

RESEARCH

Open Access



# Trajectories of long-term exposure to PCB153 and Benzo[a]pyrene (BaP) air pollution and risk of breast cancer

Pauline Desnavailles<sup>1†</sup>, Delphine Praud<sup>2,3†</sup>, Blandine Le Provost<sup>1</sup>, Hidetaka Kobayashi<sup>1</sup>, Floriane Deygas<sup>2,3</sup>, Amina Amadou<sup>2,3</sup>, Thomas Coudon<sup>2,3</sup>, Lény Grassot<sup>2,3</sup>, Elodie Faure<sup>6</sup>, Florian Couvidat<sup>4</sup>, Gianluca Severi<sup>5,6</sup>, Francesca Romana Mancini<sup>6</sup>, Béatrice Fervers<sup>2,3</sup>, Cécile Proust-Lima<sup>1†</sup> and Karen Leffondré<sup>1\*†</sup>

## Abstract

**Background** While genetic, hormonal, and lifestyle factors partially elucidate the incidence of breast cancer, emerging research has underscored the potential contribution of air pollution. Polychlorinated biphenyls (PCBs) and benzo[a]pyrene (BaP) are of particular concern due to endocrine-disrupting properties and their carcinogenic effect.

**Objective** To identify distinct long term trajectories of exposure to PCB153 and BaP, and estimate their associations with breast cancer risk.

**Methods** We used data from the XENAIR case–control study, nested within the ongoing prospective French E3N cohort which enrolled 98,995 women aged 40–65 years in 1990–1991. Cases were incident cases of primary invasive breast cancer diagnosed from cohort entry to 2011. Controls were randomly selected by incidence density sampling, and individually matched to cases on delay since cohort entry, and date, age, department of residence, and menopausal status at cohort entry. Annual mean outdoor PCB153 and BaP concentrations at residential addresses from 1990 to 2011 were estimated using the CHIMERE chemistry-transport model. Latent class mixed models were used to identify profiles of exposure trajectories from cohort entry to the index date, and conditional logistic regression to estimate their association with the odds of breast cancer.

**Results** 5058 cases and 5059 controls contributed to the analysis. Five profiles of trajectories of PCB153 exposure were identified. The class with the highest PCB153 concentrations had a 69% increased odds of breast cancer compared to the class with the lowest concentrations (95% CI 1.08, 2.64), after adjustment for education and matching factors. The association between identified BaP trajectories and breast cancer was weaker and suffered from large CI.

**Conclusions** Our results support an association between long term exposure to PCB153 and the risk of breast cancer, and encourage further studies to account for lifetime exposure to persistent organic pollutants.

**Keywords** Breast cancer, Long-term exposure, Polychlorobiphenyl, Benzo[a]pyrene, Trajectories of exposure, Environmental carcinogens

<sup>†</sup>Pauline Desnavailles and Delphine Praud share first authorship.

<sup>†</sup>Cécile Proust-Lima and Karen Leffondré share last authorship.

\*Correspondence:

Karen Leffondré

Karen.leffondre@u-bordeaux.fr

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Background

Breast cancer represents the most prevalent cancer globally, accounting for 2.26 million new cases in 2020, and being the leading cause of cancer-related deaths in women [47]. While substantial progress has been made in understanding the etiology of breast cancer, encompassing genetic, hormonal, and lifestyle factors, these elements only partially explain its incidence. Air pollution classified as carcinogenic to humans by the International Agency for Research on Cancer (IARC) [35], has emerged as a potential contributor to breast cancer risk [45, 50]. Air pollution is a complex mixture of multiple pollutants including compounds with endocrine-disrupting properties such as polychlorinated biphenyls (PCBs) and polycyclic aromatic hydrocarbons (PAHs) [10]. Increased risks for breast cancer associated with exposure to PCBs, and benzo[a]pyrene (BaP) the most well-known PAH, have been reported [3, 15, 31, 34]. However, to date, the epidemiologic evidence on the effect of environmental exposure to these agents has been inconsistent.

PCBs are a group of synthetic chlorinated chemicals composed of various congeners. They were widely used in industry between the 1930s and the 1970s for their thermal, electrical, and non-flammable properties until their prohibition in the 1980s in many countries, including France [4, 31]. While they were mainly used as dielectric fluid in capacitors and transformers, they have also been used in construction materials and were released into the environment by leakage, volatilization or erosion. As persistent organic pollutants, PCBs are stable in various environmental matrices, i.e., biota, air, water, soil or sediment, and have been shown to bioaccumulate in the human body [29]. Among PCBs, PCB153 is the most abundantly detected in the environment and the primary contributor to the PCBs body burden estimate [6, 14, 21, 52, 54]. Most previous epidemiological studies on breast cancer have estimated exposure to PCBs from blood or adipose tissue samples, particularly from breast tissue, with inconsistent results [31, 32]. While some studies found association for some PCBs only [56], others found positive associations for a group of PCB congeners [28, 37, 44, 52], or even inverse association [30]. A positive correlation between PCB153 and total PCBs has been observed in several studies [6, 14]. In the XENAIR case-control study nested with the French prospective E3N cohort [2], we assessed air pollutant exposure at the residential addresses up to 22 years prior to breast cancer diagnosis and found a positive association with long term exposure to airborne PCB153 [15]. For each one standard deviation (SD, 55 pg/m<sup>3</sup>) increase of the cumulative of airborne PCB153 concentration levels, the adjusted odds ratio (OR) was 1.19 (95% confidence interval (CI): 1.08–1.31) [15].

PAHs are mainly produced by domestic and vehicular emissions as well as industrial and natural combustion [5]. They have both estrogenic and anti-estrogenic effects [50, 51]. Benzo[a]pyrene (BaP) is one of the best known and well characterized PAH components [25], and is often used as a surrogate for estimating total PAH exposure [7]. The general population is exposed to it through ambient air, tobacco smoke, water, and food [18, 41]. Two epidemiological studies have found a positive association between breast cancer and BaP exposure from traffic [38, 39]. In the XENAIR case-control study, we also highlighted a positive association with long term exposure to airborne BaP concentrations at the residential addresses up to 22 years prior to breast cancer diagnosis: adjusted OR of 1.15 (95% CI: 1.04–1.27) for each increase of 1 interquartile range (IQR) (1.42 ng/m<sup>3</sup>) in the cumulative airborne BaP concentration levels [1].

None of the previous studies have explored the longitudinal trajectories of exposure to PCB153 or BaP preceding the diagnosis of breast cancer. Studies predominately considered summary estimates of the exposure history, such as time-averaged or cumulative exposure metrics which only partly reflect the evolution of the subjects' exposure over time. Even in our two previous studies using the XENAIR case-control data [1, 15], the cumulative exposure metric did not reflect the temporal dynamic of exposure. Yet, breast cancer risk may vary depending on the timing of exposure and according to exposure trajectories for different agents [16, 20, 49]. Investigating the exposure trajectories may accelerate the understanding of the effects of airborne PCBs and BaP exposures on breast cancer risk.

In the present study, we address this need by identifying distinct trajectory patterns of outdoor exposure to PCB153 and BaP at the residential addresses, over up to 22 years before breast cancer diagnosis, and estimate their association with the risk of breast cancer, using XENAIR case-control data and a latent class mixed model (LCMM) approach.

## Methods

### The E3N cohort

E3N is an ongoing French prospective cohort which aims to investigate risk factors for severe chronic conditions in women [11]. It enrolled 98 995 women living in France between June 1990 and November 1991, aged 40–65 years, and insured with a national health insurance covering workers from the French National Education System (MGEN). Participants were followed with a self-administrated mailed questionnaire every 2 or 3 years, collecting data on their lifestyle (smoking and physical activity), reproductive factors (age at menarche and menopause, number of children, age at first full-term

pregnancy, and breastfeeding), anthropometry (height, weight), medical history (benign breast disease) and familial history of cancer. To date, a total of 13 questionnaires were sent to participants, with a participation rate around 80% [11]. Breast cancers were first self-declared in each questionnaire, and then confirmed by pathology reports for 93% of the cases. The proportion of false-positive was less than 5%. In addition to the questionnaires, blood samples were collected from 25 000 participants between 1994 and 1999, and saliva samples from 47 000 participants between 2009 and 2011. Residential addresses were collected from questionnaires along the follow-up. Place of birth (postal code and municipality) was collected in the first questionnaire and used to categorize each participant's birthplace as either urban or rural status based on data from the closest national census [9].

#### The XENAIR nested case–control study

XENAIR is a case–control study nested within the E3N cohort [1]. The main objective of this study was to investigate chronic long-term effects of exposure to multiple ambient air pollutants on the risk of breast cancer. Cases were all incident cases of primary invasive breast cancer diagnosed from entry into the E3N cohort to the 10th follow-up questionnaire (2011), excluding Paget's disease and phyllodes tumors as well as women who had any cancer diagnosis before cohort entry. Cases not ascertained by pathology report were not excluded because of the low false-positive proportion in self-reports (<5%). For each case, one control was randomly selected using incidence density sampling, among E3N cohort participants at risk of breast cancer at the time of case's diagnosis, using time since cohort entry as the time axis. Cases who gave a blood sample were matched to controls who also gave one on the following factors measured at blood collection: calendar date ( $\pm 3$  months), age ( $\pm 1$  year), menopausal status, and French department of residence. Cases who did not provide a blood sample were matched to controls on the same criteria but were collected at cohort entry, as well as on the presence or not of a saliva sample.

All cases' and controls' residential addresses that were collected in the E3N questionnaires, from cohort entry to the index date (diagnosis for cases, selection for controls), were geocoded (X and Y coordinates, addresses) using the ArcGIS Software (ArcGIS Locator version 10.0, Environmental System Research Institute – ESRI, Redlands, CA, USA) and its reference street network database (BD Adresse<sup>®</sup>) from the National Geographic Institute. Geocoding was performed by a trained technician blinded to the case–control status of the participants, using a validated method [19]. Management of missing and incomplete addresses has been described

previously [15]. In brief, women who had at least one missing address in questionnaires, or at least one address outside the continental mainland France, were excluded, as well as their matched cases or controls. In the present study, we further excluded all cases and controls with a follow-up shorter than one year, as our statistical analysis relied on mean annual exposure.

#### PCB153 and BaP exposure assessment

Annual atmospheric concentration levels of PCB153 and BaP were estimated at the subjects' consecutive geocoded residential addresses from their cohort entry (1990–1991) to their index date (date of diagnosis for case and selection for her individually matched control, which could occur at any time between 1991 and 2011) using the CHIMERE model [1, 15], a regional chemistry-transport model developed by the National Institute for Industrial Environment and Risks [12, 26]. CHIMERE was designed to simulate pollutant transport using emission data, meteorological fields, and boundary conditions, at a spatial resolution of  $0.125^\circ \times 0.0625^\circ$  (approximately  $7 \times 7$  km). This allowed us to compute for each woman, individual exposure trajectories for each pollutant (Figure S1 in Supplementary Material). For BaP, we further described how much of annual concentrations were above the European target value of 10  $\text{pg.10/m}^3$ , or above a previous estimated acceptable environmental risk for cancer of 1.2  $\text{ng.10/m}^3$  [26].

#### Statistical analysis

To identify distinct profiles of exposure trajectories to PCB153 and BaP, we used a latent class mixed model (LCMM) for each pollutant [33, 42]. The LCMM assumes that the population is composed of  $G$  distinct non-observed subgroups (latent classes) of subjects, each subgroup being characterized by a specific pattern of trajectory. The LCMM was defined as follows: The woman membership to each latent class followed a categorical distribution with probabilities to be estimated. Given a specific class-membership, the trajectory of the annual average air pollutant concentration estimates from 1990 to 2011 was modelled by a mixed model with class-specific fixed effects. Because neither PCB153 nor BaP annual average concentrations estimates were normally distributed (Figure S2 in Supplementary Material), the mixed model included an internal transformation using I-splines functions with four knots at quintiles of the pollutant distribution (4.71, 6.93, 9.26 and 12.29  $\text{ng/m}^3$  for PCB153; and 0.88, 1.25, 1.67 and 2.35  $\text{pg.10/m}^3$  for BaP). Because trajectories of exposure to each pollutant were nonlinear in time (Figure S1 in Supplementary Material), the class-specific mean trajectories were modelled using a natural spline function of time, with two knots

at tertiles of follow-up (4.30 and 8.97 years). All spline functions and knots were chosen to maximize the fit to data according to the Akaike Information Criteria (AIC). Multivariate normal random effects on the intercept and each spline coefficient of the time function were added to account for the correlation between repeated pollutant measures of each woman. Their variance–covariance matrix was unstructured and common across classes. Each LCMM was estimated by maximum likelihood using 100 different sets of initial parameter values to ensure convergence to the global maximum [27]. Once the model was estimated, each woman was a posteriori classified in the class of exposure trajectory to which she had the highest probability to belong, given her own observed trajectory. To estimate the number of classes  $G$  for each pollutant, we estimated five LCMM for each pollutant, each with a given number of classes ( $G = 1, 2, 3, 4, 5$ ). We selected the best LCMM for each pollutant with respect to i) the fit to data using the Bayesian Information criterion (BIC), ii) the discrimination between classes using entropy, and iii) a mix of both fit and discrimination using the Integrated Completed Likelihood (ICL) criterion [8].

Once we selected the best LCMM for each pollutant, we described demographic and lifestyle characteristics in each identified classes. We further used conditional logistic regression to estimate the adjusted association between exposure trajectory class membership and the odds of being diagnosed with breast cancer. Because the entropy of the best LCMM was excellent (0.94 for PCB153 and 0.96 for BaP) and the sample size was large, we did not account for classification uncertainty [17]. In addition to matching factors, each model was adjusted for the level of education (secondary, 1 to 2-year university degree,  $\geq 3$ -year university degree) which was the minimally sufficient adjustment set identified from the directed acyclic graph (Figure S3 in Supplementary Material [15]). We also adjusted for age at the index date (in years) to account for potential residual differences in age within each case–control pair. In a sensitivity analysis, we further adjusted for known and potential risk factors of breast cancer: urban residence at birth (yes, no), age at menarche ( $< 12$ ,  $12$ – $13$ ,  $\geq 14$  years), total physical activity (quartiles of METs-h/week), smoking status (never, current, former smoker), alcohol drinking (0,  $1$ – $6.6$ ,  $\geq 6.7$  g/day), body mass index ( $< 25$ ,  $25$ – $29$ ,  $\geq 30$  kg/m<sup>2</sup>), parity and age at the first full-term pregnancy (no child/ $1$ – $2$  children and age  $< 30$  years,  $1$ – $2$  children and age  $\geq 30$  years,  $> 3$  children), ever breastfeeding (yes, no), use of oral contraceptives (yes, no), use of menopausal hormone replacement therapy (yes, no), and history of personal benign breast disease (yes, no). All variables were taken at cohort entry except for alcohol

consumption which was collected at questionnaire 3 through the dietary questionnaire and for hormonal treatments, we used the information available in the last questionnaire before the index date. Missing data were imputed by the modal class for risk factors with less than 5% of missing data, and a 'missing' category was created for risk factors with more missing data [15, 24].

All analyses were performed using the R software version 3.6.1, and the *lcmm* R package version 1.9.2 for the identification of classes of exposure trajectories [42].

## Results

### Cases and controls selection and characteristics

Among the 98,995 women included in the E3N cohort study, 6,298 incident breast cancer diagnoses were identified between 1990 and 2011 (Figure S4 in Supplementary Material [15]). After excluding 19 cases of Paget disease and phyllodes tumors, 3 cases with missing data on matching factors, and 1,054 cases or controls with missing data on addresses or addresses outside metropolitan France, the XENAIR case–control study included 5,222 cases and 5,222 matched controls [15]. For the present study, we further excluded 331 women with follow-up shorter than one year, leaving 5,058 cases and 5,059 controls for the trajectory analysis, and 5,058 case–control pairs for the conditional logistic regression analysis.

For both, cases and controls, the median age at cohort entry was close to 49 years (IQR: 44, 54) and the median follow-up time from cohort entry to the index date was 11.5 years (IQR: 7.0, 15.7) (Table 1). The median of the mean annual concentrations estimates over the follow-up period was 8.1 pg/m<sup>3</sup> (IQR: 5.3, 11.5) for PCB153 and 1.5 pg.10/m<sup>3</sup> (IQR: 1.1, 2.1) for BaP in cases, and almost identical in controls. Almost all annual BaP concentrations estimates were below the European target value of 10 pg.10/m<sup>3</sup>, yet about 62% were above the estimated acceptable environmental risk for cancer of 1.2 ng.10/m<sup>3</sup>. As expected, cases tended to have fewer children than controls and more often personal history of benign breast disease and family history of breast cancer (Table 1).

### Classes of PCB153 trajectories and association with breast cancer

The best LCMM for PBC153 in terms of both fit to data and discrimination capacity between classes had five latent classes (Figure S5, Table S1 Supplementary materials). The mean exposure trajectory estimated in each of the five identified classes are shown in the left panel of Fig. 1. The right panel shows individual observed trajectories of women classified in each class, to illustrate the variability of exposure within each class. Except for Class 4, all classes had a decreasing mean trajectory of the annual average PCB153 concentrations, with varying

**Table 1** Characteristics of cases and controls, XENAIR case–control study nested in the E3N cohort, France, 1990–2011

Characteristics	Cases (n = 5058) Median (IQR) or n (%)	Controls (n = 5059) Median (IQR) or n (%)
Age at cohort entry (years)	48.8 (44.1, 54.0)	48.7 (44.0, 54.0)
Age at index date (years)	60.7 (55.3, 66.3)	60.8 (55.2, 66.3)
Follow-up time in XENAIR (years)	11.5 (7.0, 15.7)	11.5 (7.0, 15.7)
Annual PCB153 concentration (in pg/m <sup>3</sup> ) over the follow-up (number of annual concentrations: 54,997 in cases, 55,045 in controls)	8.1 (5.3, 11.5)	8.1 (5.3, 11.4)
Annual BaP concentration (in ng.10/m <sup>3</sup> ) over the follow-up (number of annual concentrations: 54,997 in cases, 55,045 in controls)	1.5 (1.0, 2.1)	1.4 (1.0, 2.1)
< 1.2	20,538 (37.3)	20,868 (37.9)
1.2 – < 10.0	34,335 (62.4)	34,094 (61.9)
≥ 10.0	124 (0.2)	83 (0.2)
Urban residence at birth		
Rural	1328 (26.3)	1382 (27.3)
Urban	3196 (63.2)	3091 (61.1)
Missing	534 (10.6)	586 (11.6)
Urban residence at cohort entry		
Rural	1456 (28.8)	1456 (28.8)
Urban	3068 (60.7)	3017 (59.6)
Missing	534 (10.6)	586 (11.6)
Number of residential addresses change during follow-up		
Never	2940 (58.1)	2974 (58.8)
Once	1674 (33.1)	1648 (32.6)
Twice or more	444 (8.8)	437 (8.6)
Education		
Secondary	774 (15.3)	846 (16.7)
1- to 2-year university degree	2408 (47.6)	2540 (50.2)
≥ 3 year university degree	1876 (37.1)	1673 (33.1)
Body Mass Index (kg/m <sup>2</sup> ) at cohort entry		
< 25	4196 (83.0)	4207 (83.2)
25 – 29	719 (14.2)	695 (13.7)
≥ 30	143 (2.8)	157 (3.1)
Total physical activity (METs-h/week) at cohort entry		
< 25.3	1262 (25.0)	1188 (23.5)
25.3 – 35.6	1343 (26.6)	1318 (26.1)
35.6 – 51.8	1301 (25.7)	1292 (25.5)
≥ 51.8	1152 (22.8)	1261 (24.9)
Smoking status at cohort entry		
Never smoker	2717 (53.7)	2786 (55.1)
Current smoker	757 (15.0)	715 (14.1)
Former smoker	1584 (31.3)	1558 (30.8)
Alcohol drinking (g/day) at Q3		
0	402 (7.9)	473 (9.3)
1 – 6.6	1292 (25.5)	1391 (27.5)
≥ 6.7	2002 (39.6)	1867 (36.9)
Missing	1362 (26.9)	1328 (26.3)
Age at menarche (years)		
< 12	1053 (20.8)	1025 (20.3)
12–13	2631 (52.0)	2616 (51.7)
≥ 14	1374 (27.2)	1418 (28.0)

**Table 1** (continued)

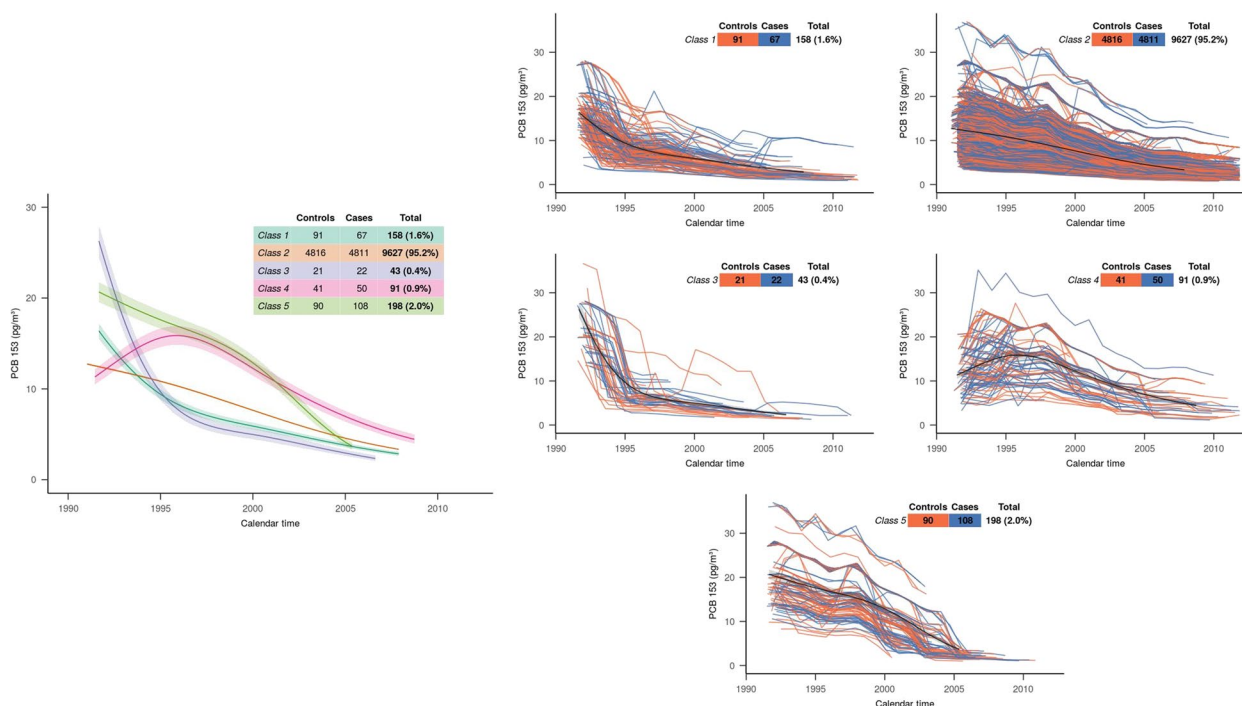
Characteristics	Cases (n = 5058) Median (IQR) or n (%)	Controls (n = 5059) Median (IQR) or n (%)
Parity at cohort entry		
No child	650 (12.9)	542 (10.7)
1–2	3099 (61.3)	3006 (59.4)
≥ 3	1309 (25.9)	1511 (29.9)
Age at first full-term pregnancy (years)		
No child	650 (12.9)	542 (10.7)
< 30	3751 (74.2)	3977 (78.6)
≥ 30	657 (13.0)	540 (10.7)
Ever breastfeeding at cohort entry		
No child	650 (12.9)	542 (10.7)
No	1656 (32.7)	1736 (34.3)
Yes	2752 (54.4)	2781 (55.0)
Ever use of oral contraceptives		
No	2070 (40.9)	2084 (41.2)
Yes	2988 (59.1)	2975 (58.8)
Menopausal status		
Premenopausal at cohort entry	2019 (39.9)	2008 (39.7)
Menopause during follow-up	2253 (44.5)	2334 (46.1)
Postmenopausal at index date	786 (15.5)	717 (14.2)
Ever use of menopausal hormone replacement therapy at cohort entry		
No	1915 (37.9)	2065 (40.8)
Yes	3038 (60.1)	2883 (57.0)
Missing	105 (2.1)	111 (2.2)
Ever mammography at cohort entry		
No	1160 (22.9)	1378 (27.2)
Yes	3898 (77.1)	3681 (72.8)
Previous history of personal benign breast disease at cohort entry		
No	3587 (70.9)	3925 (77.6)
Yes	1471 (29.1)	1134 (22.4)
Previous family history of breast cancer at cohort entry		
No	4199 (83.0)	4521 (89.4)
Yes	859 (17.0)	538 (10.6)

IQR interquartile range

slopes (Fig. 1). Class 1 included 158 (1.6%) women (67 cases, 91 controls) who were overall exposed to low PCB153 concentrations except in the early 90 s (median at 6.6 pg/m<sup>3</sup>, IQR: 4.2, 10.5, Table 2). Class 2 was by far the largest class with 9,627 (95.2%) women (4,816 cases, 4,811 controls) and had a linear and slowly decreasing mean trajectory of PCB153 concentrations over the whole study period (1991–2011) (Fig. 1) (median at 8.0 pg/m<sup>3</sup>, IQR: 5.3, 11.3, Table 2). Class 3 was the smallest class with only 43 (0.4%) women (22 cases, 21 controls). This class had the highest concentration levels in the early 1990s (more than 25 pg/m<sup>3</sup>, Fig. 1), but exhibited the deepest decline from 1990 to 1996, resulting overall in the lowest median PCB153 concentrations

over the whole study period (5.6 pg/m<sup>3</sup>, IQR: 3.6, 12.5, Table 2). Class 4 was made of 91 (0.9%) women (50 cases, 41 controls) and was the only class with an increasing mean PCB153 trajectory from 1991 to 1996, likely due to a large number of residential moves (Table 2). With a mean trajectory that decreased thereafter but stayed at high concentration levels (Fig. 1), this class presented the second highest median concentrations over the whole period (11.3 pg/m<sup>3</sup>, IQR: 7.4, 16.1, Table 2). Class 5 was made of 198 (2.0%) women (108 cases, 90 controls) with high concentrations until the early 2000s (Fig. 1), and thus exhibited the highest median PCB153 concentrations levels over the whole study period (16.7 pg/m<sup>3</sup>, IQR: 10.6, 22.9, Table 2).





**Fig. 1** Trajectories of PCB153 concentrations, XENAIR case–control study, France, 1990–2011 ( $n = 5\,058$  cases,  $5\,059$  controls). The left panel shows the estimated mean trajectory of the average annual PCB153 concentrations over the follow-up in the five classes identified by LCMM model. The right-hand panel shows the individual trajectories of subjects classified in each class, with the black bold line representing the estimated mean trajectory in the class with its 95% CI. The representation of each estimated mean trajectory is truncated at the 95th percentile of the distribution of observed exposure times in women a posteriori classified in the class

Mean age was similar in the five classes of PBC153 trajectories (Table 2). As expected, Class 5 which had the highest median PCB153 concentrations over the whole-time window (1990–2011), also had the highest proportion of women who lived in urban area at cohort entry (1990–1991), but also at birth (Table 2). Compared to other classes, women from Class 5 were more likely to have higher education levels, lower physical activity, an older age at their first full-term pregnancy, and to report a family history of breast cancer (Table 2). Class 4 which had the only non-monotone mean exposure trajectory over time, had a large proportion of women with residential mobility over the study period (Table 2).

Classes 2–5 tended to have a higher risk of breast cancer compared to Class 1, even if the 95% CI was large for Class 3 because of fewer cases and controls (Table 3). In particular, Classes 4 and 5 with the strongest concentrations had more than 60% higher odds of breast cancer than Class 1 after adjustment for education level, age at the index date, as well as matching factors (e.g. OR = 1.69, 95% CI: 1.08, 2.64 for Class 5, Table 3). All results were very similar after further adjustment for risk factors of breast cancer (Table 3).

### Classes of BaP trajectories and association with breast cancer

The best LCMM for BaP in terms of both fit to data and discrimination capacity between the classes had four latent classes (Figure S4 and Table S1 in supplementary materials). All the four mean trajectories of BaP exposure were declining from 1990 to 2011, but at different rates and with different average levels, ranging from the lowest concentrations for Class 1 (91.9% of women) to the highest concentrations for Class 4 which showed very high concentrations in the early 1990s (Fig. 2). The median average annual BaP concentrations over the study period was  $1.4 \text{ ng}\cdot 10/\text{m}^3$  (IQR: 0.9, 2.0) in Class 1,  $1.6 \text{ ng}\cdot 10/\text{m}^3$  (IQR: 0.7, 2.7) in Class 2,  $3.0 \text{ ng}\cdot 10/\text{m}^3$  (IQR: 1.9, 4.6) in Class 3, and  $4.4 \text{ ng}\cdot 10/\text{m}^3$  (IQR: 2.7, 7.1) in Class 4 (Table 4). In classes 3 and 4, more than 91% of all annual concentration levels were above the estimated acceptable risk level of  $1.2 \text{ ng}\cdot 10/\text{m}^3$  compared to less than 62% in Classes 1 and 2.

Mean age at cohort entry was similar in all classes of BaP trajectories (Table 4). Like PCB153-Class 5, BaP-Class 4 which had the highest BaP concentrations, also had the highest proportion of women having lived in urban areas at birth and in 1990–1991. Compared to

**Table 2** Characteristics of the participants according to their PCB153 exposure trajectory, in the XENAIR case–control study nested in the E3N cohort, France, 1990–2011 ( $n = 5\,058$  cases,  $5\,059$  controls)

Characteristics	Class 1 ( $n = 158$ , 67 cases, 91 controls) Median (IQR) or $n$ (%)	Class 2 ( $n = 9,627$ , 4,811 cases, 4,816 controls) Median (IQR) or $n$ (%)	Class 3 ( $n = 43$ , 22 cases, 21 controls) Median (IQR) or $n$ (%)	Class 4 ( $n = 91$ , 50 cases, 41 controls) Median (IQR) or $n$ (%)	Class 5 ( $n = 198$ , 108 cases, 90 controls) Median (IQR) or $n$ (%)
Average annual PCB153 concentration over the follow-up (pg/m <sup>3</sup> )	6.6 (4.2, 10.5)	8.0 (5.3, 11.3)	5.6 (3.6, 12.5)	11.3 (7.4, 16.1)	16.7 (10.6, 22.9)
Age at cohort entry (years)	49.6 (43.6, 54.4)	48.7 (44.1, 54.0)	51.5 (45.0, 54.6)	47.4 (43.0, 52.6)	48.7 (44.7, 54.1)
Age at index date (years)	61.3 (55.8, 66.4)	60.7 (55.2, 66.4)	62.0 (56.7, 68.3)	61.3 (57.1, 65.9)	60.1 (55.5, 64.9)
Follow-up time (years)	10.5 (8.0, 15.4)	11.5 (6.9, 15.8)	11.5 (6.7, 14.8)	13.0 (9.3, 17.2)	10.5 (7.7, 13.2)
Urban residence at birth <sup>a</sup>	108 (74.0)	5960 (69.6)	23 (59.0)	53 (67.1)	143 (81.7)
Urban residence at cohort entry <sup>a</sup>	116 (79.5)	5711 (66.7)	34 (87.2)	58 (73.4)	166 (94.9)
Number of residential addresses changes during follow-up					
None	0 (0.0)	5821 (60.5)	0 (0.0)	0 (0.0)	93 (47.0)
One	112 (70.9)	3054 (31.7)	32 (74.4)	50 (54.9)	74 (37.4)
Two or more	46 (29.1)	752 (7.8)	11 (25.6)	41 (45.1)	31 (15.7)
Education					
Secondary	19 (12.0)	1566 (16.3)	7 (16.3)	8 (8.8)	20 (10.1)
1- to 2-year university degree	84 (53.2)	4738 (49.2)	17 (39.5)	32 (35.2)	77 (38.9)
≥ 3 year university degree	55 (34.8)	3323 (34.5)	19 (44.2)	51 (56.0)	101 (51.0)
Body Mass Index (kg/m <sup>2</sup> ) at cohort entry					
< 25	127 (80.4)	8004 (83.1)	37 (86.0)	74 (81.3)	161 (81.3)
25–29	25 (15.8)	1345 (14.0)	4 (9.3)	12 (13.2)	28 (14.1)
≥ 30	6 (3.8)	278 (2.9)	2 (4.7)	5 (5.5)	9 (4.5)
Total physical activity (METs-h/week) at cohort entry					
< 25.3	49 (31.0)	2307 (24.0)	12 (27.9)	24 (26.4)	58 (29.3)
25.3 – 35.6	42 (26.6)	2512 (26.1)	13 (30.2)	28 (30.8)	66 (33.3)
35.6 – 51.8	37 (23.4)	2477 (25.7)	7 (16.3)	25 (27.5)	47 (23.7)
≥ 51.8	30 (19.0)	2331 (24.2)	11 (25.6)	14 (15.4)	27 (13.6)
Smoking status at cohort entry					
Never smoker	82 (51.9)	5264 (54.7)	19 (44.2)	47 (51.6)	91 (46.0)
Current smoker	26 (16.5)	1378 (14.3)	8 (18.6)	18 (19.8)	42 (21.2)
Former smoker	50 (31.6)	2985 (31.0)	16 (37.2)	26 (28.6)	65 (32.8)
Alcohol drinking (g/day) at Q3 <sup>a</sup>					
0	4 (3.7)	855 (12.1)	2 (7.4)	5 (8.5)	9 (6.0)
1 – 6.6	34 (31.5)	2569 (36.3)	6 (22.2)	21 (35.6)	53 (35.1)
≥ 6.7	70 (64.8)	3658 (51.7)	19 (70.4)	33 (55.9)	89 (58.9)
Age at menarche (years)					
< 12	29 (18.4)	1970 (20.5)	7 (16.3)	18 (19.8)	54 (27.3)
12–13	83 (52.5)	4994 (51.9)	21 (48.8)	50 (54.9)	99 (50.0)
≥ 14	46 (29.1)	2663 (27.7)	15 (34.9)	23 (25.3)	45 (22.7)
Parity at cohort entry					
No child	21 (13.3)	1115 (11.6)	5 (11.6)	13 (14.3)	38 (19.2)
1–2	86 (54.4)	5829 (60.5)	26 (60.5)	54 (59.3)	110 (55.6)
≥ 3	51 (32.3)	2683 (27.9)	12 (27.9)	24 (26.4)	50 (25.3)



**Table 2** (continued)

Characteristics	Class 1 (n = 158, 67 cases, 91 controls) Median (IQR) or n (%)	Class 2 (n = 9,627, 4,811 cases, 4,816 controls) Median (IQR) or n (%)	Class 3 (n = 43, 22 cases, 21 controls) Median (IQR) or n (%)	Class 4 (n = 91, 50 cases, 41 controls) Median (IQR) or n (%)	Class 5 (n = 198, 108 cases, 90 controls) Median (IQR) or n (%)
Age at first full-term pregnancy (years)					
No child	21 (13.3)	1115 (11.6)	5 (11.6)	13 (14.3)	38 (19.2)
< 30	124 (78.5)	7374 (76.6)	34 (79.1)	70 (76.9)	126 (63.6)
≥ 30	13 (8.2)	1138 (11.8)	4 (9.3)	8 (8.8)	34 (17.2)
Ever breastfeeding at cohort entry	80 (50.6)	5278 (54.8)	26 (60.5)	45 (49.5)	104 (52.5)
Ever use of oral contraceptives at cohort entry	95 (60.1)	5646 (58.6)	25 (58.1)	65 (71.4)	132 (66.7)
Menopausal status					
Premenopausal at cohort entry	73 (46.2)	3819 (39.7)	21 (48.8)	29 (31.9)	85 (42.9)
Menopause during follow-up	66 (41.8)	4366 (45.4)	16 (37.2)	55 (60.4)	84 (42.4)
Postmenopausal at index date	19 (12.0)	1442 (15.0)	6 (14.0)	7 (7.7)	29 (14.6)
Ever use of menopausal hormone replacement therapy up to the index date <sup>a</sup>	90 (58.4)	5625 (59.7)	28 (68.3)	58 (63.7)	120 (62.2)
Ever mammography at cohort entry	123 (77.8)	7185 (74.6)	38 (88.4)	64 (70.3)	169 (85.4)
Previous history of personal benign breast disease at cohort entry	48 (30.4)	2476 (25.7)	9 (20.9)	16 (17.6)	56 (28.3)
Previous family history of breast cancer at cohort entry	27 (17.1)	1317 (13.7)	4 (9.3)	9 (9.9)	40 (20.2)

IQR interquartile range

<sup>a</sup> percentage derived on available data

other BaP trajectories classes, they had the highest physical activity levels, less frequently a history of personal benign breast disease, and were more likely to having breastfed (Table 4). Among the 178 women in BaP-Class 4, only 6 were also in PCB153-Class 5, while 159 were in the large PCB153-Class 2 (Table S2, Supplementary materials).

The association between the classes of BaP trajectories and the risk of breast cancer was by far less clear than for PCB153, whatever the set of adjustment factors (Table 5). Confidence intervals were indeed large, even if they may indicate a tendency to a slight stronger risk for Classes 3 and 4 compared to Class 1 (e.g. OR = 1.21, 95% CI: 0.91; 1.62 for Class 3, Table 5).

## Discussion

Our trajectory-based analysis suggested an association between the risk of breast cancer and the trajectories of the outdoor PCB153 concentrations levels estimated at the residential addresses over a period of up to 22 years prior to diagnosis. More specifically, we found five distinct classes of PCB153 exposure trajectories which mostly declined at different rates and levels over the

1990–2011 period, with higher risk of breast cancer for trajectories with highest concentrations over the whole follow-up compared to the class with the lowest concentrations. For BaP, we found four distinct classes of exposure trajectories that all declined over the 1990–2011 period at different rates, which were weaker than rates of decline for PCB153. The confidence intervals for the association with breast cancer were too large to argue that higher trajectories of BaP were at higher risk, even if a few signals emerged.

To our knowledge, our study is the first study to investigate trajectories of long-term atmospheric pollutant exposure and their association with the risk of breast cancer. Overall, the evidence regarding the association between PCB exposure and breast cancer risk is not entirely conclusive, while some studies reported statistically significant associations between PCB exposure and increased breast cancer risk, conflicting results exist. Our results are consistent with previous studies that found associations between breast cancer and measures of PCBs in blood or adipose tissue samples [28, 37, 44, 52], and in particular in studies examining separately group II PCBs, including PCB153 [31, 53, 56]. Our results are

**Table 3** Association between classes of PCB153 trajectories and breast cancer, XENAIR case–control study nested in the E3N cohort, France, 1990–2011 ( $n = 5,058$  case–control pairs)

Class	<i>n</i> cases / controls	aOR <sup>a</sup> (95% CI)	aOR <sup>b</sup> (95% CI)
1	67 / 91	1	1
2	4811 / 4816	1.36 (0.99; 1.89)	1.38 (0.99; 1.93)
3	22 / 21	1.45 (0.73; 2.86)	1.41 (0.71; 2.81)
4	50 / 41	1.66 (0.98; 2.81)	1.76 (1.03; 3.00)
5	108 / 90	1.69 (1.08; 2.64)	1.64 (1.05; 2.59)

CI confidence interval, aOR adjusted odds ratio from conditional logistic regression

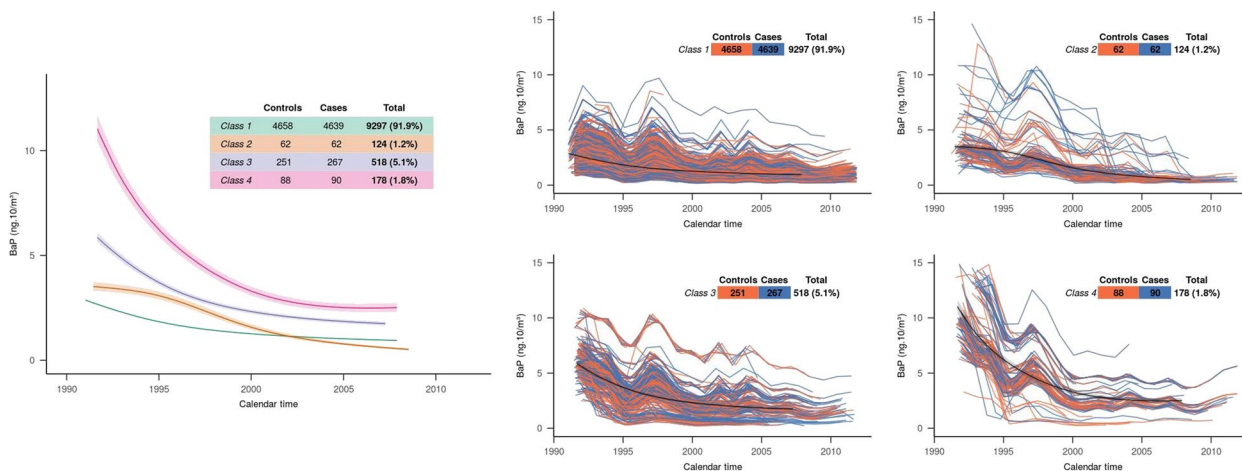
<sup>a</sup> OR adjusted for level of education, age at index date (in years), and matching factors

<sup>b</sup> OR further adjusted for known and potential risk factors of breast cancer: urban residence at birth (yes, no or missing), age at menarche (< 12, 12–13, ≥ 14 years), total physical activity (quartiles of METs-h/week), smoking status (never, current, former smoker), alcohol drinking (0, 1–6.6, ≥ 6.7 g/day or missing), body mass Index (< 25, 25–29, ≥ 30 kg/m<sup>2</sup>), parity and age at the first full-term pregnancy (no child/1–2 children and age < 30 years, 1–2 children and age ≥ 30 years, > 3 children), ever breastfeeding (yes/no), use of oral contraceptives (yes, no), use of menopausal hormone replacement therapy (yes, no or missing), previous history of personal benign breast disease (yes, no), all taken at cohort entry, except alcohol consumption and hormone replacement therapy use

also consistent with those of our previous study using the cumulative exposure estimates of PCB153 based on the same XENAIR case–control data [15]. Nevertheless, several other studies showed no evidence of the effect of exposure to PCB on breast cancer risk [32]. However, our

trajectory approach did not confirm results of previous studies supporting an association between breast cancer and BaP exposure from traffic [38, 39], nor the results of our previous study using the cumulative dose of BaP based on XENAIR case–control data [1]. The median pollutant concentrations over the whole follow-up were indeed more similar between the classes of BaP than between the classes of PCB153. Of note, women in the class exposed to highest levels of BaP concentrations over the study period had different characteristics; they had the highest physical activity level, less frequently a history of personal benign breast disease, and were more likely to having breastfed. However, adjustment for these factors in the sensitivity analysis did not change the results.

Our study has three major strengths. The first strength comes from its large number of breast cancer cases and the extensive data collected prospectively in the E3N cohort. The second major strength lies in its comprehensive exposure assessment approach, considering participants’ residential histories for a period up to 22 years prior to breast cancer diagnosis, accommodating residential mobility and spatio-temporal variations in pollutant emissions. This approach provides a more accurate representation of long-term exposure patterns and distinguishes this work from other studies that have relied on simpler exposure assessments, such as those that only consider current residential addresses. With the reconstruction of emissions data carried out by the reference institute in France (INERIS) and the use of a recognised chemistry-transport model [48], we are fortunate to have a dataset covering France and going back to 1990, which is unprecedented on a European scale. The third



**Fig. 2** Trajectories of BaP concentrations, XENAIR case–control study, France, 1990–2011 ( $n = 5\,058$  cases,  $5\,059$  controls). The left panel shows the estimated mean trajectory of the average annual BaP concentrations over the follow-up in the four classes identified by LCMM model. The right-hand panel shows the individual trajectories of subjects classified in each class, with the black bold line representing the estimated mean trajectory in the class with its 95% CI. The representation of each estimated mean trajectory is truncated at the 95th percentile of the distribution of observed exposure times in subjects a posteriori classified in the class

**Table 4** Characteristics of the participants according to their BaP exposure trajectory, in the XENAIR case–control study nested in the E3N cohort, France, 1990–2011 ( $n = 5\,058$  cases,  $5\,059$  controls)

Characteristics	Class 1 ( $n = 9,297$ , 4,639 cases, 4,658 controls)		Class 2 ( $n = 124$ , 62 cases, 62 controls)		Class 3 ( $n = 518$ , 267 cases, 251 controls)		Class 4 ( $n = 178$ , 90 cases, 88 controls)	
	Median (IQR) or $n$ (%)		Median (IQR) or $n$ (%)		Median (IQR) or $n$ (%)		Median (IQR) or $n$ (%)	
<b>Average annual BaP concentration over the follow-up (in ng.10/m<sup>3</sup>)</b>								
Number of annual concentrations	100,611		1859		5662		1910	
Median (IQR)	1.4 (0.9, 2.0)		1.6 (0.7, 2.7)		3.0 (1.9, 4.6)		4.4 (2.7, 7.1)	
< 1.2	40,137	(39.9)	715	(38.5)	490	(8.7)	64	(3.4)
[1.2- 10.0[	60,474	(60.1)	1130	(60.8)	5153	(91.0)	1672	(87.5)
≥ 10.0	0	(0.0)	14	(0.8)	19	(0.3)	174	(9.1)
<b>Age at cohort entry (years)</b>								
	48.8 (44.1, 54.1)		48.3 (44.0, 51.8)		48.0 (43.5, 53.2)		48.9 (44.1, 53.3)	
<b>Age at index date (years)</b>								
	60.7 (55.2, 66.4)		63.5 (60.1, 67.7)		60.2 (54.9, 65.2)		60.6 (54.9, 65.8)	
<b>Follow-up time (years)</b>								
	11.4 (6.9, 15.7)		15.3 (12.8, 17.9)		11.6 (7.9, 14.9)		11.2 (6.9, 14.8)	
<b>Urban residence at birth<sup>a</sup></b>								
	5717 (69.2)		78 (70.3)		360 (77.4)		132 (81.0)	
<b>Urban residence at cohort entry<sup>a</sup></b>								
	5546 (67.2)		86 (77.5)		321 (69.0)		132 (81.0)	
<b>Number of residential addresses changes during follow-up</b>								
Never	5487 (59.0)		4 (3.2)		302 (58.3)		121 (68.0)	
Once	3034 (32.6)		66 (53.2)		173 (33.4)		49 (27.5)	
Twice or more	776 (8.3)		54 (43.5)		43 (8.3)		8 (4.5)	
<b>Education</b>								
Secondary	1465 (15.8)		19 (15.3)		108 (20.8)		28 (15.7)	
1- to 2-year university degree	4537 (48.8)		66 (53.2)		253 (48.8)		92 (51.7)	
≥ 3 year university degree	3295 (35.4)		39 (31.5)		157 (30.3)		58 (32.6)	
<b>Body Mass Index (kg/m<sup>2</sup>) at cohort entry</b>								
< 25	7755 (83.4)		98 (79.0)		402 (77.6)		148 (83.1)	
25—29	1286 (13.8)		15 (12.1)		92 (17.8)		21 (11.8)	
≥ 30	256 (2.8)		11 (8.9)		24 (4.6)		9 (5.1)	
<b>Total physical activity (METs-h/week) at cohort entry</b>								
< 25.3	2276 (24.5)		24 (19.4)		121 (23.4)		29 (16.3)	
25.3 – 35.6	2457 (26.4)		43 (34.7)		114 (22.0)		47 (26.4)	
35.6 – 51.8	2381 (25.6)		31 (25.0)		135 (26.1)		46 (25.8)	
≥ 51.8	2183 (23.5)		26 (21.0)		148 (28.6)		56 (31.5)	
<b>Smoking status at cohort entry</b>								
Never smoker	5034 (54.1)		72 (58.1)		292 (56.4)		105 (59.0)	
Current smoker	1361 (14.6)		17 (13.7)		74 (14.3)		20 (11.2)	
Former smoker	2902 (31.2)		35 (28.2)		152 (29.3)		53 (29.8)	
<b>Alcohol drinking (g/day) at Q3<sup>a</sup></b>								
0	806 (11.8)		6 (6.2)		47 (12.4)		16 (12.9)	
1 – 6.6	2464 (36.1)		43 (44.3)		135 (35.6)		41 (33.1)	
≥ 6.7	3557 (52.1)		48 (49.5)		197 (52.0)		67 (54.0)	
<b>Age at menarche (years)</b>								
< 12	1926 (20.7)		31 (25.0)		92 (17.8)		29 (16.3)	

**Table 4** (continued)

Characteristics	Class 1 (n = 9,297, 4,639 cases, 4,658 controls)	Class 2 (n = 124, 62 cases, 62 controls)	Class 3 (n = 518, 267 cases, 251 controls)	Class 4 (n = 178, 90 cases, 88 controls)
	Median (IQR) or n (%)	Median (IQR) or n (%)	Median (IQR) or n (%)	Median (IQR) or n (%)
12–13	4832 (52.0)	57 (46.0)	265 (51.2)	93 (52.2)
≥ 14	2539 (27.3)	36 (29.0)	161 (31.1)	56 (31.5)
<b>Parity at cohort entry</b>				
No child	1097 (11.8)	15 (12.1)	59 (11.4)	21 (11.8)
1–2	5622 (60.5)	76 (61.3)	310 (59.8)	97 (54.5)
≥ 3	2578 (27.7)	33 (26.6)	149 (28.8)	60 (33.7)
<b>Age at first full-term pregnancy (years)</b>				
No child	1097 (11.8)	15 (12.1)	59 (11.4)	21 (11.8)
< 30	7098 (76.3)	94 (75.8)	401 (77.4)	135 (75.8)
≥ 30	1102 (11.9)	15 (12.1)	58 (11.2)	22 (12.4)
<b>Ever breastfeeding at cohort entry</b>				
No child	1097 (11.8)	15 (12.1)	59 (11.4)	21 (11.8)
No	3139 (33.8)	40 (32.3)	172 (33.2)	41 (23.0)
Yes	5061 (54.4)	69 (55.6)	287 (55.4)	116 (65.2)
<b>Use of oral contraceptives at cohort entry</b>				
	5471 (58.8)	83 (66.9)	310 (59.8)	99 (55.6)
<b>Menopausal status</b>				
Premenopausal at baseline	3731 (40.1)	43 (34.7)	190 (36.7)	63 (35.4)
Menopause during follow-up	4176 (44.9)	75 (60.5)	248 (47.9)	88 (49.4)
Postmenopausal at index date	1390 (15.0)	6 (4.8)	80 (15.4)	27 (15.2)
<b>Use of menopausal hormone replacement therapy up to the index date</b>				
	5407 (59.5)	81 (66.9)	322 (62.6)	111 (63.4)
<b>Ever mammography at cohort entry</b>				
	6969 (75.0)	90 (72.6)	371 (71.6)	149 (83.7)
<b>Previous history of personal benign breast disease at cohort entry</b>				
	2435 (26.2)	29 (23.4)	120 (23.2)	21 (11.8)
<b>Previous family history of breast cancer at cohort entry</b>				
	1284 (13.8)	18 (14.5)	74 (14.3)	21 (11.8)

IQR interquartile range

<sup>a</sup> percentage derived on available data

major strength of our study lies in its original trajectory-based analytical approach which has several advantages. This approach provides a comprehensive visual description of all available individual exposure history. It also allows similar exposure trajectory patterns to be grouped together, without making any assumptions about these groupings, and thus can distinguish for instance women who were exposed to high concentrations of PCB153 in the early 1990s only, from those who were exposed to high concentrations until the 2000s. In addition, among group-based approaches, the LCMM technique has the asset to account for correlation between the repeated exposure measures of each woman thanks to individual random effects, which is likely to provide much more robust results than group-based approaches that ignore

such correlation [22, 36]. Finally, the grouping allowed us to estimate the association between the profile of exposure trajectory and the risk of breast cancer, while fully accounting for the dynamic and the timing of exposures via the trajectory modeling, as opposed to standard statistical analyses relying on individual average exposure over time, or on individual standard cumulative dose of exposure.

Our results should yet be interpreted in light of certain limitations. The first major limitation is the lack of data on air pollution exposure prior to E3N cohort entry [55]. Because women entered the cohort at the age of 45–60 years with a median at nearly 50, this did not allow us to explore critical windows of susceptibility like puberty, pregnancy, and even menopause, which are yet

**Table 5** Association between classes of BaP trajectories and breast cancer, XENAIR case–control study nested in the E3N cohort, France, 1990–2011 ( $n = 5,058$  case–control pairs)

	n cases / controls	aOR <sup>a</sup> (95% CI)	aOR <sup>b</sup> (95% CI)
Class			
1	4639 / 4658	1	1
2	62 / 62	1.05 (0.73; 1.51)	1.05 (0.73, 1.52)
3	267 / 251	1.21 (0.91; 1.62)	1.16 (0.86, 1.56)
4	90 / 88	1.23 (0.73; 2.08)	1.24 (0.73, 2.10)

CI confidence interval, aOR adjusted odds ratio from conditional logistic regression

<sup>a</sup> OR adjusted for level of education, age at index date (in years), and matching factors

<sup>b</sup> OR further adjusted for known and potential risk factors of breast cancer: urban residence at birth (yes, no or missing), age at menarche (< 12, 12–13, ≥ 14 years), total physical activity (quartiles of METs-h/week), smoking status (never, current, former smoker), alcohol drinking (0, 1–6.7, ≥ 6.7 g/day or missing), body mass index (< 25, 25–29, ≥ 30 kg/m<sup>2</sup>), parity and age at the first full-term pregnancy (no child/1–2 children and age < 30 years, 1–2 children and age ≥ 30 years, > 3 children), ever breastfeeding (yes/no), use of oral contraceptives (yes, no), use of menopausal hormone replacement therapy (yes, no or missing), previous history of personal benign breast disease (yes, no), all taken at cohort entry, except alcohol consumption and hormone replacement therapy use

of major importance for understanding breast cancer etiology and latency [49]. Some of the women had menopause during the study period, but the sample size of most identified classes of exposure trajectories was too small to further distinguish them and draw any conclusion. However, we believe that the trajectory approach used in the present study suffers less from the lack of exposure data prior to cohort entry than an analytical approach based on a standard cumulative dose derived from the cohort entry. Indeed, such a cumulative dose fully depends on the length of follow-up, which varied from 1 to 22 years depending on the time at breast cancer diagnosis. The trajectory approach based on the LCMM accounted for these differences in the duration of follow-up, thanks to the linear mixed modeling approach. The mixed modelling approach also allows all available exposure data to be accounted for, without requiring any lagging or weighting procedure as when investigating the cumulative dose. Moreover, using calendar time as the time axis, which actually corresponded to the time since the first exposure assessment at cohort entry in 1990, allows a natural and interpretable description of exposure trajectories over time, as well as a fair comparison between cases and their individually matched controls. Indeed, exposure level highly depends on calendar year as shown in our results, making calendar time an appropriate time axis when regrouping exposure trajectories into latent classes of similar patterns. In addition, the

association between the identified classes of exposures and breast cancer was adjusted for age at the index date and follow-up duration by design, resulting in an adjustment for age at cohort entry in 1990, year of birth, and thus total duration of exposure to air pollution. The association was therefore estimated in cases and controls who were exposed to air pollution at the same ages in the same calendar years over lifetime. Future research could yet much leverage our study by utilizing exposure proxies for past periods for which direct exposure data are unavailable. We also experienced a second limitation due to the absence of data on workplace, indoor, and dietary exposure, which are crucial for a comprehensive exposure assessment. Residential exposures indeed represent only a part of the overall daily exposure. Indoor air pollution also contributes due to diverse indoor sources, and has been shown, via PAHs, to be potentially associated with breast cancer risk [23, 40, 51]. All of these sources should be accounted for in future studies. A third limitation lies in the exposure assessment. While the CHIMERE model is a valuable tool for estimating pollutant concentrations over long periods [12], it has limited spatial resolution (7 × 7 km in the present study), making it less effective in capturing variations in pollutant levels in urban and suburban areas where exposure can vary significantly [13] particularly for BaP for which traffic is an important source of emission [5]. This may have potentially resulted in exposure misclassification. However, even if the spatial resolution is coarse we may assume it as non-differential since all the external emission sources are included and moreover, approximately the same proportion of cases and controls (60%) lived in urban area at cohort entry. Yet, it would be important to confirm our results using more refined exposure assessment approaches. The fourth limitation is related to our analytical strategy. We indeed investigated each pollutant separately, without adjusting for other pollutants, nor considering any potential effect modification by other pollutants. Yet, atmospheric pollution is a complex mixture of substances with likely synergistic effects. Several statistical approaches exist to investigate multi-exposures [46]. However, most of them have not been adapted to longitudinal data. An extension of the trajectory-based approach that we used in the present study could actually be used to explore simultaneously the trajectories of several pollutants [43]. However, such modelling is complex and requires further work. Despite our effort to reduce the variability between individual trajectories within each class, we also observed important variability in terms of levels of exposure within each identified class, which might partly explain why most OR comparing each class of trajectories to the least exposed class did not reach statistical significance. We have to develop a

method that would give at least as much importance to the level than to the shape of trajectories. In addition, we used a two-stage analytical approach where we first identified the classes of trajectories, and then estimated their association with the risk of breast cancer. In the second stage, we neglected the uncertainty of the classification obtained in the first stage because the discrimination between the classes was very good (entropy above 0.94). However, this would be of interest to confirm our results using a joint modelling approach allowing the two steps to be performed simultaneously [43], when the method will be implemented for case–control data in the *lcmm* R package. Our analysis also assumed constant measures of association between the identified classes and the odds of breast cancer over the follow-up period. It would be important to investigate how the association might differ across calendar times, using for example interaction terms between the latent classes and time. The last potential limitation of our study lies in the specificity of the studied population. The E3N cohort included women ensured by a national health insurance plan that mainly covers teachers at all levels (primary, secondary, and higher education). This implies a higher level of education than the general population of women [11], and maybe a tendency towards healthier behaviors including diet. This may have potentially induced an underrepresentation of strongly exposed women and thus an underestimation of the association with breast cancer.

## Conclusions

This study provides a comprehensive description of the trajectories of residential outdoor exposure to PCB153 and BaP over a period of up to 22 years prior to breast cancer diagnosis. Our findings are consistent with a significantly increased risk of breast cancer among women who had high level of exposure to PCB153 between 1991 and 2011, compared to women less exposed during that period. There was no association for BaP, but the analysis should be replicated by incorporating all sources of dietary and indoor exposures, as well as a proxy for earlier air pollution exposure. From a methodological perspective, our study illustrates how a trajectory-based approach may be used to describe and investigate protracted environmental exposures varying over time, and also open avenues for new methodological developments.

### Abbreviations

BaP	Benzo[a]pyrene
CI	Confidence Interval
IQR	Interquartile Range
LCMM	Latent Class Mixed Models
OR	Odds Ratio

PAHs	Polycyclic aromatic hydrocarbons
PCBs	Polychlorinated biphenyls
SD	Standard Deviation

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12940-024-01106-x>.

Supplementary Material 1.

### Acknowledgements

The authors would like to thank all participants for providing data and physicians for providing pathology reports. They also thank Charlotte Carretero, Maxime Guillou and Matthieu Dubuis for geocoding addresses and Céline Kernaleguen, Maxime Valdenaire, Laureen Dartois, Roselyn Gomes-Rima, and Amandine Gelot for data management and assistance for cohort data collection.

### Authors' contributions

PD: main data analysis; tables and figures preparation; draft of the methods and results sections of the manuscript text. DP: funding acquisition; data collection; data curation; draft and substantial revision of the manuscript text. BLP: sensitivity data analysis; draft and substantial revision of the manuscript text and tables. HK: preliminary data analysis. FD, AA, LG: data curation. TC, EF, FC, GS, FRM, BF: funding acquisition; case–control study design; data collection. CPL: co-supervision of the main data analysis; creation of the software (*lcmm* R package). KL: funding acquisition; case–control study design; supervision of all data analysis; draft and substantial revision of the manuscript text. All authors reviewed the manuscript.

### Funding

This work was carried out in partnership with the ARC Foundation for Cancer Research (CANCAIR201601245). The E3N cohort receives financial support from the the National French Cancer League, the Mutuelle Générale de l'Éducation Nationale (MGEN), the Institut Gustave Roussy, the Institut National de la Santé et de la Recherche (INSERM). BLP were supported by the French National Agency for Food, Environmental and Occupational Health and Safety (ANSES-22-EST-195) and the National French Cancer League. AA was supported by the Savoie Regional Committee of the National French Cancer League. DP was supported by a post-doctoral fellowship from the National French Cancer League. TC was supported by a post-doctoral fellowship from the Fondation de France. The funding bodies had no role in the conception, design, planning, or writing of the study.

### Availability of data and materials

Data are available from the authors upon reasonable request.

### Declarations

#### Ethics approval and consent to participate

Our research is based on the existing French national cohort E3N. Each participant provided written informed consent, in compliance with the French National Commission for Computerized Data and Individual Freedom (CNIL).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Bordeaux Population Health Center (BPH Inserm U1219), Université de Bordeaux, 146 Rue Leo Saignat, Bordeaux 33000, France. <sup>2</sup>Department of Prevention Cancer Environment, Centre Léon Bérard, Lyon, France. <sup>3</sup>Inserm U1296 Radiations : Défense, Santé, Environnement, Lyon, France. <sup>4</sup>National Institute for Industrial Environment and Risks (INERIS), Verneuil-en-Halatte, France. <sup>5</sup>Centre de Recherche en Épidémiologie Et Santé Des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS UVSQ,



Gustave Roussy, Villejuif, France. <sup>6</sup>Department of Statistics, Computer Science and Applications (DISIA), University of Florence, Florence, Italy.

Received: 19 March 2024 Accepted: 24 July 2024  
Published online: 07 September 2024

## References

- Amadou A, Coudon T, Praud D, Salizzoni P, Leffondré K, Lévêque E, Boutron-Ruault MC, Danjou A, Morelli X, Le Cornet C, Perrier L, Couvidat F, Bessagnet B, Caudeville J, Faure E, Mancini FR, Gulliver J, Severi G, Fervers B. Chronic low-dose exposure to xenoestrogen ambient air pollutants and breast cancer risk: XENAIR Protocol. *JMIR Research Protocols* 2020; 9(9):e15167. <https://doi.org/10.2196/15167>.
- Amadou A, Praud D, Coudon T, Danjou AMN, Faure E, Leffondré K, Le Romancer M, Severi G, Salizzoni P, Mancini FR, Fervers B. Chronic long-term exposure to cadmium air pollution and breast cancer risk in the French E3N cohort. *Int J Cancer*. 2020;146:341–51. <https://doi.org/10.1002/ijc.32257>.
- Amadou A, Praud D, Coudon T, Deygas F, Grassot L, Faure E, Couvidat F, Caudeville J, Bessagnet B, Salizzoni P, Gulliver J, Leffondré K, Severi G, Mancini FR, Fervers B. Risk of breast cancer associated with long-term exposure to benzo[a]pyrene (BaP) air pollution: evidence from the French E3N cohort study. *Environ Int*. 2021;149: 106399. <https://doi.org/10.1016/j.envint.2021.106399>.
- Arnich N, Tard A, Leblanc J-C, Le Bizet B, Narbonne J-F, Maximilien R. Dietary intake of non-dioxin-like PCBs (NDL-PCBs) in France, impact of maximum levels in some foodstuffs. *Regul Toxicol Pharmacol* 2009;54:287–93. <https://doi.org/10.1016/j.yrtph.2009.05.010>.
- Baan, R., Grosse, Y., Straif, K., Secretan, B., El Ghissassi, F., Bouvard, V., Benbrahim-Tallaa, L., Guha, N., Freeman, C., Galichet, L., Coglian, V., WHO International Agency for Research on Cancer Monograph Working Group. A review of human carcinogens—Part F: chemical agents and related occupations. *Lancet Oncol*. 2009;10:1143–4. [https://doi.org/10.1016/S1470-2045\(09\)70358-4](https://doi.org/10.1016/S1470-2045(09)70358-4).
- Bachelet D, Truong T, Verner M-A, Arveux P, Kerbrat P, Charlier C, Guihenneuc-Jouyau C, Guénel P. Determinants of serum concentrations of 1,1-dichloro-2,2-bis(p-chlorophenyl)ethylene and polychlorinated biphenyls among French women in the CECILE study. *Environ Res*. 2011;111:861–70. <https://doi.org/10.1016/j.envres.2011.06.001>.
- Beyea J, Hatch M, Stellman SD, Santella RM, Teitelbaum SL, Prokopczyk B, Camann D, Gammon MD. Validation and calibration of a model used to reconstruct historical exposure to polycyclic aromatic hydrocarbons for use in epidemiologic studies. *Environ Health Perspect*. 2006;114:1053–8. <https://doi.org/10.1289/ehp.8659>.
- Biernacki C, Celeux G, Govaert G. Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Trans Pattern Anal Mach Intell*. 2000;22:719–25. <https://doi.org/10.1109/34.865189>.
- Binachon B, Dossus L, Danjou AMN, Clavel-Chapelon F, Fervers B. Life in urban areas and breast cancer risk in the French E3N cohort. *Eur J Epidemiol*. 2014;29:743–51. <https://doi.org/10.1007/S10654-014-9942-Z>.
- Brody JG, Moysich KB, Humblet O, Attfield KR, Beehler GP, Rudel RA. Environmental pollutants and breast cancer: epidemiologic studies. *Cancer*. 2007;109:2667–711. <https://doi.org/10.1002/cncr.22655>.
- Clavel-Chapelon F, E3N Study Group. Cohort Profile: The French E3N Cohort Study. *Int J Epidemiol*. 2015;44:801–9. <https://doi.org/10.1093/ije/dyu184>.
- Couvidat F, Bessagnet B, Garcia-Vivanco M, Real E, Menut L, Colette A. Development of an inorganic and organic aerosol model (CHIMERE 2017 β v1.0): seasonal and spatial evaluation over Europe. *Geosci Model Dev*. 2018;11:165–94. <https://doi.org/10.5194/gmd-11-165-2018>.
- Cyrys J, Eeftens M, Heinrich J, Ampe C, Armengaud A, Beelen R, Bellander T, Beregszaszi T, Birk M, Cesaroni G, Cirach M, de Hoogh K, De Nazelle A, de Vocht F, Declercq C, Dedelè A, Dimakopoulou K, Eriksen K, Galassi C, Grąlewičienė R, Grivas G, Gruzieva O, Gustafsson AH, Hoffmann B, Iakovides M, Ineichen A, Krämer U, Lanki T, Lozano P, Madsen C, Meliefste K, Modig L, Mølter A, Mosler G, Nieuwenhuijsen M, Nonnemacher M, Oldenwening M, Peters A, Pontet S, Probst-Hensch N, Quass U, Raaschou-Nielsen O, Ranzi A, Sugiri D, Stephanou EG, Taimisto J, Tsai M-Y, Vaskövi É, Villani S, Wang M, Brunekreef B, Hoek G. Variation of NO<sub>2</sub> and NO<sub>x</sub> concentrations between and within 36 European study areas: Results from the ESCAPE study. *Atmos Environ*. 2012;62:374–90. <https://doi.org/10.1016/j.atmosenv.2012.07.080>.
- Demers A, Ayotte P, Brisson J, Dodin S, Robert J, Dewailly E. Risk and aggressiveness of breast cancer in relation to plasma organochlorine concentrations. *Cancer Epidemiol Biomarkers Prev*. 2000;9:161–6.
- Deygas F, Amadou A, Coudon T, Grassot L, Couvidat F, Bessagnet B, Faure E, Salizzoni P, Gulliver J, Caudeville J, Severi G, Mancini FR, Leffondré K, Fervers B, Praud D. Long-term atmospheric exposure to PCB153 and breast cancer risk in a case-control study nested in the French E3N cohort from 1990 to 2011. *Environ Res*. 2021;195:110743–110743. <https://doi.org/10.1016/j.envres.2021.110743>.
- Donat-Vargas C, Guerrero-Zotano Á, Casas A, Baena-Cañada JM, Lope V, Antolín S, García-Saénz JÁ, Bermejo B, Muñoz M, Ramos M, de Juan A, Jara Sánchez C, Sánchez-Rovira P, Antón A, Brunet J, Gavilá J, Salvador J, Arriola Arellano E, Bezares S, Fernández de Larrea-Baz N, Pérez-Gómez B, Martín M, Pollán M. Trajectories of alcohol consumption during life and the risk of developing breast cancer. *Br J Cancer*. 2021;125:1168–76. <https://doi.org/10.1038/s41416-021-01492-w>.
- Elliott MR, Zhao Z, Mukherjee B, Kanaya A, Needham BL. Methods to account for uncertainty in latent class assignments when using latent classes as predictors in regression models, with application to acculturation strategy measures. *Epidemiol Camb Mass*. 2020;31:194–204. <https://doi.org/10.1097/EDE.0000000000001139>.
- Famiyeh L, Chen K, Xu J, Sun Y, Guo Q, Wang C, Lv J, Tang Y-T, Yu H, Snape C, He J. A review on analysis methods, source identification, and cancer risk evaluation of atmospheric polycyclic aromatic hydrocarbons. *Sci Total Environ*. 2021;789: 147741. <https://doi.org/10.1016/j.scitotenv.2021.147741>.
- Faure E, Danjou AMN, Clavel-Chapelon F, Boutron-Ruault M-C, Dossus L, Fervers B. Accuracy of two geocoding methods for geographic information system-based exposure assessment in epidemiological studies. *Environ Health*. 2017;16: 15. <https://doi.org/10.1186/s12940-017-0217-5>.
- Fenton SE, Birnbaum LS. Timing of environmental exposures as a critical element in breast cancer risk. *J Clin Endocrinol Metab*. 2015;100:3245–50. <https://doi.org/10.1210/jc.2015-2848>.
- Ferrante MC, Amero P, Santoro A, Monnolo A, Simeoli R, Di Guida F, Mattace Raso G, Meli R. Polychlorinated biphenyls (PCB 101, PCB 153 and PCB 180) alter leptin signaling and lipid metabolism in differentiated 3T3-L1 adipocytes. *Toxicol Appl Pharmacol*. 2014;279:401–8. <https://doi.org/10.1016/j.taap.2014.06.016>.
- Fitzmaurice GM, Laird NM, Ware JH. *Applied longitudinal analysis*, 2nd edition, Wiley Series in Probability and Statistics. New Jersey: John Wiley & Sons, Inc.; 2011. p.740.
- Gamboa-Loira B, López-Carrillo L, Mar-Sánchez Y, Stern D, Cebrián ME. Epidemiologic evidence of exposure to polycyclic aromatic hydrocarbons and breast cancer: a systematic review and meta-analysis. *Chemosphere*. 2022;290: 133237. <https://doi.org/10.1016/j.chemosphere.2021.133237>.
- García-Acosta S, Clavel-Chapelon F. Dealing with missing, abnormal and incoherent data in E3N cohort study. *Epidemiol Public Health Rev Epidemiologie Santé Publique*. 1999;47:515–23.
- Gianelle V, Colombi C, Caserini S, Ozgen S, Galante S, Marongiu A, Lanzani G. Benzo(a)pyrene air concentrations and emission inventory in Lombardy region, Italy. *Atmospheric Pollut Res*. 2013;4:257–66. <https://doi.org/10.5094/APR.2013.028>.
- Guerreiro CBB, Horálek J, de Leeuw F, Couvidat F. Benzo(a)pyrene in Europe: ambient air concentrations, population exposure and health effects. *Environ Pollut*. 2016;214:657–67. <https://doi.org/10.1016/j.envpol.2016.04.081>.
- Hipp JR, Bauer DJ. Local solutions in the estimation of growth mixture models. *Psychol Methods*. 2006;11:36–53. <https://doi.org/10.1037/1082-989X.11.1.36>.
- Huang W, He Y, Xiao J, Huang Y, Li A, He M, Wu K. Risk of breast cancer and adipose tissue concentrations of polychlorinated biphenyls and organochlorine pesticides: a hospital-based case-control study in Chinese women. *Environ Sci Pollut Res Int*. 2019;26:32128–36. <https://doi.org/10.1007/s11356-019-06404-3>.
- IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. *Polychlorinated Biphenyls and Polybrominated Biphenyls*. Lyon (FR): International Agency for Research on Cancer; 2016. (IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, No. 107). p. 510.

30. Itoh H, Iwasaki M, Hanaoka T, Kasuga Y, Yokoyama S, Onuma H, Nishimura H, Kusama R, Tsugane S. Serum organochlorines and breast cancer risk in Japanese women: a case-control study. *Cancer Causes Control CCC*. 2009;20:567–80. <https://doi.org/10.1007/s10552-008-9265-z>.
31. Lauby-Secretan B, Loomis D, Grosse Y, El GF, Bouvard V, Benbrahim-Tallaa L, Guha N, Baan R, Mattock H, Straif K. Carcinogenicity of polychlorinated biphenyls and polybrominated biphenyls. *Lancet Oncol*. 2013;14:287–8. [https://doi.org/10.1016/S1470-2045\(13\)70104-9](https://doi.org/10.1016/S1470-2045(13)70104-9).
32. Leng L, Li J, Luo X-M, Kim J-Y, Li Y-M, Guo X-M, Chen X, Yang Q-Y, Li G, Tang N-J. Polychlorinated biphenyls and breast cancer: a congener-specific meta-analysis. *Environ Int*. 2016;88:133–41. <https://doi.org/10.1016/j.envint.2015.12.022>.
33. Lévêque E, Lacourt A, Philipps V, Luce D, Guénel P, Stücker I, Proust-Lima C, Leffondré K. A new trajectory approach for investigating the association between an environmental or occupational exposure over lifetime and the risk of chronic disease: Application to smoking, asbestos, and lung cancer. *PLoS ONE*. 2020;15: e0236736. <https://doi.org/10.1371/journal.pone.0236736>.
34. Liu H, Sun Y, Ran L, Li J, Shi Y, Mu C, Hao C. Endocrine-disrupting chemicals and breast cancer: a meta-analysis. *Front Oncol*. 2023;13: 1282651. <https://doi.org/10.3389/fonc.2023.1282651>.
35. Loomis D, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, Guha N, Baan R, Mattock H, Straif K, International Agency for Research on Cancer Monograph Working Group IARC. The carcinogenicity of outdoor air pollution. *Lancet Oncol*. 2013;14:1262–3.
36. Mésidor M, Rousseau M-C, O'Loughlin J, Sylvestre M-P. Does group-based trajectory modeling estimate spurious trajectories? *BMC Med Res Methodol*. 2022;22:194. <https://doi.org/10.1186/s12874-022-01622-9>.
37. Millikan R, DeVoto E, Duell EJ, Tse CK, Savitz DA, Beach J, Edmiston S, Jackson S, Newman B. Dichlorodiphenyldichloroethene, polychlorinated biphenyls, and breast cancer among African-American and white women in North Carolina. *Cancer Epidemiol Biomarkers Prev*. 2000;9:1233–40.
38. Mordukhovich I, Beyea J, Herring AH, Hatch M, Stellman SD, Teitelbaum SL, Richardson DB, Millikan RC, Engel LS, Shantakumar S, Steck SE, Neugut AI, Rossner P, Santella RM, Gammon MD. Vehicular traffic-related polycyclic aromatic hydrocarbon exposure and breast cancer incidence: the long Island Breast Cancer Study Project (LIBCSP). *Environ Health Perspect*. 2016;124:30–8. <https://doi.org/10.1289/ehp.1307736>.
39. Nie J, Beyea J, Bonner MR, Han D, Vena JE, Rogerson P, Vito D, Muti P, Trevisan M, Edge SB, Freudenheim JL. Exposure to traffic emissions throughout life and risk of breast cancer: the Western New York Exposures and Breast Cancer (WEB) study. *Cancer Causes Control CCC*. 2007;18:947–55. <https://doi.org/10.1007/s10552-007-9036-2>.
40. Parikh PV, Wei Y. PAHs and PM2.5 emissions and female breast cancer incidence in metro Atlanta and rural Georgia. *Int J Environ Health Res*. 2016;26:458–66. <https://doi.org/10.1080/09603123.2016.1161178>.
41. Patel AB, Shaikh S, Jain KR, Desai C, Madamwar D. Polycyclic aromatic hydrocarbons: sources, toxicity, and remediation approaches. *Front Microbiol*. 2020;11:562813.
42. Proust-Lima C, Philipps V, Liquet B. Estimation of extended mixed models using latent classes and latent processes: the R package lcmm. *J Stat Softw*. 2017;78:1–56. <https://doi.org/10.18637/jss.v078.i02>.
43. Proust-Lima C, Saulnier T, Philipps V, Traon AP-L, Péran P, Rascol O, Meissner WG, Foubert-Samier A. Describing complex disease progression using joint latent class models for multivariate longitudinal markers and clinical endpoints. *Stat Med*. 2023;42:3996–4014. <https://doi.org/10.1002/sim.9844>.
44. Recio-Vega R, Velazco-Rodríguez V, Ocampo-Gómez G, Hernandez-Gonzalez S, Ruiz-Flores P, Lopez-Marquez F. Serum levels of polychlorinated biphenyls in Mexican women and breast cancer risk. *J Appl Toxicol JAT*. 2011;31:270–8. <https://doi.org/10.1002/jat.1672>.
45. Rodgers KM, Udesky JO, Rudel RA, Brody JG. Environmental chemicals and breast cancer: An updated review of epidemiological literature informed by biological mechanisms. *Environ Res*. 2018;160:152–82. <https://doi.org/10.1016/j.envres.2017.08.045>.
46. Stafoggia M, Breitner S, Hampel R, Basagaña X. Statistical approaches to address multi-pollutant mixtures and multiple exposures: the state of the science. *Curr Env Health Rep*. 2017;4:481–90. <https://doi.org/10.1007/s40572-017-0162-z>.
47. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71:209–49. <https://doi.org/10.3322/caac.21660>.
48. Terrenoire E, Bessagnet B, Rouil L, Tognet F, Pirovano G, Létinois L, Beauchamp M, Colette A, Thunis P, Amann M, Menut L. High-resolution air quality simulation over Europe with the chemistry transport model CHIMERE. *Geosci Model Dev*. 2015;8:21–42. <https://doi.org/10.5194/gmd-8-21-2015>.
49. Terry MB, Michels KB, Brody JG, Byrne C, Chen S, Jerry DJ, Malecki KMC, Martin MB, Miller RL, Neuhausen SL, Silk K, Trentham-Dietz A. Breast Cancer and the Environment Research Program (BCERP). Environmental exposures during windows of susceptibility for breast cancer: a framework for prevention research. *Breast Cancer Res BCR*. 2019;21:96. <https://doi.org/10.1186/s13058-019-1168-2>.
50. White AJ, Bradshaw PT, Hamra GB. Air pollution and Breast Cancer: A Review. *Curr Epidemiol Rep*. 2018;5:92–100. <https://doi.org/10.1007/s40471-018-0143-2>.
51. White AJ, Chen J, Teitelbaum SL, McCullough LE, Xu X, Hee Cho Y, Conway K, Beyea J, Stellman SD, Steck SE, Mordukhovich I, Eng SM, Beth Terry M, Engel LS, Hatch M, Neugut AI, Hibshoosh H, Santella RM, Gammon MD. Sources of polycyclic aromatic hydrocarbons are associated with gene-specific promoter methylation in women with breast cancer. *Environ Res*. 2016;145:93–100. <https://doi.org/10.1016/j.envres.2015.11.033>.
52. Wielsøe M, Kern P, Bonefeld-Jørgensen EC. Serum levels of environmental pollutants is a risk factor for breast cancer in Inuit: a case control study. *Environ Health Glob Access Sci Source*. 2017;16:56. <https://doi.org/10.1186/s12940-017-0269-6>.
53. Wolff MS, Britton JA, Teitelbaum SL, Eng S, Deych E, Ireland K, Liu Z, Neugut AI, Santella RM, Gammon MD. Improving organochlorine biomarker models for cancer research. *Cancer Epidemiol Biomarkers Prev*. 2005;14:2224–36. <https://doi.org/10.1158/1055-9965.EPI-05-0173>.
54. World Health Organization, Food and Agriculture Organization of the United Nations & Joint FAO/WHO Expert Committee on Food Additives. Meeting ( 80th: 2015, Rome, Italy) . ( 2016) . Evaluation of certain food additives and contaminants: eightieth report of the Joint FAO/WHO Expert Committee on Food Additives. 109 p. World Health Organization. <https://iris.who.int/handle/10665/204410>.
55. Zhai Y, Amadou A, Mercier C, Praud D, Faure E, Iwaz J, Severi G, Mancini FR, Coudon T, Fervers B, Roy P. The impact of left truncation of exposure in environmental case-control studies: evidence from breast cancer risk associated with airborne dioxin. *Eur J Epidemiol*. 2021. <https://doi.org/10.1007/s10654-021-00776-y>.
56. Zhang J, Huang Y, Wang X, Lin K, Wu K. Environmental polychlorinated biphenyl exposure and breast cancer risk: a meta-analysis of observational studies. *PLoS ONE*. 2015;10: e0142513. <https://doi.org/10.1371/journal.pone.0142513>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.