Environment International 123 (2019) 189–200

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

The early-life exposome: Description and patterns in six European countries

Ibon Tamayo-Uria[a,b,c,d], Léa Maitre[a,b,c], Cathrine Thomsen[e], Mark J. Nieuwenhuijsen[a,b,c], Leda Chatzi[f], Valérie Siroux[g], Gunn Marit Aasvang[e], Lydiane Agier[g], Sandra Andrusaityte[h], Maribel Casas[a,b,c], Montserrat de Castro[a,b,c], Audrius Dedele[h], Line S. Haug[e], Barbara Heude[i,j], Regina Grazuleviciene[h], Kristine B. Gutzkow[e], Norun H. Krog[e], Dan Mason[k], Rosemary R.C. McEachan[k], Helle M. Meltzer[e], Inga Petraviciene[h], Oliver Robinson[a,b,c,l], Theano Roumeliotaki[m], Amrit K. Sakhi[e], Jose Urquiza[a,b,c], Marina Vafeiadi[m], Dagmar Waiblinger[k], Charline Warembourg[a,b,c], John Wright[k], Rémy Slama[g], Martine Vrijheid[a,b,c,1], Xavier Basagaña[a,b,c,*,1]

[a] ISGlobal, Barcelona, Spain
[b] Universitat Pompeu Fabra (UPF), Barcelona, Spain
[c] CIBER Epidemiología y Salud Pública (CIBERESP), Spain
[d] Department of Statistics, Faculty of Arts and Sciences, Harvard University, Cambridge, MA, USA
[e] Norwegian Institute of Public Health, Oslo, Norway
[f] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, USA
[g] Team of Environmental Epidemiology, IAB, Institute for Advanced Biosciences, Inserm, CNRS, CHU-Grenoble-Alpes, University Grenoble-Alpes, Grenoble, France
[h] Department of Environmental Sciences, Vytautas Magnus University, Kaunas, Lithuania
[i] INSERM, UMR1153 Epidemiology and Biostatistics Sorbonne Paris Cité Center (CRESS), Team 'Early Origin of the child's Health and Development' (ORCHAD), Villejuif, France
[j] Paris Descartes University, Paris, France
[k] Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK
[l] MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK
[m] Department of Social Medicine, Faculty of Medicine, University of Crete, Heraklion, Crete, Greece

ARTICLE INFO

Handling Editor: Lesa Aylward

Keywords:
Exposome
Environmental exposures
Early life
Pregnancy
Children

ABSTRACT

Characterization of the "exposome", the set of all environmental factors that one is exposed to from conception onwards, has been advocated to better understand the role of environmental factors on chronic diseases.

Here, we aimed to describe the early-life exposome. Specifically, we focused on the correlations between multiple environmental exposures, their patterns and their variability across European regions and across time (pregnancy and childhood periods). We relied on the Human Early-Life Exposome (HELIX) project, in which 87 environmental exposures during pregnancy and 122 during the childhood period (grouped in 19 exposure groups) were assessed in 1301 pregnant mothers and their children at 6–11 years in 6 European birth cohorts.

Some correlations between exposures in the same exposure group reached high values above 0.8. The median correlation within exposure groups was > 0.3 for many exposure groups, reaching 0.69 for water disinfection by products in pregnancy and 0.67 for the meteorological group in childhood. Median correlations between different exposure groups rarely reached 0.3. Some correlations were driven by cohort-level associations (e.g. air pollution and chemicals). Ten principal components explained 45% and 39% of the total variance in the pregnancy and childhood exposome, respectively, while 65 and 90 components were required to explain 95% of the exposome variability. Correlations between maternal (pregnancy) and childhood exposures were high (> 0.6) for most exposures modeled at the residential address (e.g. air pollution), but were much lower and even close to zero for some chemical exposures.

In conclusion, the early life exposome was high dimensional, meaning that it cannot easily be measured by or reduced to fewer components. Correlations between exposures from different exposure groups were much lower than within exposure groups, which have important implications for co-exposure confounding in multiple

* Corresponding author at: ISGlobal, Institute for Global Health, C/Doctor Aiguader 88, 08003 Barcelona, Spain.
E-mail address: xavier.basagana@isglobal.org (X. Basagaña).
[1] Equally contributing authors.

https://doi.org/10.1016/j.envint.2018.11.067
Received 23 July 2018; Received in revised form 26 November 2018; Accepted 26 November 2018
0160-4120/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

exposure studies. Also, we observed the early life exposome to be variable over time and to vary by cohort, so measurements at one time point or one place will not capture its complexities.

1. Introduction

A complex interplay between genetic and environmental factors is assumed to contribute to the development of chronic diseases in humans (Willett, 2002). The World Health Organization (WHO) has attributed nearly half of global mortality to a handful of environmental exposures, illustrating the relevance of environmental factors for health (Lim et al., 2012). To date, most research on environmental determinants of disease has focused on single exposures, with the exceptions of a few studies considering simultaneously more than a couple of families of exposures, focusing on health outcomes such as birth weight (Dadvand et al., 2014; Hystad et al., 2014; Lenters et al., 2015) or type II diabetes mellitus (Patel et al., 2010). A more comprehensive approach, studying all environmental exposures that can influence health from conception onward, defined as the "exposome", has been advocated to better understand the role of environmental factors on multifactorial and chronic pathologies (Wild, 2005). Early life is a key starting point for the development of a lifetime exposome, particularly due to heightened vulnerability and potentially lifelong impact of exposure during this period (Vrijheid et al., 2014). In utero and childhood exposures, including environmental contaminant exposures (Vrijheid et al., 2016), but also factors in the urban and built environment (Gascon et al., 2016), can permanently change the body's structure, physiology, and metabolism, predisposing individuals to the development of chronic pathologies later in life, a hypothesis based on the Developmental Origins of Health and Disease (DOHaD) paradigm (Heindel et al., 2015).

The exposome potentially consists of hundreds of exposures, many of which may be inter-related. An important challenge in studies associating the exposome with health is the simultaneous consideration of these many correlated exposures (Slama and Vrijheid, 2015). Understanding what a typical exposome looks like, including the structure of correlations between and within groups of exposures, is important for multiple steps in exposome research, including the planning of which exposures to measure and the development of statistical analysis protocols. Moreover, transparent knowledge of the correlation structure of an exposome data set is required for interpretation of associations (Patel et al., 2015; Robinson and Vrijheid, 2015). Additionally, the

Table 1
Overview of the exposure variables included in the pregnancy and childhood exposome.

Exposure group	Individual exposure variables	Number of variables	
		Pregnancy	Childhood
Atmospheric pollutants	NO_2, $PM_{2.5}$, PM_{10}, PM_{abs} at home address. Averaged over pregnancy and year before visit in childhood	4	4
UV	Ambient ultraviolet radiation (UVR) levels at home address. Averaged over month before visit during childhood. Not available during pregnancy.	0	1
Surrounding natural spaces	Average Normalized Difference Vegetation Index (NDVI) within a 100 m buffer; presence of a major green space in a distance of 300 m; presence of a major blue space in a distance of 300 m. Home and school address during childhood.	3	6
Meteorology	Temperature, humidity, pressure at home address. Pressure not available during childhood. Averaged over pregnancy and month before visit during childhood.	3	2
Built environment	Population density, building density, street connectivity, access to bus public transport (lines and stops), facility richness (pregnancy only) and density, Land Use Evenness Index and walkability index in a 300 m buffer. Home and school address during childhood (walkability only at home address).	9	15
Traffic	Total traffic load of roads in a 100 m buffer (pregnancy and home), total traffic load of major roads in a 100 m buffer (home and school), traffic density on nearest road (pregnancy and home) and inverse distance to nearest road (pregnancy and home).	3	5
Road traffic noise	24-hour road noise levels (pregnancy, and home and school during childhood). Night time noise levels for home during childhood.	1	3
OCs	Blood concentrations of DDE, DDT, HCB, PCB (118, 138, 153, 170, 180)	8	8
PBDE	Blood concentrations of PBDE47, PBDE153	2	2
PFASs	Blood concentrations of PFOA, PFNA, PFUNDA, PFHxS, PFOS	5	5
Metals and elements	Blood concentrations of As, Cd, Co, Cs, Cu, Hg, Mn, Mo, Pb, Tl	10	10
Phthalate metabolites	Urine concentrations of MEP, MiBP, MnBP, MBzP, MEHP, MEHHP, MEOHP, MECPP, OHMiNP, OXOMiNP	10	10
Phenols	Urine concentrations of MEPA, ETPA, BPA, PRPA, BUPA, OXBE, TRCS	7	7
OP pesticide metabolites	Urine concentrations of DMP, DMTP, DMDTP (childhood only), DEP, DETP	4	5
Tobacco smoking	Urine concentration of cotinine (pregnancy and childhood), active/passive smoking during pregnancy, number of cigarettes during pregnancy, parental smoking and passive smoking during childhood.	3	3
Water DBPs	THM, chloroform, brominated THMs	3	0
Indoor air pollution	NO_2, TEX, Benzene, $PM_{2.5}$, PM_{abs}	0	5
Lifestyle	Diet, physical activity, breastfeeding duration, alcohol consumption, sleep duration, pets	12	27
Socio-economic capital	Family affluence score, social contact, social participation, house crowding	0	4
		87	122

Abbreviations used: **Home**: childhood home address; **School**: childhood school address; **NO_2**, nitrogen dioxide; **$PM_{2.5}$**, particulate matter with an aerodynamic diameter of < 2.5 μm; **PM_{10}**, particulate matter with an aerodynamic diameter of < 10 μm; **PM_{abs}**, absorbance of $PM_{2.5}$filters; **UV**, ultraviolet; **NDVI**: Normalized Difference Vegetation Index; **DDE**, 4,4′dichlorodiphenyl dichloroethylene; **DDT**, 4,4′dichlorodiphenyltrichloroethane; **HCB**, hexachlorobenzene; **PCB**, polychlorinated biphenyl; **PBDE**, polybrominated diphenyl ether; **PFOA**, perfluorooctanoate; **PFNA**, perfluorononanoate; **PFUNDA**, perfluoroundecanoate; **PHFxS**, perfluorohexane sulfonate; **PFOS**, perfluorooctane sulfonate; **As**, arsenic; **Cd**, cadmium; **Co**, cobalt; **Cs**, caesium; **Cu**, copper; **Hg**, mercury; **Mn**, manganese; **Mo**, molybdenum; **Pb**, lead; **Tl**, thallium; **MEP**, monoethyl phthalate; **MiBP**, mono-isobutyl phthalate; **MnBP**, mono-*n*-butyl phthalate; **MBzP**, mono benzyl phthalate; **MEHP**, mono-2-ethylhexyl phthalate; **MEHHP**, mono-2-ethyl-5-hydroxyhexyl phthalate; **MEOHP**, mono-2-ethyl-5-oxohexyl phthalate; **MECPP**, mono-2-ethyl 5-carboxypentyl phthalate; **OHMiNP**, mono-4-methyl-7-hydroxyoctyl phthalate; **OXOMiNP**, mono-4-methyl-7-oxooctyl phthalate; **MEPA**, methyl paraben; **ETPA**, ethyl paraben; **BPA**, bisphenol A; **PRPA**, propyl paraben; **BUPA**, *N*-butyl paraben; **OXBE**, oxybenzone; **TRCS**, triclosan; **DMP**, dimethyl phosphate; **DMTP**, dimethyl thiophosphate; **DMDTP**, dimethyl dithiophosphate; **DEP**, diethyl phosphate; **DETP**, diethyl thiophosphate; **THMs**, trihalomethanes; **TEX**, toluene-ethylbenzene-xylene.

description of multiple exposure patterns can aid the identification of population groups at highest environmental health risk.

A few studies have recently described correlation structures of a large set of exposures, including the NHANES dataset (Patel and Ioannidis, 2014; Patel and Manrai, 2015), the Spanish INMA birth cohort (Robinson et al., 2015), and the LIFE study (Chung et al., 2018). However, the available information is still scarce. Wider studies, including a broader range of measured exposures, multiple lifetime periods of exposure assessment, and larger populations covering multiple regions are needed to more fully understand the complexity of the human exposome.

Here, we aim to describe the early-life exposome using data from the Human Early-Life Exposome (HELIX) project, in which > 200 environmental exposures of concern for child health were assessed in 1301 pregnant women and their children at 6–11 years in 6 European birth cohorts. Specifically, we focus on the description and analyses of correlations between multiple environmental exposures, their patterns and their variability across European regions and across time (pregnancy and childhood periods).

2. Material and methods

2.1. The study population

This study is part of the HELIX project, which aims to characterize the exposome during early-life and evaluate its relationship to molecular signatures and child health outcomes (Vrijheid et al., 2014). HELIX is based on six European birth cohort studies: BiB (Born in Bradford), United Kingdom (Wright et al., 2013); EDEN (Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant), France (Heude et al., 2016); INMA (INfancia y Medio Ambiente), Spain (Guxens et al., 2012); KANC (Kaunus Cohort), Lithuania (Grazuleviciene et al., 2009); MoBa (Norwegian Mother and Child Cohort Study), Norway (Magnus et al., 2016); and Rhea, Greece (Chatzi et al., 2017). In this paper we used data from the HELIX subcohort of 1301 mother-child pairs (MoBa (272), KANC (204), BIB (205), EDEN (198), INMA (223) and RHEA (199)) with information on environmental exposures during pregnancy and between 6 and 11 years of age (median 8.1 years; hereafter referred to as “childhood period”).

Follow-up visits of the children in the subcohort took place in 2014 and 2015 in the 6 study centers, and questionnaire information and biological samples for biomarker determination were collected. More details about the subcohort, including recruitment and data collection methods, are available in Maitre et al. (2018). Approval was obtained from the national ethics committees for every cohort. All participating women provided informed, written consent.

2.2. Exposome variables

In the current analysis we included 87 environmental exposure variables for the pregnancy period and 122 for the childhood period (Table 1). Exposure assessment methods are described in detail for each exposure group in Annex 1. More exposure variables were available in the project (see Annex 1), but were not included in the current analysis for the following reasons: they had < 30 subjects in one exposure category without possibility to recode; they had a correlation of 0.9 or higher with another similar variable that was likely measuring the same thing (Annex 1, pg 3) (e.g., buffers for spatial variables, home and school estimates for meteorological variables, erythemal UV, Vitamin-D and DNA damaging UV variables, facility richness and density and dichotomous and location-specific ETS in childhood); or they were calculated for several exposure windows and only the longest exposure window was included (e.g. pregnancy average instead of trimester averages). Exposure levels and distributions for these variables are shown in Annex 2.

In brief, exposure to outdoor factors (atmospheric pollutants, ultra violet radiation – UV-, surrounding natural spaces, meteorology, built environment, traffic, and road traffic noise) was estimated using geospatial models, monitoring stations, satellite data and land use databases, and was assigned to study participants according to their geocoded home and school addresses using GIS platforms (Annex 1 and Robinson et al. (2018)). Chemical exposures were measured in serum, plasma, blood or urine samples using maternal samples collected during pregnancy or at birth stored by the cohorts and samples newly collected from the children during childhood (Haug et al. (2018) and Annex 1). Chemical biomarkers measured included organochlorine compounds (OCs), polybrominated diphenyl eters (PBDEs), per- and polyfluoroalkyl substances (PFAS), metals and elements, phthalate metabolites, phenols, organophosphate (OP) pesticide metabolites and cotinine. OCs and PBDEs were adjusted for serum lipid concentrations, and phthalate metabolites, phenols, OP pesticide metabolites and cotinine were adjusted for urinary creatinine. Details on the laboratories, limits of quantification, limits of detection and quality control can be found in Haug et al. (2018). Table A1.5 in the Annex shows the percent of quantifiable samples, which were higher than 90% for 77% of the childhood exposures and for 82% of the pregnancy exposures. Information on active and passive tobacco smoking was collected through questionnaires. For the pregnancy period, we also assigned exposure to water disinfection by-products (DBPs) based on measurements and models for the water supply of the participant's residency (Jeong et al., 2012). For the childhood period, we estimated exposure to indoor air pollutants (NO_2, benzene, and TEX-toluene, ethylbenzene, xylene) by combining measurements in the homes of a subgroup of 150 children during the two time periods with questionnaire data (Annex 1). During childhood, questionnaire information was collected on socio and economic capital of the family based on the Family Affluence Score (FAS) and on social participation, social contact and house crowding data (Annex 1). Finally, information on other lifestyle factors, including maternal and child diet, breastfeeding, maternal and child physical activity, alcohol consumption, pets, and child sleep duration, was collected through questionnaires (Annex 1).

2.3. Statistical analysis

Missing values of exposures and adjustment variables were imputed using the method of chained equations (White et al., 2011), using the mice package in R (van Buuren and Groothuis-Oudshoorn, 2011). Prior to imputation, skewed exposure variables were transformed to achieve normality or categorized if no transformation worked (Annex 2). Missing values, ranged from no missing values for some child phthalate metabolites to 65% for fast-food intake during pregnancy (Annex 2). The mean percentage of missing values per exposure was 12% (first quartile 0.9% and third quartile 16.8%). None of the participants had complete data on all exposures. Yet, for 98% of individuals < 30% of exposure variables had missing values. Annex 3 includes a full description of the imputation process. As the analyses conducted in this paper do not provide standard errors, confidence intervals or p-values, a single imputed dataset was used, as single imputation can already correct for potential biases (Donders et al., 2006).

2.3.1. Correlations between exposures

Given the multicentre design of our study and the potential influence of cohort in the correlations between exposures, we calculated correlations firstly using the entire imputed entire dataset combining the 6 cohorts (referred to as “overall” correlations) and secondly using exposure variables centered and standardized within each cohort (referred to as “within–cohort” correlations). The latter were computed by first centering each exposure by cohort, i.e. subtracting the cohort means and dividing by the cohort-specific standard deviations, and then computing the correlations. We calculated Pearson, polychoric or polyserial correlations, as appropriate, between all pairs of exposures using the polycor and rexposome R Packages (Hernandez-Ferrer and

Gonzalez, 2018). Very similar correlation coefficients were obtained if Spearman correlations were computed instead of Pearson correlations (data not shown). Correlation matrices were calculated separately for the pregnancy and childhood exposures. In order to characterize the heterogeneity of correlations between exposures in the same exposure group by cohort, we i) calculated correlation coefficients separately for each cohort; ii) conducted a meta-analysis of the cohort-specific correlations using the metacor command of the meta R package; and iii) reported the I^2 index of heterogeneity. Overall and within-cohort correlation matrices were displayed and compared using heat maps. In order to visualize the complex relations between exposures, we conducted a network analysis based on the within-cohort correlation matrix, separately for the pregnancy and childhood periods. Additionally, we calculated Pearson's correlations between exposures of the mother during pregnancy and the same exposures measured in the child at 6–11 years when available.

2.3.2. Dimensionality of exposome

We performed two different analyses using Principal Component Analysis (PCA) to achieve two different goals. In the first analysis, the aim was to compare the levels of exposures in the different cohorts.

Given the large number of exposures included in the analysis, we first reduced the dimensionality by conducting a separate PCA within each of the 19 pre-defined exposure groups, and retained only the first principal component for all of them. This way, we created a composite index variable (principal component scores) for each exposure group, and then averaged the scores by cohort to compare the levels. We followed this strategy instead of conducting a PCA with all exposures from all exposure groups together in order to have a better interpretability of the resulting principal components, and to ensure a better comparability of the resulting components in the pregnancy and childhood datasets.

In the second analysis, the aim was to quantify the underlying dimensionality of the data. In this case, we did include all the exposures from the different exposure families in the same analysis (separately for pregnancy and postnatal period), but we only focused on the number of components needed to explain 70% or 95% of the variation. The second analysis was conducted separately on the overall data and on the centered within-cohort data. In both analyses, we conducted PCA with *varimax rotation*.

All statistical analyses were conducted in the R software environment (R3.3.0; http://www.r-project.org). The network analysis was conducted using the Gephi software (Bastian et al., 2009).

3. Results

Table 2 shows the number of participants by cohort, as well as the age and sex distribution. Mothers had an average age of 31 years at child birth, with some variations by cohort. There were also variations by cohort in the age of the children, with EDEN having the oldest children (average age 11) and KANC the youngest (average age 6).

3.1. Exposure levels

Distributions of all exposure variables are detailed in Annex 2 by individual cohorts (Fig. A2.1 for pregnancy and Fig. A2.2 for childhood). Further details on chemical exposure biomarkers are found in Haug et al. (2018). In order to summarize these data, we computed a single principal component separately for each exposure group. Details on these principal component analyses are available in Annex 4, Tables A4.1 and A4.2. Fig. 1 plots the principal component scores by cohort to characterize the cohorts' profiles. The highest values for outdoor exposures (i.e. high pollution, traffic and low greenness) were found in the INMA cohort. In terms of chemicals, the highest exposure to OCs and phthalates were found in the EDEN cohort during pregnancy and in KANC and RHEA, respectively, during childhood; PBDEs were highest in RHEA during pregnancy, while fewer differences by cohort were observed during childhood; phenols were highest in MOBA during pregnancy and in BIB during childhood; and metals were highest in MOBA during pregnancy and in EDEN during childhood. Smoking was highest in RHEA, and water DBPs were highest in INMA. In terms of lifestyle, characterized mainly by consumption of fruit, MOBA showed the highest levels.

3.2. Within- and between-exposure group correlations

Table 3 shows the correlations within exposure groups for the pregnancy and childhood periods. Both overall and within-cohort correlations are shown. Overall correlations tend to be higher than within-cohort correlations, except for the atmospheric pollutants group. This is because, in this case, cohort-level averages do not follow the same correlation pattern than within-cohort correlations. For example, despite the fact that PM_{10} and NO_2 are positively correlated within each cohort, the RHEA cohort had the highest levels of PM_{10} but the lowest levels of NO_2 (Robinson et al., 2018). The highest median overall correlations within exposure groups were observed among water DBPs (r = 0.69) and among meteorological variables in childhood (r = 0.67) and in pregnancy (r = 0.54), followed by tobacco smoking exposure in pregnancy (r = 0.54), PFASs in pregnancy (r = 0.49) and in childhood (r = 0.45), and air pollutants in childhood (r = 0.44). All other exposure groups showed median overall correlations below 0.42, with metals, lifestyle and surrounding natural spaces showing the lowest within-group correlations (below 0.12). In general, correlations within exposure groups were quite heterogeneous by cohort, as shown by the high I^2 (Table 3), except for the metals, lifestyle and socio-economic capital groups.

During pregnancy, the median absolute correlations between exposure groups were much lower (0.08) than within exposure groups (0.20). Fig. 2 shows the median within-group (diagonal) and between-group (off-diagonal) absolute correlations for the pregnancy period. For overall correlations (Fig. 1a), the highest absolute median correlations between exposure groups were observed between noise and meteorological variables (r = 0.45), between water DPBs and meteorological variables (0.39), between water DBPs and atmospheric pollutants (r = 0.32), and noise and traffic (r = 0.29). The rest of the median absolute correlations between groups of chemical exposures were all equal to or below 0.24. The maximum correlation of lifestyle variables was with atmospheric pollutants (r = 0.19). Similar patterns were observed for the within-cohort correlations (Fig. 1b), but the correlations between groups were considerably reduced (r = 0.03 (0.01–0.03)).This was especially the case for water DBPs, and for the correlations between outdoor exposures and chemicals, which were very close to 0. All correlations between noise and the other exposure groups were strongly reduced, except for the correlation between atmospheric pollutants and noise (r = 0.13).

The pattern of correlations between exposures measured during

Table 2
Population characteristics.

Cohort	N	Pregnancy		Childhood		
		Maternal age, years		Age, years		Sex
		Mean	(25th; 75th percentile)	Mean	(25th; 75th percentile)	% female
BiB	205	28.6	(25.0; 33.0)	6.6	(6.5; 6.8)	44.9
EDEN	198	30.5	(27.5; 34.0)	10.9	(10.4; 11.2)	42.9
INMA	223	32.0	(29.5; 34.8)	8.8	(8.4; 9.2)	46.2
KANC	204	29.2	(25.7; 32.3)	6.4	(6.1; 6.8)	45.6
MoBa	272	32.7	(30.0; 35.0)	8.5	(8.2; 8.8)	47.8
RHEA	199	30.8	(27.3; 33.5)	6.5	(6.4; 6.6)	44.2
Total	1301	30.8	(27.2; 34.0)	8.1	(6.5; 8.9)	45.4

childhood was generally similar to the patterns observed during pregnancy (Annex 4, Fig. A4.1). One notable difference was that correlations within meteorological variables were much stronger in the childhood period, where the last month averages were used, compared to the pregnancy period, where we used 9-month averages. Other interesting correlations were found between PFAS and OCs (within: 0.20, overall: 0.19) and between tobacco smoke and indoor air pollution (within: 0.18, overall: 0.24) where within-cohort correlations remained similar to overall correlations. Annex 4, Fig. A4.3 shows histograms and cumulative density functions of all correlations.

Annex 4, Figs. A4.4–A.4.7 display the correlations within and between exposure groups by sex. No notable differences in correlations were observed between the two sexes.

3.3. Network visualization

The exposome correlation structure, using individual exposures and within-cohort correlations, is visualized as a network in Fig. 3. Exposures that are close together in the network are more correlated than more distant ones. The pregnancy network (Fig. 3A) shows a cluster with the outdoor exposome variables, dominated by the built environment variables but also including air pollutants, traffic variables, natural spaces, meteorological variables, and water DBPs. A few lifestyle variables, such as meat consumption and physical activity, were close to the outdoor exposome cluster. On the top-left part of the network there is another cluster made of PFASs and OCs, with contributions from metals, PBDEs, tobacco smoke and diet variables. Phenols and phthalate metabolites form separate clusters linked together by BPA and MEP. A few variables, like folic acid and certain metals, are not connected to other exposures. The childhood network (Fig. 3B) appears more compact than the pregnancy network, with more links between exposure families. The outdoor exposome still appears as a cluster. It is connected to OCs via indoor air pollution and tobacco smoke. OCs continue to be close to PFASs. Breastfeeding and some lifestyle variables such as physical activity and organic food consumption appear close to OCs. Diet variables form their own cluster, with some connections to OCs and PFASs. On the left hand side there is a cluster of phenols, phthalate metabolites, meteorological variables and OP pesticides. Meteorological variables in childhood represent last month values, while in pregnancy they represent pregnancy averages. Thus, in the childhood period meteorological variables represent recent exposure, and this may explain why they are closer to phenols than in the pregnancy period (phenols are present in cosmetics such as sunscreen, and their use may be influenced by recent meteorology). The main connections of the cluster formed by phenols, phthalate metabolites, meteorological variables and OP pesticides are phthalate metabolites and PBDEs with metals, and meteorology with indoor pollution.

3.4. Dimensionality of exposome

When conducting PCA analyses with all exposures together, 10 principal components explained 45% of the total variance in the pregnancy exposome, while 26 components explained 70% and 65 components explained 95%. These numbers of principal components represent 11%, 30% and 75% of the original dimensionality of the pregnancy exposome (87). In the childhood exposome, 10 components explained 39%, 42 explained 70% and 90 components explained 95% of the total variance (Fig. A4.2). These numbers of principal components represent 8%, 34% and 74% of the original dimensionality of the childhood exposome (87), so the reduction in dimensionality is of

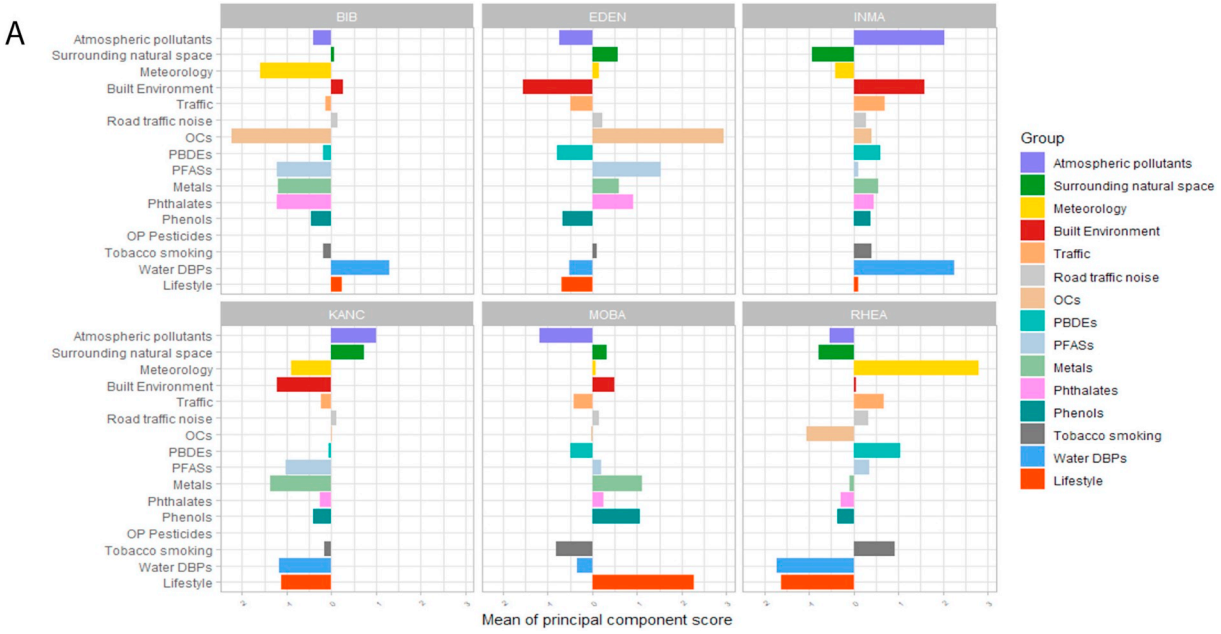


Fig. 1. Profile of pregnancy* (A) and childhood** (B) exposures in the 6 cohorts according to the first component identified by PCA applied separately to each exposure group. Positive values indicate values above the overall mean, while negative values indicate values below the overall mean.
The loadings of each PCA analyses for all exposures are presented in Tables A.4.2 and A.4.3 in Annex 4. The exposures with highest loadings in each component were the following:
*For the pregnancy period (exposure, loading): atmospheric pollutants (NO_2, 0.95), surrounding natural space (green spaces, 0.99), meteorology (temperature, 0.94), built environment (facility richness, 0.94), traffic (inverse distance, 0.99), OCs (PCB180, 0.93), PBDEs (PBDE47, 0.99), PFASs (PFHXS, 0.92), metals (As, 0.8), phthalates (MEOHP, 0.93), phenols (ETPA, 0.95), OP pesticides (DMP, 0.93), tobacco smoking (Cotinine, 0.96), water DBPs (brominated THMs, 0.94), lifestyle (fruit 0.69).
**For the childhood period (exposure, loading): atmospheric pollutants ($PM_{2.5}$ 0.87), surrounding natural space (NDVI school, 0.92), meteorology (temperature 0.93), built environment (population density, 0.89), traffic (inverse distance, 0.95), road traffic noise (noise all day, 0.33), OCs (PCB180, 0.97), PBDEs (PBDE153, 0.98), PFASs (PFUNDA, 0.92), metals (As, 0.96), phthalates (MEHHP, 0.97), phenols (PRPA, 0.91), OP pesticides (DMP, 0.96), tobacco smoking (ETS, 0.96), lifestyle (KIDMED score, 0.80), indoor air (indoor $PM_{2.5}$, 0.96), socio-eco capital (social participation, 0.99).

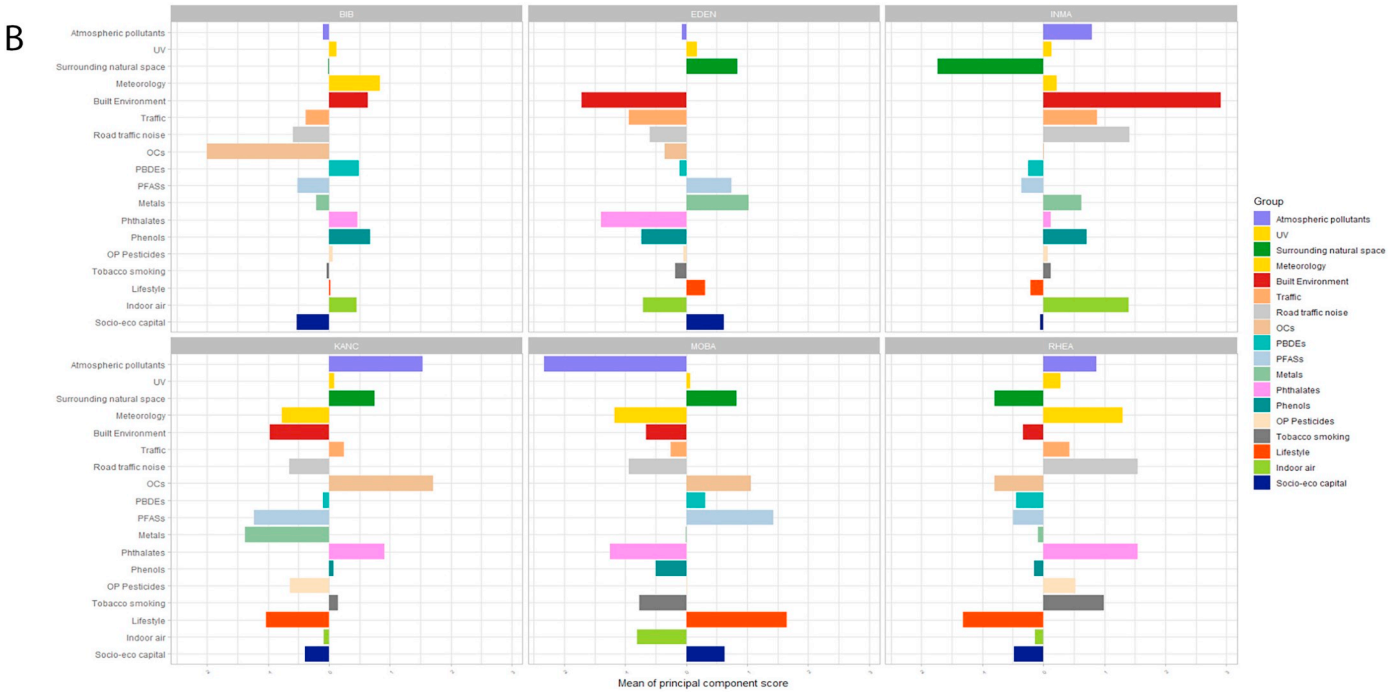


Fig. 1. (*continued*)

similar magnitude in the two periods. The percent of variance explained was slightly lower when using the cohort-centered data (Fig. A4.2).

3.5. Correlations between pregnancy and childhood exposures

We calculated the correlation between the same exposure measured during pregnancy (mother) and childhood. Fig. 4 shows that correlation coefficients are high (> 0.6) for many of the exposures estimated at the residential address, such as atmospheric pollutants, green space, noise, traffic, and certain built environment variables (connectivity, building density). Other built environment factors such as facility density had lower correlations (r < 0.40). The correlation between pregnancy and childhood estimates was also lower for meteorological variables (r = 0.40 for temperature and r = 0.10 for humidity) because the pregnancy period includes nine months while childhood period includes 1-month information and is therefore more subject to seasonality. Within the exposures measured by biomarkers, DDE showed the highest correlation (r = 0.62) between the pregnancy and childhood period. Among OCs, DDE had the highest correlations while HCB showed almost no correlation between pregnancy and childhood samples (r = 0.02). Among PFASs, PFHXS and PFOS showed correlations above 0.4, while the correlation was lower (r < 0.25) for other PFASs. Cs and Hg showed correlations > 0.4, but the other metals showed lower correlations (r < 0.25), even negative in the case of Cd (r = −0.05). Phthalate metabolites showed low correlation between maternal and child sample (between −0.1 and 0.1). The correlations

Table 3
Median absolute correlation within exposure groups for the pregnancy and childhood periods. Overall and centered within-cohort correlations are provided.

Exposure group	Pregnancy period				Childhood period			
	N	Overall Median (min-max)	Within-cohort Median (min-max)	I^2 Median (min-max)	N	Overall Median (min-max)	Within-cohort Median (min-max)	I^2 Median (min-max)
All exposure variables	87	0.08 (0.00–0.99)	0.03 (0.00–0.94)	0.54 (0.00–0.99)	122	0.07 (0.00–0.98)	0.03 (0.00–0.98)	0.00 (0.00–0.99)
Atmospheric pollutants	4	0.30 (0.11–0.74)	0.53 (0.47–0.63)	0.96 (0.94–0.97)	4	0.44 (0.21–0.76)	0.55 (0.47–0.62)	0.97 (0.89–0.98)
UV	0	–	–	–	1	–	–	–
Surrounding natural spaces	3	0.07 (0.02–0.22)	0.05 (0.04–0.32)	0.51 (0.00–0.79)	6	0.11 (0.01–0.67)	0.09 (0.01–0.34)	0.26 (0.00–0.90)
Meteorology	3	0.54 (0.27–0.70)	0.11 (0.10–0.59)	0.95 (0.95–0.98)	2	0.67 (0.67–0.67)	0.47 (0.47–0.47)	0.97 (0.97–0.97)
Built environment	9	0.27 (0.01–0.89)	0.28 (0.01–0.89)	0.86 (0.00–0.98)	15	0.27(0.00–0.72)	0.17(0.00–0.68)	0.71 (0.00–0.97)
Traffic	3	0.26 (0.06–0.31)	0.21 (0.10–0.35)	0.84 (0.82–0.93)	5	0.20 (0.11–0.57)	0.13 (0.02–0.56)	0.70 (0.00–0.99)
Road traffic noise	1	–	–		3	0.34 (0.23–0.59)	0.11 (0.05–0.32)	0.24 (0.00–0.99)
OCs	8	0.38 (0.01–0.97)	0.42 (0.11–0.95)	0.92 (0.63–0.99)	8	0.42 (0.05–0.96)	0.56 (0.18–0.94)	0.83 (0.00–0.97)
PBDE	2	0.21 (0.21–0.21)	0.11 (0.11–0.11)	0.87 (0.87–0.87)	2	0.36 (0.36–0.36)	0.34 (0.34–0.34)	0.37 (0.37–0.37)
PFASs	5	0.49 (0.12–0.67)	0.52 (0.29–0.73)	0.91 (0.54–0.94)	5	0.45 (0.30–0.60)	0.50 (0.26–0.55)	0.67 (0.00–0.92)
Metals and elements	10	0.07 (0.00–0.99)	0.03 (0.00–0.30)	0.00 (0.00–0.91)	10	0.06 (0.01–0.51)	0.09 (0.01–0.42)	0.07 (0.00–0.89)
Phthalate metabolites	10	0.23 (0.00–0.91)	0.21 (0.00–0.91)	0.65 (0.00–0.99)	10	0.25 (0.08–0.98)	0.25 (0.06–0.97)	0.23 (0.00–0.92)
Phenols	7	0.18 (0.01–0.57)	0.14 (0.03–0.53)	0.31 (0.00–0.99)	7	0.21 (0.07–0.64)	0.17 (0.07–0.58)	0.40 (0.00–0.95)
OP pesticide metabolites	4	0.32 (0.24–0.63)	0.33 (0.24–0.59)	0.85 (0.74–0.97)	5	0.26 (0.18–0.50)	0.27 (0.17–0.51)	0.61 (0.00–0.93)
Tobacco smoking	3	0.54 (0.27–0.62)	0.41(0.33–0.45)	0.91 (0.83–0.91)	3	0.42 (0.38–0.51)	0.43 (0.37–0.51)	0.88 (0.85–0.90)
Water DBPs	3	0.69 (0.39–0.76)	0.38 (0.19–0.39)	0.97 (0.86–0.99)	0	–	–	–
Indoor air pollution	0	–	–	–	5	0.20 (0.05–0.43)	0.14 (0.01–0.53)	0.60 (0.00–0.83)
Lifestyle	12	0.11 (0–0.42)	0.03 (0.00–0.14)	0.00 (0.00–0.58)	27	0.07 (0.00–0.59)	0.04 (0.00–0.70)	0.00 (0.00–0.91)
Socio-economic capital	0	–	–	–	4	0.06 (0.02–0.15)	0.04 (0.00–0.13)	0.00 (0.00–0.46)

N: number of compounds assessed in the exposure family. **min-max**: minimum and maximum values observed across all pairs of exposures in the exposure group.

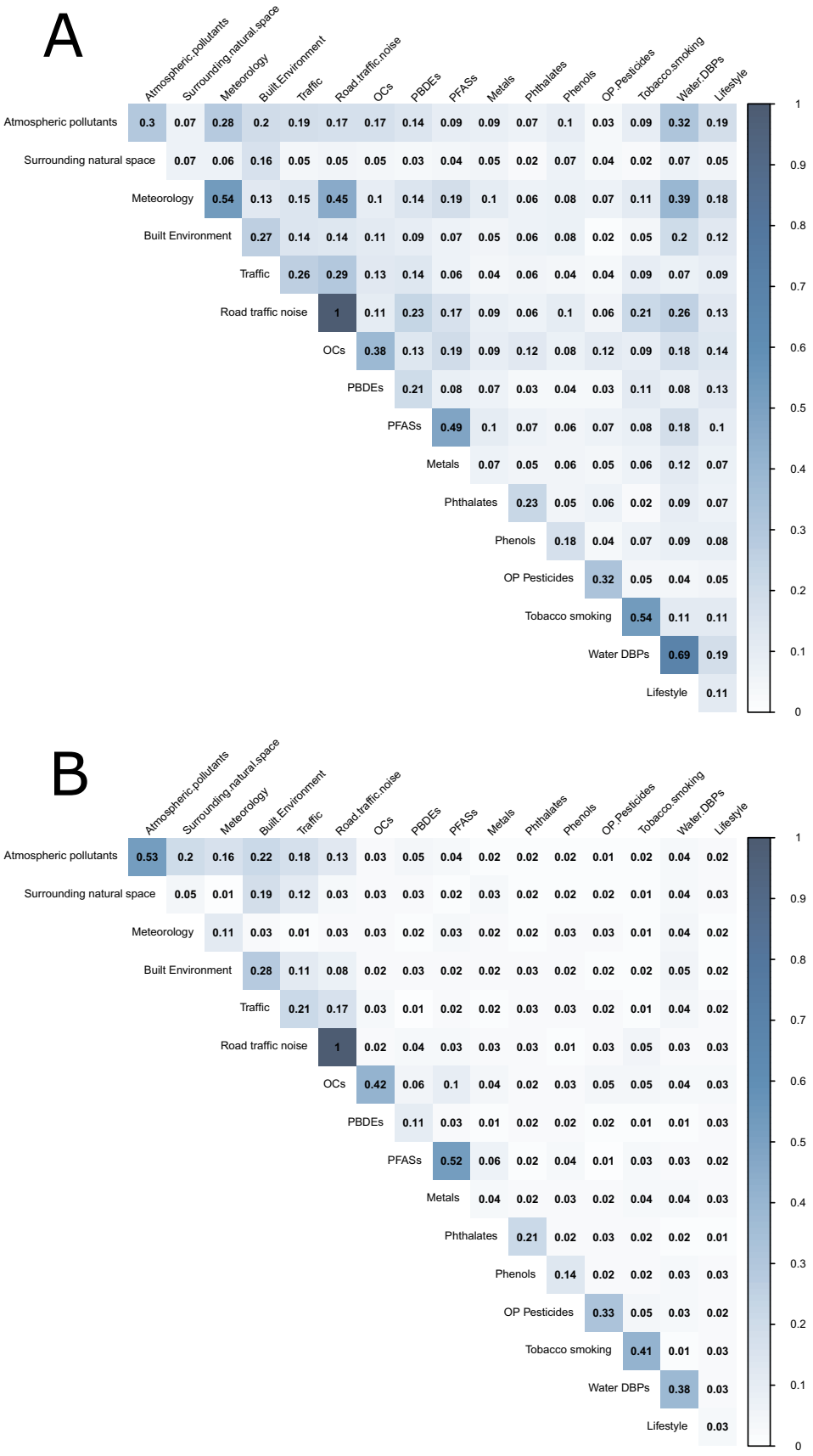


Fig. 2. Median absolute correlations within exposure groups (diagonal) and between exposure groups (off-diagonal) in the pregnancy period. Panel (A) shows overall correlations and panel (B) shows within-cohort correlations.

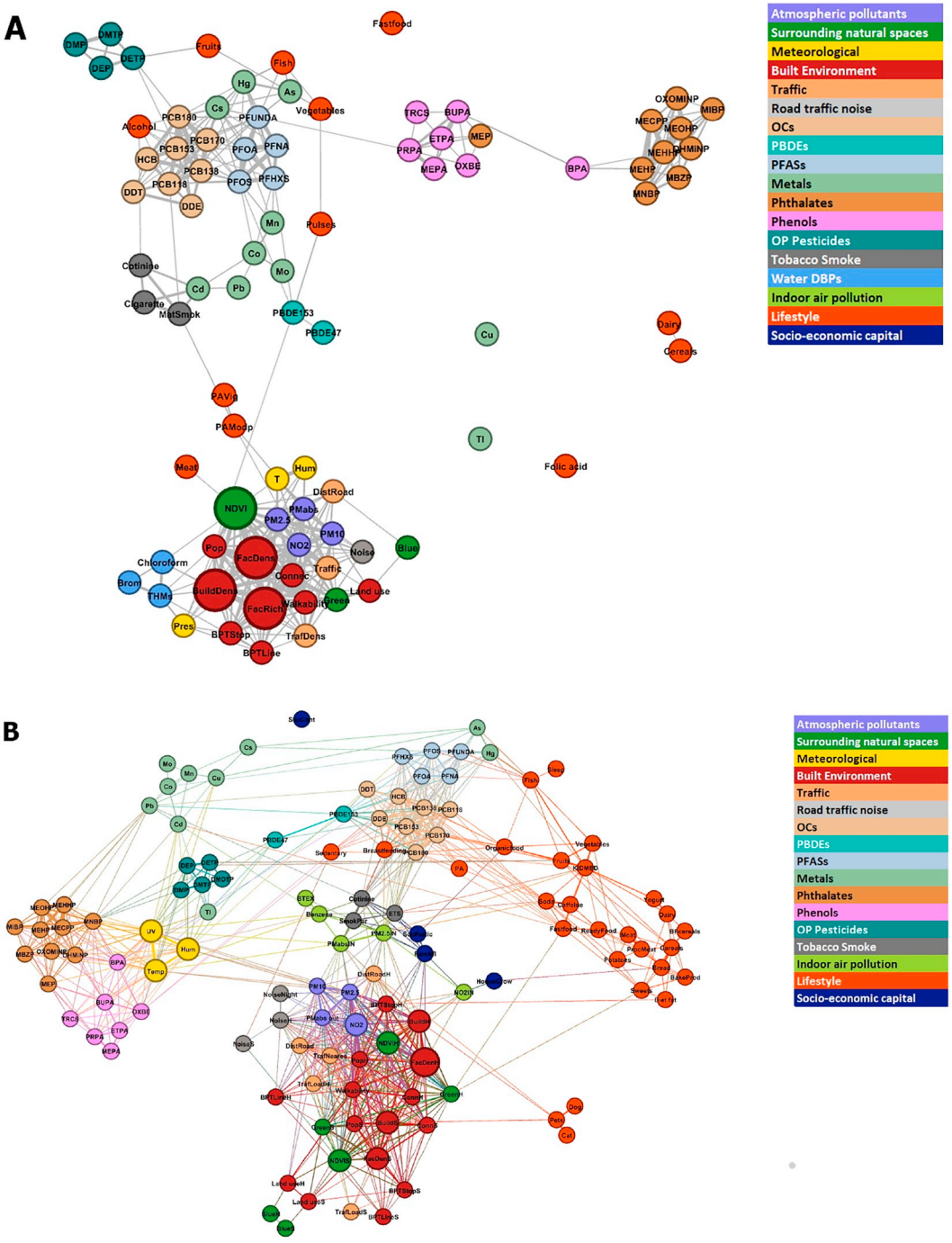


Fig. 3. Network visualization of the exposome. The size of the nodes is proportional to the number of correlations were > 0.5 outside the exposure group and the length of the edges is proportional to the inverse of the correlation (the higher the correlation, the shorter the edge length) between exposures. The colour of the nodes represents the pre-defined exposure groups. The minimum absolute correlation to create an edge was 0.10. Panel A shows the pregnancy exposome, and panel B shows the childhood exposome. Networks were built using within-cohort correlations.

were also small for phenols and OP pesticides (between 0 and 0.20). The temporal changes for chemicals in HELIX have been reported by Haug et al. (2018), with slight differences due to different imputation methods used. Diet variables showed low correlations ($r < 0.1$) with the exception of fish and cereal intake with somewhat greater values ($r = 0.22$ and 0.35 respectively).

Fig. A.4.8 displays the correlations between pregnancy and childhood exposures by sex. No notable differences were observed between the two sexes.

4. Discussion

The HELIX project has measured exposure to many environmental factors of concern for child health, allowing a detailed analysis of the structure of the early life exposome, including its correlations, patterns and variability. Some very high correlations (> 0.8) between particular pairs of exposures were observed and some exposure groups included exposures with high correlations between all exposures in the group (e.g. reaching median correlations 0.69). Overall, however, the median

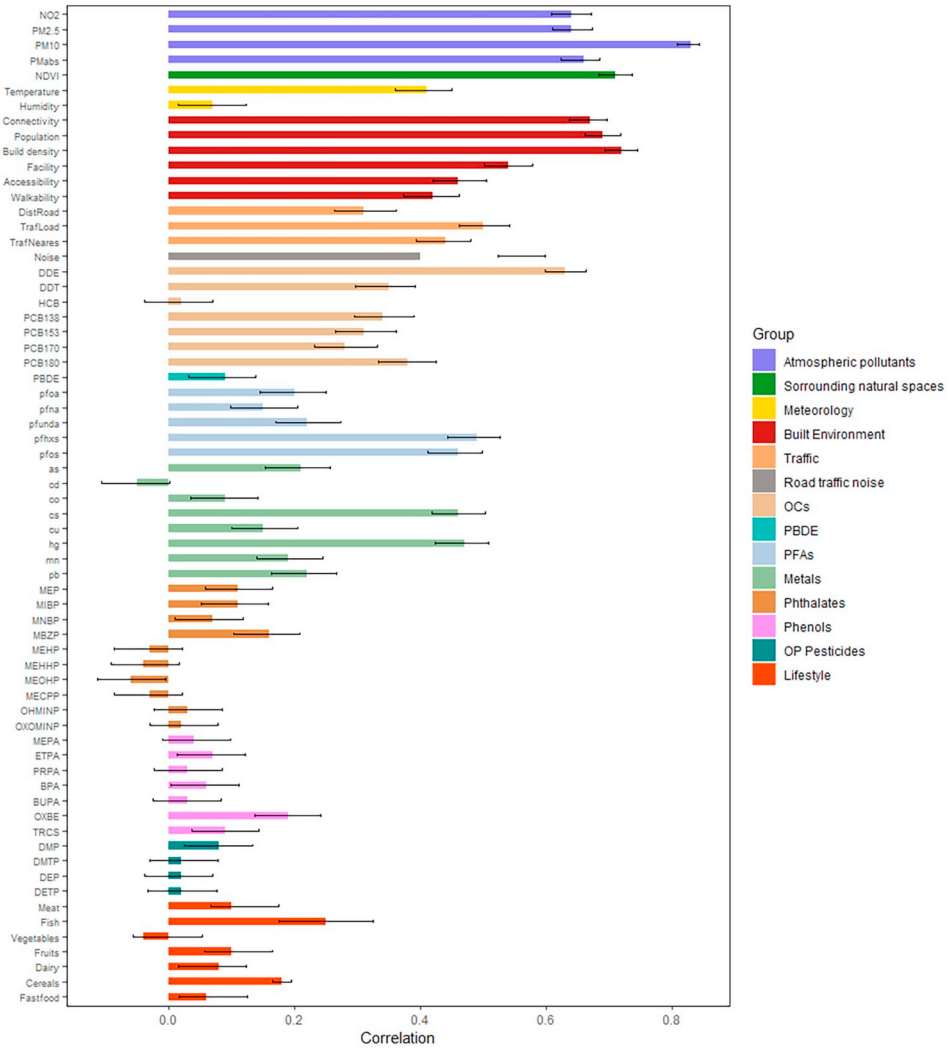


Fig. 4. Correlation of exposures levels between the pregnancy and childhood periods.

correlation between exposures within the same exposure group was only 0.2. Correlations were much lower between exposures from different exposure groups, with a median correlation of 0.08. Our results show that the early life exposome is high dimensional and cannot be summarized by a handful of principal components, and that it varies spatially and temporally.

4.1. Correlation between exposures

Our results show that correlations within the same exposure group can be high, but that correlations between exposures from different groups were low. High correlations between exposures from the same exposure group have been described previously (Lenters et al., 2015; Patel and Ioannidis, 2014; Robinson et al., 2015). The correlations found in our study are very similar than those reported in the U.S. for similar exposure groups (e.g. OCs, phthalates, phenols, metals), even though we reported higher correlations for OP pesticides (Patel and Ioannidis, 2014). The finding of generally low correlations between exposures from different groups, especially after removing cohort effects, is important as it would support the notion that, if this finding is generalizable to all populations, epidemiological studies focusing on a single family of exposures may not be confounded by exposures from other groups. Thus, results from epidemiological studies that focus on a single exposure or on a subset of exposures from the same group are not expected to be greatly affected by not having included exposures from other groups. However, one should be careful; although many between-group correlations were low, they were not negligible. In addition, although the median correlations between exposure groups were low, for particular pairs of exposures, high correlations were observed (e.g. between meteorology and road traffic noise or meteorology and water DBPs).

Some of the high correlations may be due to exposures being measured using a common methodology (e.g. GIS variables) or a common biological medium (e.g. variables measured in urine) or because some variables were used to create others. We made efforts to use the same procedures in all the different cohorts to maximize comparability. Variables belonging to the outdoor exposome (i.e. atmospheric pollutants, traffic noise, natural environments or built environment) tended to be more correlated than others in our data. However, it should be taken into account that some of these exposures were modeled and estimated at a static point, i.e. at the residential and school addresses, and they were capturing annual means. If instead one considers personal measurements that account for time activity patterns or estimates exposure for shorter time periods, correlations can vary. For example, some of the correlations reported by a sub-study based on personal measurements within the HELIX project found higher correlations between urban exposures than the ones reported here (Donaire-Gonzalez et al., unpublished results). The correlation between exposures has important implications when one tries to identify a causal factor for disease among a set of correlated exposures (Agier et al.,

2016; Barrera-Gómez et al., 2017). A high correlation can drastically reduce the sensitivity of the methods to identify the true exposure and increases the probability of obtaining false positives. Thus, according to our results, it will be more difficult to correctly identify the true exposure causing a health effect among the outdoor exposome variables, if such a true association exists.

It is important to note that despite some high correlations, redundancy in our exposome data was still low. To capture 95% of the total variance in the exposome set, 65 principal components were needed and 72 if the cohort effects were removed. Thus, the early-life exposome is complex and high-dimensional, and it is not easy to synthesize it in a few exposure scores. Measurement errors affecting all exposures can also contribute to the observed low correlations between exposures and therefore to the fact that exposures cannot be summarized using a handful of principal components. It is also noteworthy that, in general, the correlations between environmental pollutant exposures and other lifestyle variables were low after controlling for cohort. The fact that exposures show low correlations with lifestyle factors means a lower potential for confounding by lifestyle in studies on the effect of environmental exposures on health.

Overall, it is recommended that investigators conducting exposome research conduct a thorough exploration of the structure of the exposome before evaluating exposome-health associations. Then, even in a context of a high dimensional exposome, there are several statistical methods that can be useful to explore exposome-health relationships (Agier et al., 2016; Stafoggia et al., 2017).

4.2. Cohort effects

When we summarized exposure groups by their first PCA score, we observed strong differences in some of the exposures by cohort. For example, BIB (UK) had the lowest values of OCs while EDEN (France) had the highest, INMA (Spain) had the highest values on atmospheric pollutants and water DBPs, RHEA (Greece) had the highest temperature and lowest water DBPs, MOBA (Norway) had the highest metal values and KANC (Lithuania) the lowest for the same variable. The geographical pattern of environmental exposures can be explained by multiple factors, such as different meteorology, city configuration, country-specific environmental policy or predominant diet. Multicenter studies have the advantage of capturing more varied exposure patterns and different ranges of exposure levels, but between-city differences can also drive many of the correlations between variables. In our data, we observed how, after removing the cohort effect, correlations between exposures from different exposure groups were reduced. Thus, adjustment for cohort is very important in studies linking the exposome with health outcomes, especially if the number of locations is small. Analyses that do not adjust for cohort may be driven by ecological associations obtained from a small number of locations, and although they can reflect true associations, they are more prone to be driven by confounding by other cohort or city level variables (Basagaña et al., 2018). In addition, the distribution of exposures across SES (socioeconomic status) strata can be strongly city-dependent (Robinson et al., 2015; Temam et al., 2017).

4.3. Mother-child correlations

One of the special characteristics of the exposome is that, unlike the genome, it changes over time. This makes the characterization of the exposome challenging. In our data, comparing two points that were 6 to 11 years apart, only one exposure (modeled PM_{10}) had a correlation > 0.8, and a few others, also related to atmospheric pollutants or the built environment, had correlations > 0.6. Chemical exposure can have different half-life in the human body, which influence the temporal correlations. As expected, concentration of persistent chemicals (i.e., PBDEs, and PFASs) showed high temporal correlations whereas non-persistent chemicals (i.e. phthalate metabolites, phenols and OP pesticide metabolites), with a short biological half-life and large within-subject variability (Casas et al., 2018), showed low temporal correlations for phthalate metabolites, phenols and OPs were low. As a result, to reliably estimate associations of these non-persistent chemicals with disease we cannot rely on a measure taken at a single time period, as it may not represent exposure in other periods of life.

Part of the observed variability between periods in our study may be explained by how the exposures were measured. Variations in the outdoor exposome can for example be due to the fact that we used the nine months of pregnancy for the pregnancy period and monthly averages for the childhood period. Moving residences is also a factor introducing temporal variations in the outdoor exposome.

4.4. Missing data

Missing data is a common problem in most epidemiological studies. Often, studies rely on complete case analyses, i.e. they discard participants with missing exposure data even if this approach only provides valid results under the assumption that missing values occurred completely at random or else it may introduce selection bias (Donders et al., 2006). An exposome paradigm provides additional complications as the higher the number of variables that need to be examined jointly, the lower the number of complete cases (in our data, there were no complete cases). Thus, the use of techniques like multiple imputation, which provides valid results under less restrictive assumptions (Donders et al., 2006), is imperative, for example, to implement model selection techniques to build regression models.

Applying multiple imputation to large datasets involves extra difficulties (Stuart et al., 2009). In general, it is recommended that imputation models should not include > 25 variables, as adding more predictors usually provides little gain and can bring problems of convergence (van Buuren and Groothuis-Oudshoorn, 2011). To this end, we built the imputation model for each variable by selecting a set of < 25 predictors that avoided collinearity problems (Annex 3).

4.5. Limitations

Although we collected data on hundreds of variables there are many unmeasured exposures that also conform the early-life exposome. E.g. a 2006 review paper reported 100,000 chemicals registered for commercial use, 200 of them being neurotoxic for humans (Grandjean and Landrigan, 2006). We only had data from a limited number of locations and therefore our results are not generalizable, but they provide a good picture of some of the variations of the early-life exposome in Europe. Likewise, we only observed cross-sections of the exposome and only assessed differences across two time points separated by at least 6 years, leaving the early childhood period unexplored. Our childhood period covered a wide age range, from 6 to 11, in which children go through considerable growth. Previous studies have shown that body burdens of exposure to persistent organic chemicals decrease with age mainly because of an increase in dilution of these chemicals in the total blood volume (Gascon et al., 2015). For non-persistent chemicals measured in urine, such changes are less relevant, but behavioral changes over age may of course lead to changes in exposure. In the present study it was difficult to conduct analyses by age because the age ranges of the children were strongly determined by cohort, which has a strong relationship with levels of chemicals.

5. Conclusions

This multicenter study with over 200 single environmental exposures measured showed the early life exposome to be high dimensional in terms of having little redundancy. Correlations between exposures from different exposure groups were much lower than within exposure groups. This suggests that, in many cases, studies that focus on a single exposure family should not suffer from strong unmeasured

confounding by omitted exposures, although this can be different in other geographical settings or age groups. In addition, the early-life exposome varies strongly by region and by life periods, so measurements at one time point or one place will not capture its complexities.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2018.11.067.

Acknowledgements

We are grateful to all the participating children, parents, practitioners and researchers in the six countries who took part in this study. We thank the following entities that made the collection of many exposure data possible: the ESCAPE project for the air pollution modeling, HiWate for providing part of the water DPB models. We also thank the Yorkshire Water, UK, for the provision of water supply zone boundaries and routine water quality monitoring data, including Trihalomethanes (chloroform, bromoform, dibromochloromethane, and bromodichloromethane), from which all disinfection by product exposures estimates were derived for the Born in Bradford cohort.

This work was supported by the European Commission Seventh Framework Programme (FP7/2007–2013) [grant number: 308333–the HELIX project].

INMA data collections were supported by grants from the Instituto de Salud Carlos III, CIBERESP, and the Generalitat de Catalunya-CIRIT. KANC was funded by the grant of the Lithuanian Agency for Science Innovation and Technology (6-04-2014_31V-66). The Norwegian Mother and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research, NIH/NIEHS (contract no N01-ES-75558), NIH/NINDS (grant no.1 UO1 NS 047537-01 and grant no.2 UO1 NS 047537-06A1). The Rhea project was financially supported by European projects (EU FP6-2003-Food-3-NewGeneris, EU FP6. STREP Hiwate, EU FP7 ENV.2007.1.2.2.2. Project No 211250 Escape, EU FP7-2008-ENV-1.2.1.4 Envirogenomarkers, EU FP7-HEALTH-2009-single stage CHICOS, EU FP7 ENV.2008.1.2.1.6. Proposal No 226285 ENRIECO, EUFP7-HEALTH-2012 Proposal No 308333 HELIX, FP7 European Union project, No. 264357 MeDALL), and the Greek Ministry of Health (Program of Prevention of obesity and neurodevelopmental disorders in preschool children, in Heraklion district, Crete, Greece: 2011–2014; “Rhea Plus”: Primary Prevention Program of Environmental Risk Factors for Reproductive Health, and Child Health: 2012–15).

Dr. Chatzi received additional funding by the Southern California Environmental Health Sciences Center (grant # P30ES007048) funded by the National Institute of Environmental Health Sciences. Dr. Maribel Casas received funding from Instituto de Salud Carlos III (Ministry of Economy and Competitiveness) (MS16/00128). Dr. Charline Warembourg received funding from the Fondation de France (grant # 00069251).

ISGlobal is a member of the Agency for the Research Centers of Catalonia (CERCA) Programme, Generalitat de Catalunya.

References

Agier, L., Portengen, L., Chadeau-Hyam, M., Basagaña, X., Giorgis-Allemand, L., Siroux, V., Robinson, O., Vlaanderen, J., González, J.R., Nieuwenhuijsen, M.J., Vineis, P., Vrijheid, M., Slama, R., Vermeulen, R., 2016. A systematic comparison of linear regression–based statistical methods to assess exposome-health associations. Environ. Health Perspect. 124, 1848–1856. https://doi.org/10.1289/EHP172.

Barrera-Gómez, J., Agier, L., Portengen, L., Chadeau-Hyam, M., Giorgis-Allemand, L., Siroux, V., Robinson, O., Vlaanderen, J., González, J.R., Nieuwenhuijsen, M., Vineis, P., Vrijheid, M., Vermeulen, R., Slama, R., Basagaña, X., 2017. A systematic comparison of statistical methods to detect interactions in exposome-health associations. Environ. Health 16, 74. https://doi.org/10.1186/s12940-017-0277-6.

Basagaña, X., Pedersen, M., Barrera-Gómez, J., Gehring, U., Giorgis-Allemand, L., Hoek, G., Stafoggia, M., Nieuwenhuijsen, M.J., Brunekreef, B., Slama, R., ESCAPE Birth Outcomes Working Group, 2018. Analysis of multicentre epidemiological studies: contrasting fixed or random effects modelling and meta-analysis. Int. J. Epidemiol. https://doi.org/10.1093/ije/dyy117.

Bastian, M., Heymann, S., Jacomy, M., 2009. Gephi: an open source software for exploring and manipulating networks. In: Int. AAAI Conf. Weblogs Soc. Media.

Casas, M., Basagaña, X., Sakhi, A.K., Haug, L.S., Philippat, C., Granum, B., Manzano-Salgado, C.B., Brochot, C., Zeman, F., de Bont, J., Andrusaityte, S., Chatzi, L., Donaire-Gonzalez, D., Giorgis-Allemand, L., Gonzalez, J.R., Gracia-Lavedan, E., Grazuleviciene, R., Kampouri, M., Lyon-Caen, S., Pañella, P., Petraviciene, I., Robinson, O., Urquiza, J., Vafeiadi, M., Vernet, C., Waiblinger, D., Wright, J., Thomsen, C., Slama, R., Vrijheid, M., 2018. Variability of urinary concentrations of non-persistent chemicals in pregnant women and school-aged children. Environ. Int. 121, 561–573. https://doi.org/10.1016/j.envint.2018.09.046.

Chatzi, L., Leventakou, V., Vafeiadi, M., Koutra, K., Roumeliotaki, T., Chalkiadaki, G., Karachaliou, M., Daraki, V., Kyriklaki, A., Kampouri, M., Fthenou, E., Sarri, K., Vassilaki, M., Fasoulaki, M., Bitsios, P., Koutis, A., Stephanou, E.G., Kogevinas, M., 2017. Cohort profile: the mother-child cohort in Crete, Greece (Rhea Study). Int. J. Epidemiol. 46, 1392–1393k. https://doi.org/10.1093/ije/dyx084.

Chung, M.K., Kannan, K., Louis, G.M., Patel, C.J., 2018. Toward capturing the exposome: exposure biomarker variability and co-exposure patterns in the shared environment. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.8b01467.

Dadvand, P., Ostro, B., Figueras, F., Foraster, M., Basagaña, X., Valentín, A., Martinez, D., Beelen, R., Cirach, M., Hoek, G., Jerrett, M., Brunekreef, B., Nieuwenhuijsen, M.J., 2014. Residential proximity to major roads and term low birth weight. Epidemiology 25, 518–525. https://doi.org/10.1097/EDE.0000000000000107.

Donaire-Gonzalez, D., Valentin, A., Curto, A., Casas, M., Basagaña, X., Manzano-Salgado, C., de Castro, M., Tamayo-Uria, I., Robinson, O., Pañella, P., de Bont, J., Waiblinger, D., McEachan, R., Wright, J., Granum, B., Schwarze, P.E., Aasvang, G.M., Meinhard-Kjellstad, C.H., Andrusaityte, S., Petraviciene, I., Dedele, A., Grazuleviciene, R., Lyon-Caen, S., Slama, R., Vafeiadi, M., Chatzi, L., Vrijheid, M., Nieuwenhuijsen, M., Michalaki, E., Kampouri, M., 2018. Personal Assessment of the External Exposome During Pregnancy and Childhood in Europe. (unpublished results).

Donders, A.R.T., van der Heijden, G.J.M.G., Stijnen, T., Moons, K.G.M., 2006. Review: a gentle introduction to imputation of missing values. J. Clin. Epidemiol. 59, 1087–1091. https://doi.org/10.1016/j.jclinepi.2006.01.014.

Gascon, M., Vrijheid, M., Garí, M., Fort, M., Grimalt, J.O., Martinez, D., Torrent, M., Guxens, M., Sunyer, J., 2015. Temporal trends in concentrations and total serum burdens of organochlorine compounds from birth until adolescence and the role of breastfeeding. Environ. Int. 74, 144–151. https://doi.org/10.1016/j.envint.2014.10.010.

Gascon, M., Vrijheid, M., Nieuwenhuijsen, M.J., 2016. The built environment and child health: an overview of current evidence. Curr. Environ. Health Rep. 3, 250–257. https://doi.org/10.1007/s40572-016-0094-z.

Grandjean, P., Landrigan, P., 2006. Developmental neurotoxicity of industrial chemicals. Lancet 368, 2167–2178. https://doi.org/10.1016/S0140-6736(06)69665-7.

Grazuleviciene, R., Danileviciute, A., Nadisauskiene, R., Vencloviene, J., 2009. Maternal smoking, GSTM1 and GSTT1 polymorphism and susceptibility to adverse pregnancy outcomes. Int. J. Environ. Res. Public Health 6, 1282–1297. https://doi.org/10.3390/ijerph6031282.

Guxens, M., Ballester, F., Espada, M., Fernández, M.F., Grimalt, J.O., Ibarluzea, J., Olea, N., Rebagliato, M., Tardón, A., Torrent, M., Vioque, J., Vrijheid, M., Sunyer, J., Project, I.N.M.A., 2012. Cohort profile: the INMA—INfancia y Medio Ambiente—(environment and childhood) project. Int. J. Epidemiol. 41, 930–940. https://doi.org/10.1093/ije/dyr054.

Haug, L.S., Sakhi, A.K., Cequier, E., Casas, M., Maitre, L., Basagaña, X., Andrusaityte, S., Chalkiadaki, G., Chatzi, L., Coen, M., de Bont, J., Dedele, A., Ferrand, J., Grazuleviciene, R., Gonzalez, J.R., Gutzkow, K.B., Keun, H.C., McEachan, R.R.C., Meltzer, H.M., Petraviciene, I., Robinson, O., Saulnier, P., Slama, R., Sunyer, J., Urquiza, J., Vafeiadi, M., Wright, J., Vrijheid, M., Thomsen, C., 2018. In-utero and childhood chemical exposome in six European mother-child cohorts. Environ. Int. 121 (Pt 1), 751–763. https://doi.org/10.1016/j.envint.2018.09.056.

Heindel, J.J., Balbus, J., Birnbaum, L., Brune-Drisse, M.N., Grandjean, P., Gray, K., Landrigan, P.J., Sly, P.D., Suk, W., Cory Slechta, D., Thompson, C., Hanson, M., 2015. Developmental origins of health and disease: integrating environmental influences. Endocrinology 156, 3416–3421. https://doi.org/10.1210/EN.2015-1394.

Hernandez-Ferrer, C., Gonzalez, J.R., 2018. rexposome: Exposome Exploration and Outcome Data Analysis. R Package Version 1.2.0.

Heude, B., Forhan, A., Slama, R., Douhaud, L., Bedel, S., Saurel-Cubizolles, M.-J., Hankard, R., Thiebaugeorges, O., De Agostini, M., Annesi-Maesano, I., Kaminski, M., Charles, M.-A., EDEN Mother-Child Cohort Study Group, 2016. Cohort profile: the EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. Int. J. Epidemiol. 45, 353–363. https://doi.org/10.1093/ije/dyv151.

Hystad, P., Davies, H.W., Frank, L., Van Loon, J., Gehring, U., Tamburic, L., Brauer, M., 2014. Residential greenness and birth outcomes: evaluating the influence of spatially correlated built-environment factors. Environ. Health Perspect. https://doi.org/10.1289/ehp.1308049.

Jeong, C.H., Wagner, E.D., Siebert, V.R., Anduri, S., Richardson, S.D., Daiber, E.J., McKague, A.B., Kogevinas, M., Villanueva, C.M., Goslan, E.H., Luo, W., Isabelle, L.M., Pankow, J.F., Grazuleviciene, R., Cordier, S., Edwards, S.C., Righi, E., Nieuwenhuijsen, M.J., Plewa, M.J., 2012. Occurrence and toxicity of disinfection byproducts in European drinking waters in relation with the HIWATE epidemiology study. Environ. Sci. Technol. 46, 12120–12128. https://doi.org/10.1021/es3024226.

Lenters, V., Portengen, L., Smit, L.A.M., Jönsson, B.A.G., Giwercman, A., Rylander, L., Lindh, C.H., Spanò, M., Pedersen, H.S., Ludwicki, J.K., Chumak, L., Piersma, A.H., Toft, G., Bonde, J.P., Heederik, D., Vermeulen, R., 2015. Phthalates, perfluoroalkyl acids, metals and organochlorines and reproductive function: a multipollutant assessment in Greenlandic, Polish and Ukrainian men. Occup. Environ. Med. 72, 385–393. https://doi.org/10.1136/oemed-2014-102264.

Lim, S.S., Vos, T., Flaxman, A.D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M.,

Anderson, H.R., Andrews, K.G., Aryee, M., Atkinson, C., Bacchus, L.J., Bahalim, A.N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M.L., Blore, J.D., Blyth, F., Bonner, C., Borges, G., Bourne, R., Boussinesq, M., Brauer, M., Brooks, P., Bruce, N.G., Brunekreef, B., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Bull, F., Burnett, R.T., Byers, T.E., Calabria, B., Carapetis, J., Carnahan, E., Chafe, Z., Charlson, F., Chen, H., Chen, J.S., Cheng, A.T.-A., Child, J.C., Cohen, A., Colson, K.E., Cowie, B.C., Darby, S., Darling, S., Davis, A., Degenhardt, L., Dentener, F., Des Jarlais, D.C., Devries, K., Dherani, M., Ding, E.L., Dorsey, E.R., Driscoll, T., Edmond, K., Ali, S.E., Engell, R.E., Erwin, P.J., Fahimi, S., Falder, G., Farzadfar, F., Ferrari, A., Finucane, M.M., Flaxman, S., Fowkes, F.G.R., Freedman, G., Freeman, M.K., Gakidou, E., Ghosh, S., Giovannucci, E., Gmel, G., Graham, K., Grainger, R., Grant, B., Gunnell, D., Gutierrez, H.R., Hall, W., Hoek, H.W., Hogan, A., Hosgood, H.D., Hoy, D., Hu, H., Hubbell, B.J., Hutchings, S.J., Ibeanusi, S.E., Jacklyn, G.L., Jasrasaria, R., Jonas, J.B., Kan, H., Kanis, J.A., Kassebaum, N., Kawakami, N., Khang, Y.-H., Khatibzadeh, S., Khoo, J.-P., Kok, C., Laden, F., Lalloo, R., Lan, Q., Lathlean, T., Leasher, J.L., Leigh, J., Li, Y., Lin, J.K., Lipshultz, S.E., London, S., Lozano, R., Lu, Y., Mak, J., Malekzadeh, R., Mallinger, L., Marcenes, W., March, L., Marks, R., Martin, R., McGale, P., McGrath, J., Mehta, S., Mensah, G.A., Merriman, T.R., Micha, R., Michaud, C., Mishra, V., Mohd Hanafiah, K., Mokdad, A.A., Morawska, L., Mozaffarian, D., Murphy, T., Naghavi, M., Neal, B., Nelson, P.K., Nolla, J.M., Norman, R., Olives, C., Omer, S.B., Orchard, J., Osborne, R., Ostro, B., Page, A., Pandey, K.D., Parry, C.D.H., Passmore, E., Patra, J., Pearce, N., Pelizzari, P.M., Petzold, M., Phillips, M.R., Pope, D., Pope, C.A., Powles, J., Rao, M., Razavi, H., Rehfuess, E.A., Rehm, J.T., Ritz, B., Rivara, F.P., Roberts, T., Robinson, C., Rodriguez-Portales, J.A., Romieu, I., Room, R., Rosenfeld, L.C., Roy, A., Rushton, L., Salomon, J.A., Sampson, U., Sanchez-Riera, L., Sanman, E., Sapkota, A., Seedat, S., Shi, P., Shield, K., Shivakoti, R., Singh, G.M., Sleet, D.A., Smith, E., Smith, K.R., Stapelberg, N.J.C., Steenland, K., Stöckl, H., Stovner, L.J., Straif, K., Straney, L., Thurston, G.D., Tran, J.H., Van Dingenen, R., van Donkelaar, A., Veerman, J.L., Vijayakumar, L., Weintraub, R., Weissman, M.M., White, R.A., Whiteford, H., Wiersma, S.T., Wilkinson, J.D., Williams, H.C., Williams, W., Wilson, N., Woolf, A.D., Yip, P., Zielinski, J.M., Lopez, A.D., Murray, C.J.L., Ezzati, M., AlMazroa, M.A., Memish, Z.A., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380, 2224–2260. https://doi.org/10.1016/S0140-6736(12)61766-8.

Magnus, P., Birke, C., Vejrup, K., Haugan, A., Alsaker, E., Daltveit, A.K., Handal, M., Haugen, M., Høiseth, G., Knudsen, G.P., Paltiel, L., Schreuder, P., Tambs, K., Vold, L., Stoltenberg, C., 2016. Cohort profile update: the Norwegian Mother and child Cohort Study (MoBa). Int. J. Epidemiol. 45, 382–388. https://doi.org/10.1093/ije/dyw029.

Maitre, L., de Bont, J., Casas, M., Robinson, O., Aasvang, G.M., Agier, L., Andrušaitytė, S., Ballester, F., Basagaña, X., Borràs, E., Brochot, C., Bustamante, M., Carracedo, A., de Castro, M., Dedele, A., Donaire-Gonzalez, D., Estivill, X., Evandt, J., Fossati, S., Giorgis-Allemand, L., Gonzalez, J.R., Granum, B., Grazuleviciene, R., Bjerve Gützkow, K., Småstuen Haug, L., Hernandez-Ferrer, C., Heude, B., Ibarluzea, J., Julvez, J., Karachaliou, M., Keun, H.C., Hjertager Krog, N., Lau, C.E., Leventakou, V., Lyon-Caen, S., Manzano, C., Mason, D., McEachan, R., Meltzer, H.M., Petraviciene, I., Quentin, J., Roumeliotaki, T., Sabido, E., Saulnier, P.J., Siskos, A.P., Siroux, V., Sunyer, J., Tamayo, I., Urquiza, J., Vafeiadi, M., van Gent, D., Vives-Usano, M., Waiblinger, D., Warembourg, C., Chatzi, L., Coen, M., van den Hazel, P., Nieuwenhuijsen, M.J., Slama, R., Thomsen, C., Wright, J., Vrijheid, M., 2018. Human Early Life Exposome (HELIX) study: a European population-based exposome cohort. BMJ Open 8 (9), e021311. https://doi.org/10.1136/bmjopen-2017-021311.

Patel, C.J., Ioannidis, J.P.A., 2014. Placing epidemiological results in the context of multiplicity and typical correlations of exposures. J. Epidemiol. Community Health 68, 1096–1100. https://doi.org/10.1136/jech-2014-204195.

Patel, C.J., Manrai, A.K., 2015. Development of exposome correlation globes to map out environment-wide associations. Pac. Symp. Biocomput. 20, 231–242. https://doi.org/10.1142/9789814644730_0023.

Patel, C.J., Bhattacharya, J., Butte, A.J., 2010. An Environment-Wide Association Study (EWAS) on type 2 diabetes mellitus. PLoS One 5, e10746. https://doi.org/10.1371/journal.pone.0010746.

Patel, C.J., Ioannidis, J.P.A., Cullen, M.R., Rehkopf, D.H., 2015. Systematic assessment of the correlations of household income with infectious, biochemical, physiological, and environmental factors in the United States, 1999–2006. Am. J. Epidemiol. 181, 171–179. https://doi.org/10.1093/aje/kwu277.

Robinson, O., Vrijheid, M., 2015. The pregnancy exposome. Curr. Environ. Health Rep. 2, 204–213. https://doi.org/10.1007/s40572-015-0043-2.

Robinson, O., Basagaña, X., Agier, L., de Castro, M., Hernandez-Ferrer, C., Gonzalez, J.R., Grimalt, J.O., Nieuwenhuijsen, M., Sunyer, J., Slama, R., Vrijheid, M., 2015. The pregnancy exposome: multiple environmental exposures in the INMA-Sabadell birth cohort. Environ. Sci. Technol. 49, 10632–10641. https://doi.org/10.1021/acs.est.5b01782.

Robinson, O., Tamayo, I., de Castro, M., Valentin, A., Giorgis-Allemand, L., Hjertager Krog, N., Marit Aasvang, G., Ambros, A., Ballester, F., Bird, P., Chatzi, L., Cirach, M., Dėdelė, A., Donaire-Gonzalez, D., Gražuleviciene, R., Iakovidis, M., Ibarluzea, J., Kampouri, M., Lepeule, J., Maitre, L., McEachan, R., Oftedal, B., Siroux, V., Slama, R., Stephanou, E.G., Sunyer, J., Urquiza, J., Vegard Weyde, K., Wright, J., Vrijheid, M., Nieuwenhuijsen, M., Basagaña, X., 2018. The urban exposome during pregnancy and its socioeconomic determinants. Environ. Health Perspect. 126, 77005. https://doi.org/10.1289/EHP2862.

Slama, R., Vrijheid, M., 2015. Some challenges of studies aiming to relate the exposome to human health. Occup. Environ. Med. 72, 383–384. https://doi.org/10.1136/oemed-2014-102546.

Stafoggia, M., Breitner, S., Hampel, R., Basagaña, X., 2017. Statistical approaches to address multi-pollutant mixtures and multiple exposures: the state of the science. Curr. Environ. Health Rep. 4 (4), 481–490. https://doi.org/10.1007/s40572-017-0162-z.

Stuart, E.A., Azur, M., Frangakis, C., Leaf, P., 2009. Multiple imputation with large data sets: a case study of the children's mental health initiative. Am. J. Epidemiol. 169, 1133–1139. https://doi.org/10.1093/aje/kwp026.

Temam, S., Burte, E., Adam, M., Antó, J.M., Basagaña, X., Bousquet, J., Carsin, A.-E., Galobardes, B., Keidel, D., Künzli, N., Le Moual, N., Sanchez, M., Sunyer, J., Bono, R., Brunekreef, B., Heinrich, J., de Hoogh, K., Jarvis, D., Marcon, A., Modig, L., Nadif, R., Nieuwenhuijsen, M., Pin, I., Siroux, V., Stempfelet, M., Tsai, M.-Y., Probst-Hensch, N., Jacquemin, B., 2017. Socioeconomic position and outdoor nitrogen dioxide (NO2) exposure in Western Europe: a multi-city analysis. Environ. Int. 101, 117–124. https://doi.org/10.1016/j.envint.2016.12.026.

van Buuren, S., Groothuis-Oudshoorn, K., 2011. mice: multivariate imputation by chained equations in R. J. Stat. Softw. 45, 1–67. https://doi.org/10.18637/jss.v045.i03.

Vrijheid, M., Slama, R., Robinson, O., Chatzi, L., Coen, M., van den Hazel, P., Thomsen, C., Wright, J., Athersuch, T.J., Avellana, N., Basagaña, X., Brochot, C., Bucchini, L., Bustamante, M., Carracedo, A., Casas, M., Estivill, X., Fairley, L., van Gent, D., Gonzalez, J.R., Granum, B., Gražulevičienė, R., Gutzkow, K.B., Julvez, J., Keun, H.C., Kogevinas, M., McEachan, R.R.C., Meltzer, H.M., Sabidó, E., Schwarze, P.E., Siroux, V., Sunyer, J., Want, E.J., Zeman, F., Nieuwenhuijsen, M.J., 2014. The human early-life exposome (HELIX): project rationale and design. Environ. Health Perspect. 122, 535–544. https://doi.org/10.1289/ehp.1307204.

Vrijheid, M., Casas, M., Gascon, M., Valvi, D., Nieuwenhuijsen, M., 2016. Environmental pollutants and child health—a review of recent concerns. Int. J. Hyg. Environ. Health 219, 331–342. https://doi.org/10.1016/j.ijheh.2016.05.001.

White, I.R., Royston, P., Wood, A.M., 2011. Multiple imputation using chained equations: issues and guidance for practice. Stat. Med. 30, 377–399. https://doi.org/10.1002/sim.4067.

Wild, C.P., 2005. Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol. Biomark. Prev. 14, 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456.

Willett, W.C., 2002. Balancing life-style and genomics research for disease prevention. Science 296, 695–698. https://doi.org/10.1126/science.1071055.

Wright, J., Small, N., Raynor, P., Tuffnell, D., Bhopal, R., Cameron, N., Fairley, L., Lawlor, D.A., Parslow, R., Petherick, E.S., Pickett, K.E., Waiblinger, D., West, J., Born in Bradford Scientific Collaborators Group, 2013. Cohort profile: the Born in Bradford multi-ethnic family cohort study. Int. J. Epidemiol. 42, 978–991. https://doi.org/10.1093/ije/dys112.