





## Early Life Exposures

# Association between the pregnancy exposome and fetal growth

Lydiane Agier,<sup>1</sup> Xavier Basagaña,<sup>2,3,4</sup> Carles Hernandez-Ferrer,<sup>2,3,4</sup>  
Léa Maitre,<sup>2,3,4</sup> Ibon Tamayo Uria,<sup>2,3,4</sup> Jose Urquiza,<sup>2,3,4</sup>  
Sandra Andrusaityte,<sup>5</sup> Maribel Casas,<sup>2,3,4</sup> Montserrat de Castro,<sup>2,3,4</sup>  
Enrique Cequier,<sup>6</sup> Leda Chatzi,<sup>7</sup> David Donaire-Gonzalez,<sup>8,9</sup>  
Lise Giorgis-Allemand,<sup>1</sup> Juan R Gonzalez,<sup>2,3,4</sup> Regina Grazuleviciene,<sup>5</sup>  
Kristine B Gützkow,<sup>6</sup> Line S Haug,<sup>6</sup> Amrit K Sakhi,<sup>6</sup>  
Rosemary RC McEachan,<sup>10</sup> Helle M Meltzer,<sup>6</sup> Mark Nieuwenhuijsen,<sup>2,3,4</sup>  
Oliver Robinson ,<sup>11</sup> Theano Roumeliotaki,<sup>12</sup> Jordi Sunyer,<sup>2,3,4</sup>  
Cathrine Thomsen,<sup>6</sup> Marina Vafeiadi,<sup>12</sup> Antonia Valentin,<sup>2,3,4</sup>  
Jane West ,<sup>10</sup> John Wright,<sup>10</sup> Valérie Siroux,<sup>1</sup> Martine Vrijheid<sup>2,3,4</sup> and  
Rémy Slama<sup>1\*</sup>

<sup>1</sup>Inserm, CNRS, University Grenoble Alpes, Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health, IAB, Grenoble, France, <sup>2</sup>ISGlobal, Barcelona, Spain, <sup>3</sup>Universitat Pompeu Fabra (UPF), Barcelona, Spain, <sup>4</sup>CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain, <sup>5</sup>Department of Environmental Sciences, Vytautas Magnus University, Kaunas, Lithuania, <sup>6</sup>Norwegian Institute of Public Health, Oslo, Norway, <sup>7</sup>Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA, <sup>8</sup>Institute for Risk Assessment Sciences (IRAS), Division of Environmental Epidemiology (EEPI), Utrecht University, Utrecht, The Netherlands, <sup>9</sup>Mary MacKillop Institute for Health Research, Australian Catholic University, Melbourne, Australia, <sup>10</sup>Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK, <sup>11</sup>MRC Centre for Environment and Health, School of Public Health, Imperial College London, London, UK and <sup>12</sup>Department of Social Medicine, University of Crete, Greece

\*Corresponding author. Inserm, CNRS, University Grenoble-Alpes, Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health, Institute for Advanced Biosciences, Joint Research Center, Rond-Point de la Chantourne, 38700 La Tronche, France. E-mail: remy.slama@univ-grenoble-alpes.fr

Editorial decision 31 January 2020; Accepted 31 January 2020

## Abstract

**Background:** Several environmental contaminants were shown to possibly influence fetal growth, generally from single exposure family studies, which are prone to publication bias and confounding by co-exposures. The exposome paradigm offers perspectives to avoid selective reporting of findings and to control for confounding by co-exposures. We aimed to characterize associations of fetal growth with the pregnancy chemical and external exposomes.

**Methods:** Within the Human Early-Life Exposome project, 131 prenatal exposures were assessed using biomarkers and environmental models in 1287 mother–child pairs from six European cohorts. We investigated their associations with fetal growth using a deletion-substitution-addition (DSA) algorithm considering all exposures simultaneously, and an exposome-wide association study (ExWAS) considering each exposure independently. We corrected for exposure measurement error and tested for exposure–exposure and sex–exposure interactions.

**Results:** The DSA model identified lead blood level, which was associated with a 97 g birth weight decrease for each doubling in lead concentration. No exposure passed the multiple testing-corrected significance threshold of ExWAS; without multiple testing correction, this model was in favour of negative associations of lead, fine particulate matter concentration and absorbance with birth weight, and of a positive sex-specific association of parabens with birth weight in boys. No two-way interaction between exposure variables was identified.

**Conclusions:** This first large-scale exposome study of fetal growth simultaneously considered >100 environmental exposures. Compared with single exposure studies, our approach allowed making all tests (usually reported in successive publications) explicit. Lead exposure is still a health concern in Europe and parabens health effects warrant further investigation.

**Key words:** Biomarkers, cohort, chemical exposures, environment, exposome, fetal growth, mixtures

#### Key Messages

- We conducted the first exposome study considering the possible association of fetal growth with >100 exposures, some, such as mono-4-methyl-7-hydroxyoctyl (OHMiNP) and mono-4-methyl-7-oxooctyl (OXOMiNP) phthalate metabolites, being investigated in humans for the first time.
- We considered possible sex-specific effects and exposures interactions effects on fetal growth, and accounted for exposure measurement error in >1200 children from the Human Early-Life Exposome European cohorts.
- Lead maternal pregnancy exposure was associated with decreased fetal growth.
- Parabens health effects on fetal growth in male offspring were suggested, with weaker evidence.
- Associations of all exposures tested are provided for future exposure-specific meta-analyses.

## Introduction

Humans live in an environment that includes chemical, physical, biological and social factors that can influence health. The ‘exposome’ concept, encompassing the totality of human environmental exposures from conception onward, calls for a complete consideration of these environmental exposures.<sup>1</sup> It covers a very large number of factors: the chemical exposome alone includes tens of thousands of identified natural and man-made chemicals. Research is currently at an early stage of characterizing these exposures and their associations with human health, in isolation or considering possible interactions. Several disciplines, in particular toxicology and epidemiology, contribute to this effort. Most epidemiological research aimed at characterizing associations of environmental

factors with health have so far relied on the assessment of exposure to a single compound or compound family (e.g. atmospheric pollutants). A few studies have simultaneously considered more than a couple of families of exposures, relating them to outcomes such as birth weight,<sup>2–6</sup> fecundity,<sup>7,8</sup> type II diabetes mellitus,<sup>9</sup> respiratory health<sup>10</sup> or mortality.<sup>11</sup>

The developmental period (from the prenatal period to the first years of life) is considered a particularly relevant exposure window. Exposures during this period could affect the body structure, physiology, epigenetic marks and metabolism. These alterations may in turn lead to adverse health effects in the short- and long-terms.<sup>12</sup> One of the first health parameters that can be studied in relation to the early-life exposome is fetal growth (i.e. birth weight

corrected for gestational duration), which has relevance for later health.<sup>13,14</sup>

Human studies based on a single exposure family have reported fetal growth to be probably sensitive to particulate matter,<sup>15</sup> altitude (or atmospheric pressure)<sup>16</sup> and maternal active and passive smoking.<sup>17</sup> Associations with fetal growth were also reported for polychlorinated biphenyls (PCBs),<sup>18</sup> metals such as cadmium<sup>19</sup> and lead,<sup>20</sup> and per- and poly-fluoroalkyl substances,<sup>21</sup> with lower and varying levels of evidence. Among non-persistent chemicals with strong within-subject temporal variability, compounds such as parabens<sup>22</sup> or organophosphate pesticides<sup>5</sup> have so far been considered by very few studies: in the case of bisphenol A and of some phthalates, several studies exist but lack consistency,<sup>23</sup> which may be partly attributed to studies generally using a single spot urine sample to assess exposure to these non-persistent compounds. Indeed, reliance on spot biospecimens causes attenuation bias in the exposure–health association under the hypothesis of classical-type measurement error.<sup>24</sup>

At least five exposome-wide studies have been conducted in relation to fetal growth,<sup>2–6</sup> considering up to 57 chemicals from 6 families.<sup>5</sup> From a methodological standpoint, such exposome research raises many challenges. These include the ability to consider a large number of exposures, measurement error (which is expected to be differential across exposures, i.e. its amplitude varies depending on the biological persistence of each compound<sup>24</sup>), the correction for confounding by co-exposures, low statistical power, and the identification of statistical or biological interactions between exposures.<sup>25–27</sup> Among the existing studies of birth weight sensitivity to multiple environmental contaminant families,<sup>2–6</sup> two considered possible sex-specific effects on birth weight,<sup>2,4</sup> and one additionally considered possible interactions between exposures.<sup>2</sup> With one possible exception relying on the pooling of two urine samples,<sup>5</sup> most of these studies are likely to suffer from strong exposure measurement error for non-persistent compounds, which are the chemicals most produced today.<sup>5</sup>

In this study, we aimed to evaluate the relationship between multiple environmental exposures from both the internal (including e.g. urinary and blood biomarkers, diet) and urban (urban environment, meteorological factors, water disinfection by-products and atmospheric pollutants) exposomes and fetal growth: we considered issues such as possible confounding by co-exposures, exposure measurement error, statistical interactions between exposures, and between exposures and offspring sex.

## Methods

### Study population

We relied on mother–child pairs from six European birth cohorts [Born in Bradford (BiB; UK), Étude des Déterminants Pré et Postnatals du Développement et de la Santé de l'Enfant (EDEN; France), Infancia y Medio Ambiente (INMA; Spain), Kaunas Cohort (KANC; Lithuania), Norwegian Mother, Father and Child Cohort Study (MoBa; Norway) (see [Supplementary Material S1](#), available as [Supplementary data](#) at *IJE* online) and Mother Child Cohort in Crete (RHEA; Greece)] for whom 131 exposures were assessed during pregnancy ([Table 1](#)) as part of the Human Early-Life Exposome (HELIX) project.

HELIX is one of the first large prospective exposome projects on early-life exposures.<sup>28,29</sup> It aims to characterize the early-life exposure to multiple environmental factors and its association with child health. The cohorts included ~32 000 mother–child pairs with harmonized information on the urban exposome, among which 1301 pairs were characterized for their internal exposome. From these, we obtained birth weight and gestational duration data in 1287 mother–child pairs, who constitute our study population. The study was approved by the relevant ethical committees from each country and an informed consent form was signed by all participants or the parents of the children.

### Birth-weight data

Birth weight was collected as part of the study protocol of each cohort and harmonized in the context of the European Study of Cohorts for Air Pollution Effects (ESCAPE) project.<sup>15</sup> Whenever possible, gestational duration was defined as the interval between the start of the last menstruation and delivery; when the date of the last menstruation was missing, ultrasound-based estimates were used; when both measures were missing, obstetrician estimates were used.

### Characterization of the pregnancy exposome

We assessed environmental exposures (i) using geographic information systems, remote sensing and spatio-temporal modeling; (ii) from questionnaires; and (iii) from exposure biomarkers assessed in urine and blood samples collected during pregnancy. The exposure assessment is described in [Supplementary Material S2](#), available as [Supplementary data](#) at *IJE* online,<sup>30</sup> and exposure levels are described in [Supplementary Table S1](#), available as [Supplementary data](#)

at *IJE* online. Most of the exposure biomarkers had high detection frequencies (78% had >90% detected levels). Values below the limits of detection were imputed using the quantile regression approach for the imputation of left-censored missing data.<sup>31</sup> After transforming exposures to approach normality, missing data for exposures and adjustment factors were imputed using the chained equations method.<sup>32</sup> Twenty datasets were imputed in order to account for the uncertainty associated with the imputation procedure (see [Supplementary Material S3](#), available as [Supplementary data](#) at *IJE* online, for more details).<sup>30</sup> All continuous exposure variables were standardized by the inter-quartile range (IQR). We used version 2.2 of the HELIX exposome dataset.

For biomarker-based exposures, the structure of measurement error is expected to be of classical-type, a situation in which the impact of exposure measurement error on the dose–response estimates can be limited by statistical modelling if information on the within-subject compound variability is available (it can either be estimated from repeated measurements or be provided by external sources).<sup>24</sup> Exposures assessed from questionnaires or environmental models are also measured with error, but their structure is unlikely to be of classical type. We corrected exposure measurement error of classical-type by applying regression calibration, a regression method that aims to estimate the true exposure value based on the exposure within-subject temporal variability, and on the information provided by the other exposures (see [Supplementary Material S4](#), available as [Supplementary data](#) at *IJE* online).<sup>33</sup> Since no repeated biospecimens were collected in our study population, we had to rely on external estimates of intraclass-correlation coefficients (ICCs) issued from other studies in pregnant women; ICCs for 26 exposure variables could be identified ([Supplementary Table S2](#), available as [Supplementary data](#) at *IJE* online).

### Overall statistical analyses strategy

Our primary analysis relied on the deletion-substitution-addition (DSA) variable selection algorithm; we additionally used an exposure-by-exposure exposome-wide association study (ExWAS) analysis. In previous simulation studies investigating an exposome context similar to ours, DSA showed a lower false discovery rate compared with other families of linear regression-based methods.<sup>27</sup> This model can be expanded to consider interaction terms, although the expected sensitivity is lowered.<sup>25</sup> ExWAS, which was expected to have a greater sensitivity than DSA, at the cost of a much higher false discovery rate,<sup>25,27</sup> was secondarily used to allow comparisons with former single exposure studies.

All models were adjusted for a set of pre-defined adjustment factors: gestational duration (simple and quadratic terms), sex of the newborn (determined by clinical examination at birth), parity, maternal height, maternal weight before pregnancy (using a broken stick model with a knot at 60 kg), number of cigarettes smoked per day by the mother during the second trimester of gestation, maternal education and season of conception. We also adjusted for the cohort using a fixed effect variable.<sup>34</sup>

### Multi-exposure DSA analysis

DSA is an iterative linear regression variable selection algorithm that, at each iteration, tests for the removal of a variable, the replacement of one variable by another, or the addition of a variable to the model.<sup>35</sup> Only linear terms were considered in the main model, and the maximum model size was set to 50, a number that was never reached. The final model was selected by minimizing the value of the root mean squared error of predictions using 5-fold cross-validated data.

We adapted the DSA model in two ways: (i) by stacking the imputed datasets and running DSA on this extended dataset, using weights not to artificially inflate the number of observations.<sup>36</sup> This method provides unbiased estimates if the estimates based on a single data set are unbiased.<sup>37</sup> (ii) To cope with model instability due to cross-validation, we ran DSA on our stacked dataset 100 times, and included in a final linear regression model all the exposure variables that were selected in at least five DSA runs.

We first applied DSA with terms of degree one (i.e., no polynomial of degree two or interaction terms) only. Second, we applied DSA with all exposure–exposure, sex–exposure and cohort–exposure interaction terms, by allowing all quadratic terms and two-way interaction terms between all exposures and adjustment factors to be selected.

To limit any impact of collinearity, we included in the DSA procedures only one a priori selected variable for compounds/factors that were estimated over different time windows or in different buffers, and for groups of variables whose absolute correlation coefficients were >0.90 (the variable with the smallest proportion of imputed values was selected). Following these criteria, 85 out of the 131 exposure variables entered the DSA selection method (see [Table 1](#)).

### Exposome-wide association study

The ExWAS approach consists of a covariate-by-covariate estimation of the exposure–outcome association through independent linear regression models.<sup>9</sup> For each exposure variable, results from the 20 imputed datasets were aggregated using Meng and Rubin’s rule for multiple imputed

**Table 1.** List of exposures assessed during pregnancy

Exposure family	Compound/factor	Unit	Exposure window/categories
<b>Urban exposome</b>			
Built environment	Bus lines	None; at least one	100m <sup>a</sup> , 300 m, 500 m buffer <sup>a</sup>
	Bus stops	None; at least one	100 m buffer <sup>a</sup>
	Building density	Number of stops/km <sup>2</sup>	300 m, 500 m buffer <sup>a</sup>
	Connectivity density	m <sup>2</sup> with buildings/km <sup>2</sup>	100 m <sup>a</sup> , 300 m buffer
	Number of facilities	Number of intersections/ km <sup>2</sup>	100 m <sup>a</sup> , 300 m buffer
	Number of facility types	Number of facilities/km <sup>2</sup>	300 m buffer
	Land use, Shannon's evenness index	Number of facility types/km <sup>2</sup>	300 m buffer
	Population density	Index	300 m buffer
	Walkability index	Inhabitants/km <sup>2</sup>	At pregnancy address
Atmospheric pollutants	NO <sub>2</sub>	Index	At pregnancy address
	PM <sub>10</sub> mass concentration	µg/m <sup>3</sup>	T1 <sup>a</sup> , T2 <sup>a</sup> , T3 <sup>a</sup> , pregnancy
	PM <sub>2.5</sub> mass concentration	µg/m <sup>3</sup>	T1 <sup>a</sup> , T2 <sup>a</sup> , T3 <sup>a</sup> , pregnancy
	PM <sub>2.5</sub> absorbance	10 <sup>-5</sup> /m	T1 <sup>a</sup> , T2 <sup>a</sup> , T3 <sup>a</sup> , pregnancy
	Road traffic noise	Day and night sound pressure level	A-weighted dB, tertiles (<55; 55–60; 60–65; >65)
Night sound pressure level		A-weighted dB, tertiles (<50; 50–55; 55–60; >60)	At pregnancy address <sup>a</sup>
Meteorological variables	Humidity	Percentage	T1 <sup>a</sup> , T2 <sup>a</sup> , T3 <sup>a</sup> , pregnancy
	Land surface temperature from satellite image	°C	At pregnancy address <sup>a</sup>
	Pressure	Bar	T1 <sup>a</sup> , T2 <sup>a</sup> , T3 <sup>a</sup> , pregnancy
	Temperature	°C	T1 <sup>a</sup> , T2 <sup>a</sup> , T3 <sup>a</sup> , pregnancy
Surrounding natural space	Blue space	Absence; presence	300m buffer
	Green space	Absence; presence	300m buffer
	NDVI	Index	100m, 300 m <sup>a</sup> , 500 m buffer <sup>a</sup>
Road traffic	Inverse distance to nearest road	m <sup>-1</sup>	From pregnancy address
	Traffic load of major roads	None; at least one	100 m buffer <sup>a</sup>
	Traffic load of all roads	Number of vehicles × m of road	100m buffer
	Traffic density on nearest road	Number of vehicles/day	From pregnancy address
Water disinfection by-products	Total brominated trihalomethanes	µg/l of water	T1 <sup>a</sup> , T2 <sup>a</sup> , T3 <sup>a</sup> , pregnancy
	Chloroform	µg/l of water	T1 <sup>a</sup> , T2 <sup>a</sup> , T3 <sup>a</sup> , pregnancy
	Total trihalomethanes	µg/l of water	T1 <sup>a</sup> , T2 <sup>a</sup> , T3 <sup>a</sup> , pregnancy
<b>Internal exposome</b>			
Metals and essential elements	Arsenic	µg/l of blood	Spot pregnancy sample
	Cadmium	µg/l of blood	Spot pregnancy sample
	Cobalt	µg/l of blood	Spot pregnancy sample
	Caesium	µg/l of blood	Spot pregnancy sample
	Copper	µg/l of blood	Spot pregnancy sample
	Mercury	µg/l of blood	Spot pregnancy sample
	Manganese	µg/l of blood	Spot pregnancy sample
	Molybdenum	µg/l of blood	Spot pregnancy sample
	Lead	µg/l of blood	Spot pregnancy sample
	Thallium	Undetected; detected in blood	Spot pregnancy sample
Lifestyle	Alcohol consumption	No; yes	During pregnancy
	Cereal consumption	Number of times/week, tertiles (<9.0; 9.0–27.3; >27.3)	During pregnancy
	Dairy products consumption	Number of times/week, tertiles (<17.1; 17.1–27.1; >27.1)	During pregnancy
	Fast food consumption	Number of times/week, tertiles (<17.1; 17.1–27.1; >27.1)	During pregnancy
	Fish and seafood consumption	Number of times/week, tertiles (<0.25; 0.25–0.83; >0.83)	During pregnancy
	Folic acid supplementation	Number of times/week, tertiles (<0.25; 0.25–0.83; >0.83)	During pregnancy
	Fruit consumption	Number of times/week, tertiles (<0.25; 0.25–0.83; >0.83)	During pregnancy

(Continued)

Table 1. Continued

Exposure family	Compound/factor	Unit	Exposure window/categories
	Legume consumption	Number of times/week, tertiles	During pregnancy
	Meat consumption	(<1.9; 1.9–4.1; >4.1)	During pregnancy
	Vegetables consumption	0: no; 1: yes	During pregnancy
	Walking and/or cycling activity	Number of times/week, tertiles	T1 <sup>a</sup> , T3
	Exercise or sport activity	(<9.6; 9.6–18.2; >18.2)	T1 <sup>a</sup> , T3
		Number of times/week, tertiles	
		(<0.5; 0.5–2.0; >2.0)	
		Number of times/week, tertiles	
		(<6.5; 6.5–10.0; >10.0)	
		Number of times/week, tertiles	
		(<8.8; 8.8–16.5; >16.5)	
		None; sometimes; often; very often	
		Low; medium or high	
Organochlorine compounds (OCs)	DDE (Dichlorodiphenyldichloroethylene)	ng/g of lipids in serum or plasma	Spot pregnancy sample
	DDT (Dichlorodiphenyltrichloroethane)	ng/g of lipids in serum or plasma	Spot pregnancy sample
	HCB (Hexachlorobenzene)	ng/g of lipids in serum or plasma	Spot pregnancy sample
	PCBs (Polychlorinated biphenyls)	ng/g of lipids in serum or plasma	PCB118
		ng/g of lipids in serum or plasma	PCB138
		ng/g of lipids in serum or plasma	PCB153
		ng/g of lipids in serum or plasma	PCB170
		ng/g of lipids in serum or plasma	PCB180 <sup>a</sup>
		ng/g of lipids in serum or plasma	Sum of the PCBs 118, 138, 153, 170 and 180
Polybrominated diphenyl ethers (PBDEs)	PBDE47	ng/g of lipids in serum or plasma	Spot pregnancy sample
	PBDE153	ng/g of lipids in serum or plasma	Spot pregnancy sample
Organophosphate (OP) pesticide metabolites	Diethyl dithiophosphate (DEDTP)	Detected; undetected in urine	Spot pregnancy sample
	DEP (Diethyl phosphate)	µg/g of creatinine in urine	Spot pregnancy sample
	DETP (Diethyl thiophosphate)	µg/g of creatinine in urine	Spot pregnancy sample
	DMDTP (Dimethyl dithiophosphate)	µg/g of creatinine in urine	Spot pregnancy sample <sup>a</sup>
	DMP (Dimethyl phosphate)	µg/g of creatinine in urine	Spot pregnancy sample
	DMTP (Dimethyl thiophosphate)	µg/g of creatinine in urine	Spot pregnancy sample
Per- and poly-fluoroalkyl substances (PFASs)	PFHxS (Perfluorohexane sulfonate)	µg/l in serum or plasma or whole blood	Spot pregnancy sample
	PFNA (Perfluorononanoate)	µg/l in serum or plasma or whole blood	Spot pregnancy sample
	PFOA (Perfluorooctanoate)	µg/l in serum or plasma or whole blood	Spot pregnancy sample
	PFOS (Perfluorooctane sulfonate)	µg/l in serum or plasma or whole blood	Spot pregnancy sample
	PFUnDA (Perfluoroundecanoate)	µg/l in serum or plasma or whole blood	Spot pregnancy sample
Phenols	BPA (Bisphenol A)	µg/g of creatinine in urine	Spot pregnancy sample
	MEPA (Methyl paraben)	µg/g of creatinine in urine	Spot pregnancy sample
	ETPA (Ethyl paraben)	µg/g of creatinine in urine	Spot pregnancy sample
	OXBE (Oxybenzone)	µg/g of creatinine in urine	Spot pregnancy sample
	PRPA (Propyl paraben)	µg/g of creatinine in urine	Spot pregnancy sample
	BUPA (N-Butyl paraben)	µg/g of creatinine in urine	Spot pregnancy sample
	TRCS (Triclosan)	µg/g of creatinine in urine	Spot pregnancy sample
Phthalate metabolites	Sum of DEHP (Di-ethylhexyl phthalate) metabolites	µg/g of creatinine in urine	Spot pregnancy sample
	MBzP (Mono benzyl phthalate)	µg/g of creatinine in urine	Spot pregnancy sample <sup>a</sup>
	MECPP (Mono-2-ethyl 5-carboxypentyl phthalate)	µg/g of creatinine in urine	Spot pregnancy sample
	MEHHP (Mono-2-ethyl-5-hydroxyhexyl phthalate)	µg/g of creatinine in urine	Spot pregnancy sample
		µg/g of creatinine in urine	Spot pregnancy sample

(Continued)

**Table 1.** Continued

Exposure family	Compound/factor	Unit	Exposure window/categories
	MEHP (Mono-2-ethylhexyl phthalate)	µg/g of creatinine in urine	Spot pregnancy sample
	MEOHP (Mono-2-ethyl-5-oxohexyl phthalate)	µg/g of creatinine in urine	Spot pregnancy sample
	MEP (Monoethyl phthalate)	µg/g of creatinine in urine	Spot pregnancy sample
	MiBP (Mono-iso-butyl phthalate)	µg/g of creatinine in urine	Spot pregnancy sample
	MnBP (Mono-n-butyl phthalate)		
	OHMiNP (Mono-4-methyl-7-hydroxyoctyl phthalate)		
	OXOMiNP (Mono-4-methyl-7-oxooctyl phthalate)		

NDVI, Normalized difference vegetation index; NO<sub>2</sub>, nitrogen dioxide; PM<sub>2.5</sub>, particulate matter in the ambient air with an aerodynamical diameter <2.5 µm; PM<sub>10</sub>, particulate matter in the ambient air with an aerodynamical diameter <10 µm; T1, averaged over the first trimester of pregnancy; T2, averaged over the second trimester of pregnancy; T3, averaged over the third trimester of pregnancy.

<sup>a</sup>The variable was excluded from the DSA procedure for collinearity reasons, i.e. either another variable was included measuring the same compound/factor estimated over a different time window/buffer, or it displayed an absolute pairwise (Pearson, polyserial or polychoric coefficient, as appropriate) correlation coefficient >0.90 with another variable (PCB180 and MECCPP displayed a high correlation with PCB153 and MEOHP, respectively).

data.<sup>38</sup> We applied both the Benjamini and Yekutieli false discovery rate correction<sup>39</sup> and the Li *et al.* family wise error rate correction<sup>40</sup> approaches for multiple hypothesis testing.

Investigating for interactions in ExWAS was done in three independent steps. We first performed an ExWAS including a cohort–exposure interaction term; we relied on the  $I^2$  statistic to measure the between-cohort heterogeneity of the exposure association with birth weight (the lower the  $I^2$ , the more consistent the association across cohorts).<sup>41</sup> Second, we performed an ExWAS incorporating a sex–exposure interaction term. Finally, we incorporated successively all exposure–exposure interaction terms (including quadratic terms for each exposure).

### Sensitivity analyses

Analyses were repeated (i) without correcting biomarkers for exposure measurement error, and (ii) excluding the observations related to the mothers who smoked during pregnancy, in order to account for the fact that tobacco smoke contains several metals, including cadmium, arsenic and lead, as well as particulate matter.

All analyses were performed using the R software version 3.4 (www.r-project.org). The R code is provided in [Supplementary Material S5](#), available as [Supplementary data](#) at *IJE* online.

## Results

### Study population

At inclusion, mothers were on average (standard deviation) 30.8 (4.9) years old, with a pre-pregnancy weight of 67.7

(14.3) kg ([Supplementary Table S3](#), available as [Supplementary data](#) at *IJE* online). Among the 1287 children, birth weight was on average 3380 g (5<sup>th</sup>–95<sup>th</sup> centiles, 2550–4240 g, [Figure 1](#)). Children were born on average after 39.6 weeks of gestation; 63 (5%) were born before 37 completed gestational weeks.

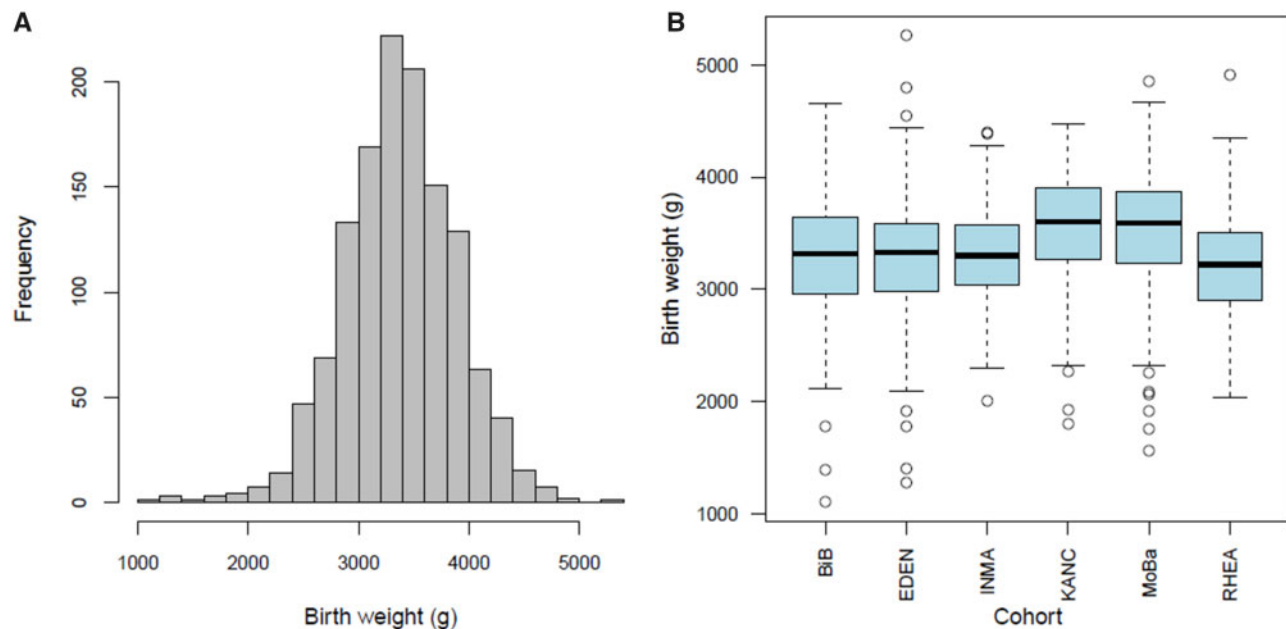
### Association of the exposome with fetal growth from the main model (DSA)

Maternal lead blood concentration was selected by the DSA method; the birth weight change was –98 g [95% confidence interval (CI): –182; –14; exposure selected in 14 out of 100 DSA runs] for a unit increase in log<sub>2</sub>-transformed lead exposure, i.e. for each doubling in lead concentration. When we allowed for two-way interactions, the DSA model did not select any exposure–exposure, sex–exposure or cohort–exposure interaction term.

No exposure displayed an absolute correlation value with lead above 0.35, and the lead–birth weight association was robust to co-exposure adjustment (when adjusting on one other exposure, birth weight decrease was in the 90–104 g range for each doubling in lead concentration).

### Association of the exposome with fetal growth from ExWAS approach

No exposure–outcome association was detected in ExWAS when correcting for multiple hypotheses testing. [Table 2](#) reports the three exposures with an (uncorrected) *P*-value below 5%: lead (mean birth weight change for each doubling in lead concentration: –98 g, 95% CI: –182; –14), particulate matter in the (ambient) air with aerodynamical diameter <2.5 µm (PM<sub>2.5</sub>) absorbance in the third



**Figure 1** Birth weight distribution displayed overall among all cohorts (A) and as boxplots by cohort (B).

**Table 2.** Adjusted associations between the exposome and fetal growth (ExWAS approach)

Exposure variable	Exposure family	Transformation	IQR	ICC	ExWAS	
					Estimate (95% CI) <sup>a</sup>	P-value
Lead	Metals and essential elements	Log <sub>2</sub>	0.5	0.73	-48.6 (-90.5; -6.7)	0.023
PM <sub>2.5</sub> absorbance, 3 <sup>rd</sup> trimester of pregnancy	Atmospheric pollutants	Ln	0.4	<sup>b</sup>	-49.9 (-94.4; -5.5)	0.028
PM <sub>2.5</sub> mass concentration, 3 <sup>rd</sup> trimester of pregnancy	Atmospheric pollutants	None	4.5	<sup>b</sup>	-33.4 (-66.3; -0.6)	0.046

CI, Confidence interval of the coefficient estimate; ICC, intra-class coefficient of correlation; IQR, inter-quartile range of the (normalized and corrected for measurement error) exposure variable.

<sup>a</sup>Estimates are given as a change in mean birth weight (g) for each inter-quartile range (defined over all observations) increase in (normalized and corrected for measurement error) exposure. Only exposures with an uncorrected P-value < 5% are reported. Associations were adjusted for gestational duration (simple and quadratic terms), sex of the newborn, parity, maternal height, maternal weight before pregnancy (using a broken stick model with a knot at 60 kg), maternal smoking during the second trimester of pregnancy, maternal education, season of conception and cohort (fixed effect variable).

<sup>b</sup>Atmospheric pollutants were not assumed to suffer from classical-type measurement error; no measurement error correction based on the ICC was applied.

trimester of pregnancy (mean birth weight change for a log-transformed exposure increase of 0.4: -50 g, 95% CI: -94; -5) and PM<sub>2.5</sub> mass concentration in the third trimester of pregnancy (mean birth weight change for an exposure increase of 4.5 µg/m<sup>3</sup>: -33 g, 95% CI: -66; -1).

Associations of these three exposures with birth weight were homogeneous across cohorts ( $I^2 < 10\%$ , Figure 2, Supplementary Table S4, available as Supplementary data at *IJE* online); they were stronger, both in terms of P-value and of effect size, when tested only on subjects whose exposure had not been imputed (Supplementary Table S5, available as Supplementary data at *IJE* online).

Modification by offspring sex of the effect measure of exposure was suggested (i.e. interaction P-value < 0.05) in

ExWAS for ethyl-paraben ( $P = 0.0060$ ), propyl-paraben ( $P = 0.0062$ ) and for molybdenum ( $P = 0.046$ ), with positive estimated parameters in male births and trends for a negative parameter in female births (Table 3 and Figure 3). No exposure-exposure interaction was detected in ExWAS when correcting for multiple hypotheses testing.

### Sensitivity analyses

When analyses were repeated without correcting biomarkers for exposure measurement error, the differences were that: (i) lead was not identified in the DSA analysis; (ii) dimethyl thiophosphate (DMTP), a non-persistent biomarker of organophosphate pesticides exposure (ICC, 0.20) was identified in ExWAS (adjusted birth weight



**Table 3.** Adjusted effect measure of exposures on birth weight by offspring sex (ExWAS approach)

Exposure	Exposure family	Transformation	IQR	ICC	Interaction P-value	Sex	Estimate (95% CI) <sup>a</sup>	P-value
Ethyl-paraben	Phenols	Log <sub>2</sub>	3.7	0.44	0.0060	Female	-43 (-135; 49)	0.36
						Male	112 (31; 193)	0.007
Propyl-paraben	Phenols	Log <sub>2</sub>	3.0	0.44	0.0062	Female	-61 (-142; 20)	0.14
						Male	92 (11; 173)	0.026
Molybdenum	Metals and essential elements	Log <sub>2</sub>	0.4	<sup>b</sup>	0.046	Female	-19 (47; 8.2)	0.15
						Male	24 (-8.6; 56)	0.17
N-Butyl-paraben	Phenols	Log <sub>2</sub>	4.3	0.51	0.055	Female	-14 (-114; 86)	0.78
						Male	100 (5.6; 195)	0.038
Cadmium <sup>b</sup>	Metals and essential elements	Log <sub>2</sub>	0.7	0.78	0.060	Female	1.0 (-57; 59)	0.97
						Male	-55 (-107; 3.6)	0.036
Lead <sup>b</sup>	Metals and essential elements	Log <sub>2</sub>	0.5	0.73	0.092	Female	-6.0 (-73; 61)	0.86
						Male	-76 (-128; -24)	0.0041
Arsenic	Metals and essential elements	Log <sub>2</sub>	2.3	0.42	0.11	Female	-133 (-258; -6.6)	0.039
						Male	-24 (-149; 101)	0.71
PFOS <sup>b</sup>	Per- and poly-fluoroalkyl substances	Log <sub>2</sub>	0.7	<sup>b</sup>	0.14	Female	-54 (-98; -10)	0.015
						Male	-15 (-57; 28)	0.50
PM <sub>2.5</sub> absorbance in 3rd trimester of pregnancy	Atmospheric pollutants	Ln	0.4	<sup>b</sup>	0.38	Female	-35 (-91; 20)	0.21
						Male	-62 (-114; -10)	0.019
DDT <sup>b</sup>	Organochlorine compounds	Log <sub>2</sub>	1.4	<sup>b</sup>	0.31	Female	-15 (-66; 36)	0.56
						Male	-46 (-90; -2.8)	0.037

CI, Confidence interval of the coefficient estimate; DDT, dichlorodiphenyltrichloroethane; IQR, inter-quartile range of the (normalized and corrected for measurement error) exposure variable; PFOS, perfluorooctane sulfonate.

<sup>a</sup>Estimates are given as a change in mean birth weight (g) for each inter-quartile range (defined over all observations) increase in (normalized and corrected for measurement error) exposure. Only exposures with an uncorrected sex interaction or sex-specific *P*-value < 5% are reported. Associations were adjusted for gestational duration (simple and quadratic terms), sex of the newborn, parity, maternal height, maternal weight before pregnancy (using a broken stick model with a knot at 60 kg), maternal smoking during the second trimester of pregnancy, maternal education, season of conception and cohort (fixed effect variable).

<sup>b</sup>Atmospheric pollutants were not assumed to suffer from classical-type measurement error; for biomarker-based exposures, no ICC was available in the literature.

change, 33.9 g, 95% CI: 2.7; 65.1, *P* = 0.033, [Supplementary Table S5](#), [Supplementary Figure S1](#), available as [Supplementary data](#) at *IJE* online); for this exposure, there was some evidence of an effect measure modification by sex in favour of a stronger positive slope in male births (data not shown); (iii) modification of the effect measure by sex was weakened for ethyl-paraben (interaction *P*-value = 0.20).

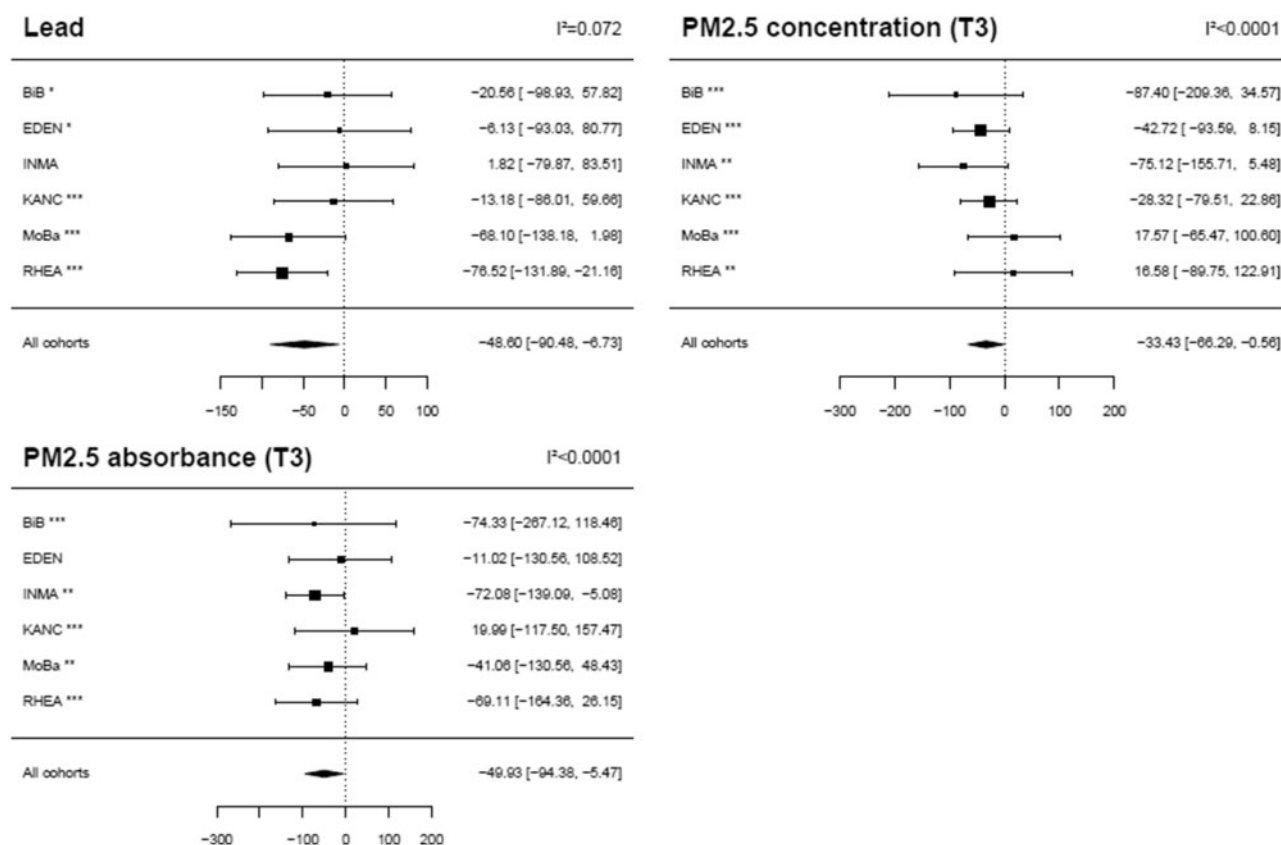
When excluding the observations related to the women who smoked during pregnancy (i.e. restricting to 1124 mother-child pairs, [Supplementary Figure S2](#), available as [Supplementary data](#) at *IJE* online), the ExWAS results yielded similar coefficient values, which was not in favour of a strong residual confounding bias due to active smoking.

## Discussion

To our knowledge, this study is the first to simultaneously consider the possible associations of fetal growth with about 100 exposures from 15 families of environmental factors. The statistical analysis of this cohort of 1287

mother-child pairs pointed towards a decreased fetal growth in association with lead maternal exposure. With a more moderate strength of evidence, we confirmed associations of PM<sub>2.5</sub> absorbance and mass concentration with fetal growth; we provided some evidence for modification by sex of the effects of ethyl- and propyl-parabens, which tended to be positively associated with birth weight in male births only. All of these associations had, to varying extents, some a priori plausibility based on the epidemiological or toxicological literature. There was no evidence of interaction between any pair of exposures, in a context of low statistical power to detect such interactions.<sup>25</sup>

The exposome approach that we adopted relies on validated and sensitive exposure metrics for most exposures, and aimed at better characterizing factors possibly affecting fetal growth and at considering simultaneously exposures that had generally been considered on a compound-by-compound basis in humans. For many exposures, such as some organophosphate or phthalate metabolites, our study was the largest, and for compounds such as mono-4-methyl-7-hydroxyoctyl (OHMiNP) and mono-4-methyl-7-



**Figure 2** Adjusted effect measure of exposures on birth weight by cohort (ExWAS approach). Estimates are given as a change in mean birth weight (g) for each inter-quartile range (defined over all observations) increase in exposure (normalized and corrected for measurement error). Only exposures with an uncorrected  $P$ -value  $< 5\%$  in the main ExWAS (i.e. without cohort-exposure interaction) are reported. Black squares display the coefficient estimates, and the horizontal lines their 95% CIs. The values of the coefficients (95% CI) are given on the right-hand side of the graphs; on the left-hand side, a symbol displays the proportion of missing values that were imputed for the given exposure variable in each cohort (\*\*\*  $< 10\%$  of imputed values, \*\*  $10\text{--}50\%$  of imputed values, \*  $50\text{--}80\%$  of imputed values, no symbol indicates more than  $80\%$  of imputed values). The exposures distribution is displayed in [Supplementary Figure S3](#), available as [Supplementary data](#) at *IJE* online. Associations were adjusted for gestational duration (simple and quadratic terms), sex of the newborn, parity, maternal height, maternal weight before pregnancy (using a broken stick model with a knot at 60 kg), maternal smoking during the second trimester of pregnancy, maternal education, season of conception and cohort (fixed effect variable). T3, averaged over the third trimester of pregnancy.

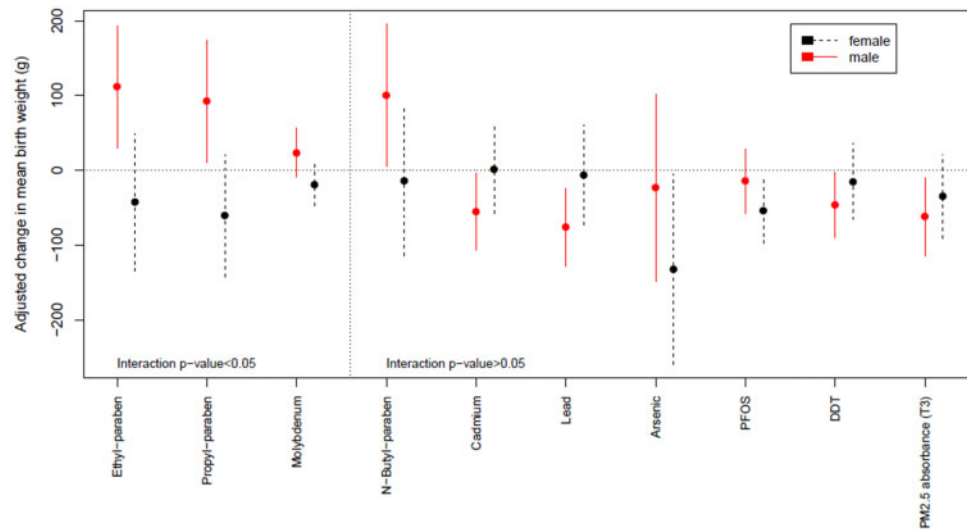
oxooctyl (OXOMiNP) phthalate metabolites, the first ever conducted in humans.

### Strengths and limitations

The main strengths of our study are the large number of exposures considered, the prospective design, the correction for exposure measurement error due to classical-type error, and the fact that biomarkers were assessed generally with very low limits of detection. We had a priori selected the statistical approaches to be used through simulation studies mimicking the situation expected in HELIX in terms of sample size, number of exposures considered and correlation structure within the exposome;<sup>25,27</sup> a difference was that the DSA model considered in these simulation studies had not been stabilized as we did here, which is expected to impact model performances.

Limitations relate to sample size (small given the large number of exposures investigated), limiting the statistical power to detect associations and, to a larger extent, interactions. Chung *et al.* recently estimated that typical existing cohort studies with hundreds of participants were underpowered (power  $< 0.8$ ) for EWAS-related investigations.<sup>42</sup> Consequently, our study should not be interpreted as providing evidence that only lead could influence birth weight. The estimates associated with all exposures are provided ([Supplementary Table S4](#), available as [Supplementary data](#) at *IJE* online) so that they can be used in future exposure-specific meta-analyses. Regarding the statistical analyses, we attempted identifying quadratic dose-response functions, and may therefore have missed exposures displaying a complex non-monotonic dose-response pattern.

In terms of study population, we relied on subgroups selected from six cohorts of pregnant women that are not



**Figure 3** Adjusted effect measure of exposures on birth weight by offspring sex (ExWAS approach). Estimates are given as a change in mean birth weight (g) for each inter-quartile range (defined over all observations) increase in exposure (normalized and corrected for measurement error). Only exposures with an uncorrected sex interaction or sex-specific  $P$ -value  $< 5\%$  are reported. The dot displays the coefficient estimate, and the vertical line its 95% CI. Associations were adjusted for gestational duration (simple and quadratic terms), sex of the newborn, parity, maternal height, maternal weight before pregnancy (using a broken stick model with a knot at 60 kg), maternal smoking during the second trimester of pregnancy, maternal education, season of conception and cohort (fixed effect variable). DDT, dichlorodiphenyltrichloroethane; PFOS, perfluorooctane sulfonate; T3, averaged over the third trimester of pregnancy.

representative of the general population. Such cohorts typically have a participation rate of  $\sim 20\text{--}50\%$  and generally over-represent specific population subgroups, such as subjects with high education level or with interest in health issues. Yet, representativeness is in principle not a validity requirement for aetiological studies.<sup>43</sup>

### Environmental influences on fetal growth

From a statistical perspective, and disregarding the strength of previously published evidence, the most likely association highlighted was that between lead exposure and birth weight.

Such an association was previously reported.<sup>5,20,44</sup> Average lead exposure was  $11\ \mu\text{g/L}$  (95% CI: 4; 26). Without imputation of missing data and correction for exposure measurement error so as to allow comparison across studies, for a unit increase in square root-transformed lead exposure, fetal growth decreased by 45 g (95% CI: 15; 75) in our study population, and by 69 g (95% CI: 46; 183) after restriction to mothers with a lead concentration in the  $0\text{--}10\ \mu\text{g/L}$  range. In a study with a median exposure of  $32\ \mu\text{g/L}$ , Xie *et al.*<sup>44</sup> reported a 148 g (95% CI: 12; 286) birth-weight decrease while Zhu *et al.*<sup>20</sup> reported a 27 g (95% CI: 17; 38) decrease after restriction to women in the  $0\text{--}10\ \mu\text{g/L}$  exposure range. Lead is a recognized reprotoxicant that is readily transferred from maternal blood to the fetus. Its effect on fetal growth could be explained by lead competing with calcium, an essential

component of bones, which might result in alterations of fetal bone formation and consequently restrictions of fetal growth.<sup>20</sup>

From our ExWAS approach, we also observed some evidence for other associations with birth weight. However these need to be considered with more caution given the expected high rate of false findings displayed by the ExWAS approach<sup>27</sup> and the fact that these associations had significance levels way above the multiple hypothesis testing-corrected significance threshold. The use of multiple testing correction procedures is debated in environmental epidemiology;<sup>45</sup> one reason is that these procedures were developed under the null hypothesis that the outcome is associated with none of the exposures considered. This is a priori quite unlikely, given that the compounds that we tested were, for the vast majority, selected because of existing toxicological or epidemiological evidence of adverse health effects (not necessarily related to fetal growth). Among the associations highlighted by the ExWAS approach without multiple testing correction, a negative association between  $\text{PM}_{2.5}$  mass concentration and birth weight is clearly supported by previous evidence,<sup>15</sup> whereas evidence is weaker for an association of  $\text{PM}_{2.5}$  absorbance, a less commonly assessed metric, in relation to fetal growth.<sup>15,46</sup> Several exposure–birth weight associations were previously reported in a small number of studies but were not identified in the present study, such as with PCB.<sup>47</sup> Our confidence intervals were however broad and did not provide strong evidence against such associations.

Some sex-specific exposure association with birth weight were a priori expected from the literature,<sup>48</sup> but in the absence of studies systematically reporting sex-specific estimates, publication bias could explain such apparent sex-specific effects. Trends for positive associations with birth weight were previously reported for ethyl-, propyl- and butyl-parabens (whose concentrations are correlated) in a study among 520 male newborns (consistent with our results), in which female newborns had not been considered.<sup>22</sup> Parabens, used as preservatives in cosmetics, are known to have oestrogenic activity and to promote adipocyte differentiation *in vitro*.<sup>49</sup>

### Exposure assessment

We considered a large number of exposures, and acknowledge that for some of these the exposure metric may have been suboptimal.

We measured the total concentration of metals and essential elements in blood. For arsenic, this does not allow distinguishing inorganic (assumed to be more toxic) from organic arsenic,<sup>50</sup> which would have been more informative. Moreover, arsenic was measured from blood, although urine is considered a more relevant matrix for measuring the internal dose.<sup>51</sup> For these two reasons, our results for arsenic should be interpreted with great caution. Two other limitations affecting the exposure–health assessment of metals and essential elements are cellular homeostasis and lipid levels in blood. Manganese, in particular, is regulated by homeostasis.<sup>52</sup> Consequently, manganese circulating levels are probably a poor biomarker of human exposure (in the sense of the amount of manganese entering the body).<sup>53</sup> Second, because the handling of lipid levels in blood biomarkers is a matter of debate,<sup>54</sup> we tested all blood biomarkers, both adjusted and unadjusted for lipids, without change in the conclusions (results not shown); similarly, results for urinary biomarkers differed little with or without adjustment for creatinine levels.

In addition, because our biomarker-based exposure estimates relied on spot biospecimens collected during pregnancy, we expect measurement error, which may be particularly large for the least persistent compounds considered (see [Supplementary Table S2](#), available as [Supplementary data](#) at *IJE* online, for ICCs reported in the literature). Classical-type exposure measurement error is expected to lead to attenuation bias in exposure–health relations. This applies in particular for bisphenol A, for some phthalate metabolites, such as di-ethylhexyl phthalate (DEHP) metabolites, as well as organophosphate pesticide metabolites: for these compounds, the literature

reports ICCs during pregnancy typically in the 0.1–0.3 range ([Supplementary Table S2](#), available as [Supplementary data](#) at *IJE* online), which is expected to translate into an attenuation bias of 70–90%.<sup>24,55</sup> For many other compounds, the situation is somewhat better but far from optimal; this is in particular the case for parabens, triclosan, and metals such as arsenic and manganese. Regarding other compounds such as lead and cadmium, ICCs in the 0.7–0.8 range have been reported, so that the spot biospecimen that we relied on may provide an estimate of exposure over a longer period than a few days or weeks. Given this issue when assessing the exposure–health associations, we attempted to correct for classical-type error in the statistical analyses through our regression calibration approach relying on ICCs. In the absence of repeated measures of exposure biomarkers during pregnancy, we relied on external estimates of ICCs ([Supplementary Table S2](#), available as [Supplementary data](#) at *IJE* online), resulting in a likely lower efficiency of regression calibration than would have been obtained with study-specific ICCs. Moreover, this method does not correct for any effect of measurement error other than of classical-type, such as expected for outdoor atmospheric pollutant levels for example.

Finally, some exposure variables had a large proportion of missing values ([Supplementary Table S6](#), available as [Supplementary data](#) at *IJE* online). Yet, for the exposures most strongly associated with fetal growth in ExWAS, associations were stronger after restriction to the population with non-imputed exposure values ([Supplementary Table S5](#), available as [Supplementary data](#) at *IJE* online), which is in favour of our missing data imputation procedure not biasing estimates away from the null.

### Conclusions

Our targeted approach focused on a large number of exposures with some a priori evidence for a health effect, based on the human or toxicological literature. A relevant use of our results would be for meta-analyses on specific exposures; all estimates from our ExWAS analyses are provided for such a purpose in [Supplementary Table S4](#), available as [Supplementary data](#) at *IJE* online.

We have illustrated some potential challenges facing exposome studies. Our study allowed bringing an information equivalent to that generated by about 100 single exposure studies, avoiding selective reporting of findings and controlling to some extent the false discovery rate, but expectedly at the cost of reduced power. Our study paves the way for future prospective exposome studies. These should possibly consider much larger populations and rely on

repeated biospecimens collection to limit exposure measurement error for the compounds with the strongest within-subject variability.<sup>24,56,57</sup> From a public health perspective, lead exposure during pregnancy (at the levels encountered in the years 1999–2010, when these pregnancies occurred) may still be a health concern in the EU while the effects of pregnancy exposure to parabens warrant further investigation.

## Supplementary data

Supplementary data are available at *IJE* online.

## Funding

This work was supported by the European Community's Seventh Framework Program (FP7/2007–2013) [grant number 308333 – the HELIX project to M. Vrijheid].

## Acknowledgements

We acknowledge the input of the HELIX consortium. We thank Marta Cirach for her advice regarding atmospheric pollutants and the ESCAPE study consortium (Principal Investigator, Professor B. Brunekreef) for providing regression models to estimate atmospheric pollutants exposure levels, the PHENOTYPE study for proving the methodology to estimate green space exposure and the TAPAS project for providing the methodology to estimate the built environment measures. We acknowledge the support of Région Auvergne Rhône-Alpes for scientific collaborations with Catalonia. We are grateful to all the participating families in the six countries who took part in this cohort study (BiB, EDEN, INMA, KANC, MoBa and RHEA cohorts), especially those families who came in for a clinical examination of their child, who in addition donated blood and urine to this specific study. We are equally grateful to all the fieldworkers for their dedication and efficiency in this study. The HELIX program built on six existing cohorts that received previous funding, including the major cohorts listed here. Born in Bradford (BiB) is only possible because of the enthusiasm and commitment of the children and parents in BiB. We are grateful to all the participants, health professionals and researchers who have made BiB happen. We thank Eleonora P. Uphoff for her role in data collection. The study has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement number 308333—the HELIX project—for data collection and analyses. The EDEN cohort has been supported by grants from MGEN, ANR, ANSES, Fondation de France. We thank Sonia Brishoual, Angélique Serre and Michèle Grosdenier (Poitiers Biobank, CRB BB-0033–00068, Poitiers, France) for biological sample management and Frédéric Millot (principal investigator), Pierre-Jean Saulnier, Elodie Migault, Manuela Grellier Boue and Sandy Bertin (Clinical Investigation Center, Inserm CIC1402, CHU de Poitiers, France) for planning and investigational actions. We are also grateful to Véronique Ferrand-Rigalleau and Noella Gorry (CHU de Poitiers, Poitiers, France) for administrative assistance. The Norwegian Mother, Father and Child Cohort Study (MoBa) is supported by the Norwegian Ministry of Health, and the US National Institutes of Health (NIH) National Institute of Environmental Health Sciences (contract no N01-ES-75558) and National Institute

of Neurological Disorders and Stroke [grant numbers UO1 NS 047537–01, 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 number 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 number 226285 ENRIECO, EUPF7- HEALTH-2012 proposal number 308333 HELIX, FP7 European Union project number 264357 MeDALL), and the Greek Ministry of Health (Program of Prevention of Obesity and Neurodevelopmental Disorders in Preschool Children, Heraklion district, Crete, Greece: 2011– 2014; 'RHEA Plus': Primary Prevention Program of Environmental Risk Factors for Reproductive Health, and Child Health: 2012–2015). L.C. received additional funding from the Southern California Environmental Health Sciences Center [grant number P30ES007048 to L.C.] funded by the National Institute of Environmental Health Sciences. A full roster of the INMA and RHEA projects investigators can be found at [http://www.proyectoinma.org/presentacion-inma/listado-investigadores/en\\_listadoinvestigadores.html](http://www.proyectoinma.org/presentacion-inma/listado-investigadores/en_listadoinvestigadores.html) and <http://www.rhea.gr/en/about-rhea/the-rhea-team/>, respectively.

## Author contributions

M.Vr., M.N., R.S., L.C., C.T., J.Wr. and R.G. designed the HELIX project; M.Vr. was responsible for overall coordination of the project. L.M., I.T.U., J.U., S.A., M.C., L.C., D.D.-G., L.G.-A., R.G., K.B.G., R.R.C.M.E., H.M.M., M.N., O.R., T.R., J.S., C.T., M.Va., J.We., J.Wr., M.Vr. and R.S. contributed to the data collection in the cohorts. E.C., L.S.H., A.K.S. and C.T. analysed chemical biomarkers. I.T.U., M.dC., D.D.G., M.N. and A.V. generated the geo-spatial data. L.A., R.S., X.B., M.N., I.T.U., O.R., J.R.G. and M.Vr. designed the statistical analysis protocol; L.A. analysed the data, with support from X.B., R.S. and V.S. L.A., R.S., X.B., M.Vr., J.R.G. and C.H.-F. contributed analytic tools for statistical analyses. All authors contributed to the writing, critical interpretation of the data, and approved the manuscript.

## Conflict of interest

None declared.

## References

1. Wild CP. Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 2005;14:1847–50.
2. Govarts E, Remy S, Bruckers L *et al.* Combined effects of prenatal exposures to environmental chemicals on birth weight. *Int J Environ Res Public Health* 2016;13:495.
3. Lenters V, Portengen L, Rignell-Hydbom A *et al.* Prenatal phthalate, perfluoroalkyl acid, and organochlorine exposures and term birth weight in three birth cohorts: multi-pollutant models based on elastic net regression. *Environ Health Perspect* 2016;124:365–72.
4. Govarts E, Iszatt N, Trnovec T *et al.* Prenatal exposure to endocrine disrupting chemicals and risk of being born small for gestational age: pooled analysis of seven European birth cohorts. *Environ Int* 2018;115:267–78.

5. Woods MM, Lanphear BP, Braun JM, McCandless LC. Gestational exposure to endocrine disrupting chemicals in relation to infant birth weight: a Bayesian analysis of the HOME Study. *Environ Health* 2017;16:115.
6. Dadvand P, Ostro B, Figueras F *et al.* Residential proximity to major roads and term low birth weight. *Epidemiology* 2014;25: 518–25.
7. Buck Louis GM, Sundaram R, Schisterman EF *et al.* Persistent environmental pollutants and couple fecundity: the LIFE study. *Environ Health Perspect* 2013;121:231–6.
8. Lenters V, Portengen L, Smit L *et al.* Phthalates, perfluoroalkyl acids, metals and organochlorines and reproductive function: a multipollutant assessment in Greenlandic, Polish and Ukrainian men. *Occup Environ Med* 2015;72:385–93.
9. Patel CJ, Bhattacharya J, Butte AJ. An Environment-Wide Association Study (EWAS) on type 2 diabetes mellitus. *PLoS One* 2010;5:e10746.
10. Agier L, Basagaña X, Maitre L *et al.* Early life exposome and lung function in children from the HELIX cohort. *Lancet Planet Health* 2019;3:e81–e92.
11. Patel CJ, Rehkopf DH, Leppert JT *et al.* Systematic evaluation of environmental and behavioural factors associated with all-cause mortality in the United States National Health and Nutrition Examination Survey. *Int J Epidemiol* 2013;42:1795–810.
12. Barouki R, Mel en E, Herceg Z *et al.* Epigenetics as a mechanism linking developmental exposures to long-term toxicity. *Environ Int* 2018;114:77–86.
13. Wilcox AJ. Fertility and pregnancy: an epidemiologic perspective. Oxford: Oxford University Press, 2010, pp. 324.
14. Gluckman PD, Hanson MA. The consequences of being born small - an adaptive perspective. *Horm Res Paediatr* 2006; 65(Suppl 3):5–14.
15. Pedersen M, Giorgis-Allemand L, Bernard C *et al.* Ambient air pollution and low birthweight: a European cohort study (ESCAPE). *Lancet Respir Med* 2013;1:695–704.
16. Beltran AJ, Wu J, Laurent O. Associations of meteorology with adverse pregnancy outcomes: a systematic review of preeclampsia, preterm birth and birth weight. *Int J Environ Res Public Health* 2013;11:91–172.
17. Salmasi G, Grady R, Jones J, McDonald SD; Knowledge Synthesis Group. Environmental tobacco smoke exposure and perinatal outcomes: a systematic review and meta-analyses. *Acta Obstet Gynecol Scand* 2010;89:423–41.
18. Govarts E, Nieuwenhuijsen M, Schoeters G *et al.* Birth weight and prenatal exposure to polychlorinated biphenyls (PCBs) and dichlorodiphenyldichloroethylene (DDE): a meta-analysis within 12 European birth cohorts. *Environ Health Perspect* 2012;120: 162–70.
19. Menai M, Heude B, Slama R *et al.* Association between maternal blood cadmium during pregnancy and birth weight and the risk of fetal growth restriction: the EDEN mother-child cohort study. *Reprod Toxicol* 2012;34:622–27.
20. Zhu M, Fitzgerald EF, Gelberg KH, Lin S, Druschel CM. Maternal low-level lead exposure and fetal growth. *Environ Health Perspect* 2010;118:1471–75.
21. Steenland K, Barry V, Savitz D. Serum perfluorooctanoic acid (PFOA) and birthweight: an updated meta-analysis with bias analysis. *Epidemiology* 2018;29:765–76.
22. Philippat C, Botton J, Calafat AM, Ye X, Charles M-A, Slama R. Prenatal exposure to phenols and growth in boys. *Epidemiology* 2014;25:625–35.
23. Vrijheid M, Casas M, Gascon M, Valvi D, Nieuwenhuijsen M. Environmental pollutants and child health—a review of recent concerns. *Int J Hyg Environ Health* 2016;219:331–42.
24. Perrier F, Giorgis-Allemand L, Slama R, Philippat C. Within-subject pooling of biological samples to reduce exposure misclassification in biomarker-based studies. *Epidemiology* 2016;27:378–88.
25. Barrera-G omez J, Agier L, Portengen L *et al.* A systematic comparison of statistical methods to detect interactions in exposome-health associations. *Environ Health* 2017;16:74.
26. Slama R, Vrijheid M. Some challenges of studies aiming to relate the exposome to human health. *Occup Environ Med* 2015;72: 383–84.
27. Agier L, Portengen L, Chadeau-Hyam M *et al.* A systematic comparison of linear regression-based statistical methods to assess exposome-health associations. *Environ Health Perspect* 2016; 124:1848–56.
28. Maitre L, de Bont J, Casas M *et al.* Human early life exposome (HELIX) study: a European population-based exposome cohort. *BMJ Open* 2018;8:e021311.
29. Vrijheid M, Slama R, Robinson O *et al.* The human early-life exposome (HELIX): project rationale and design. *Environ Health Perspect* 2014;122:535–44.
30. Tamayo-Uria I, Maitre L, Thomsen C *et al.* The early-life exposome: description and patterns in six European countries. *Environ Int* 2019;123:189–200.
31. Haug LS, Sakhi AK, Cequier E *et al.* In-utero and childhood chemical exposome in six European mother-child cohorts. *Environ Int* 2018;121(Pt 1):751–63.
32. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377–99.
33. Carroll R, Ruppert D, Stefanski L, Crainiceanu C. *Measurement Error in Nonlinear Models: A Modern Perspective*. 2nd edn. Boca Raton, Florida: Chapman and Hall, CRC Press, 2006.
34. Basagaña X, Pedersen M, Barrera-G omez J *et al.* Analysis of multicentre epidemiological studies: contrasting fixed or random effects modelling and meta-analysis. *Int J Epidemiol* 2018;47: 1343–54.
35. Sinisi SE, van der Laan MJ. Deletion/substitution/addition algorithm in learning with applications in genomics. *Stat Appl Genet Mol Biol* 2004;3:1–38.
36. Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? *Stat Med* 2008;27: 3227–46.
37. Cohen J, Cohen P, West S.G, Aiken L. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Mahwah, NJ: Laurence Erlbaum Associates, 2003.
38. Meng X-L, Rubin DB. Performing likelihood ratio tests with multiply-imputed data sets. *Biometrika* 1992;79:103–11.
39. Benjamini Y, Yekutieli D. The control of the False Discovery rate in multiple testing under dependency. *Ann Stat* 2001;29: 1165–88.
40. Li M-X, Yeung JMY, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant p-value

- thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet* 2012;131:747–56.
41. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539–58.
  42. Chung M, Buck Louis G, Kannan K, Patel C. Exposome-wide association study of semen quality: Systematic discovery of endocrine disrupting chemical biomarkers in fertility require large sample sizes. *Environ Int* 2019;125:505–14.
  43. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol* 2013;42:1012–14.
  44. Xie X, Ding G, Cui C *et al.* The effects of low-level prenatal lead exposure on birth outcomes. *Environ Pollut* 2013;175:30–34.
  45. Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology* 1990;1:43–46.
  46. Slama R, Morgenstern V, Cyrus J *et al.* Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: a study relying on a land-use regression exposure model. *Environ Health Perspect* 2007;115:1283–92.
  47. Casas M, Nieuwenhuijsen M, Martínez D *et al.* Prenatal exposure to PCB-153, p, p'-DDE and birth outcomes in 9000 mother-child pairs: exposure-response relationship and effect modifiers. *Environ Int* 2015;74:23–31.
  48. Callan AC, Hinwood AL, Heyworth J, Phi DT, Odland JØ. Sex specific influence on the relationship between maternal exposures to persistent chemicals and birth outcomes. *Int J Hyg Environ Health* 2016;219:734–41.
  49. Hu P, Chen X, Whitener RJ *et al.* Effects of parabens on adipocyte differentiation. *Toxicol Sci* 2013;131:56–70.
  50. Hughes MF. Biomarkers of exposure: a case study with inorganic arsenic. *Environ Health Perspect* 2006;114:1790–96.
  51. Marchiset-Ferlay N, Savanovitch C, Sauvart-Rochat M-P. What is the best biomarker to assess arsenic exposure via drinking water? *Environ Int* 2012;39:150–71.
  52. Chen P, Chakraborty S, Mukhopadhyay S *et al.* Manganese homeostasis in the nervous. *J Occup Environ Hyg* 2014;11:210–17.
  53. Ge X, Wang F, Zhong Y *et al.* Manganese in blood cells as an exposure biomarker in manganese-exposed workers healthy cohort. *J Trace Elem Med Biol* 2018;45:41–47.
  54. Schisterman EF, Whitcomb BW, Louis GMB, Louis TA. Lipid adjustment in the analysis of environmental contaminants and human health risks. *Environ Health Perspect* 2005;113:853–57.
  55. Rappaport SM, Symanski E, Yager JW, Kupper LL. The relationship between environmental monitoring and biological markers in exposure assessment. *Environ Health Perspect* 1995;103(Suppl 3):49–53.
  56. Casas M, Basagaña X, Sakhi AK *et al.* Variability of urinary concentrations of non-persistent chemicals in pregnant women and school-aged children. *Environ Int* 2018;121:561–73.
  57. Vernet C, Philippat C, Agier L *et al.* An empirical validation of the within-subject biospecimens pooling approach to minimize exposure misclassification in biomarker-based studies. *Epidemiology* 2019;30:756–67.