gene 2 (BRCA2) [7, 10], X-ray repair cross complementing (XRCC) genes (specifically, XRCC1, XRCC2, and XRCC3) [7, 10-14], the excision repair cross complementing group 2 (ERCC2) [7, 10], and ataxia telangiectasia mutated (ATM) gene [7].

Air pollutants include a variety of compounds with endocrine disrupting and carcinogenic properties resulting in formation of DNA adducts, disruption of DNA damage repair, induction of carcinogen-activating enzymes, and DNA methylation of tumor suppressor genes in breast tissue [15-17]. Thus, women with different polymorphisms in DNA repair genes could potentially have different susceptibility to the effects of air pollution, but their interactions have never been examined. To our knowledge, this is the first study to investigate if selected polymorphisms in breast cancer-relevant DNA repair genes modify the association between air pollutants and breast cancer risk. For this investigation, after literature review of the existing evidence, we selected SNPs in BRCA1, BRCA2, ATM, XRCC1, XRCC2, XRCC3, and XPD (ERCC2) genes that consistently demonstrated at least moderate associations with breast cancer risk, had minor allele frequency prevalence of at least 10% in the general populations or controls, and were available in the UK Biobank. Our study specifically focused on air pollutants for which we have previously found associations with postmenopausal breast cancer in UK Biobank (2007 PM_{10} and cumulative average PM_{10}) [18].

Material and methods

Study population

We used data from the UK Biobank, an established population-based prospective cohort. Volunteers aged 40-69 years were recruited during 2006-2010 from England, Scotland and Wales via National Health Service (NHS) patient registers [19]. The enrollment process has been previously described [20]. Briefly, at enrollment participants provided health, lifestyle, and socio-demographic information, blood, urine and saliva samples, and agreed to participate in follow-ups. About 20% of participants were invited and completed their first repeat assessment during 2012-2013. Outcomes were assessed via linkages to the NHS Central Registers, with latest completed on February 29, 2020 for England and Wales, and on August 31, 2021 for Scotland.

Since the follow-up data did not include updates on women's menopausal or reproductive status, and premenopausal and postmenopausal breast cancers are epidemiologically distinct entities, this study included only postmenopausal women at baseline. Women were considered postmenopausal if they reported (i) having menopause, (ii) bilateral oophorectomy, or (iii) hysterectomy with one or both ovaries retained and being 54 years or older for an ever smoker or 56 years or older for a never smoker [18, 21]. Women with hysterectomy who had at least one ovary retained and were younger than 54 (for ever smokers) or younger than 56 (for never smokers) were excluded from the analysis as being premenopausal as done previously [22-26]. All included women had no history of breast cancer or any other type of cancer (except non-melanoma skin) at recruitment. A breast cancer diagnosis (invasive or in-situ) was identified based on diagnostic codes according to the 9th or the 10th revision of the International Classification of Diseases (174, 2330, C50, D05).

Out of 158,979 eligible women, we further excluded participants with missing information on all air pollution assessments, SNPs, or selected breast cancer risk factors (**Figure 1**). The final sample included 150,929 women (95% of all eligible women in UK Biobank) of which 5,969 developed breast cancer during the follow-up. UK Biobank protocol was approved by the NorthWest Multi-centre Research Ethics Committee (MREC), which covers the UK. All participants of UK Biobank provided written consent at recruitment.

Genotype data information

DNA was extracted from blood samples collected from participants at enrollment in UK Biobank (2006-2010) [27]. Genotyping was

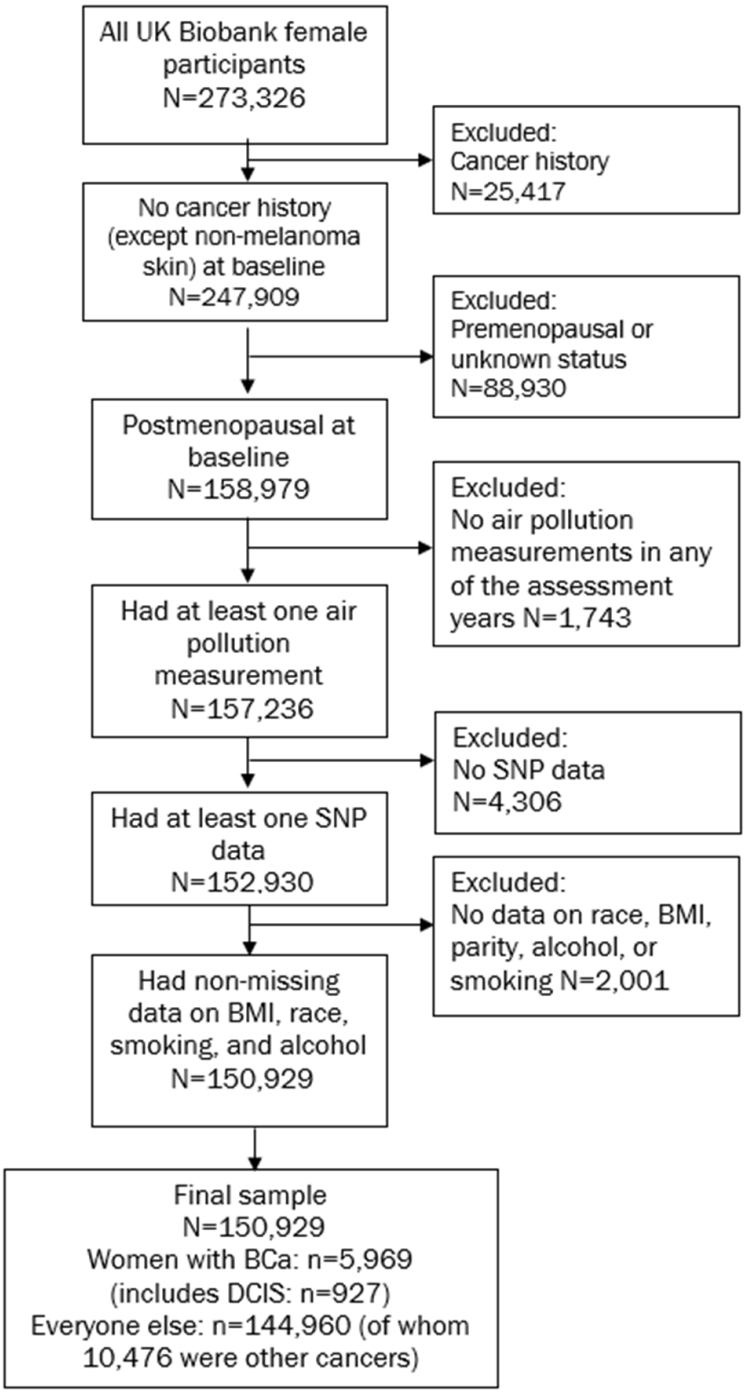

Figure 1. Study sample selection diagram

performed with two similar arrays. A subset of 49,950 participants in the UK Biobank Lung Exome Variant Evaluation (UK BiLEVE) [28] study were genotyped at 807,411 markers using the Applied Biosystems UK BiLEVE Axiom Array by Affymetrix (now part of Thermo Fisher Scientific) [29]. Next, 438,427 participants were genotyped using the closely related Applied Biosystems UK Biobank Axiom Array (825,927 markers) that shares 95% of marker content with the UK BiLEVE Axiom Array [30]. The marker content of the UK Biobank Axiom array was chosen to capture genome-wide genetic variation (single nucleotide polymorphism [SNP] and short insertions and deletions [indels]). We initially included 23 SNPs based on existing evidence of potential associations with breast cancer risk. The final set of SNPs was then limited to those with > 1% prevalence in the study sample (**Table 1**).

Air pollution data

In UK Biobank, the annual averages of particulate matter PM_{10} were available for 2007 and 2010, during the baseline assessment period. Air pollution estimates for 2007 were derived from EU-wide air pollution maps (resolution 100 m × 100 m) [31]. The x, y-coordinates of UK Biobank participants' residential locations were overlaid with these maps (projected to British National Grid) and the corresponding air pollution concentration of the 100 m × 100 m grid cell were assigned to the coordinate. EU-wide air pollution maps were modelled based on a Land Use Regression (LUR) model for Europe which also includes satellite derived air pollution estimates to improve the model performance [31]. Air pollution estimates for the year 2010 were modelled for each address using a LUR model developed as part of the European Study of Cohorts for Air Pollution Effects (ESCAPE) [32, 33]. The LUR model was based on ESCAPE monitoring done between January 26th, 2010, and January 18th, 2011, and air pollution esti-

Table 1. SNPs included in the study

Chromosome	Gene	SNP rs ID	Affy SNP ID
14	XRCC3	rs861539	s_affy10138453
17	BRCA1	rs16942	s_affy13890975
17	BRCA1	rs4986852	s_affy13890986
19	XRCC1	rs25487	s_affy15998738
19	XRCC1	rs25489	s_affy15998758
19	XRCC1	rs1799782	s_affy15998777
19	XPD (ERCC2)	rs13181	s_affy16026426
19	XPD (ERCC2)	rs1799793	s_affy16026647
7	XRCC2	rs3218536	s_affy29998477
11	ATM	rs1800056	s_affy4166044
11	ATM	rs1800057	s_affy4166144
13	BRCA2	rs144848	s_affy9226186
13	BRCA2	rs4987117	s_affy9226375
13	BRCA2	rs11571833	s_affy9227257

Abbreviations: s_aff#, Affymetrix identifier; SNP rs, unique RefSNP ID number identifying the "reference SNP cluster" containing this SNP in dbSNP; SNP, single nucleotide polymorphisms.

mates were representative for the year 2010. The cumulative average PM_{10} exposure was calculated as the average across all available measures in UK Biobank (2007 and 2010). Of all PM_{10} measures in UK Biobank, we have previously reported significant positive associations for 2007 PM_{10} and cumulative average PM_{10} which we included in the current analysis [18].

Covariates information

Information on breast cancer risk factors, such as age at menopause, age at menarche, race/ethnicity, body mass index (BMI), postmenopausal hormone therapy, parity and age at first child's birth, smoking status, alcohol use, and family history of breast cancer was collected at baseline and repeated follow-up assessments. A small percentage (< 0.5%) of women had missing information on race, parity, BMI, smoking, and alcohol, and were excluded from this study. However, for variables with relatively larger proportion of missingness (> 3%), women were classified as "Unknown" for the missing risk factor.

As we reported previously, modifiable risk factors from baseline and repeat assessments were highly correlated in a subset of women with first, second, or third subsequent assessments and thus all analyses included the baseline data on these risk factors [18].

Statistical analysis

First, Cox proportional hazards regression models were used to analyze the association between each SNP and breast cancer risk while adjusting for known breast cancer risk factors: age at recruitment (years, continuous), age at menarche (years, continuous), BMI (kg/m^2, continuous), race (Caucasian [reference], other), parity/age at first child's birth (nulliparous [reference], parous with age at first birth ≤ 25 years, parous with age at first birth > 25 years, parous with unknown age at first birth), family history of breast cancer in first degree relatives (none [reference], any, unknown), age at menopause (< 46 [reference], 46 to < 50, 50 to < 55, ≥ 55 years, unknown), postmenopausal hormone use (never [reference], past, current, unknown), smoking (never [reference], past, current), and alcohol consumption (never [reference], past, current). The follow-up time for women began at study entry and ended at the time of breast cancer diagnosis for women with breast cancer, at the time of death for women who died during the follow-up, at the time of cancer diagnosis for women who developed other cancer types during the follow-up, or at the last linkage to the cancer registries. The risk estimates were presented as hazard ratios (HRs) and their corresponding 95% confidence intervals (CIs). For each SNP, the reference genotype was the homozygous reference listed in the NIH dbSNP database.

Proportional hazards assumption was tested prior to regression analyses. Because age at recruitment violated this assumption, the interaction term between age at recruitment and time was included in the models [34].

Next, Cox proportional hazards regression was used to examine interactions between each SNP and air pollutants for which we previously found associations with BCa in this cohort [18] (2007 PM_{10} and cumulative average PM_{10}) by

Table 2. Characteristics of the study participants at baseline, by breast cancer status

Characteristic	Breast cancer (n=5,969)	No Breast cancer (n=144,960)
Mean (SD)		
Age at enrollment, years[a]	60.61 (5.11)	60.07 (5.53)
Follow-up time, years[a]	5.32 (3.14)	10.72 (1.99)
Age at menarche, years[b]	12.93 (1.62)	12.96 (1.59)
Age at menopause, years[a]	50.34 (4.93)	49.72 (5.08)
BMI, kg/m^2[a]	27.60 (5.00)	27.14 (5.06)
N (%)		
Race[#]		
White	5,757 (96.45)	138,712 (95.69)
Other	212 (3.55)	6,248 (4.31)
Parity/AFB, years[b]		
Nulliparous	1,013 (16.97)	23,476 (16.19)
Any children with AFB ≤ 25 years	2,266 (37.96)	58,179 (40.13)
Any children with AFB > 25 years	1,863 (31.21)	45,055 (31.08)
Any children with unknown AFB	827 (13.85)	18,250 (12.59)
Postmenopausal hormone therapy[a]		
Never used hormones	2,841 (50.07)	72,541 (52.62)
Past	2,141 (37.73)	52,475 (38.07)
Current	692 (12.20)	12,828 (9.30)
Family history[a]		
Breast cancer	664 (11.42)	10,484 (7.43)
No breast cancer	5,146 (88.57)	130,553 (92.57)
Smoking status[b]		
Never	3,388 (56.76)	84,990 (58.63)
Past	2,075 (34.76)	48,421 (33.4)
Current	506 (8.48)	11,549 (7.97)
Alcohol intake		
Non-drinker	342 (5.73)	8,848 (6.1)
Past drinker	209 (3.5)	5,456 (3.76)
Current drinker	5,418 (90.77)	130,656 (90.13)

Abbreviations: AFB, age at first birth; BMI, Body Mass Index; kg, kilogram; m, meters; N, number of participants; Q, quartiles; SD, standard deviation; [a]Difference statistically significant at 0.0001 level; [b]Difference statistically significant at 0.05 level. *P*-values calculates using the Pearson's Chi-square test for categorical data and the Wilcoxon Rank Sum test for continuous data.

including an interaction term in the models. Two different approaches were implemented to test the interaction between each SNP and PM_{10}: (i) using continuous air pollutant variables; and (ii) using the respective medians within each of the exposure categories. Finally, the associations between PM_{10} and breast cancer risk were stratified by genotype. Each analysis included only incident breast cancer cases diagnosed in the same year when the PM_{10} measurements were taken or those diagnosed later. The follow-up start time and the subset of women included in each analysis were based upon the measure of PM_{10} included in the given model. The follow-up start times were 2007 for the model examining PM_{10} exposure from 2007, and the last available exposure assessment year 2010 for the model analyzing the cumulative average exposure to PM_{10}. The follow-up time ended at the time of breast cancer diagnosis, of death or of other cancer diagnosis during the follow-up, or at the time of the last linkage to the cancer registries, whichever occurred first.

In the stratified analyses by genotype, PM_{10} measures were modeled as continuous variables or as quartiles based on the distribution in the study sample, specific to assessment year or cumulative average exposure for PM_{10}. The quartiles cut-offs, in μg/m^3, were as follows: 20.14, 21.72, and 23.54 for 2007 PM_{10}; and 17.92, 19.04, and 20.25 for cumulative average PM_{10}. All the tests were two-sided and the significance of the effects was assessed at 5% level of significance. All analyses were performed using SAS (SAS Institute Inc. version 9.4).

Results

In this prospective study of 150,929 postmenopausal women, 5,969 developed breast cancer and 144,960 women remained breast cancer-free during the follow-up. The mean age of the study population at enrollment was 60.1 years (range 40-71 years). The average follow-up time calculated for analyses was 5.3 years (standard deviation SD = 3.1 years) for breast cancer cases and 10.7 years (SD = 2.0 years) for cancer-free women. The distribution of baseline characteristics by breast cancer status are presented in **Table 2**. As compared to women

Table 3. Associations of SNPs with postmenopausal breast cancer risk

SNP ID	Genotype	Genotype prevalence in cases/controls (5,969/150,929)	HR[a]	P for trend
rs861539 (s_affy10138453)				
	A A	768/18598	1.019 (0.939, 1.106)	0.337
	A G	2815/66838	1.044 (0.988, 1.103)	
	G G	2326/57928	1.00	
rs16942 (s_affy13890975)				
	C C	639/15546	0.988 (0.906, 1.077)	0.628
	C T	2595/63326	0.984 (0.933, 1.039)	
	T T	2728/65962	1.00	
rs4986852 (s_affy13890986)				
	T T	2/81	0.572 (0.143, 2.289)	0.398
	T C	298/3736	1.064 (0.947, 1.196)	
	C C	5663/137940	1.00	
rs25487 (s_affy15998738)				
	C C	2340/59616	0.967 (0.892, 1.050)	0.052
	T C	2856/66293	1.060 (0.979, 1.048)	
	T T	768/18889	1.00	
rs25489 (s_affy15998758)				
	T T	9/271	0.842 (0.438, 1.619)	0.280
	T C	534/12283	1.060 (0.970, 1.158)	
	C C	5357/130747	1.00	
rs1799782 (s_affy15998777)				
	A A	33/624	1.277 (0.906, 1.798)	0.361
	A G	684/16429	1.018 (0.940, 1.102)	
	G G	5248/127771	1.00	
rs13181 (s_affy16026426)				
	G G	790/19301	0.986 (0.909, 1.068)	0.711
	G T	2750/66759	0.994 (0.941, 1.049)	
	T T	2423/58641	1.00	
rs1799793 (s_affy16026647)				
	T T	625/15816	0.934 (0.857, 1.019)	0.078
	T C	2603/63995	0.965 (0.914, 1.018)	
	C C	2721/64735	1.00	
rs3218536 (s_affy29998477)				
	T T	35/1025	0.834 (0.598, 1.163)	0.263
	T C	884/21942	0.972 (0.905, 1.044)	
	C C	5040/121810	1.00	
rs1800056 (s_affy4166044)				
	C C	1/28	0.916 (0.129, 6.508)	0.661
	C T	160/3561	1.078 (0.922, 1.262)	
	T T	5803/141243	1.00	
rs1800057 (s_affy4166144)				
	G G	2/82	0.598 (0.120, 2.393)	0.679
	G C	294/7161	0.984 (0.875, 1.107)	
	C C	5669/137535	1.00	
rs144848 (s_affy9226186)				
	C C	468/11523	0.990 (0.898, 1.091)	0.874
	C A	2408/58380	0.999 (0.947, 1.054)	
	A A	3065/74437	1.00	

rs4987117 (s_affy9226375)				
	T T	7/128	1.315 (0.627, 2.761)	0.907
	T C	378/9128	0.997 (0.899, 1.107)	
	C C	5580/135642	1.00	
rs11571833 (s_affy9227257)				
	T T	0/9	NE	0.033
	T A	125/2535	1.222 (1.023, 1.458)	
	A A	5836/142205	1.00	

Abbreviations: HR, hazard ratio; NE, non-estimable; s_aff#, Affymetrix identifier; rs, RefSNP ID number; SNP, single nucleotide polymorphisms; [a]Adjusted for age, body mass index, race, age at menopause, age at menarche, parity/age at first birth, postmenopausal hormone use, family history of breast cancer, alcohol consumption, and smoking.

without breast cancer, women with a breast cancer diagnosis were, on average, older (60.6 vs. 60.1 years, p for difference < 0.0001), had higher BMI (27.6 vs. 27.1 kg/m^2, p for difference < 0.001), were more likely to be current postmenopausal hormone therapy users at the time of enrollment (12.2% vs. 9.3%, p for difference < 0.001), and more likely to have a family history of breast cancer (11.4% vs. 7.4%, p for difference < 0.001, **Table 2**).

Associations of individual SNPs with breast cancer risk

In our sample, heterozygous carriers of BRCA2-rs11571833 T allele had an increased risk of breast cancer (HR = 1.22, 95% CI 1.02, 1.46, p-trend = 0.033) as compared to the homozygous carriers of the BRCA2-rs11571833 A allele (**Table 3**). We found no additional significant associations between other SNPs and breast cancer risk (**Table 3**).

Interactions of individual SNPs with PM_{10}

When PM_{10} was modeled as continuous measure, we found a significant interaction of SNP XRCC2-rs3218536 with 2007 PM_{10} and cumulative average PM_{10} (**Table 4**) with stronger positive associations of these exposures with breast cancer risk in heterozygous carriers of the rs3218536 T allele (2007 PM_{10}: HR per 10 μg/m^3 = 1.54, 95% CI 1.22, 1.95; cumulative average PM_{10}: HR per 10 μg/m^3 = 2.80, 95% CI 1.99, 3.96) as compared to homozygous carriers of rs3218536 C allele (2007 PM_{10}: HR per 10 μg/m^3 = 1.14, 95% CI 1.03, 1.30; cumulative average PM_{10}: HR per 10 μg/m^3 = 1. 89, 95% CI 1.64, 2.18) (**Table 4**).

When PM_{10} was modeled as quartiles, we observed similar association patterns for XRCC2-rs3218536 (**Tables 5** [for 2007 PM_{10}] and **6** [for cumulative PM_{10}]). Among heterozygous carriers of the XRCC2-rs3218536 T allele as well as among the homozygous carriers of rs3218536 C allele, women who were in the 4th quartile for 2007 PM_{10} had significant higher risk of breast cancer as compared to those in the 1st quartile (HR for 4th vs. 1st quartile = 1.37, 95% CI 1.12, 1.66 and HR = 1.13, 95% CI 1.04, 1.22, respectively) (**Table 5**). Among heterozygous carriers of the XRCC2-rs3218536 T allele as well as among the homozygous carriers of rs3218536 C allele, women in the 4th quartile for cumulative average PM_{10} had higher risk of breast cancer as compared to those in the 1st exposure quartile (HR for 4th vs. 1st quartile = 1.45, 95% CI 1.21, 1.74 and HR = 1.34, 95% CI 1.24, 1.45, respectively) (**Table 6**).

Discussion

In this large prospective cohort of postmenopausal women enrolled in the UK Biobank, we investigated the association of selected SNPs in DNA damage signaling and repair genes with postmenopausal breast cancer risk, as well as whether the association of PM_{10} with postmenopausal breast cancer risk differed by selected SNPs genotypes. The associations of 2007 PM_{10} and cumulative average PM_{10} with breast cancer risk differed across genotypes in XRCC2-rs3218536, with stronger associations in heterozygous carriers XRCC2-rs3218536 T allele as well as homozygous carriers of rs3218536 C.

The findings on a positive association of SNP BRCA2-rs11571833 with breast cancer are consistent with results from a meta-analysis of 9 genome-wide association studies, including 10,052 breast cancer cases and 12,575 con-

Table 4. Association of PM_{10} (continuous measure) with postmenopausal breast cancer risk, by genotype

SNP/Gene Genotype	PM_{10} 2007	Cumulative average PM_{10}
	HR per 10 μg/m³ (95% CI)[a]	HR per 10 μg/m³ (95% CI)[a]
rs861539 (s_affy10138453)/XRCC3		
A A	1.120 (0.870, 1.450)	2.030 (1.410, 2.930)
A G	1.180 (1.003, 1.350)	2.050 (1.700, 2.480)
G G	1.230 (1.060, 1.420)	1.930 (1.570, 2.390)
P for interaction	0.825	0.502
rs16942 (s_affy13890975)/BRCA1		
C C	1.200 (0.910, 1.590)	1.872 (1.252, 2.799)
C T	1.170 (1.020, 1.340)	1.994 (1.635, 2.432)
T T	1.200 (1.005, 1.370)	2.012 (1.657, 2.433)
P for interaction	0.868	0.932
rs4986852 (s_affy13890986)/BRCA1		
T T	NE	NE
T C	0.830 (0.540, 1.260)	1.590 (0.870, 2.900)
C C	1.0121 (1.010, 1.330)	2.010 (1.760, 2.300)
P for interaction	0.094	0.456
rs25487 (s_affy15998738)/XRCC1		
C C	1.120 (0.970, 1.290)	1.836 (1.490, 2.261)
T C	1.260 (1.110, 1.440)	2.176 (1.799, 2.632)
T T	1.130 (0.880, 1.460)	1.809 (1.254, 2.610)
P for interaction	0.406	0.460
rs25489 (s_affy15998758)/XRCC1		
T T	1.219 (0.927, 1.603)	NE
T C	1.260 (0.930, 1.720)	1.864 (1.195, 2.906)
CC	1.170 (1.070, 1.290)	1.968 (1.714, 2.260)
P for interaction	0.836	0.064
rs1799782 (s_affy15998777)/XRCC1		
A A	0.680 (0.190, 2.450)	0.430 (0.070, 2.740)
A G	0.980 (0.740, 1.290)	1.010 (1.360, 2.970)
G G	1.020 (1.010, 1.030)	2.000 (1.740, 2.300)
P for interaction	0.413	0.672
rs13181 (s_affy16026426)/XPD (ERCC2)		
G G	1.210 (0.940, 1.560)	1.990 (1.390, 2.860)
G T	1.130 (0.990, 1.290)	1.980 (1.630, 2.400)
T T	1.0230 (1.070, 1.410)	1.970 (1.600, 2.410)
P for interaction	0.656	0.185
rs1799793 (s_affy16026647)/XPD (ERCC2)		
T T	1.220 (0.920, 1.620)	1.850 (1.230, 2.780)
T C	1.022 (1.060, 1.390)	2.120 (1.740, 2.590)
C C	1.014 (1.000, 1.300)	1.870 (1.540, 2.280)
P for interaction	0.753	0.663
rs3218536 (s_affy29998477)/XRCC2		
T T	0.520 (0.160, 1.750)	0.450 (0.080, 2.370)
T C	1.540 (1.220, 1.950)	2.800 (1.990, 3.930)
C C	1.140 (1.030, 1.250)	1.890 (1.640, 2.180)
P for interaction	0.023	0.048

rs1800056 (s_affy4166044)/ATM		
C C	NE	NE
C T	1.510 (0.870, 2.600)	2.090 (0.950, 4.630)
T T	1.170 (1.070, 1.290)	1.980 (1.730, 2.260)
P for interaction	0.651	0.622
rs1800057 (s_affy4166144)/ATM		
G G	NE	NE
G C	1.230 (0.810, 1.850)	1.800 (1.300, 3.260)
C C	1.180 (1.070, 1.290)	1.990 (1.740, 2.270)
P for interaction	0.361	0.591
rs144848 (s_affy9226186)/BRCA2		
C C	1.240 (0.900, 1.700)	2.205 (1.401, 3.469)
C A	1.240 (1.070, 1.430)	1.969 (1.602, 2.421)
A A	1.140 (1.000, 1.290)	1.970 (1.640, 2.368)
P for interaction	0.685	0.925
rs4987117 (s_affy9226375)/BRCA2		
T T	NE	NE
T C	1.320 (0.920, 1.900)	2.440 (1.430, 4.150)
C C	1.180 (1.070, 1.290)	1.960 (1.710, 2.240)
P for interaction	0.178	0.106
rs11571833 (s_affy9227257)/BRCA2		
T T	NE	NE
T A	1.180 (0.630, 2.210)	1.250 (0.900, 5.590)
A A	1.180 (1.080, 1.300)	1.980 (1.740, 2.260)
P for interaction	0.999	0.539

Abbreviations: CI, Confidence interval; HR, Hazard ratio; NE, Non-estimable; rs, RefSNP ID number; [a]Adjusted for age, body mass index, race, age at menopause, age at menarche, parity/age at first birth, postmenopausal hormone use, family history of breast cancer, alcohol consumption, and smoking.

Table 5. Association of quartiles of 2007 PM_{10} with postmenopausal breast cancer risk, by genotype

SNP Genotype	PM_{10} 2007 2nd quartile	PM_{10} 2007 3rd quartile	PM_{10} 2007 4th quartile	P for trend[b]
	HR (95% CI)[a]	HR (95% CI)[a]	HR (95% CI)[a]	
rs861539 (s_affy10138453)/XRCC3				
A A	1.065 (0.866, 1.309)	1.291 (1.060, 1.573)	1.042 (0.843, 1.289)	0.044
A G	1.158 (1.042, 1.287)	1.126 (1.012, 1.253)	1.149 (1.031, 1.281)	0.026
G G	1.254 (1.116, 1.408)	1.123 (0.996, 1.266)	1.210 (1.073, 1.363)	< 0.001
P for interaction	0.841			
rs16942 (s_affy13890975)/BRCA1				
C C	1.194 (0.954, 1.494)	1.140 (0.907, 1.433)	1.155 (0.918, 1.453)	0.450
T C	1.252 (1.121, 1.399)	1.217 (1.088, 1.361)	1.186 (1.058, 1.329)	< 0.001
T T	1.108 (0.996, 1.233)	1.066 (0.957, 1.187)	1.128 (1.011, 1.257)	0.137
P for interaction	0.753			
rs4986852 (s_affy13890986)/BRCA1				
T T	NE	NE	NE	NE
T C	1.027 (0.750, 1.407)	1.141 (0.840, 1.550)	0.842 (0.594, 1.193)	0.383
C C	1.187 (1.101, 1.279)	1.141 (1.058, 1.231)	1.172 (1.086, 1.265)	< 0.001
P for interaction	0.131			

rs25487 (s_affy15998738)/XRCC1				
C C	1.224 (1.091, 1.374)	1.134 (1.008, 1.275)	1.116 (0.990, 1.258)	0.008
T C	1.135 (1.021, 1.261)	1.131 (1.017, 1.258)	1.189 (1.068, 1.324)	0.012
T T	1.198 (0.978, 1.468)	1.177 (0.960, 1.444)	1.140 (0.925, 1.406)	0.307
P for interaction	0.490			
rs25489 (s_affy15998758)/XRCC1				
T T	0.220 (0.017, 2.804)	1.222 (0.169, 8.858)	2.153 (0.328, 1.135)	0.351
T C	1.276 (1.003, 1.623)	1.102 (0.860, 1.413)	1.192 (0.927, 1.532)	0.228
CC	1.168 (1.082, 1.261)	1.141 (1.056, 1.233)	1.145 (1.058, 1.239)	< 0.001
P for interaction	0.879			
rs1799782 (s_affy15998777)/XRCC1				
A A	1.289 (0.493, 3.370)	0.615 (0.199, 1.901)	0.985 (0.354, 2.742)	0.641
A G	1.017 (0.822, 1.259)	1.032 (0.835, 1.276)	0.926 (0.742, 1.156)	0.780
G G	1.200 (1.110, 1.297)	1.158 (1.071, 1.253)	1.187 (1.096, 1.285)	< 0.001
P for interaction	0.285			
rs13181 (s_affy16026426)/XPD (ERCC2)				
G G	1.222 (1.000, 1.494)	1.285 (1.051, 1.572)	1.229 (0.997, 1.514)	0.079
G T	1.200 (1.079, 1.334)	1.114 (0.999, 1.241)	1.130 (1.012, 1.261)	0.009
T T	1.132 (1.009, 1.270)	1.118 (0.997, 1.255)	1.143 (1.017, 1.284)	0.088
P for interaction	0.718			
rs1799793 (s_affy16026647)/XPD (ERCC2)				
T T	1.193 (0.953, 1.494)	1.295 (1.035, 1.620)	1.181 (0.934, 1.493)	0.154
T C	1.246 (1.116, 1.391)	1.190 (1.064, 1.331)	1.201 (1.072, 1.347)	< 0.001
C C	1.118 (1.004, 1.245)	1.061 (0.952, 1.182)	1.100 (0.986, 1.227)	0.188
P for interaction	0.604			
rs3218536 (s_affy29998477)/XRCC2				
T T	1.322 (0.583, 2.996)	0.687 (0.250, 1.886)	0.532 (0.180, 1.570)	0.307
T C	1.327 (1.096, 1.607)	1.224 (1.007, 1.488)	1.366 (1.124, 1.661)	0.008
C C	1.156 (1.068, 1.251)	1.130 (1.043, 1.223)	1.126 (1.038, 1.221)	0.002
P for interaction	0.068			
rs1800056 (s_affy4166044)/ATM				
C C	NE	NE	NE	
C T	2.899 (1.753, 4.792)	2.641 (1.594, 4.377)	1.559 (0.886, 2.742)	< 0.001
T T	1.149 (1.068, 1.237)	1.113 (1.033, 1.199)	1.144 (1.061, 1.233)	< 0.001
P for interaction	0.881			
rs1800057 (s_affy4166144)/ATM				
G G	NE	NE	NE	NE
G C	1.956 (1.384, 2.764)	1.833 (1.297, 2.590)	1.345 (0.923, 1.959)	< 0.001
C C	1.148 (1.065, 1.236)	1.112 (1.031, 1.199)	1.144 (1.060, 1.235)	< 0.001
P for interaction	0.632			
rs144848 (s_affy9226186)/BRCA2				
C C	1.102 (0.847, 1.435)	1.282 (0.991, 1.658)	1.136 (0.871, 1.482)	0.299
C A	1.194 (1.064, 1.341)	1.211 (1.078, 1.360)	1.225 (1.089, 1.378)	0.002
A A	1.173 (1.061, 1.297)	1.069 (0.964, 1.184)	1.102 (0.993, 1.223)	0.178
P for interaction	0.303			
rs4987117 (s_affy9226375)/BRCA2				
T T	NE	NE	NE	NE
T C	1.173 (0.880, 1.564)	1.124 (0.839, 1.507)	1.287 (0.963, 1.721)	0.392
C C	1.177 (1.091, 1.268)	1.140 (1.056, 1.229)	1.144 (1.059, 1.235)	< 0.001
P for interaction	0.306			

rs11571833 (s_affy9227257)/BRCA2				
T T	NE	NE	NE	NE
T A	1.270 (0.755, 2.137)	1.543 (0.929, 2.562)	1.267 (0.741, 2.617)	0.422
A A	1.173 (1.090, 1.263)	1.132 (1.051, 1.263)	1.149 (1.066, 1.239)	< 0.001
P for interaction	0.931			

Abbreviations: CI, Confidence interval; HR, Hazard ratio; NE, Non-estimable; rs, RefSNP ID number; s_aff#, Affymetrix identifier; SNP, single nucleotide polymorphisms; [a]Adjusted for age, body mass index, race, age at menopause, age at menarche, parity/age at first birth, postmenopausal hormone use, family history of breast cancer, alcohol consumption, and smoking. [b]P for trend using median air pollutant level within each genotype; Note: Quartiles defined as: 1st: ≤ 20.14, 2nd: > 20.14- ≤ 21.72, 3rd: > 21.72- ≤ 23.54, 4th: > 23.54.

Table 6. Association of quartiles of cumulative average PM_{10} with breast cancer risk, by SNPs alleles

SNP Genotype	Cum average PM_{10} 2nd quartile	Cum average PM_{10} 3rd quartile	Cum average PM_{10} 4th quartile	P for trend[b]
	HR (95% CI)[a]	HR (95% CI)[a]	HR (95% CI)[a]	
rs861539 (s_affy10138453)/XRCC3				
A A	0.983 (0.799, 1.209)	1.029 (0.837, 1.264)	1.363 (1.118, 1.662)	0.002
A G	1.064 (0.955, 1.185)	1.034 (0.928, 1.153)	1.369 (1.234, 1.520)	< 0.001
G G	1.084 (0.962, 1.222)	1.092 (0.969, 1.231)	1.344 (1.196, 1.510)	< 0.001
P for interaction	0.542			
rs16942 (s_affy13890975)/BRCA1				
C C	1.078 (0.859, 1.353)	1.010 (0.801, 1.273)	1.306 (1.047, 1.630)	0.053
C T	1.113 (0.995, 1.244)	1.042 (0.930, 1.168)	1.345 (1.206, 1.501)	< 0.001
T T	0.996 (0.892, 1.113)	1.070 (0.959, 1.193)	1.369 (1.232, 1.522)	< 0.001
P for interaction	0.908			
rs4986852 (s_affy13890986)/BRCA1				
T T	NE	NE	NE	NE
T C	0.994 (0.713, 1.386)	1.365 (0.998, 1.866)	1.206 (0.864, 1.684)	0.135
C C	1.058 (0.980, 1.142)	0.035 (0.958, 1.117)	1.360 (1.264, 1.464)	< 0.001
P for interaction	0.636			
rs25487 (s_affy15998738)/XRCC1				
C C	1.046 (0.929, 1.178)	1.071 (0.951, 1.205)	1.315 (1.172, 1.476)	< 0.001
T C	1.044 (0.937, 1.162)	1.037 (0.930, 1.155)	1.393 (1.256, 1.544)	< 0.001
T T	1.131 (0.922, 1.388)	1.043 (0.846, 1.286)	1.316 (1.075, 1.609)	0.360
P for interaction	0.515			
rs25489 (s_affy15998758)/XRCC1				
T T	NE	NE	NE	NE
T C	1.124 (0.881, 1.434)	1.070 (0.835, 1.369)	1.243 (0.975, 1.586)	0.346
C C	1 047 (0.968, 1.133)	1.044 (0.965, 1.129)	1.355 (1.256, 1.462)	< 0.001
P for interaction	0.159			
rs1799782 (s_affy15998777)/XRCC1				
A A	1.017 (0.391, 2.648)	0.647 (0.210, 1.997)	0.865 (0.317, 2.358)	0.864
A G	1.021 (0.818, 1.275)	1.023 (0.821, 1.277)	1.295 (1.045, 1.605)	0.045
G G	1.061 (0.980, 1.148)	1.059 (0.978, 1.147)	1.361 (1.261, 1.470)	< 0.001
P for interaction	0.940			
rs13181 (s_affy16026426)/XPD (ERCC2)				
G G	0.996 (0.815, 1.217)	1.002 (0.818, 1.228)	1.350 (1.111, 1.639)	0.003
G T	1.074 (0.963, 1.198)	1.049 (0.940, 1.171)	1.349 (1.213, 1.500)	< 0.001
T T	1.050 (0.934, 1.181)	1.062 (0.945, 1.195)	1.345 (1.201, 1.506)	< 0.001
P for interaction	0.252			

rs1799793 (s_affy16026647)/XPD (ERCC2)				
T T	1.056 (0.843, 1.323)	1.080 (0.862, 1.354)	1.356 (1.086, 1.693)	0.350
T C	1.072 (0.957, 1.199)	1.081 (0.965, 1.210)	1.390 (1.247, 1.550)	< 0.001
C C	1.039 (0.930, 1.160)	1.011 (0.905, 1.130)	1.310 (1.178, 1.458)	< 0.001
P for interaction	0.738			
rs3218536 (s_affy29998477)/XRCC2				
T T	1.449 (0.595, 3.529)	0.128 (0.445, 2.858)	0.607 (0.204, 1.806)	0.475
T C	0.904 (0.741, 1.103)	1.096 (0.905, 1.326)	1.448 (1.205, 1.740)	< 0.001
C C	1.083 (0.999, 1.174)	1.046 (0.964, 1.135)	1.341 (1.240, 1.450)	< 0.001
P for interaction	0.205			
rs1800056 (s_affy4166044)/ATM				
C C	NE	NE	NE	NE
C T	1.663 (1.053, 2.626)	1.573 (0.994, 2.492)	1.370 (0.838, 2.240)	0.145
T T	1.042 (0.967, 1.124)	1.039 (0.963, 1.120)	1.350 (1.255, 1.452)	< 0.001
P for interaction	0.612			
rs1800057 (s_affy4166144)/ATM				
G G	NE	NE	NE	NE
G C	1.178 (0.845, 1.643)	1.183 (0.852, 1.643)	1.278 (0.913, 1.788)	0.545
C C	1.047 (0.970, 1.130)	1.043 (0.966, 1.126)	1.352 (1.256, 1.455)	< 0.001
P for interaction	0.617			
rs144848 (s_affy9226186)/BRCA2				
C C	1.203 (0.920, 1.573)	1.231 (0.942, 1.610)	1.472 (1.135, 1.910)	< 0.001
C A	1.130 (1.005, 1.269)	1.088 (0.966, 1.225)	1.393 (1.243, 1.561)	< 0.001
A A	0.976 (0.880, 1.083)	1.001 (0.903, 1.110)	1.300 (1.177, 1.437)	< 0.001
P for interaction	0.946			
rs4987117 (s_affy9226375)/BRCA2				
T T	NE	NE	NE	NE
T C	1.055 (0.785, 1.419)	1.001 (0.744, 1.347)	1.540 (1.162, 2.040)	0.004
C C	1.055 (0.977, 1.139)	1.056 (0.978, 1.141)	1.338 (1.242, 1.442)	< 0.001
P for interaction	0.157			
rs11571833 (s_affy9227257)/BRCA2				
T T	NE	NE	NE	NE
T A	1.738 (1.000, 3.019)	1.839 (1.054, 3.207)	1.906 (1.092, 3.325)	0.050
A A	1.044 (0.968, 1.125)	1.040 (0.964, 1.121)	1.344 (1.250, 1.445)	0.261
P for interaction	0.362			

Abbreviations: CI, Confidence interval; HR, Hazard ratio; NE, Non-estimable; rs, RefSNP ID number; s_aff#, Affymetrix identifier; SNP, single nucleotide polymorphisms; [a]Adjusted for age, body mass index, race, age at menopause, age at menarche, parity/age at first birth, postmenopausal hormone use, family history of breast cancer, alcohol consumption, and smoking. [b]P for trend using median air pollutant level within each genotype; Note: Quartiles defined as: 1st: ≤ 17.92, 2nd: > 17.92- ≤ 19.04, 3rd: > 19.04- ≤ 20.25, 4th: > 20.25

trols of European ancestry (odds ratio [OR] = 1.39, 95% CI 1.13, 1.71). The results are also consistent with the genome-wide association study that identified a significant association between the variants and the risk of breast cancer (OR = 1.26, 95% CI 1.14, 1.39, P = 4.9 × 10^{-8}) [35]. Similar positive association between rs11571833 and breast cancer risk were also found in other previous studies [36-38].

We reported significant interactions of PM_{10} with XRCC2-rs3218536, a C to T polymorphism located in exon 3 of the XRCC2 gene that has been investigated for its potential impact on breast cancer susceptibility. Consistent with previous studies showing that XRCC2-rs3218536 T allele has a protective effect on breast cancer risk [39-43], or epithelial ovarian cancer risk [44], we found stronger

associations of 2007 PM_{10} and cumulative average PM_{10} with breast cancer risk in heterozygous carriers of the T allele and in women homozygous for rs3218536 C allele. Further studies are needed to confirm our findings which could potentially offer some new insights into breast cancer risk management in risk allele carriers living in the areas with high air pollution levels.

To our best knowledge, this is the first study to examine the interactions of PM_{10} with SNPs in DNA repair genes with respect to postmenopausal breast cancer risk. We utilized a population-based prospective cohort with a stringent protocol, rigorous ascertainment of the health outcomes through continuous linkages to the national registries, and well-validated methods for air pollution assessment. Our study has a few notable limitations. First, even though information on covariates was used only at baseline, correlations with values from follow-up assessments were high. Therefore, evaluating participant risk factors at baseline is unlikely to introduce misclassification. Second, since air pollution exposures were estimated based on a single home address, contribution of external factors unrelated to residential locations to the exposure misclassification cannot be addressed. Nonetheless, recent studies demonstrated very small contribution of commuting to total weekly exposure with negligible underestimation of health effects as compared to the models combining home and workplace exposures [45, 46]. Finally, as this is the first study to examine interactions of SNPs in DNA repair genes with air pollution in relation to breast cancer risk, the findings require further external validation in subsequent investigations.

In conclusion, we found, for the first time, significant interactions of SNPs XRCC2-rs3218536 with PM_{10} exposure with stronger associations among risk allele carriers. Future studies are warranted to confirm our results and to elucidate the potential biological mechanisms underlying the observed associations. Future studies would also benefit from inclusion of polymorphisms in genes involved in other potential biological pathways that could interact with air pollution given underlying mechanisms behind the specific PM_{10} constituents, for example, endocrine disruption.

Acknowledgements

This research has been conducted using the UK Biobank Resource under Application Number 67356.

Disclosure of conflict of interest

None.

Address correspondence to: Carmen Smotherman, Department of Epidemiology, University of Florida, College of Public Health and Health Professions and College of Medicine, 2004 Mowry Road, Gainesville, FL 32610, USA. Tel: 904-962-1884; ORCID: 0000-0002-6544-7757; E-mail: carmen.msmotherman@jax.ufl.edu

References

[1] Kitagishi Y, Kobayashi M and Matsuda S. Defective DNA repair systems and the development of breast and prostate cancer (review). Int J Oncol 2013; 42: 29-34.

[2] Santarpia L, Iwamoto T, Di Leo A, Hayashi N, Bottai G, Stampfer M, Andre F, Turner NC, Symmans WF, Hortobagyi GN, Pusztai L and Bianchini G. DNA repair gene patterns as prognostic and predictive factors in molecular breast cancer subtypes. Oncologist 2013; 18: 1063-1073.

[3] Jackson SP and Bartek J. The DNA-damage response in human biology and disease. Nature 2009; 461: 1071-1078.

[4] Smith TR, Miller MS, Lohman KK, Case LD and Hu JJ. DNA damage and breast cancer risk. Carcinogenesis 2003; 24: 883-889.

[5] Blondeaux E, Arecco L, Punie K, Graffeo R, Toss A, De Angelis C, Trevisan L, Buzzatti G, Linn SC, Dubsky P, Cruellas M, Partridge AH, Balmaña J, Paluch-Shimon S and Lambertini M. Germline TP53 pathogenic variants and breast cancer: a narrative review. Cancer Treat Rev 2023; 114: 102522.

[6] Evans DG, Woodward ER, Bajalica-Lagercrantz S, Oliveira C and Frebourg T. Germline TP53 testing in breast cancers: why, when and how? Cancers (Basel) 2020; 12: 3762.

[7] Costa S, Medeiros R and Schmitt F. DNA signalling/repair genetic polymorphisms and breast cancer risk: a review. Applied Cancer Research 2005; 25: 161-180.

[8] Kwei KA, Kung Y, Salari K, Holcomb IN and Pollack JR. Genomic instability in breast cancer: pathogenesis and clinical implications. Mol Oncol 2010; 4: 255-266.

[9] Duijf PHG, Nanayakkara D, Nones K, Srihari S, Kalimutho M and Khanna KK. Mechanisms of

genomic instability in breast cancer. Trends Mol Med 2019; 25: 595-611.
[10] Goode EL, Ulrich CM and Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol Biomarkers Prev 2002; 11: 1513-1530.
[11] Chacko P, Rajan B, Joseph T, Mathew BS and Pillai MR. Polymorphisms in DNA repair gene XRCC1 and increased genetic susceptibility to breast cancer. Breast Cancer Res Treat 2005; 89: 15-21.
[12] Figueiredo JC, Knight JA, Briollais L, Andrulis IL and Ozcelik H: Polymorphisms XRCC1-R399Q and XRCC3-T241M and the risk of breast cancer at the Ontario site of the Breast Cancer Family Registry. Cancer Epidemiol Biomarkers Prev 2004; 13: 583-591.
[13] Smith TR, Miller MS, Lohman K, Lange EM, Case LD, Mohrenweiser HW and Hu JJ. Polymorphisms of XRCC1 and XRCC3 genes and susceptibility to breast cancer. Cancer Lett 2003; 190: 183-190.
[14] Sterpone S and Cozzi R. Influence of XRCC1 genetic polymorphisms on ionizing radiation-induced DNA damage and repair. J Nucleic Acids 2010; 2010: 780369.
[15] Qu SX and Stacey NH. Formation and persistence of DNA adducts in different target tissues of rats after multiple administration of benzo[a]pyrene. Carcinogenesis 1996; 17: 53-59.
[16] Stowers SJ and Anderson MW. Formation and persistence of benzo(a)pyrene metabolite-DNA adducts. Environ Health Perspect 1985; 62: 31-39.
[17] Callahan CL, Bonner MR, Nie J, Han D, Wang Y, Tao MH, Shields PG, Marian C, Eng KH, Trevisan M, Beyea J and Freudenheim JL: Lifetime exposure to ambient air pollution and methylation of tumor suppressor genes in breast tumors. Environ Res 2017; 161:418-424.
[18] Smotherman C, Sprague B, Datta S, Braithwaite D, Qin H and Yaghjyan L. Association of air pollution with postmenopausal breast cancer risk in UK Biobank. Breast Cancer Res 2023; 25: 83.
[19] Stanczyk FZ, Ploszaj S, Gentzschein E, Qian D and Mishell DR Jr. Effect of oral contraceptives containing 20 and 35 micrograms ethinyl estradiol on urinary prostacyclin and thromboxane metabolite levels in smokers and nonsmokers. Contraception 1999; 59: 17-23.
[20] Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T and Collins R. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015; 12: e1001779.
[21] Willett W, Stampfer MJ, Bain C, Lipnick R, Speizer FE, Rosner B, Cramer D and Hennekens CH. Cigarette smoking, relative weight, and menopause. Am J Epidemiol 1983; 117: 651-658.
[22] Yaghjyan L, Colditz GA, Collins LC, Schnitt SJ, Rosner B, Vachon C and Tamimi RM. Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to tumor characteristics. J Natl Cancer Inst 2011; 103: 1179-1189.
[23] Emaus A, Dieli-Conwright C, Xu X, Lacey JV Jr, Ingles SA, Reynolds P, Bernstein L and Henderson KD. Increased long-term recreational physical activity is associated with older age at natural menopause among heavy smokers: the California Teachers Study. Menopause 2013; 20: 282-290.
[24] Yaghjyan L, Colditz GA, Rosner B and Tamimi RM. Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to the time since the mammogram. Cancer Epidemiol Biomarkers Prev 2013; 22: 1110-1117.
[25] Yaghjyan L, Wang Z, Warner ET, Rosner B, Heine J and Tamimi RM. Reproductive factors related to childbearing and a novel automated mammographic measure, V. Cancer Epidemiol Biomarkers Prev 2024; 33: 804-811.
[26] Yaghjyan L, Austin-Datta RJ, Oh H, Heng YJ, Vellal AD, Sirinukunwattana K, Baker GM, Collins LC, Murthy D, Rosner B and Tamimi RM. Associations of reproductive breast cancer risk factors with breast tissue composition. Breast Cancer Res 2021; 23: 70.
[27] Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P and Marchini J. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018; 562: 203-209.
[28] Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, Soler Artigas M, Billington CK, Kheirallah AK, Allen R, Cook JP, Probert K, Obeidat M, Bossé Y, Hao K, Postma DS, Paré PD and Ramasamy A; UK Brain Expression Consortium (UKBEC); Mägi R, Mihailov E, Reinmaa E, Melén E, O'Connell J, Frangou E and Delaneau O; OxGSK Consortium; Freeman C, Petkova D, McCarthy M, Sayers I, Deloukas P, Hubbard R, Pavord I, Hansell AL, Thomson NC, Zeggini E, Morris AP, Marchini J, Strachan DP, Tobin MD and Hall IP. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic as-

sociation study in UK Biobank. Lancet Respir Med 2015; 3:769-781.
[29] Zielinska E, Niewiarowski W, Bodalski J, Stanczyk A, Bolanowski W and Rebowski G. Arylamine N-acetyltransferase (NAT2) gene mutations in children with allergic diseases. Clin Pharmacol Ther 1997; 62: 635-642.
[30] Vatten LJ, Ursin G, Ross RK, Stanczyk FZ, Lobo RA, Harvei S and Jellum E. Androgens in serum and the risk of prostate cancer: a nested case-control study from the Janus serum bank in Norway. Cancer Epidemiol Biomarkers Prev 1997; 6: 967-969.
[31] Vienneau D, de Hoogh K, Bechle MJ, Beelen R, van Donkelaar A, Martin RV, Millet DB, Hoek G and Marshall JD. Western European Land Use Regression Incorporating Satellite- and Ground-Based Measurements of NO2 and PM10. Environ Sci Technol 2013; 47: 13555-13564.
[32] Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, Declercq C, Dėdelė A, Dons E, de Nazelle A, Dimakopoulou K, Eriksen K, Falq G, Fischer P, Galassi C, Gražulevičienė R, Heinrich J, Hoffmann B, Jerrett M, Keidel D, Korek M, Lanki T, Lindley S, Madsen C, Mölter A, Nádor G, Nieuwenhuijsen M, Nonnemacher M, Pedeli X, Raaschou-Nielsen O, Patelarou E, Quass U, Ranzi A, Schindler C, Stempfelet M, Stephanou E, Sugiri D, Tsai MY, Yli-Tuomi T, Varró MJ, Vienneau D, Klot Sv, Wolf K, Brunekreef B and Hoek G. Development of Land Use Regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. Environ Sci Technol 2012; 46: 11195-11205.
[33] Beelen R, Hoek G, Vienneau D, Eeftens M, Dimakopoulou K, Pedeli X, Tsai MY, Künzli N, Schikowski T, Marcon A, Eriksen KT, Raaschou-Nielsen O, Stephanou E, Patelarou E, Lanki T, Yli-Tuomi T, Declercq C, Falq G, Stempfelet M, Birk M, Cyrys J, von Klot S, Nádor G, Varró MJ, Dėdelė A, Gražulevičienė R, Mölter A, Lindley S, Madsen C, Cesaroni G, Ranzi A, Badaloni C, Hoffmann B, Nonnemacher M, Krämer U, Kuhlbusch T, Cirach M, de Nazelle A, Nieuwenhuijsen M, Bellander T, Korek M, Olsson D, Strömgren M, Dons E, Jerrett M, Fischer P, Wang M, Brunekreef B and de Hoogh K. Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe - The ESCAPE project. Atmos Environ 2013; 72: 10-23.
[34] Therneau T and Grambsch P. Modeling survival data: extending the Cox model. 2000; 48.
[35] Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, Lee A, Turnbull C and Rahman N; Breast and Ovarian Cancer Susceptibility Collaboration; Fletcher O, Peto J, Gibson L, Dos Santos Silva I, Nevanlinna H, Muranen TA, Aittomäki K, Blomqvist C, Czene K, Irwanto A, Liu J, Waisfisz Q, Meijers-Heijboer H and Adank M; Hereditary Breast and Ovarian Cancer Research Group Netherlands (HEBON); van der Luijt RB, Hein R, Dahmen N, Beckman L, Meindl A, Schmutzler RK, Müller-Myhsok B, Lichtner P, Hopper JL, Southey MC, Makalic E, Schmidt DF, Uitterlinden AG, Hofman A, Hunter DJ, Chanock SJ, Vincent D, Bacot F, Tessier DC, Canisius S, Wessels LF, Haiman CA, Shah M, Luben R, Brown J, Luccarini C, Schoof N, Humphreys K, Li J, Nordestgaard BG, Nielsen SF, Flyger H, Couch FJ, Wang X, Vachon C, Stevens KN, Lambrechts D, Moisse M, Paridaens R, Christiaens MR, Rudolph A, Nickels S, Flesch-Janys D, Johnson N, Aitken Z, Aaltonen K, Heikkinen T, Broeks A, Veer LJ, van der Schoot CE, Guénel P, Truong T, Laurent-Puig P, Menegaux F, Marme F, Schneeweiss A, Sohn C, Burwinkel B, Zamora MP, Perez JI, Pita G, Alonso MR, Cox A, Brock IW, Cross SS, Reed MW, Sawyer EJ, Tomlinson I, Kerin MJ, Miller N, Henderson BE, Schumacher F, Le Marchand L, Andrulis IL, Knight JA, Glendon G and Mulligan AM; kConFab Investigators; Australian Ovarian Cancer Study Group; Lindblom A, Margolin S, Hooning MJ, Hollestelle A, van den Ouweland AM, Jager A, Bui QM, Stone J, Dite GS, Apicella C, Tsimiklis H, Giles GG, Severi G, Baglietto L, Fasching PA, Haeberle L, Ekici AB, Beckmann MW, Brenner H, Müller H, Arndt V, Stegmaier C, Swerdlow A, Ashworth A, Orr N, Jones M, Figueroa J, Lissowska J, Brinton L, Goldberg MS, Labrèche F, Dumont M, Winqvist R, Pylkäs K, Jukkola-Vuorinen A, Grip M, Brauch H, Hamann U and Brüning T; GENICA (Gene Environment Interaction and Breast Cancer in Germany) Network; Radice P, Peterlongo P, Manoukian S, Bonanni B, Devilee P, Tollenaar RA, Seynaeve C, van Asperen CJ, Jakubowska A, Lubinski J, Jaworska K, Durda K, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM, Bogdanova NV, Antonenkova NN, Dörk T, Kristensen VN, Anton-Culver H, Slager S, Toland AE, Edge S, Fostira F, Kang D, Yoo KY, Noh DY, Matsuo K, Ito H, Iwata H, Sueta A, Wu AH, Tseng CC, Van Den Berg D, Stram DO, Shu XO, Lu W, Gao YT, Cai H, Teo SH, Yip CH, Phuah SY, Cornes BK, Hartman M, Miao H, Lim WY, Sng JH, Muir K, Lophatananon A, Stewart-Brown S, Siriwanarangsan P, Shen CY, Hsiung CN, Wu PE, Ding SL, Sangrajrang S, Gaborieau V, Brennan P, McKay J, Blot WJ, Signorello LB, Cai Q, Zheng W, Deming-Halverson S, Shrubsole M, Long J, Simard J, Garcia-Closas M, Pharoah PD, Chenevix-

Trench G, Dunning AM, Benitez J and Easton DF. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet 2013; 45: 353-361, 361e351-352.

[36] Lilyquist J, Ruddy KJ, Vachon CM and Couch FJ. Common genetic variation and breast cancer risk-past, present, and future. Cancer Epidemiol Biomarkers Prev 2018; 27: 380-394.

[37] Meeks HD, Song H, Michailidou K, Bolla MK, Dennis J, Wang Q, Barrowdale D and Frost D; MBRACE; McGuffog L, Ellis S, Feng B, Buys SS, Hopper JL, Southey MC and Tesoriero A; kConFab Investigators; James PA, Bruinsma F and Campbell IG; Australia Ovarian Cancer Study Group; Broeks A, Schmidt MK, Hogervorst FB; HEBON; Beckman MW, Fasching PA, Fletcher O, Johnson N, Sawyer EJ, Riboli E, Banerjee S, Menon U, Tomlinson I, Burwinkel B, Hamann U, Marme F, Rudolph A, Janavicius R, Tihomirova L, Tung N, Garber J, Cramer D, Terry KL, Poole EM, Tworoger SS, Dorfling CM, van Rensburg EJ, Godwin AK, Guénel P and Truong T; GEMO Study Collaborators; Stoppa-Lyonnet D, Damiola F, Mazoyer S, Sinilnikova OM, Isaacs C, Maugard C, Bojesen SE, Flyger H, Gerdes AM, Hansen TV, Jensen A, Kjaer SK, Hogdall C, Hogdall E, Pedersen IS, Thomassen M, Benitez J, González-Neira A, Osorio A, Hoya Mde L, Segura PP, Diez O, Lazaro C, Brunet J, Anton-Culver H, Eunjung L, John EM, Neuhausen SL, Ding YC, Castillo D, Weitzel JN, Ganz PA, Nussbaum RL, Chan SB, Karlan BY, Lester J, Wu A, Gayther S, Ramus SJ, Sieh W, Whittermore AS, Monteiro AN, Phelan CM, Terry MB, Piedmonte M, Offit K, Robson M, Levine D, Moysich KB, Cannioto R, Olson SH, Daly MB, Nathanson KL, Domchek SM, Lu KH, Liang D, Hildebrant MA, Ness R, Modugno F, Pearce L, Goodman MT, Thompson PJ, Brenner H, Butterbach K, Meindl A, Hahnen E, Wappenschmidt B, Brauch H, Brüning T, Blomqvist C, Khan S, Nevanlinna H, Pelttari LM, Aittomäki K, Butzow R, Bogdanova NV, Dörk T, Lindblom A, Margolin S, Rantala J, Kosma VM, Mannermaa A, Lambrechts D, Neven P, Claes KB, Maerken TV, Chang-Claude J, Flesch-Janys D, Heitz F, Varon-Mateeva R, Peterlongo P, Radice P, Viel A, Barile M, Peissel B, Manoukian S, Montagna M, Oliani C, Peixoto A, Teixeira MR, Collavoli A, Hallberg E, Olson JE, Goode EL, Hart SN, Shimelis H, Cunningham JM, Giles GG, Milne RL, Healey S, Tucker K, Haiman CA, Henderson BE, Goldberg MS, Tischkowitz M, Simard J, Soucy P, Eccles DM, Le N, Borresen-Dale AL, Kristensen V, Salvesen HB, Bjorge L, Bandera EV, Risch H, Zheng W, Beeghly-Fadiel A, Cai H, Pylkäs K, Tollenaar RA, Ouweland AM, Andrulis IL and Knight JA; OCGN; Narod S, Devilee P, Winqvist R, Figueroa J, Greene MH, Mai PL, Loud JT, García-Closas M, Schoemaker MJ, Czene K, Darabi H, McNeish I, Siddiquil N, Glasspool R, Kwong A, Park SK, Teo SH, Yoon SY, Matsuo K, Hosono S, Woo YL, Gao YT, Foretova L, Singer CF, Rappaport-Feurhauser C, Friedman E, Laitman Y, Rennert G, Imyanitov EN, Hulick PJ, Olopade OI, Senter L, Olah E, Doherty JA, Schildkraut J, Koppert LB, Kiemeney LA, Massuger LF, Cook LS, Pejovic T, Li J, Borg A, Öfverholm A, Rossing MA, Wentzensen N, Henriksson K, Cox A, Cross SS, Pasini BJ, Shah M, Kabisch M, Torres D, Jakubowska A, Lubinski J, Gronwald J, Agnarsson BA, Kupryjanczyk J, Moes-Sosnowska J, Fostira F, Konstantopoulou I, Slager S and Jones M; PRostate cancer AssoCiation group To Investigate Cancer Associated aLterations in the genome; Antoniou AC, Berchuck A, Swerdlow A, Chenevix-Trench G, Dunning AM, Pharoah PD, Hall P, Easton DF, Couch FJ, Spurdle AB and Goldgar DE. BRCA2 polymorphic stop codon K3326X and the risk of breast, prostate, and ovarian cancers. J Natl Cancer Inst 2016; 108: djv315.

[38] Thompson ER, Gorringe KL, Rowley SM, Li N, McInerny S, Wong-Brown MW, Devereux L and Li J; Lifepool Investigators; Trainer AH, Mitchell G, Scott RJ, James PA and Campbell IG. Reevaluation of the BRCA2 truncating allele c.9976A > T (p.Lys3326Ter) in a familial breast cancer context. Sci Rep 2015; 5: 14800.

[39] Loizidou MA, Michael T, Neuhausen SL, Newbold RF, Marcou Y, Kakouri E, Daniel M, Papadopoulos P, Malas S, Kyriacou K and Hadjisavvas A. Genetic polymorphisms in the DNA repair genes XRCC1, XRCC2 and XRCC3 and risk of breast cancer in Cyprus. Breast Cancer Res Treat 2008; 112: 575-579.

[40] He Y, Zhang Y, Jin C, Deng X, Wei M, Wu Q, Yang T, Zhou Y and Wang Z. Impact of XRCC2 Arg188His polymorphism on cancer Susceptibility: a meta-analysis. PLoS One 2014; 9: e91202.

[41] Silva SN, Tomar M, Paulo C, Gomes BC, Azevedo AP, Teixeira V, Pina JE, Rueff J and Gaspar JF. Breast cancer risk and common single nucleotide polymorphisms in homologous recombination DNA repair pathway genes XRCC2, XRCC3, NBS1 and RAD51. Cancer Epidemiol 2010; 34: 85-92.

[42] Pooley KA, Baynes C, Driver KE, Tyrer J, Azzato EM, Pharoah PD, Easton DF, Ponder BA and Dunning AM. Common single-nucleotide polymorphisms in DNA double-strand break repair genes and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2008; 17: 3482-3489.

[43] Garcia-Closas M, Egan KM, Newcomb PA, Brinton LA, Titus-Ernstoff L, Chanock S, Welch R, Lissowska J, Peplonska B, Szeszenia-Dabrowska N, Zatonski W, Bardin-Mikolajczak A and

Struewing JP. Polymorphisms in DNA double-strand break repair genes and risk of breast cancer: two population-based studies in USA and Poland, and meta-analyses. Hum Genet 2006; 119: 376-388.

[44] Auranen A, Song H, Waterfall C, Dicioccio RA, Kuschel B, Kjaer SK, Hogdall E, Hogdall C, Stratton J, Whittemore AS, Easton DF, Ponder BA, Novik KL, Dunning AM, Gayther S and Pharoah PD. Polymorphisms in DNA repair genes and epithelial ovarian cancer risk. Int J Cancer 2005; 117: 611-618.

[45] Ragettli MS, Phuleria HC, Tsai MY, Schindler C, de Nazelle A, Ducret-Stich RE, Ineichen A, Perez L, Braun-Fahrländer C, Probst-Hensch N and Künzli N. The relevance of commuter and work/school exposure in an epidemiological study on traffic-related air pollution. J Expo Sci Environ Epidemiol 2015; 25: 474-481.

[46] Dibben C and Clemens T. Place of work and residential exposure to ambient air pollution and birth outcomes in Scotland, using geographically fine pollution climate mapping estimates. Environ Res 2015; 140: 535-541.