IJC
International Journal of Cancer

# Polymorphisms in DNA repair genes, traffic-related polycyclic aromatic hydrocarbon exposure and breast cancer incidence

Irina Mordukhovich[1], Jan Beyea[2], Amy H. Herring[3,4], Maureen Hatch[5], Steven D. Stellman[6], Susan L. Teitelbaum[7], David B. Richardson[1], Robert C. Millikan[1*], Lawrence S. Engel[1], Sumitra Shantakumar[8], Susan E. Steck[9], Alfred I. Neugut[6,10], Pavel Rossner, Jr.[11,12], Regina M. Santella[11] and Marilie D. Gammon[1]

[1] Department of Epidemiology, Lambertville, NJ
[2] Consulting in the Public Interest, Lambertville, NJ
[3] Biostatistics, University of North Carolina, Chapel Hill, NC
[4] Carolina Population Center, University of North Carolina, Chapel Hill, NC
[5] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD
[6] Department of Epidemiology, Columbia University, New York, NY
[7] Department of Preventive Medicine, Mount Sinai School of Medicine, New York, NY
[8] Glaxo-Smith Kline, Inc., Singapore
[9] Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, South Carolina, CA
[10] Departments of Medicine, Columbia University, New York, NY
[11] Environmental Health Sciences, Columbia University, New York, NY
[12] Laboratory of Genetic Ecotoxicology, Institute of Experimental Medicine as CR, Prague, Czech Republic



**Vehicular traffic polycyclic aromatic hydrocarbons (PAHs) have been associated with breast cancer incidence in epidemiologic studies, including our own. Because PAHs damage DNA by forming adducts and oxidative lesions, genetic polymorphisms that alter DNA repair capacity may modify associations between PAH-related exposures and breast cancer risk. Our goal was to examine the association between vehicular traffic exposure and breast cancer incidence within strata of a panel of nine biologically plausible nucleotide excision repair (NER) and base excision repair (BER) genotypes. Residential histories of 1,508 cases and 1,556 controls were assessed in the Long Island Breast Cancer Study Project between 1996 and 1997 and used to reconstruct residential traffic exposures to benzo[a]pyrene, as a proxy for traffic-related PAHs. Likelihood ratio tests from adjusted unconditional logistic regression models were used to assess multiplicative interactions. A gene-traffic interaction was evident ($p = 0.04$) for *ERCC2* (Lys751); when comparing the upper and lower tertiles of 1995 traffic exposure estimates, the odds ratio (95% confidence interval) was 2.09 (1.13, 3.90) among women with homozygous variant alleles. Corresponding odds ratios for 1960–1990 traffic were also elevated nearly 2–3-fold for *XRCC1*(Arg194Trp), *XRCC1*(Arg399Gln) and *OGG1*(-Ser326Cys), but formal multiplicative interaction was not evident. When DNA repair variants for *ERCC2*, *XRCC1* and *OGG1* were combined, among women with 4–6 variants, the odds ratios were 2.32 (1.22, 4.49) for 1995 traffic and 2.96 (1.06, 8.21) for 1960–1990 traffic. Our study is first to report positive associations between traffic-related PAH exposure and breast cancer incidence among women with select biologically plausible DNA repair genotypes.**

**Key words:** traffic, DNA repair, polycyclic aromatic hydrocarbons, breast cancer

**Abbreviations:** BER: base excision repair; CCA: complete case analysis; CI: confidence interval; ERCC1/ERCC2/ERCC4/ERCC5: excision repair cross-complementing groups 1 2, 4, and 5; LD: linkage disequilibrium; LIBCSP: Long Island Breast Cancer Study Project; MI: multiple imputation; NER: nucleotide excision repair; OGG1: 8-oxoguanine DNA glycosylase 1; OR: odds ratio; PAHs: polycyclic aromatic hydrocarbons; RERI: relative excess risk due to interaction; SNP: single nucleotide polymorphism; XPA: xeroderma pigmentosum group A; XRCC1: x-ray repair cross complementing group 1

Additional Supporting Information may be found in the online version of this article.

*Deceased as of 10/7/2012.

**Grant sponsor:** U.S. Department of Defense; **Grant number:** BC100414, BC972772; **Grant sponsor:** National Cancer Institute and the National Institute of Environmental Health Sciences; **Grant number:** CA/ES66572, P30ES009089, P30ES10126, T32CA09330 and T32ES007018; **Grant sponsor:** Breast Cancer Research Foundation, Columbia University Women at Risk Program

**DOI:** 10.1002/ijc.30079

**History:** Received 27 July 2015; Accepted 4 Feb 2016; Online 4 Mar 2016

**Correspondence to**: Irina Mordukhovich, Exposure, Epidemiology and Risk, Program, Department of Environmental Health, Harvard T.H. Chan School of Public Health, Landmark Center, 401 Park Dr., Boston, MA 02215, USA, Tel.: [617-384-8754], E-mail: imordukh@hsph.harvard.edu

 

**What's new?**

**Polycyclic aromatic hydrocarbons are known carcinogens generated by vehicular traffic. They damage DNA by forming adducts or inducing oxidative stress. Here the authors examined the association between traffic exposure and breast cancer incidence while taking into account different DNA repair genotypes. Cancer risks were elevated in women carrying certain DNA repair variants especially when some variants were combined. These findings may help identify women who are particularly susceptible to the carcinogenic effects of traffic pollution on the breast.**



Breast cancer is the most common malignancy among women in the United States,[1] northern Europe, England, Canada, and Australia.[2] Polycyclic aromatic hydrocarbons (PAHs) are incomplete combustion by-products found in air pollution.[3,4] In particular, vehicular traffic is a major contributor to ambient PAH levels, especially near cities.[5,6] PAHs are confirmed human lung carcinogens and potent mammary carcinogens in animal models.[4] PAH-DNA adducts[7–9] and air pollution exposure[10,11] have been associated with breast cancer in prior epidemiological studies, including in our own.[8,12]

PAHs damage DNA by forming bulky adducts[13] and by inducing oxidative stress.[4] If unrepaired, PAH-induced DNA damage can lead to somatic mutations in tumor suppressor genes or proto-oncogenes, thereby contributing to carcinogenesis.[7] Hence, genetic polymorphisms that alter DNA repair capacity may modify associations between PAH-related exposures and breast cancer risk.[14] Of the four major DNA repair pathways,[15] nucleotide excision repair (NER) and base excision repair (BER) are most likely to mend PAH-induced DNA damage, since the NER pathway repairs bulky DNA adducts and the BER pathway repairs oxidative DNA damage.[13,16,17]

The few studies conducted to date suggest that polymorphisms in NER and BER genes may interact with PAH-related exposures, such as cigarette smoke or PAH-DNA adducts, with respect to breast cancer risk.[18–21] However, the impact of DNA repair polymorphisms on the relationship between air pollution and breast cancer has not been evaluated previously. Our study examines the association between residential vehicular traffic exposure and breast cancer incidence within strata of NER and BER genotypes, using the resources of a population-based investigation.[22] Findings of this first study to examine these associations, if confirmed, may increase understanding of breast cancer etiology and help identify women who may be particularly susceptible to the carcinogenic effects of traffic pollution on the breast.

## Material and Methods

Our study used resources from the case-control component of the Long Island Breast Cancer Study Project (LIBCSP), a population-based investigation conducted among adult women (ages 20–98) residing in Nassau and Suffolk counties in Long Island, NY.[22] This study was approved by the institutional review boards of all participating institutions.

### Study population

Case participants were diagnosed with a first primary invasive or *in situ* breast cancer between July 1996 and August 1997, and were recruited via rapid case ascertainment through contact with the pathology departments of local hospitals. Control participants were women with no personal history of breast cancer. Controls under age 65 were identified using random digit dialing,[23] and those aged 65 years or older were identified via Health Care Finance Administration rosters. Control participants were frequency matched based on the expected age distribution among the cases. In total, 1,508 cases and 1,556 controls (82.1% and 62.7% of eligible participants, respectively) completed the case-control study interview. Most participants were post-menopausal and white, which is consistent with the underlying racial distribution of the study counties at the time of data collection.[22]

Previous studies using LIBCSP resources reported positive associations between several PAH-related exposures and overall breast cancer incidence: elevated PAH-DNA adducts in mononuclear cells,[8] long-term residential exposure to environmental tobacco smoke,[24] increased consumption of grilled/smoked meat,[25] indoor stove and fireplace use,[26] and the top 5% of long-term residential vehicular traffic exposures, compared with exposures below the median.[12]

### Study questionnaire

Trained interviewers administered a 2-hr structured questionnaire in participants' homes.[22] The questionnaire evaluated detailed information regarding a wide range of factors used in our analyses, including lifetime residential history, demographic characteristics, reproductive and medical history, and body size.[22]

### Traffic exposure assessment

Participants' historical residential traffic-related PAH exposures were reconstructed using a model which incorporated lifetime residential history in the study counties, historical U.S. vehicular PAH emissions and roadway-specific traffic patterns, local meteorological data, pollutant dispersion factors, and excess emissions at intersections.[27–29] In a validation study, the model accurately predicted residential soil PAH measurements as well as participants' levels of PAH-DNA adducts in mononuclear cells.[27]

Exposure estimates were reconstructed for the years 1960–1990 and 1995. Some participants moved to the study area

after 1960, and others had periods of incomplete address data while residing in the study counties. Hence, an exposure surrogate was developed to estimate pre-arrival exposures, with variance added during multiple imputation (MI), and missing post-arrival exposures were calculated via MI based on participants' known residential addresses.[29,34] The percentage of imputation was restricted to ≤20% of total 1960–1990 exposure.[12] Thus, the imputed portion of the dose was relatively minor, with an average of <10% among those with any imputed exposure information. Our research group also previously conducted sensitivity analyses in order to evaluate the impact of imputation on our study results. One sensitivity analysis restricted the imputation percentage to 0%, 20% or 30%. Another sensitivity analysis examined two imputation methods: (1) based on exposures at participants' known residences (before or after the missing exposure period) and (2) according to census place (i.e., city, town or village, which was often known even when a woman's address was incomplete). As previously reported,[12] these sensitivity analyses did not meaningfully alter results for the association between traffic PAH exposure and breast cancer risk. Hence, we are confident that imputing <20% of exposure did not introduce spurious findings. For 1995 exposure, only women with complete, unimputed exposure information were included in regressions [denoted as a complete case analysis (CCA)].

Because our exposure model predicted benzo[a]pyrene levels as a surrogate for all traffic PAHs, exposure estimates were normalized to the mean exposure level for the year 1995 and presented as relative units of average annual exposure during 1995.[27,29] Exposures in the year 1995 ranged between 0.02 and 31 relative units (mean = 1.0), and cumulative exposures during the years 1960–1990 (20% MI) ranged between 16 and 5467 relative units of average annual exposure in 1995 (mean = 212)[12]. Most LIBCSP participants had lived in their current home for at least 15 years,[22] and previously reported correlations between exposures in 1995 and 1960–1990 were strong.[12]

### Blood sample collection and DNA extraction

Study interviewers, who were either nurses or certified phlebotomists, obtained a 40 mL non-fasting blood sample at the study visit from 73% of case and 73% of control participants.[22] Blood samples were collected in EDTA-treated tubes and shipped overnight to a laboratory in Columbia University, where they were aliquoted and stored in −80°C freezers using participants' randomly assigned study identification numbers.[22]

Mononuclear cells were separated from whole blood using Ficoll (Sigma Chemical Co., St. Louis, MO), washed twice with phosphate-buffered saline, and frozen until DNA isolation.[22] DNA was extracted using phenol and chloroform isoamyl alcohol, followed by RNase and proteinase K treatment.[20,22]

### Genotyping

Our study evaluated a panel of single nucleotide polymorphisms (SNPs) in the following genes, which were selected because of their biologic plausibility for potentially modifying the association between traffic and breast cancer incidence: excision repair cross-complementing groups 1 (*ERCC1,* 8092C/A [rs3212986]), 8-oxoguanine DNA glycosylase 1 (*OGG1,* Ser326Cys [rs1052133]), 2 (*ERCC2*, Lys751Gln [rs13181] and Asp312Asn [rs1799793]), 4 (*ERCC4*, Arg415Gln [rs1800067]), and 5 (*ERCC5,* Asp1104His [rs17655]) X-ray repair cross complementing Group 1 (*XRCC1,* Arg399Gln [rs25487] and Arg194Trp [rs1799782]), and xeroderma pigmentosum group A (*XPA,* −4A/G [rs1800975]). These SNPs were previously found to be in Hardy-Weinberg equilibrium among LIBCSP controls.[18,20,21,30] Genotyping was completed for 91–97% of women who donated a blood sample, depending on the polymorphism ($n = 2{,}050$–$2{,}177$); missing genotype data was mainly due to insufficient DNA available for assays.[18,20,21,30]

Several high-throughput methods were used by the LIBCSP study investigators for genotyping the DNA repair pathway polymorphisms. SNPs in *ERCC1*, *XRCC1*, *OGG1* and *ERCC2* (Lys751Gln) were genotyped using fluorescence polarization (FP),[18,20,21,30] which identifies the polymorphic base using a template-directed, dye-labeled dideoxynucleotide.[31] SNPs in *XPA* and *ERCC2 (Asp312Asn) were* genotyped using the Taqman assay (Applied Biosystems, Foster City, CA),[18] which uses fluorogenic oligonucleotide reverse probes for allele recognition. SNPs in *ERCC4* and *ERCC5* were genotyped using Sequenom's high-throughput matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, chosen for its ability to multiplex (BioServe Biotechnologies, Laurel, MD).[18]

Quality control measures included negative and positive controls for genotype on each plate.[18] In addition, duplicates for approximately 10% of DNA samples were selected for re-sampling. As previously reported, concordance between duplicate samples was high for all SNPs ($\kappa > 90\%$).[18,20,21,30] Laboratory personnel were blinded to case-control status and duplicate sample status.

### Statistical methods

Our study examined associations between tertiles of traffic-related PAH exposure (years 1995 and 1960–1990) and breast cancer incidence within strata of NER and BER polymorphisms. We evaluated PAH exposures for the year 1995 in order to maximize statistical power for gene-environment interaction analyses. We also evaluated 1960–1990 exposure estimates, as these may be a more accurate reflection of long-term exposure opportunity. We chose a categorical exposure variable because of the previously reported non-linear relationship between traffic PAHs and breast cancer incidence in the LIBCSP[12]. While our previous study compared the top 5% of traffic PAH exposures to those below the median, the

**Table 1.** Associations between traffic polycyclic aromatic hydrocarbon (PAH) exposure in 1995 and breast cancer, stratified by DNA repair genotype

| | Associations between tertiles of traffic exposure and breast cancer within genotype strata | | | | | | |
|---|---|---|---|---|---|---|---|
| DNA repair polymorphisms | Genotype[1] | Cases (*n*) | Controls (*n*) | Tertile of exposure[2] | Age-adjusted OR (95% CI) | *p* for trend | *p*[3] |
| *XPA* −4A/G | GG | 128 | 131 | 1 | 1.0 | 0.85 | 0.60 |
| rs1800975 | | 130 | 138 | 2 | 0.94 (0.67, 1.33) | | |
| | | 148 | 153 | 3 | 0.97 (0.69, 1.35) | | |
| | GA | 149 | 137 | 1 | 1.0 | 0.87 | |
| | | 112 | 157 | 2 | 0.64 (0.46, 0.90) | | |
| | | 141 | 121 | 3 | 1.04 (0.74, 1.47) | | |
| | AA | 30 | 35 | 1 | 1.0 | 0.88 | |
| | | 28 | 49 | 2 | 0.67 (0.34, 1.32) | | |
| | | 33 | 31 | 3 | 1.12 (0.55, 2.26) | | |
| *ERCC1* 8092C/A | CC | 151 | 162 | 1 | 1.0 | 0.48 | 0.54 |
| rs3212986 | | 139 | 204 | 2 | 0.70 (0.51, 0.96) | | |
| | | 166 | 152 | 3 | 1.12 (0.81, 1.53) | | |
| | CA | 138 | 125 | 1 | 1.0 | 0.44 | |
| | | 111 | 123 | 2 | 0.82 (0.57, 1.16) | | |
| | | 127 | 130 | 3 | 0.87 (0.62, 1.23) | | |
| | AA | 17 | 18 | 1 | 1.0 | 0.68 | |
| | | 22 | 17 | 2 | 1.36 (0.54, 3.43) | | |
| | | 25 | 21 | 3 | 1.22 (0.50, 3.00) | | |
| *ERCC4* Arg415Gln | GG | 247 | 248 | 1 | 1.0 | 0.87 | 0.50 |
| rs1800067 | | 221 | 276 | 2 | 0.78 (0.61, 1.01) | | |
| | | 250 | 247 | 3 | 0.98 (0.76, 1.26) | | |
| | GA or AA | 48 | 49 | 1 | 1.0 | 0.52 | |
| | | 41 | 60 | 2 | 0.69 (0.39, 1.21) | | |
| | | 54 | 45 | 3 | 1.20 (0.68, 2.12) | | |
| *ERCC5* Asp1104His | GG | 189 | 164 | 1 | 1.0 | 0.47 | 0.23 |
| rs17655 | | 126 | 187 | 2 | 0.57 (0.42, 0.78) | | |
| | | 159 | 148 | 3 | 0.91 (0.67, 1.24) | | |
| | GC | 91 | 110 | 1 | 1.0 | 0.93 | |
| | | 111 | 115 | 2 | 1.12 (0.76, 1.64) | | |
| | | 109 | 123 | 3 | 0.99 (0.67, 1.46) | | |
| | CC | 16 | 20 | 1 | 1.0 | 0.28 | |
| | | 18 | 26 | 2 | 0.87 (0.35, 2.12) | | |
| | | 23 | 18 | 3 | 1.61 (0.65, 4.01) | | |
| *ERCC2* Lys751Gln | AA | 123 | 123 | 1 | 1.0 | 0.41 | 0.04 |
| rs13181 | | 95 | 141 | 2 | 0.66 (0.46, 0.94) | | |
| | | 110 | 123 | 3 | 0.86 (0.60, 1.24) | | |
| | AC | 143 | 129 | 1 | 1.0 | 0.77 | |
| | | 137 | 153 | 2 | 0.78 (0.56, 1.09) | | |
| | | 159 | 147 | 3 | 0.94 (0.68, 1.31) | | |
| | CC | 36 | 52 | 1 | 1.0 | 0.02 | |
| | | 39 | 50 | 2 | 1.13 (0.62, 2.05) | | |
| | | 49 | 33 | 3 | 2.09 (1.13, 3.90) | | |

**Table 1.** Associations between traffic polycyclic aromatic hydrocarbon (PAH) exposure in 1995 and breast cancer, stratified by DNA repair genotype (Continued)

| DNA repair polymorphisms | Associations between tertiles of traffic exposure and breast cancer within genotype strata | | | | | | |
|---|---|---|---|---|---|---|---|
| | Genotype[1] | Cases (*n*) | Controls (*n*) | Tertile of exposure[2] | Age-adjusted OR (95% CI) | *p* for trend | *p*[3] |
| *ERCC2* Asp312Asn | GG | 123 | 139 | 1 | 1.0 | 0.71 | 0.87 |
| rs1799793 | | 108 | 153 | 2 | 0.78 (0.55, 1.10) | | |
| | | 130 | 131 | 3 | 1.07 (0.76, 1.51) | | |
| | GA | 143 | 119 | 1 | 1.0 | 0.47 | |
| | | 124 | 143 | 2 | 0.70 (0.49, 0.98) | | |
| | | 150 | 138 | 3 | 0.87 (0.62, 1.22) | | |
| | AA | 40 | 45 | 1 | 1.0 | 0.36 | |
| | | 36 | 49 | 2 | 0.84 (0.46, 1.55) | | |
| | | 42 | 35 | 3 | 1.35 (0.73, 2.52) | | |
| *XRCC1* Arg399Gln | GG | 122 | 136 | 1 | 1.0 | 0.91 | 0.90 |
| rs25487 | | 105 | 121 | 2 | 0.95 (0.66, 1.36) | | |
| | | 120 | 135 | 3 | 0.97 (0.68, 1.37) | | |
| | GA | 154 | 138 | 1 | 1.0 | 0.66 | |
| | | 136 | 182 | 2 | 0.64 (0.47, 0.89) | | |
| | | 168 | 135 | 3 | 1.07 (0.77, 1.48) | | |
| | AA | 31 | 31 | 1 | 1.0 | 0.90 | |
| | | 34 | 44 | 2 | 0.78 (0.40, 1.53) | | |
| | | 35 | 36 | 3 | 0.95 (0.48, 1.88) | | |
| *XRCC1* Arg194Trp | CC | 272 | 263 | 1 | 1.0 | 0.77 | 0.54 |
| rs1799782 | | 244 | 303 | 2 | 0.76 (0.60, 0.96) | | |
| | | 279 | 269 | 3 | 0.96 (0.76, 1.22) | | |
| | CT or TT | 35 | 41 | 1 | 1.0 | 0.26 | |
| | | 31 | 44 | 2 | 0.81 (0.42, 1.56) | | |
| | | 43 | 36 | 3 | 1.44 (0.76, 2.75) | | |
| *OGG1* Ser326Cys | CC | 188 | 194 | 1 | 1.0 | 0.78 | 0.78 |
| rs1052133 | | 146 | 201 | 2 | 0.73 (0.55, 0.98) | | |
| | | 184 | 176 | 3 | 1.05 (0.78, 1.40) | | |
| | GC | 100 | 96 | 1 | 1.0 | 0.75 | |
| | | 108 | 120 | 2 | 0.84 (0.58, 1.24) | | |
| | | 114 | 113 | 3 | 0.94 (0.64, 1.39) | | |
| | GG | 14 | 12 | 1 | 1.0 | 0.81 | |
| | | 14 | 19 | 2 | 0.63 (0.22, 1.78) | | |
| | | 17 | 13 | 3 | 1.11 (0.38, 3.23) | | |

[1]Women with heterozygote and homozygote variant genotypes were combined to avoid cell sizes of <10.
[2]Traffic PAH exposures were calculated as units of average annual exposure in 1995. Tertiles: 0–0.70, 0.71–1.04 and ≥1.05 relative units. Tertile-specific exposure values are not presented because only relative rather than absolute traffic PAH exposures were estimated (see text for methods).
[3]*p* for interaction.

current gene-environment interaction analysis did not have sufficient statistical power to use those categories. We therefore evaluated tertiles of exposure. This maximizes study power and is similar to the effect modification analyses from our main effects paper, which also collapsed exposure categories.[12]

We used unconditional logistic regression to estimate odds ratios (ORs) and 95% confidence intervals (CIs).[32] Trend tests ($n = 45$) were conducted within stratified models, and likelihood ratio tests ($n = 18$) were used to assess multiplicative interactions between traffic PAHs and SNPs.[33] We also calculated additive interactions between tertiles of 1995

**Table 2.** Associations between traffic polycyclic aromatic hydrocarbon (PAH) exposure in 1960–1990 and breast cancer, stratified by DNA repair genotype

| | Associations between tertiles of traffic exposure and breast cancer within genotype strata | | | | | | |
|---|---|---|---|---|---|---|---|
| DNA repair polymorphisms | Genotype[1] | Cases (*n*) | Controls (*n*) | Tertile of exposure[2] | Age-adjusted OR (95% CI) | *p* for trend | *p*[3] |
| *XPA* −4A/G | GG | 51–63[4] | 63–74 | 1 | 1.0 | 0.74 | 0.41 |
| rs1800975 | | 67–67 | 53–62 | 2 | 1.32 (0.78, 2.24) | | |
| | | 49–58 | 54–59 | 3 | 1.08 (0.63, 1.87) | | |
| | GA or AA | 60–75 | 76–92 | 1 | 1.0 | 0.11 | |
| | | 57–67 | 84–97 | 2 | 0.87 (0.52, 1.43) | | |
| | | 78–90 | 60–76 | 3 | 1.50 (0.91, 2.47) | | |
| *ERCC1* 8092C/A | CC | 57–68 | 75–88 | 1 | 1.0 | 0.47 | 0.63 |
| rs3212986 | | 63–72 | 72–86 | 2 | 1.11 (0.68, 1.82) | | |
| | | 60–72 | 64–73 | 3 | 1.21 (0.73, 2.01) | | |
| | CA or AA | 56–71 | 64–78 | 1 | 1.0 | 0.24 | |
| | | 56–66 | 66–76 | 2 | 0.89 (0.53, 1.51) | | |
| | | 68–75 | 50–61 | 3 | 1.39 (0.82, 2.36) | | |
| *ERCC4* Arg415Gln | GG | 93–109 | 109–130 | 1 | 1.0 | 0.15 | 0.57 |
| rs1800067 | | 100–112 | 121–135 | 2 | 0.98 (0.66, 1.45) | | |
| | | 105–118 | 85–101 | 3 | 1.37 (0.91, 2.06) | | |
| | GA or AA | 17–25 | 24–31 | 1 | 1.0 | 0.91 | |
| | | 13–19 | 17–23 | 2 | 1.04 (0.41, 2.66) | | |
| | | 17–21 | 19–24 | 3 | 1.03 (0.42, 2.53) | | |
| *ERCC5* Asp1104His | GG | 64–78 | 68–82 | 1 | 1.0 | 0.34 | 0.55 |
| rs17655 | | 57–68 | 79–92 | 2 | 0.81 (0.49, 1.34) | | |
| | | 67–78 | 53–67 | 3 | 1.31 (0.78, 2.18) | | |
| | GC or CC | 43–56 | 60–74 | 1 | 1.0 | 0.51 | |
| | | 52–59 | 55–64 | 2 | 1.22 (0.69, 2.16) | | |
| | | 49–58 | 53–61 | 3 | 1.22 (0.69, 2.14) | | |
| *ERCC2* Lys751Gln | AA | 38–49 | 52–65 | 1 | 1.0 | 0.57 | 0.50 |
| rs13181 | | 51–58 | 55–64 | 2 | 1.24 (0.71, 2.18) | | |
| | | 47–53 | 51–59 | 3 | 1.19 (0.66, 2.12) | | |
| | AC | 59–70 | 61–73 | 1 | 1.0 | 0.42 | |
| | | 50–58 | 63–74 | 2 | 0.79 (0.46, 1.35) | | |
| | | 60–69 | 48–56 | 3 | 1.27 (0.74, 2.16) | | |
| | CC | 10–16 | 23–29 | 1 | 1.0 | 0.10 | |
| | | 17–21 | 18–23 | 2 | 1.77 (0.66, 4.74) | | |
| | | 19–24 | 16–21 | 3 | 2.24 (0.80, 6.25) | | |
| *ERCC2* Asp312Asn | GG | 47–57 | 58–74 | 1 | 1.0 | 0.61 | 0.57 |
| rs1799793 | | 51–61 | 63–73 | 2 | 1.07 (0.62, 1.83) | | |
| | | 49–57 | 52–62 | 3 | 1.16 (0.66, 2.05) | | |
| | GA or AA | 64–81 | 79–94 | 1 | 1.0 | 0.18 | |
| | | 65–77 | 75–88 | 2 | 1.02 (0.63, 1.66) | | |
| | | 77–87 | 63–74 | 3 | 1.40 (0.87, 2.26) | | |
| *XRCC1* Arg399Gln | GG | 36–45 | 56–69 | 1 | 1.0 | 0.04 | 0.59 |
| rs25487 | | 49–60 | 50–59 | 2 | 1.51 (0.85, 2.68) | | |
| | | 55–65 | 41–51 | 3 | 1.88 (1.04, 3.41) | | |

**Table 2.** Associations between traffic polycyclic aromatic hydrocarbon (PAH) exposure in 1960–1990 and breast cancer, stratified by DNA repair genotype (Continued)

| DNA repair polymorphisms | Associations between tertiles of traffic exposure and breast cancer within genotype strata | | | | | | |
|---|---|---|---|---|---|---|---|
| | Genotype[1] | Cases (*n*) | Controls (*n*) | Tertile of exposure[2] | Age-adjusted OR (95% CI) | *p* for trend | $p^3$ |
| | GA | 60–74 | 62–79 | 1 | 1.0 | 0.93 | |
| | | 55–64 | 70–88 | 2 | 0.83 (0.50, 1.40) | | |
| | | 62–69 | 59–68 | 3 | 1.02 (0.61, 1.72) | | |
| | AA | 12–19 | 16–21 | 1 | 1.0 | 0.98 | |
| | | 11–16 | 17–22 | 2 | 0.74 (0.25, 2.18) | | |
| | | 11–17 | 13–17 | 3 | 1.07 (0.36, 3.18) | | |
| *XRCC1* Arg194Trp | CC | 102–121 | 122–144 | 1 | 1.0 | 0.44 | 0.30 |
| rs1799782 | | 104–116 | 122–137 | 2 | 0.99 (0.68, 1.46) | | |
| | | 110–121 | 104–117 | 3 | 1.17 (0.79, 1.73) | | |
| | CT or TT | 10–14 | 17–24 | 1 | 1.0 | 0.05 | |
| | | 16–20 | 20–24 | 2 | 1.61 (0.54, 4.76) | | |
| | | 19–23 | 13–18 | 3 | 3.04 (0.97, 9.51) | | |
| *OGG1* Ser326Cys | CC | 65–78 | 88–106 | 1 | 1.0 | 0.02 | 0.52 |
| rs1052133 | | 64–75 | 78–90 | 2 | 1.14 (0.71, 1.84) | | |
| | | 77–90 | 58–69 | 3 | 1.77 (1.09, 2.88) | | |
| | GC or GG | 46–57 | 48–56 | 1 | 1.0 | 0.52 | |
| | | 50–60 | 57–70 | 2 | 0.87 (0.49, 1.54) | | |
| | | 49–56 | 57–68 | 3 | 0.84 (0.48, 1.47) | | |

[1]Women with heterozygote and homozygote variant genotypes were combined to avoid cell sizes of <10.
[2]Traffic PAH exposures were calculated as units of average annual exposure in 1995. Tertiles: 0–144.50, 144.51–220.35 and ≥220.36 relative units.
[3]*p* for interaction.
[4]Effect estimates and their associated confidence intervals were combined across 30 imputed data sets; sample sizes vary across data sets because some women exceeded the 20% limit on imputation percentage for certain imputation draws.

traffic exposure and DNA repair polymorphisms in relation to breast cancer.[34] When evaluating 1960–1990 exposures, effect estimates and confidence limits were combined across 30 imputed data sets using Rubin's rules.[35] Statistical analyses were conducted using SAS version 9.3 (SAS Institute, Cary, NC). No adjustment for multiple comparisons was made, given that the panel of polymorphisms we considered was selected based on biologic plausibility. To better understand the relationship between assigned PAH exposure and the availability of blood samples for genetic testing, we evaluated whether blood donation (yes *vs.* no) among eligible participants was related to the completeness of their residential histories (≤20% *vs.* > 20% missing traffic PAH exposure). We report a positive association between blood donation status and the completeness of residential traffic PAH exposure (data not shown).

Potential confounders (educational attainment (less than high school, high school graduate, some college, college graduate, or post-college), annual household income (<$15,000, $15,000–$49,999, $50,000–$69,999, $70,000–$89,999, or $90,000+), race (White, Black, or other), religion (none, Protestant, Catholic, Jewish, or other), parity (continuous), age at first birth (continuous), body mass index (kg/m$^2$, continuous), duration of oral contraceptive use (continuous), lifetime average alcohol intake (never drinkers, <15 g/day, 15–30 g/day, 30+ g/day), lifetime average physical activity (in hours per week: none, low, average, and high), and duration of breastfeeding (continuous)) were identified by a thorough literature review and analysis of causal diagrams.[12,36] Smoking and dietary sources of PAHs were not considered as confounders, because analysis of causal diagrams[36] did not support the inclusion of smoking and dietary PAH intake in minimally sufficient adjustment sets. No covariates were retained using the 10% change in estimate criterion[33] relative to age-adjusted models.[12] Hence, regression models are adjusted only for the frequency matching factor, five-year age group.

Participants with heterozygous, variant homozygous, and homozygous major genotypes were considered separately in our study. We chose this approach because there was no *a priori* reason to combine women with heterozygous and variant homozygous genotypes into a single group, and because previous LIBCSP reports evaluating interactions between DNA repair polymorphisms and PAH-related exposures

Table 3. Associations between traffic PAH exposure and breast cancer, stratified by combined DNA repair genotypes in *ERCC2*, *XRCC1* and *OGG1*

| Number of high-risk alleles[1] | Associations between tertiles of traffic PAHs and breast cancer within genotype strata | | | | | |
|---|---|---|---|---|---|---|
| | 1995 (CCA) | | | 1960–1990 (≤20% MI) | | |
| | Cases (*n*) | Controls (*n*) | Age-adjusted OR (95% CI)[2] | Cases (*n*) | Controls (*n*) | Age-adjusted OR (95% CI)[2,3] |
| 0–1 | 42 | 42 | 1.0 | 16–26 | 20–26 | 1.0 |
| | 44 | 52 | 0.84 (0.47, 1.51) | 20–26 | 17–24 | 1.25 (0.51, 3.07) |
| | 48 | 54 | 0.88 (0.49, 1.57) | 18–22 | 22–27 | 0.89 (0.36, 2.22) |
| *p* for trend | | | 0.67 | | | 0.77 |
| 2–3 | 227 | 211 | 1.0 | 82–100 | 93–110 | 1.0 |
| | 192 | 247 | 0.70 (0.54, 0.92) | 78–91 | 105–118 | 0.86 (0.56, 1.34) |
| | 218 | 214 | 0.92 (0.70, 1.20) | 86–96 | 77–88 | 1.25 (0.80, 1.95) |
| *p* for trend | | | 0.53 | | | 0.35 |
| 4–6 | 29 | 48 | 1.0 | 9–14 | 22–29 | 1.0 |
| | 30 | 40 | 1.24 (0.64, 2.41) | 13–17 | 11–17 | 2.28 (0.75, 6.92) |
| | 46 | 32 | 2.32 (1.22, 4.49) | 20–26 | 15–19 | 2.96 (1.06, 8.21) |
| *p* for trend | | | 0.011 | | | 0.031 |

CCA, complete case analysis; MI, multiple imputation; PAH, polycyclic aromatic hydrocarbon (see text for methods).
[1]Alleles were combined from the following genes: *ERCC2* Lys751Gln, rs13181 (Gln); *XRCC1* Arg194Trp, rs1799782 (Arg) and *OGG1* Ser326Cys, rs1052133 (Ser).
[2]Traffic PAH exposures were calculated as units of average annual exposure in 1995. Tertiles (1995): 0–0.70, 0.71–1.04 and ≥1.05 relative units. Tertiles (1960–1990): 0–144.50, 144.51–220.35 and ≥220.36 relative units.
[3]Effect estimates and their associated confidence intervals were combined across 30 imputed data sets; sample sizes vary across data sets because some women exceeded the 20% limit on imputation percentage for certain imputation draws.

suggested differences between participants with heterozygous and variant homozygous genotypes that were not present when evaluating the main effects of these SNPs on breast cancer incidence.[18,21] However, participants with heterozygous and variant homozygous genotypes were combined when tertile-specific cell sizes comprised fewer than 10 cases and 10 controls.[37]

Finally, we evaluated linkage disequilibrium (LD) among individual SNPs showing some evidence of interactions with traffic PAHs[38]: *ERCC2* Lys751Gln, *XRCC1* Arg194Trp and Arg399Gln, and *OGG1* Ser326Cys. Of these, only the two *XRCC1* variants were in LD with each other. Using a previously reported method,[18,19] we examined traffic-breast cancer associations stratified according to the number of 'high-risk' alleles (0–1, 2–3, ≥4) from the following SNPs: *ERCC2* Lys751Gln (Gln), *XRCC1* Arg194Trp (Trp), and *OGG1* Ser326Cys (Ser).

## Results

The number of LIBCSP participants with available exposure estimates varied according to the traffic PAH exposure definition (1995 *vs*. 1960–1990 estimates) and the imputed data set (some women were not included in all 30 imputed data sets because they exceeded the 20% limit on imputation percentage for certain imputation draws). Sample size also varied by the SNP considered (*XPA*, *ERCC2*, *ERCC4*, *ERCC5*, *ERCC1*, *OGG1*, and *XRCC1*). The study population size with SNP assays completed varied as follows: 1995 traffic estimates, 842–905 cases and 911–958 controls; 1960–1990 traffic estimates, 332–429 cases and 368–474 controls (Tables 1 and 2).

Among women with available genotyping data, ORs for the association between traffic exposure (highest *vs*. lowest tertile) and breast cancer incidence were 1.01 (95% CI: 0.81, 1.26) and 1.04 (95% CI: 0.81, 1.33) for 1995 and 1960–1990 exposures, respectively. We present associations between traffic PAHs and breast cancer incidence within strata of DNA repair polymorphisms in Tables 1 and 2.

The association between vehicular traffic exposure in 1995 and breast cancer incidence was of larger magnitude and greater precision among women with the homozygous variant genotype for the *ERCC2* Lys751Gln polymorphism: OR (highest *vs*. lowest tertile) = 2.09 (95% CI: 1.13, 3.90; $p$ for trend = 0.02) (Table 1). Among women with the homozygous major or heterozygous genotypes, the corresponding ORs were 0.86 (95% CI: 0.60, 1.24) and 0.94 (95% CI: 0.68, 1.31), respectively ($p$-interaction = 0.04). Likewise, only the *ERCC2* Lys751Gln variant showed evidence of super-additive interaction with 1995 traffic PAHs (relative excess risk due to interaction [RERI] = 0.32, 95% CI: 0.074, 0.57). We present a joint classification table showing the separate and combined effects of 1995 traffic PAH exposure and this *ERCC2* polymorphism on breast cancer risk in the Supporting Information (Table S1). Results for this SNP were similar, though less precise, when evaluating cumulative exposure for the years 1960–1990 (Table 2).

Although there was no formal evidence for multiplicative interaction, the relationship between 1960–1990 traffic exposure and breast cancer incidence was also stronger among women with: at least one variant allele for the *XRCC1* Arg194Trp polymorphism (OR = 3.04, 95% CI:0.97, 9.51; $p$ for trend = 0.049, $p$-interaction = 0.30); the homozygous major genotype for *XRCC1* Arg399Gln (OR = 1.88, 95% CI: 1.04, 3.41; $p$ for trend = 0.04, $p$-interaction = 0.59); and the homozygous major genotype for *OGG1* Ser326Cys (OR = 1.77, 95% CI: 1.09, 2.88; $p$ for trend = 0.02, $p$-interaction = 0.52) (Table 2). The results were similar, but less precise and of smaller magnitude, when evaluating 1995 exposures (Table 1).

We found no evidence of interactions between traffic PAHs and the following SNPs: *XPA* −4A/G, *ERCC4* Arg415Gln, *ERCC5* Asp1104His, *ERCC1* 8092C/A, and *ERCC2* Asp312Asn. This lack of heterogeneity within genotype strata was evident regardless of whether we considered traffic PAHs estimated for 1995 (Table 1) or 1960–1990 (Table 2).

We examined the associations between traffic PAHs (highest *vs*. lowest tertile) and breast cancer incidence stratified by the number of high-risk alleles in *ERCC2* (Lys751Gln), *XRCC1* (Arg194Trp), and *OGG1* (Ser326Cys). These SNPs were selected based on results presented in Tables 1 and 2 and a linkage disequilibrium analysis. The ORs were strongest among women with 4 or more 'high-risk' variants (1995: OR = 2.32, 95% CI: 1.22, 4.49, $p$ for trend = 0.011; 1960–1990: OR = 2.96, 95% CI: 1.06, 8.21, $p$ for trend = 0.031), with no clear evidence of a positive association among those with fewer (0–1 or 2–3) high-risk variants (Table 3).

## Discussion

In the first breast cancer study to evaluate interactions between vehicular traffic and DNA repair polymorphisms, the estimated association between traffic exposure (upper *vs*. lower tertiles) and breast cancer incidence was of two-fold greater magnitude among women with the homozygous variant genotype for the *ERCC2* Lys751Gln polymorphism, and multiplicative interaction was evident. Our finding is consistent with the existing literature regarding *ERCC2*, PAH exposures, and breast cancer risk.[18,21,39] In addition, although formal statistical evaluation did not indicate evidence for multiplicative interaction, the association between long-term traffic exposure and breast cancer incidence was of three-fold greater magnitude among those with the homozygous major genotype for *XRCC1* Arg399Gln, and of nearly two-fold greater magnitude among women with the homozygous major genotype for *OGG1* Ser326Cys or with at least one variant allele for *XRCC1* Arg194Trp. However, for the latter, the associated confidence intervals were wide. If replicated, our findings may have public health significance due to widespread exposure to traffic emissions worldwide.

Air pollution, an important PAH source,[5,6] has been associated with breast cancer incidence in other population studies.[10,11] The LIBCSP previously reported positive associations between trafic PAHs and breast cancer incidence among biologically plausible subgroups of women, including those with very high long-term exposures, hormone-receptor negative breast tumors, or low intake of fruits/vegetables.[12]

PAHs damage DNA by causing bulky adducts[13] and oxidative lesions,[16] and the resulting damage can be repaired through NER and BER pathways.[13,16,17,40] NER is a complex process involving many proteins, including XPA, which recognizes DNA damage, and ERCC2, which unwinds DNA at the lesion site; ERCC1, ERCC4 and ERCC5 excise the damaged nucleotide fragment.[40] Proteins involved in BER include OGG1, which recognizes DNA damage and removes the damaged nucleotide, and XRCC1, a scaffolding protein.[17] Functional variants in the DNA repair genes encoding these proteins may modify associations between traffic PAHs and breast cancer risk.[14]

Studies evaluating the main effects of DNA repair genotypes on breast cancer incidence report inconsistent associations for most NER and BER polymorphisms.[18,19,41] Variants in *ERCC2* are more consistently related to breast cancer, including in several meta-analyses.[42,43] The LIBCSP likewise reports modest positive associations between breast cancer incidence and polymorphisms in *ERCC2*,[18,21] but not other NER or BER genes.[18,20,30]

Our study found that the association between vehicular traffic exposure and breast cancer incidence was of two times greater magnitude among women with the homozygous variant genotype for the *ERCC2* Lys751Gln polymorphism. This is consistent with a previous report from our study population, which observed that the association between breast cancer and variant allele homozygosity for this SNP was limited to those with PAH-DNA adduct levels above the median and to current smokers.[21] The Lys751Gln polymorphism is also linked to PAH-related DNA damage in other epidemiological studies,[44,45] including increased levels of PAH-DNA adducts in breast tissue.[39] Hence, the literature regarding PAH-related exposures, the *ERCC2* Lys751Gln polymorphism, and breast cancer risk, while limited, has been consistent to date.[7]

In contrast, associations with other NER and BER variants have been inconsistently reported. For example, we observed that effect estimates for the relationship between traffic PAHs and breast cancer incidence were 3-fold higher among women with the homozygous major genotype for the *XRCC1* Arg399Gln polymorphism, consistent with a study which reported an inverse association between the variant allele for this SNP and PAH-DNA adducts in lung tissue.[46] However, the LIBCSP investigators have previously reported positive associations between the variant *XRCC1*−399Gln allele and breast cancer incidence among women with detectable PAH-DNA adducts and among never smokers.[20]

We also report associations of nearly two-fold greater magnitude between traffic exposure and breast cancer

incidence among women with at least one variant allele for the *XRCC1* Arg194Trp polymorphism and among those with the homozygous major genotype for the *OGG1* Ser326Cys polymorphism, although multiplicative interaction was not evident. Previous studies found no evidence of interactions between these polymorphisms and active smoking (*OGG1*, *XRCC1*) or PAH-DNA adducts (*XRCC1*).[20,30] Consistent with our results, epidemiologic studies reported positive associations between the variant *XRCC1* Trp194Arg SNP and micronucleus frequencies among coke oven workers, who are exposed to high levels of PAHs,[47] as well as PAH-DNA adducts in lung tissue.[46] However, only the variant *OGG1* Ser326Cys polymorphism has been linked to DNA damage endpoints, such as oxidative DNA lesions,[48] in several investigations.[49,50]

Our results for polymorphisms in *XRCC1* and *OGG1* were of greater magnitude when evaluating 1960–1990 traffic PAH estimates relative to estimates for 1995. This may be due to the smaller sample size of the 1960–1990 dataset, which could decrease the stability of results. An alternative explanation is that the 1960–1990 traffic exposure estimates likely represent participants' true long-term exposure more accurately than the 1995 estimates.

Finally, we report no evidence of multiplicative interactions between traffic PAHs and SNPs in *XPA*, *ERCC4*, *ERCC5*, *ERCC1,* and *ERCC2 (*Asp312Asn only) in relation to breast cancer incidence. These findings are consistent with a previous LIBCSP report evaluating interactions between PAH-DNA adducts, smoking, and polymorphisms in *XPA*, *ERCC4*, and *ERCC5* with respect to breast cancer incidence.[18] Our latter study reported evidence of interactions between PAH-DNA adducts, but not active or passive smoking, and the *ERCC1* 8092C/A and *ERCC2* Asp312Asn *SNPs*.[18]

These inconsistencies in the literature could be due in part to the use of different PAH exposure surrogates across studies: we evaluated air pollution, while other investigations focused on cigarette smoking and PAH-DNA adducts. This complicates interpretation of study results, as PAH sources vary with regard to a range of properties, including exposure levels and pathways, particle size, and co-pollutants.[4] Differing results across sources of PAH exposure may also suggest a more complicated role for certain polymorphisms. For example, DNA repair genes encode repair proteins and polymorphisms in these genes may therefore affect repair activity levels[14]; such differences could be of varying importance given lower- and higher-level PAH exposures. Finally, variations in study methodology, exposure and outcome definitions, and participant characteristics, as well as possible chance findings, may underlie inconsistent findings.

Interpretation of our study results requires careful consideration of our study limitations as well as our strengths. Our study had limited power for evaluating lower level interactions, particularly when examining long-term traffic exposures, which increases the likelihood of both false positive and false negative findings.[33] We also did not have sufficient statistical power to evaluate interactions between DNA repair polymorphisms and high-level residential traffic exposures among the most highly exposed 1–5% of study subjects,[12] due to sample size constraints within genetic subgroups. This may have attenuated or obscured some associations. In addition, our results were not adjusted for multiple comparisons. However, we used a targeted pathway approach in carefully selecting NER and BER polymorphisms that were likely to interact with traffic PAH exposures.[13,16]

As previously reported, blood donation among participants is positively associated with white race, alcohol intake, hormone replacement therapy use, lactation history, and mammography, negatively associated with former smoking and age, and unrelated to case-control status.[22] The current study also reports a positive association between blood donation and completeness of participants' residential histories. This may affect the generalizability of our findings, but not the internal validity of this study.

Ours is the first study to evaluate the impact of variation in DNA repair genes on the association between breast cancer incidence and air pollution, a ubiqutious and generalizable exposure source.[3] It is one of only a small number of breast cancer studies to examine interactions between DNA repair variants and any PAH-related exposure.[18–21] Thus, we believe our investigation is an important addition to the sparse literature for this topic. Furthermore, our results for the *ERCC2* Lys751Gln SNP are consistent with the growing literature regarding this DNA repair gene, PAH exposures, and breast cancer incidence, which increases confidence in our findings.[18,21,39] Other strengths of our study include the use of a sophisticated, validated model to reconstruct historical traffic exposures,[27,29] and extensive quality control procedures which were implemented for genotyping assays and questionnaire data.[18,22] Furthermore, previously reported sensitivity analyses in our study population indicate that our findings are robust to altering the percentage of imputed exposure and the specific imputation method,[12] increasing confidence in our exposure estimates.

In summary, associations between vehicular traffic and breast cancer were >2-fold among women with *ERCC2* Lys751Gln homozygous variants, with strong evidence for multiplicative interaction. In addition, we observed 2- to 3-fold greater among women with polymorphisms in *XRCC1* and *OGG1*, and when we considered *ERCC2, XRCC1*, and *OGG1* variants together, although confidence intervals were wide and there was only suggestive evidence for multiplicative interactions. This is the first breast cancer study to examine the interactions between air pollution and DNA repair polymorphisms, and our findings require confirmation in larger studies. Future breast cancer studies should also evaluate interactions between traffic exposures and genetic variants in oxidative stress and carcinogen metabolizing pathways. Our findings, if confirmed, may help identify women who are particularly susceptible to the carcinogenic effects of

traffic pollution on the breast, and may help clarify mechanisms linking traffic PAHs to breast cancer incidence.

## Acknowledgements

This study was supported in part by awards from the U.S. Department of Defense (BC100414, BC972772), the National Cancer Institute and the National Institute of Environmental Health Sciences (CA/ES66572, P30ES009089, P30ES10126, T32CA09330 and T32ES007018), the Breast Cancer Research Foundation, Columbia University Women at Risk Program, and gifts from private citizens. J.B. is employed by Consulting in the Public Interest, Inc. S.S. participated in this research during her time as a doctoral student at UNC Chapel Hill and is currently an employee and shareholder of GlaxoSmithKline, RTP, NC, USA. Neither GlaxoSmithKline nor anyone connected to litigation provided any funding for this manuscript.

## References

1. American Cancer Society. Breast cancer facts and figures. Atlanta, GA: American Cancer Society, Inc., 2011.
2. American Cancer Society. Global cancer facts & figures, 3rd edn., Atlanta: American Cancer Society, Inc., 2015.
3. Boström CE, Gerde P, Hanberg A, et al. Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. *Environ Health Perspect* 2002; 110:451–88.
4. International Agency for Research on Cancer (IARC). Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures. Monographs on the carcinogenic risk of chemicals to humans. Lyon, France: IARC. 2010.
5. Harrison RM, Smith DJT, Luhana L. Source apportionment of atmospheric polycyclic aromatic hydrocarbons collected from an urban location in Birmingham, UK. *Environ Sci Technol* 1996; 30:825–32.
6. Nielsen T. Traffic contribution of polycyclic aromatic hydrocarbons in the center of a large city. *Atmos Environ* 1996; 30:3481–90.
7. Gammon MD, Santella RM. PAH, genetic susceptibility and breast cancer risk: an update from the Long Island Breast Cancer Study Project. *Eur J Cancer* 2008; 44:636–40.
8. Gammon MD, Santella RM, Neugut AI, et al. Environmental toxins and breast cancer on Long Island. I. Polycyclic aromatic hydrocarbon (PAH)-DNA adducts. *Cancer Epidemiol Biomarkers Prev* 2002;11:677–85.
9. Rundle A, Tang D, Hibshoosh H, et al. The relationship between genetic damage from polycyclic aromatic hydrocarbons in breast tissue and breast cancer. *Carcinogenesis* 2000; 21:1281–9.
10. Bonner MR, Han D, Nie J, et al. Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. *Cancer Epidemiol Biomarkers Prev* 2005; 14:53–60.
11. Nie J, Beyea J, Bonner MR, et al. Exposure to traffic emissions throughout life and risk of breast cancer: the Western New York Exposures and Breast Cancer Study. *Cancer Causes Control* 2007; 18:947–55.
12. Mordukhovich I, Beyea J, Herring AH, et al. Vehicular traffic-related polycyclic aromatic hydrocarbon exposure and breast cancer incidence: the long island breast cancer study project (LIBCSP). *Environ Health Perspect* 2015; 124:30–38.
13. Braithwaite E, Wu X, Wang Z. Repair of DNA lesions induced by polycyclic aromatic hydrocarbons in human cell-free extracts: involvement of two excision repair mechanisms *in vitro*. *Carcinogenesis* 1998; 19:1239–46.
14. Berwick M, Vineis P. Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. *J Natl Cancer Inst* 2000; 92: 874–97.
15. Christmann M, Tomicic MT, Roos WP, et al. Mechanisms of human DNA repair: an update. *Toxicology* 2003; 193:3–34.
16. Goode EL, Ulrich CM, Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. *Cancer Epidemiol Biomarkers Prev* 2002; 11: 1513–30.
17. Robertson AB, Klungland A, Rognes T. Leiros. DNA repair in mammalian cells: base excision repair, the long and short of it. *Cell Mol Life Sci* 2009; 66:981–93.
18. Crew D, Gammon MD, Terry MB, et al. Polymorphisms in nucleotide excision repair genes, polycyclic aromatic hydrocarbon-DNA adducts, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2007; 16:2033–41.
19. Mechanic LE, Millikan RC, Player J, et al. Polymorphisms in nucleotide excision repair genes, smoking and breast cancer in African Americans and whites: a population-based case-control study. *Carcinogenesis* 2006; 27:1377–1385.
20. Shen J, Gammon MD, Terry MB, et al. Polymorphisms in *XRCC1* modify the association between polycyclic aromatic hydrocarbon-DNA adducts, cigarette smoking, dietary antioxidants, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2005; 14:336–42.
21. Terry MB, Gammon MD, Zhang FF, et al. Polymorphisms in the DNA repair gene *XPD*, polycyclic aromatic hydrocarbon-DNA adducts, cigarette smoking, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2004; 13:2053–8.
22. Gammon MD, Neugut AI, Santella RM, et al. The Long Island Breast Cancer Study Project: description of a multi-institutional collaboration to identify environmental risk factors for breast cancer. *Breast Cancer Res Treat* 2002; 74:235–54.
23. Waksberg J. Sampling methods for random digit dialing. *J Am Stat Assoc* 1978; 73:40–46.
24. Gammon MD, Eng SM, Teitelbaum SL, et al. Environmental tobacco smoke and breast cancer incidence. *Environ Res* 2004; 96:176–85.
25. Steck E, Gaudet MM, Eng SM, et al. Cooked meat and risk of breast cancer – lifetime versus recent dietary intake. *Epidemiology* 2007; 18:373–82.
26. White AJ, Teitelbaum SL, Stellman SD, et al. Indoor air pollution exposure from use of indoor stoves and fireplaces in association with breast cancer: a case-control study. *Environ Health* 2014; 13:108.
27. Beyea J, Hatch M, Stellman SD, et al. Validation and calibration of a model used to reconstruct historical exposure to polycyclic aromatic hydrocarbons for use in epidemiologic studies. *Environl Health Perspect* 2006; 114:1053–8.
28. Beyea J, Stellman SD, Hatch M, et al. Airborne emissions from 1961 to 2004 of benzo[a]pyrene from U.S. vehicles per km of travel based on tunnel studies. *Environ Sci Technol* 2008; 42:7315–20.
29. Beyea J, Stellman SD, Teitelbaum SL, et al. Imputation method for lifetime exposure assessment in air pollution epidemiologic studies. *Environ Health* 2013; 12:62.
30. Rossner P, Terry MB, Gammon MD, et al. OGG1 polymorphisms and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2006; 15:811–5.
31. Shen J, Gammon MD, Terry MB, et al. *Xeroderma pigmentosum complementation group C* genotypes/diplotypes play no independent or interaction role with polycyclic aromatic hydrocarbon-DNA adducts for breast cancer risk. *Eur J Cancer* 2008; 44:710–7.
32. Hosmer D, Lemeshow S. Applied Logistic Regression. New York, NY: John Wiley & Sons, 1989.
33. Rothman K, Lash TL, Greenland S. Modern Epidemiology. 3rd Ed. Philadelphia, PA: Lippincott, Williams & Wilkins, 2012.
34. Hosmer DW, Lemeshow S. Confidence interval estimation of interaction. *Epidemiology* 1992; 3: 452–6.
35. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc* 1996; 91:473–89.
36. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol* 2008; 8:70.
37. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996; 49:1373–9.
38. Gaunt TR, Rodriguez S, Day I. Cubic exact solutions for the estimation of pairwise haplotype frequencies: implications for linkage disequilibrium analyses and a web tool 'CubeX.' *BMC Bioinformatics* 2007; 8:428.
39. Tang D, Cho S, Rundle A, et al. Polymorphisms in the DNA repair enzyme *XPD* are associated with increased levels of PAH-DNA adducts in a case-control study of breast cancer. *Breast Cancer Res Treat* 2002; 75:159–66.
40. Wood RD. Nucleotide excision repair in mammalian cells. *J Biol Chem* 1997; 272:23465–8.
41. Choi JY, Hamajima N, Tajima K, et al. *hOGG1 Ser326Cys* polymorphisms and breast cancer risk among Asian women. *Breast Cancer Res Treat* 2003; 79:59–62.
42. Manuguerra M, Saletta F, Karagas MR, et al. *XRCC3* and *XPD/ERCC2* single nucleotide polymorphisms and the risk of cacner: a HuGE review. *Am J Epidemiol* 2006; 164:297–302.
43. Qiu LX, Yao L, Zhang J, et al. *XPD* Lys751Gln polymorphism and breast cancer susceptibility: a meta-analysis involving 28,709 subjects. *Breast Cancer Res Treat* 2010; 124:229–35.
44. Matullo G, Peluso M, Polidoro S, et al. Combination of DNA repair gene single nucleotide polymorphisms and increased levels of DNA adducts

in a population-based study. *Cancer Epidemiol Biomarkers Prev* 2003; 12:674–7.
45. Palli D, Russo A, Masala G, et al. DNA adduct levels and DNA repair polymorphisms in traffic-exposed workers and a general population sample. *Int J Cancer* 2001; 94:121–7.
46. Zienolddiny S, Campa D, Lind H, et al. Polymorphisms of DNA repair genes and risk of non-small cell lung cancer. *Carcinogenesis* 2006; 27: 560–7.
47. Leng S, Cheng J, Zhang L, et al. The association of *XRCC1* haplotypes and chromosomal damage levels in peripheral blood lymphocyte among coke-oven workers. *Cancer Epidemiol Biomarkers Prev* 2005; 14:1295–301.
48. Tarng DC, Tsai TJ, Chen WT, et al. Effect of human *OGG1* 1245C>G gene polymorphism on 8-hydroxy-2'-deoxyguanosine levels of leukocyte DNA among patients undergoing chronic hemodialysis. *J Am Soc Nephrol* 2001; 12:2338–47.
49. Aka P, Mateuca R, Buchet JP, et al. Are genetic polymorphisms in *OGG1*, *XRCC1* and *XRCC3* genes predictive for the DNA strand break repair phenotype and genotoxicity in workers exposed to low soe ionizing radiations? *Mutat Res* 2004; 556:169–81.
50. Bravard A, Vacher M, Mortiz E, et al. Oxidation status of human *OGG1*-S326C polymorphic variant determines cellular DNA repair capacity. *Cancer Res* 2009; 69:3642–9.