PRENATAL EXPOSURE TO ENVIRONMENTAL CONTAMINANTS AND CHILD BEHAVIOUR IN THE HELIX EUROPEAN COHORTS

Paulina Jedynak
Master’s Degree in Bioinformatics and Biostatistics UOC-UB
Clinical Data Analysis

Professional supervisor:
Claire Philippat

Academic supervisor:
Miguel-Ángel Mayer

Submission date:
December 2018
## THESIS WORK SHEET

<table>
<thead>
<tr>
<th>Title:</th>
<th>Prenatal exposure to environmental contaminants and child behaviour In the HELIX European cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author:</td>
<td>Paulina Jedynak</td>
</tr>
<tr>
<td>Professional supervisor:</td>
<td>Claire Philippat</td>
</tr>
<tr>
<td>Academic supervisor:</td>
<td>Miguel-Ángel Mayer</td>
</tr>
<tr>
<td>Submission date:</td>
<td>12/2018</td>
</tr>
<tr>
<td>Degree:</td>
<td>Master’s Degree in Bioinformatics and Biostatistics UOC-UB</td>
</tr>
<tr>
<td>Area:</td>
<td>Clinical Data Analysis</td>
</tr>
<tr>
<td>Language:</td>
<td>English</td>
</tr>
<tr>
<td>Key words:</td>
<td>Prenatal exposure, Environmental contaminants, Child behaviour</td>
</tr>
</tbody>
</table>

Abstract:

=============

Abstract was removed

=============
TABLE OF CONTENTS

THESIS WORK SHEET ......................................................................................................................... i
LIST OF FIGURES .............................................................................................................................. iii
LIST OF TABLES ................................................................................................................................. iii
I. Introduction ........................................................................................................................................ 1
1. Context and justification of the project .............................................................................................. 1
1.1. Child neuropsychological development ...................................................................................... 1
1.2. Risk factors for neurodevelopmental disorders .......................................................................... 1
1.3. Environmental contaminants and neuropsychological deficits .................................................... 2
1.3.1. Personal exposure concept .................................................................................................. 5
2. Objectives of the project .................................................................................................................... 5
3. Materials & Methods ....................................................................................................................... 5
3.1. Study population .......................................................................................................................... 5
3.2. Behavioural outcomes .................................................................................................................. 8
3.2.1. “Externalising” and “Internalising” scores .......................................................................... 8
3.3. Estimation of prenatal exposure to environmental chemicals ...................................................... 8
3.3.1. Dealing with missing exposure values ................................................................................. 8
3.3.2. Dealing with undetected concentrations .............................................................................. 9
3.3.3. Adjustment for lipids and creatinine ..................................................................................... 9
3.3.4. Transformations .................................................................................................................... 9
3.4. Confounders ............................................................................................................................... 9
3.4.1. Coding of confounders ....................................................................................................... 9
3.4.2. Dealing with missing confounder values ............................................................................. 10
3.5. Statistical Methods ...................................................................................................................... 10
3.6. Correction for multiple comparisons ........................................................................................... 10
3.7. Sensitivity analysis ....................................................................................................................... 10
3.8. Statistical software ...................................................................................................................... 10
4. Project scheduling ............................................................................................................................ 10
5. Summary of the results .................................................................................................................... 13
6. Summary of the chapters ............................................................................................................... 13
V. Abbreviations and acronyms list ................................................................................................... 29
VI. Bibliography .................................................................................................................................... 30
LIST OF FIGURES

Figure 1:-------------------------------------------------------------------------------------------------------------------------1
Figure 2:-------------------------------------------------------------------------------------------------------------------------7
Figure 3:-------------------------------------------------------------------------------------------------------------------------7

LIST OF TABLES

Table 1:--------------------------------------------------------------------------------------------------------------------------3
Table 2:--------------------------------------------------------------------------------------------------------------------------6
Table 3:--------------------------------------------------------------------------------------------------------------------------9
I. Introduction

1. Context and justification of the project

1.1. Child neuropsychological development

Brain development begins in the womb, with most of the structural features of the brain formed by the eighth week of conception. These structures continue to grow and develop throughout pregnancy and after birth [1] and it is accepted that the developing brain (in foetal life and early childhood) is more vulnerable to injuries caused by environmental chemicals than the brain of an adult [2], [3]. Healthy neurodevelopment manifests with gradual achievement of psychomotor, cognitive and socio-emotional functions that enable a child to tackle different developmental challenges (see Figure 1). Deficits in any of these domains may cause clinical neurodevelopmental disorders disturbing personal, social, academic or occupational functioning later in life [4]. Disorders of neurobehavioral development affect 10–15% of all births [5] and the most prevalent conditions with onset in the developmental period are: communication impairments, specific learning and motor disorders, intellectual disability, attention-deficit hyperactivity disorder (ADHD) and autism spectrum disorder (ASD) [6], [7]. More recently described impairments, such as disruptive mood and conduct disorders, receive growing attention as potentially also rooting from neurodevelopmental deficiencies [8]. In the past years the prevalence of such disorders as ADHD and ASD has been increasing alarmingly [9] creating a need to pinpoint potential risk factors for such impairments to improve their treatment and prevention.

![Figure 1: Conceptual framework of the neuropsychological development process. Abbreviations: SLI - specific language impairment; ADHD - attention deficit hyperactivity disorder; ASD - autistic spectrum disorder. Reprinted from Ref. [10]](image)

1.2. Risk factors for neurodevelopmental disorders

Root causes of the most common early childhood neurodevelopmental disorders are multifactorial and only partly understood. Research on developmental biology and human and animal physiology showed that, apart from genetic factors, the environment of foetal and early childhood have strong effects on development, health maintenance, and
incidents of disease [11] which is the concept underlying the Developmental Origin of Health and Disease (DOHaD, [12]). Environmental factors that may play a significant role in vulnerability to neurodevelopmental disorders, particularly ASD and ADHD whose heritability was estimated to be around 50% [13]. The complementing half include maternal stress, some infections during pregnancy, malnutrition and maternal exposure to environmental toxicants [14], [15]. Approximately three percent of the developmental disabilities might be the direct result of pre- and postnatal exposure to environmental contaminants [16] that are ubiquitously present in our life, e.g. in food, drinking water or air.

1.3. Environmental contaminants and neuropsychological deficits

Several groups of chemical compounds that are ubiquitously present in the environment were proved or suggested to have adverse neurodevelopmental effects on humans, manifested by behavioural, cognitive or motor deficits (Table 1). According to the most recent reviews of environmental agents affecting human neurodevelopment, a wide range of environmental chemicals, such as organochlorine compounds (OCs), polybrominated diphenyl ethers (PBDEs), perfluoroalkyl substances (PFASs), phthalates, bisphenol A, organophosphorus pesticides (OP Pesticides), metals as well as tobacco smoke are proved to disrupt endocrine function [17]–[20]. Endocrine disrupting chemicals (EDCs) are substances that may interfere with the body’s endocrine system and thereby produce adverse health effects, including developmental, reproductive, neurological and immune effects in humans and animal models [21]. EDCs interfere, temporarily or permanently, with hormonal signalling pathways in the endocrine system. One of the EDCs’ targets is thyroid signalling pathway [17], [22], [23], whose undisturbed functioning is crucial in normal brain development, including neural differentiation and migration and neural connectivity [24], [25]. Recent advances in exposure assessment and biostatistics allow a new approach in epidemiological studies in which effects of exposures to a broad spectrum of environmental contaminants from multiple sources could be evaluated in concert.
### Table 1:
**Chemical compounds studied and examples of their environmental sources**

<table>
<thead>
<tr>
<th>Exposure family</th>
<th>Exposure</th>
<th>Acronym</th>
<th>Exemplary use/source</th>
<th>Adverse effect on neuropsychological development</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCs</td>
<td>Dichlorodiphenyldichloroethylene</td>
<td>DDE</td>
<td>Food contaminated with pesticides</td>
<td>[26]</td>
</tr>
<tr>
<td></td>
<td>Dichlorodiphenyltrichloroethane</td>
<td>DDT</td>
<td></td>
<td>[27]</td>
</tr>
<tr>
<td></td>
<td>Hexachlorobenzene</td>
<td>HCB</td>
<td>Combustion fumes of diesel fuel, wood and coal burning</td>
<td>[28]</td>
</tr>
<tr>
<td></td>
<td>Polychlorinated biphenyl-118</td>
<td>PCB-118</td>
<td>Transformers and capacitors, electrical equipment, industrial oils, cable and thermal insulation materials, adhesives and tapes, some plastics</td>
<td>[26], [29], [30]</td>
</tr>
<tr>
<td></td>
<td>Polychlorinated biphenyl-138</td>
<td>PCB-138</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Polychlorinated biphenyl-153</td>
<td>PCB-153</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Polychlorinated biphenyl-170</td>
<td>PCB-170</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Polychlorinated biphenyl-180</td>
<td>PCB-180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBDEs</td>
<td>Polybrominated diphenyl ether-47</td>
<td>PBDE-47</td>
<td>Flame retardants in textiles, plastics, wire insulation, automobiles</td>
<td>[31], [32], [18]</td>
</tr>
<tr>
<td></td>
<td>Polybrominated diphenyl ether-153</td>
<td>PBDE-153</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFASs</td>
<td>Perfluorononanoate</td>
<td>PFNA</td>
<td>Food packages, stain- and water-repellent fabrics, non-stick products (e.g., Teflon), polishes, waxes, paints, cleaning products</td>
<td>Unequivocal adverse effects not proved: [33], [34], [35]</td>
</tr>
<tr>
<td></td>
<td>Perfluorohexane sulfonate</td>
<td>PFHXS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perfluorooctanoate</td>
<td>PFOA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perfluorooctane sulfonate</td>
<td>PFOS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perfluoroundecanoic acid</td>
<td>PFUnDA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phthalates</td>
<td>Mono benzyl phthalate</td>
<td>MBzP</td>
<td>Plasticizers, solvents for other materials, food containers, vinyl flooring, adhesives, detergents, lubricating oils, automotive plastics, plastic clothes (raincoats), and personal-care products (soaps, shampoos, hair sprays, and nail polishes)</td>
<td>[36]–[43], [44]</td>
</tr>
<tr>
<td></td>
<td>Mono-2-ethyl-5-carboxypentyl phthalate</td>
<td>MECPP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mono-2-ethyl-5-hydroxyhexyl phthalate</td>
<td>MEHHP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mono-2-ethylhexyl phthalate</td>
<td>MEHP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mono-2-ethyl-5-oxohexyl phthalate</td>
<td>MEOHP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Monomethyl phthalate</td>
<td>MEP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mono-iso-butyl phthalate</td>
<td>MiBP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mono-n-butyl phthalate</td>
<td>MnBP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mono-4-methyl-7-hydroxoctyl phthalate</td>
<td>ohMiNP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mono-oxo-isononyl phthalate</td>
<td>oxo-MiNP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Effect not studied</td>
</tr>
</tbody>
</table>

- **OCs**: Organochlorine compounds
- **PBDEs**: Polybrominated diphenyl ethers
- **PFASs**: Perfluoralkyl substances
- **Phthalates**: Plasticizers, solvents, plasticizers for food packaging, adhesives, detergents, lubricating oils, automotive plastics, plastic clothes (raincoats), personal-care products (soaps, shampoos, hair sprays, nail polishes)
<table>
<thead>
<tr>
<th>Phenols</th>
<th>Phenol</th>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bisphenol A</td>
<td>BPA</td>
<td>Polycarbonate containers and coatings (cans, cups), dental sealants, cash</td>
<td>[45]–[51]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>register receipts</td>
<td></td>
</tr>
<tr>
<td>Bisphenol A</td>
<td>BPA</td>
<td>Preservatives in cosmetics like makeup, moisturizers, hair care products,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>shaving products and food products to prevent spoilage</td>
<td></td>
</tr>
<tr>
<td>Parabens:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-Butyl paraben, Ethyl</td>
<td>BUPA,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>paraben, Methyl</td>
<td>ETPA,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>paraben, Propyl</td>
<td>MEPA,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>paraben</td>
<td>PRPA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxybenzone</td>
<td>OXBE</td>
<td>Sunscreens</td>
<td></td>
</tr>
<tr>
<td>OP Pesticides</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diethyl phosphate</td>
<td>DEP</td>
<td>Pesticides including chlorpyrifos, parathion, malathion and diazinon</td>
<td>[55]–[57]</td>
</tr>
<tr>
<td>Diethyl thiophosphate</td>
<td>DETP</td>
<td></td>
<td>[58]</td>
</tr>
<tr>
<td>Dimethyl dithiophosphate</td>
<td>DMDTP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimethyl phosphate</td>
<td>DMP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimethyl thiophosphate</td>
<td>DMTP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arsenic, Cadmium, Cobalt,</td>
<td>As, Cd,</td>
<td>Industrial contamination, electronic devices, pigments, batteries, plastics,</td>
<td>[59]–[64];</td>
</tr>
<tr>
<td>Caesium, Copper, Mercury,</td>
<td>Co, Cs,</td>
<td>tobacco smoke, preservatives, combustion fumes, fertilizers, and many more</td>
<td>Unequivocal</td>
</tr>
<tr>
<td>Potassium, Magnesium,</td>
<td>Cu, Hg,</td>
<td></td>
<td>adverse</td>
</tr>
<tr>
<td>Manganese, Molybdenum,</td>
<td>K, Mg,</td>
<td></td>
<td>effects not</td>
</tr>
<tr>
<td>Sodium, Lead, Selenium,</td>
<td>Mn, Mo,</td>
<td></td>
<td>proved: As,</td>
</tr>
<tr>
<td>Zinc</td>
<td>Na, Pb,</td>
<td></td>
<td>Hg, Co, Mg,</td>
</tr>
<tr>
<td></td>
<td>Se, Zn</td>
<td></td>
<td>Tl.</td>
</tr>
<tr>
<td>Cotinine</td>
<td></td>
<td></td>
<td>[22], [65]</td>
</tr>
<tr>
<td>Tobacco smoke</td>
<td></td>
<td></td>
<td>Effect not</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>studied: Cu,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cs, K, Na</td>
</tr>
</tbody>
</table>

Abbreviations: OCs - organochlorine compounds; PBDEs - polybrominated diphenyl ethers; PFASs - perfluoroalkyl substances; OP Pesticides - organophosphorus pesticides
1.3.1. Personal exposome concept

Pregnant women from the general population are exposed to a broad range of environmental chemicals some of which can cross the placental barrier and affect the foetus [70], [71]. Until now, epidemiological studies on the effects of environmental contaminants on child neurodevelopment usually consider each exposure (or family of exposures) separately, and therefore provide only a fragmented view of environment-health associations. New studies that would simultaneously take into account multiple exposures were needed to improve the understanding of the environmental risk factors for neurodevelopmental disorders. The personal exposome concept, defined as the totality of non-genetic exposures from conception throughout the life course [72], comprises a wide range of environmental exposures from the prenatal period onwards, providing a holistic consideration of many exposures simultaneously which in turn helps to understand the complex environmental component of disease aetiology. One of the first epidemiological study implementing the personal exposome concept is the European human early life exposome (HELIX) project [73]. The aim of the HELIX project was to measure and describe exposure to a multitude of environmental contaminants during early life (pregnancy and childhood) in 6 prospective European cohorts and associate these exposures with child health outcomes [71], [74]. Such approach derives from the exposome concept and aims at estimating effect of combined exposures (e.g., based on biological pathways) on health and are likely to be more realistic than the classical approach focused on each chemical separately.

2. Objectives of the project

The focus of this study was to assess the associations between prenatal exposure to a wide range of environmental chemicals, including some that have been identified as neurodevelopmental disruptors, and child externalising and internalising behaviour between three and seven years of age in the European HELIX cohorts.

3. Materials & Methods

3.1. Study population

The study population consisted of 1321 mother-child pairs selected from the Human Early Life Exposome (HELIX) study [73]. The HELIX cohorts consist of 31 472 pregnant women recruited during pregnancy between 2003 and 2010, in the six European countries: EDEN (France), RHEA (Greece), KANC (Lithuania), MoBa (Norway), INMA (Spain) and BiB (the UK). The characteristics of the subcohorts contributing to the HELIX cohort are presented in the Table 2 [75].
Table 2: The characteristics of the sub-cohorts contributing to the HELIX project. Reprinted from Ref. [75]

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Recruitment in original cohort</th>
<th>Exclusions made during recruitment</th>
<th>Years of birth</th>
<th>Region covered by HELIX</th>
<th>No. of births in HELIX entire cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIB, UK</td>
<td>All pregnant women who attended the oral glucose tolerance test clinic at Bradford Royal Infirmary in weeks 26–28 of pregnancy.</td>
<td>Women who planned to move away from Bradford before birth were excluded.</td>
<td>2007–2010</td>
<td>Bradford</td>
<td>10849</td>
</tr>
<tr>
<td>EDEN, France</td>
<td>Pregnant women who attended prenatal care at the University hospitals of Nancy and Poitiers recruited before 24 weeks of amnionrhoea.</td>
<td>Twin pregnancies, women with known diabetes before pregnancy, insufficient French language skills and intention to move away from the recruitment area were excluded.</td>
<td>2003–2006</td>
<td>Nancy and Poitiers, urban areas</td>
<td>1900</td>
</tr>
<tr>
<td>INMA, Spain</td>
<td>Pregnant women who attended a prenatal care centre in the study region during weeks 6–10 of pregnancy.</td>
<td>Women who resided or intended to deliver outside the study area, who were aged under 16 years, who had twin or multiple pregnancies, who had assisted reproduction or who had communication problems were excluded.</td>
<td>2003–2008</td>
<td>Gipuzkoa Sabadell, Valencia</td>
<td>2083</td>
</tr>
<tr>
<td>KANC, Lithuania</td>
<td>Pregnant women who attended one of four prenatal care clinics affiliated to the hospitals of the Kaunas University of Medicine during first trimester of pregnancy.</td>
<td>Women who lived outside Kaunas municipality, had medical records of pregnancy induced hypertension and/or diabetes were excluded.</td>
<td>2007–2008</td>
<td>Kaunas</td>
<td>4107</td>
</tr>
<tr>
<td>MoBa, Norway</td>
<td>Recruitment at the first ultrasound (US) scan, i.e., during the 17–18 weeks of gestation. All women who gave singleton births in the participating maternity units.</td>
<td>None</td>
<td>1999–2008</td>
<td>Oslo</td>
<td>11095</td>
</tr>
<tr>
<td>RHEA, Greece</td>
<td>Pregnant women who attended US examination before 15 week of pregnancy with residence in and near Heraklion at Crete.</td>
<td>Women who were aged under 16 years or who had communication problems were excluded.</td>
<td>2007–2008</td>
<td>Heraklion</td>
<td>1458</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>31472</td>
</tr>
</tbody>
</table>

BIB, Born in Bradford; EDEN, Étude des Déterminants pré et postnataux du développement et de la santé de l’Enfant; INMA, Infancia y Medio Ambiente; KANC, Kaunas cohort; MoBa, Norwegian Mother and Child Cohort Study.

The study design is depicted in Figure 2. For a sub-cohort of 1321 mother-child pairs, biomarkers of exposure to several contaminants were measured in pregnant women’s blood and urine. Among them, 708 had behavioural outcomes of their children assessed between three to seven years of age [74] using the Strengths and Difficulties Questionnaire (SDQ) [76]. Inclusion criteria are illustrated in Figure 3.
Figure 2: Study design for the HELIX cohorts

HELIX original study population (N = 1321)

- France (EDEN): 202
- Greece (RHEA): 202
- Lithuania (KANC): 206
- Norway (MoBa): 272
- Spain (INMA): 224
- UK (BiB): 215

SDQ scores not available (N = 591)

- France (EDEN): 6
- Greece (RHEA): 27
- Lithuania (KANC): 119
- Norway (MoBa): 272
- Spain (INMA): 3
- UK (BiB): 164

SDQ scores available (N = 721)

- France (EDEN): 196
- Greece (RHEA): 170
- Lithuania (KANC): 83
- Norway (MoBa): 0
- Spain (INMA): 221
- UK (BiB): 51

Unknown sex of child and/or age at SDQ assessment (N = 13)

- France (EDEN): 3
- Greece (RHEA): 2
- Lithuania (KANC): 0
- Norway (MoBa): 0
- Spain (INMA): 3
- UK (BiB): 5

Figure 3: Flowchart illustrating selection of the study population from the HELIX cohorts
3.2. Behavioural outcomes

Behaviour was assessed using the Strengths and Difficulties Questionnaire (SDQ) [76] between three and seven years of age. The SDQ consists of the 25 items that were combined into four difficulties sub-scores: conduct problems, hyperactivity–inattention, peer relationship problems and emotional symptoms sub-scores and one strength sub-score: prosocial behaviour. Each item can be evaluated as “Not True”, “Somewhat True” and “Certainly True”. “Somewhat True” is always scored as 1, but the scoring of “Not True” and ‘Certainly True’ varies with the item. For each of the 5 scales the score can range from 0 to 10 if all items were completed. Higher scores on the behavioural scales indicate increased behaviour problems.

3.2.1. “Externalising” and “Internalising” scores

The externalising score is the sum of the conduct and hyperactivity–inattention sub-scores and ranges from 0 to 20 and. The internalising score also ranges from 0 to 20 and is the sum of the emotional and peer problems sub-scores. Using these two amalgamated scales may be preferable to using the four separate scales in community samples [77] and therefore were used in the present study.

3.3. Estimation of prenatal exposure to environmental chemicals

We assessed 52 environmental contaminants that are ubiquitously present in our lives. We added one grouping variable consisting of the molar sum of 4 di(2-ethylhexyl) phthalate (DEHP) metabolites (ΣDEHP). The studied compounds came from the following groups: organochlorine compounds (OCs), polybrominated diphenyl ethers (PBDEs), perfluoroalkyl substances (PFASs), phthalates, phenols, organophosphorous pesticides (OP Pesticides), metals, and cotinine - a marker of smoking. OCs, PFASs PBDEs and metals are long half-life chemicals and were measured in maternal blood while concentrations of compounds with short half-life (OP Pesticides, phenols and phthalates) were assessed in maternal urine. The maternal samples from the various cohorts were collected at different time points [71]; INMA, KANC, RHEA recruited during the first trimester, EDEN recruited during the first and second trimesters and BIB recruited during the second and third trimesters. The mean (± SD) gestational ages (weeks) at blood sample collection were 26.6 (± 1.4), 26.1 (± 1.2), 13.7 (± 2.0), 39.4 (± 1.3) and 14.1 (± 3.7) for BIB, EDEN, INMA, KANC and RHEA, respectively. Urine sample collection times were the same as for blood for BIB, EDEN and RHEA cohorts. In INMA cohort urine was collected at mean (± SD) gestational age (weeks) of 34.2 (± 1.3) while for KANC cohort urine samples were not collected at all. For more details of samples processing please refer to the HELIX project please refer to Huag et al. In Table 1 (see “Introduction, 1.3 Environmental contaminants and neuropsychological deficits”) the studied exposure biomarkers are listed together with examples of their environmental sources.

3.3.1. Dealing with missing exposure values

For some sub-cohorts and/or participants within a sub-cohort, blood or urine samples were missing preventing us from measuring some biomarkers of exposure. As it has been shown that multiple imputation is superior to excluding a variable or a stratum [78], [79], exposures for which less than 70% of observations were missing due to not being measured were imputed using multiple imputation (20) relying on the method of chained equations algorithm (mice R package) [80], [81]. The algorithms used for the imputation were the following: logistic regression for binomial variables; predictive mean matching for continuous variables; polytomous logistic regression for multinomial variables. The imputation predictors list was limited to only those variables that were not highly correlated with other variables to prevent imputation errors due to multicollinearity. The condition for variable removal from the predictors list was based on the number of variables it was highly correlated with (Spearman corr. coefficient rho ≥ 0.6). If there was more than 1 variable with the same number of correlated partners, variable with higher number of complete cases number was favoured. Highly correlated variables were removed from prediction of all other imputed variables and not only from the imputation of their correlated counterparts (as such approach produced better convergence plots). An additional restriction of the predictor list was applied based on the influx/outflux plot [82]. Variables with
low predictive ability (outflux < 0.5) were removed. Passive imputation was applied to impute transformed (e.g. log2) or combined (e.g. sum DEHP) versions of the data [83]. Diagnostics was performed at the end of the imputation process in order to check for imputation convergence and plausibility of the newly generated values (see Error! Reference source not found. and Error! Reference source not found.).

3.3.2. Dealing with undetected concentrations

Exposures for which more than 80% of observations were below the limit of detection (LOD) were dichotomized into detected/undetected, as advised by the HELIX consortium. For the other variables, values below the LOD were singly imputed using a quantile regression approach for the imputation of left-censored missing data.

3.3.3. Adjustment for lipids and creatinine

To limit the impact of between-subject variations in urine and blood dilution, the biomarker concentrations were adjusted for creatinine concentration (urine) or total fat percentage (blood) [84], [85]. This adjustment was done by dividing the urine-based biomarker concentrations by the creatinine concentrations and serum-based biomarker concentrations by total fat percentage. Chemicals from the PFASs family and metals, although collected from blood, were not adjusted for lipids as they are not lipophilic.

3.3.4. Transformations

The continuous biomarker concentrations were log-transformed (base 2) prior to analysis to limit the effect of outliers.

3.4. Confounders

A confounder is defined as variable that is associated with both the exposure and the outcome but that is not a consequence of either [86]. Information on potential confounders such as maternal behaviours and socio-economic status was collected by interview during study visits and self-report questionnaires completed by the parents. Candidate confounders were selected based on literature review.

3.4.1. Coding of confounders

To establish the best coding for continuous confounders, all continuous adjustment factors were fitted in step functions to examine the relationship between the covariates and the outcomes (see Error! Reference source not found. for example of different types of coding of maternal Body Mass Index - BMI). The best coding was chosen for the confounder basing on the lower Akaike information criterion (AIC), graphical relationship between covariate and outcome and the degrees of freedom lost. Using this approach, we selected an appropriate coding for the confounders included in our multiple models as summarised in Table 3.

<table>
<thead>
<tr>
<th>Table 3: Confounders list</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confounder group</td>
</tr>
<tr>
<td>Maternal behaviour</td>
</tr>
<tr>
<td>Breastfeeding</td>
</tr>
<tr>
<td>Socio-economic status</td>
</tr>
<tr>
<td>Maternal work status</td>
</tr>
<tr>
<td>Persons in household</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>
### 3.4.2. Dealing with missing confounder values

As for the missing exposure concentrations, the missing values for confounder variables were imputed using multiple imputation based on the chained equations algorithm (mice R package), as described previously.

### 3.5. Statistical Methods

The outcome variables studied in the present project are behavioural SDQ scores that are non-negative, integer counts. Such data can possibly follow Poisson or negative binomial distribution. The negative binomial is a generalization of Poisson regression which loosens the restrictive assumption that the variance is equal to the mean made by the Poisson model [87]. In our case, the assumption of equality of the variance and the mean is not met and therefore we used the negative binomial regression model. In our study we performed separate adjusted negative binomial regressions to study the associations between each urinary or blood exposure concentration and the behavioural outcomes (SDQ scores). For SDQ scores, a positive effect estimate indicated an increased risk for behavioural problems while a negative effect estimate indicated a protective association. The effects were expressed as incidence rate ratios (IRR) for a doubling in biomarkers concentrations.

### 3.6. Correction for multiple comparisons

Since the simple regression was run for over 50 exposures a correction for multiple comparisons was applied to decrease the false discovery rate (FDR). As several of the explanatory variables were correlated we applied a modified Bonferroni p value adjustment according to an effective number of independent tests instead of the total tests number [88].

### 3.7. Sensitivity analysis

Good modelling practice requires providing an evaluation of the confidence in the proposed model [89]. In our sensitivity analysis we confronted the observed associations between exposures and SDQ scores in the imputed dataset with results obtained in complete case analysis to exclude the bias caused by the imputation process. Moreover, we run additional models with an interaction term between sex and exposure to investigate sex specific associations. Exposures whose interaction coefficients were significantly different from 1 (uncorrected p value < 0.2) were analysed in separate simple regression analyses stratified by sex. Finally, to make sure that our results were not driven by the effect observed for one cohort, we run models based on the leave-one-cohort-out method and compared these results to our main analysis’ outcomes. Exposures whose coefficients were significantly different from 1 in the main analysis (uncorrected p value < 0.05) were re-analysed for subsets of data with one cohort removed at a time.

### 3.8. Statistical software

Statistical analyses were carried out using an open source software R version 3.5.1 and R-Studio version 1.1.456.

### 4. Project scheduling

The total time reserved for the experimental part of the project was approximately 9 weeks. Below an approximate timing for each part of the study is presented.

---

<table>
<thead>
<tr>
<th>Parity</th>
<th>Categorical (nulliparous / 1 child/ ≥ 2 children)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age</td>
<td>Numerical</td>
</tr>
<tr>
<td>Maternal Body Mass Index</td>
<td>Categorical (underweight/ normal weight/ overweight/ obesity)</td>
</tr>
<tr>
<td>Child’s sex</td>
<td>Categorical (male/ female)</td>
</tr>
<tr>
<td>Child’s age at SDQ assessment</td>
<td>Numerical</td>
</tr>
</tbody>
</table>
1) Described the levels of exposure to environmental contaminants among mother and children of the HELIX mother-child cohort (2 weeks)
   a. Provided a descriptive summary of each environmental exposure (1 week)
      • Described the exposure variables – identified samples missing and below level of detection, described the distributions of the exposure concentrations, flagged potential outliers, measured correlations between exposures
   b. Prepared the exposures data for statistical analyses (1 week)
      • Dichotomised exposures with too high number of samples whose concentration is below the level of detection
      • Transformed continuous exposures to normalize distributions
2) Identified covariates that could potentially influence exposure to environmental contaminants (2.5 weeks)
   c. Selected candidate confounders (1 week)
      • Searched the literature in order to identify candidate confounders
      • Tested each potential confounder identified in the previous step in a simple regression against the behavioural outcome in order to establish the associations between the factors influencing exposure to environmental contaminants and behavioural score
   d. Chose the confounders set and their optimal coding (1.5 weeks)
      • Tried different coding for the selected confounders (continuous, step function, etc.) and used AIC in order to establish the coding best reflecting the nature of the variable structure
      • Run multiple regression of different sets of confounders against the dependent variable in order to establish the configuration of confounders that describes best the variability of the behavioural outcome
3) Determined associations between the personal exposome and behavioural outcomes (4.5 weeks)
   e. Pre-processed the final set of exposures, covariates and outcome variables (1 week)
      • Adjusted the exposure concentrations for urine and blood dilution using creatinine and total fat percentage, respectively
      • Imputed the missing values and/or values below the limit of detection; Performed diagnostics of the imputation process and outcomes
   f. Described associations between each exposure and the behavioural outcome (1.5 weeks)
      • Run exposome-wide association study (ExWAS) on outcome variable vs. each exposure adjusted for the confounders in order to establish the potential associations
   g. Provided evaluation of the proposed model (1 week)
      • Run sensitivity analyses including factors that could possibly confound the observed associations between exposome and behavioural outcome

Below a graphical representation of the project scheduling is presented:
1) Describe the levels of exposure to environmental contaminants among mother and children of the HELIX mother-child cohort
   a. Provide a descriptive summary of each environmental exposure
      • Describe the exposure variables – identify samples missing and below level of detection, describe the distributions of the exposure concentrations, flag potential outliers, measure correlations between exposures
   b. Prepare the exposures data for further analyses
      • If necessary, dichotomise exposures with too high number of samples whose concentration is below the level of detection
      • If necessary, transform continuous exposures to normalize distributions and eliminate outliers

2) Identify covariates that could potentially influence exposure to environmental contaminants
   c. Select candidate confounders
      • Search the literature in order to identify candidate confounders
      • Test each potential confounder identified in the previous step in a simple regression against the behavioural outcome in order to establish the associations between the factors influencing exposure to environmental contaminants and behavioural score
   d. Choose the confounders set and their optimal coding
      • Try different coding for the selected confounders (continuous, step function, etc.) and use AIC in order to establish the coding best reflecting the nature of the variable structure
      • Run multiple regression of different sets of confounders against the dependent variable in order to establish the configuration of confounders that describes best the variability of the behavioural outcome

3) Determine associations between the personal exposome and behavioural outcomes
   e. Pre-process the final set of exposures, covariates and outcome variables
      • Adjust the exposure concentrations for urine and blood dilution using creatinine and total, respectively
      • Impute the missing values and/or values below the limit of detection; Perform diagnostics of the imputation process and outcomes
   f. Describe associations between each exposure and the behavioural outcome
      • Run exposome-wide association study (ExWAS) on outcome variable vs. each exposure adjusted for the confounders in order to establish the potential associations
   g. Provide evaluation of the proposed model
      • Run sensitivity analysis including factors that could possibly confound the observed associations between exposome and behavioural outcome
   h. OPTIONALLY: Describe associations between a set of multiple exposures and the behavioural outcome
      • Run elastic net regularized regression model (ENET) on a set of exposures against the dependent variable adjusted for the confounders in order to establish the associations between personal exposome and the behavioural outcome (with regard to false discovery rate correction)
5. **Summary of the results**

We provided a detailed description of the studied population including maternal exposures and child behavioural outcomes; We identified confounders that could potentially influence exposure to environmental contaminants and behavioural outcomes; We described associations between prenatal exposures and postnatal behavioural outcomes; We validated the observed results by running several sensitivity analyses.

6. **Summary of the chapters**

Chapter II – Results including the description of the studied population (part I) and of the associations between prenatal exposure to environmental contaminants and child behaviour (part II).

Chapter IV – Conclusions including the highlights of the present project and proposed approaches for future work.

Chapter V – Abbreviations and acronyms list including the terms used in the present work, in alphabetical order.

Chapter VI – Bibliography including the full list of works and websites consulted for this work.

Chapter VII – Annexes including results for the imputation diagnostics, example of comparison of different coding of a covariate, distribution of behavioural outcomes in the studied population, distribution of adjusted and transformed exposure concentrations, numerical summary of all simple regressions run on each exposure against behavioural outcome, and results of the sensitivity analyses.
Pages 14 – 28 were removed
### V. Abbreviations and acronyms list

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADHD</td>
<td>Attention-deficit hyperactivity disorder</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike information criterion</td>
</tr>
<tr>
<td>As</td>
<td>Arsenic</td>
</tr>
<tr>
<td>ASD</td>
<td>Autism spectrum disorder</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>BPA</td>
<td>Bisphenol A</td>
</tr>
<tr>
<td>BUPA</td>
<td>N-Butyl paraben</td>
</tr>
<tr>
<td>CBCL</td>
<td>Children behaviour checklist</td>
</tr>
<tr>
<td>Cd</td>
<td>Cadmium</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>Co</td>
<td>Cobalt</td>
</tr>
<tr>
<td>Cs</td>
<td>Caesium</td>
</tr>
<tr>
<td>Cu</td>
<td>Copper</td>
</tr>
<tr>
<td>DDE</td>
<td>Dichlorodiphenyldichloroethylene</td>
</tr>
<tr>
<td>DDT</td>
<td>Dichlorodiphenyltrichloroethane</td>
</tr>
<tr>
<td>DEHP</td>
<td>Di-2-ethylhexyl-phthalate</td>
</tr>
<tr>
<td>DEP</td>
<td>Diethyl phosphate</td>
</tr>
<tr>
<td>DETP</td>
<td>Diethyl thiophosphate</td>
</tr>
<tr>
<td>DMTP</td>
<td>Dimethyl thiophosphate</td>
</tr>
<tr>
<td>DMP</td>
<td>Dimethyl phosphate</td>
</tr>
<tr>
<td>DOHaD</td>
<td>Developmental origin of health and disease</td>
</tr>
<tr>
<td>EDCs</td>
<td>Endocrine disrupting chemicals</td>
</tr>
<tr>
<td>ETPA</td>
<td>Ethyl paraben</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>HCB</td>
<td>Hexachlorobenzene</td>
</tr>
<tr>
<td>HELIX</td>
<td>European human early life exposome</td>
</tr>
<tr>
<td>Hg</td>
<td>Mercury</td>
</tr>
<tr>
<td>IRR</td>
<td>Incidence rate ratio</td>
</tr>
<tr>
<td>K</td>
<td>Potassium</td>
</tr>
<tr>
<td>LOD</td>
<td>Limit of detection</td>
</tr>
<tr>
<td>MBzP</td>
<td>Mono benzyl phthalate</td>
</tr>
<tr>
<td>MECPP</td>
<td>Mono-2-ethyl 5-carboxypentyl phthalate</td>
</tr>
<tr>
<td>MEHP</td>
<td>Mono-2-ethylhexyl phthalate</td>
</tr>
<tr>
<td>MEOHP</td>
<td>Mono-2-ethyl-5-oxohexyl phthalate</td>
</tr>
<tr>
<td>MEP</td>
<td>Monoethyl phthalate</td>
</tr>
<tr>
<td>MEPA</td>
<td>Methyl paraben</td>
</tr>
<tr>
<td>Mg</td>
<td>Magnesium</td>
</tr>
<tr>
<td>MiBP</td>
<td>Mono-isobutyl phthalate</td>
</tr>
<tr>
<td>Mn</td>
<td>Manganese</td>
</tr>
<tr>
<td>MnBP</td>
<td>Mono-n-butyl phthalate</td>
</tr>
<tr>
<td>Mo</td>
<td>Molybdenum</td>
</tr>
<tr>
<td>Na</td>
<td>Sodium</td>
</tr>
<tr>
<td>OCs</td>
<td>Organochlorine compounds</td>
</tr>
<tr>
<td>o xo-MINP</td>
<td>Mono-4-methyl-7-oxooctyl phthalate</td>
</tr>
<tr>
<td>OPC<strong>P</strong>es</td>
<td>Organophosphorus pesticides</td>
</tr>
<tr>
<td>OXBE</td>
<td>Oxybenzone</td>
</tr>
<tr>
<td>Pb</td>
<td>Lead</td>
</tr>
<tr>
<td>PBDE-153</td>
<td>Polybrominated diphenyl ether-153</td>
</tr>
<tr>
<td>PBDE-47</td>
<td>Polybrominated diphenyl ether-47</td>
</tr>
<tr>
<td>PBDEs</td>
<td>polybrominated diphenyl ethers</td>
</tr>
<tr>
<td>PCB-118</td>
<td>Polychlorinated biphenyl-118</td>
</tr>
<tr>
<td>PCB-138</td>
<td>Polychlorinated biphenyl-138</td>
</tr>
<tr>
<td>PCB-153</td>
<td>Polychlorinated biphenyl-153</td>
</tr>
<tr>
<td>PCB-170</td>
<td>Polychlorinated biphenyl-170</td>
</tr>
<tr>
<td>PCB-180</td>
<td>Polychlorinated biphenyl-180</td>
</tr>
<tr>
<td>PFASs</td>
<td>Perfluoroalkyl substances</td>
</tr>
<tr>
<td>PFHXS</td>
<td>Perfluorohexane sulfonate</td>
</tr>
<tr>
<td>PFNA</td>
<td>Perfluoronoranoate</td>
</tr>
<tr>
<td>PFOA</td>
<td>Perfluorooctanoate</td>
</tr>
<tr>
<td>PFOS</td>
<td>Perfluorooctane sulfonate</td>
</tr>
<tr>
<td>PFUnDA</td>
<td>Perfluroundecanoate</td>
</tr>
<tr>
<td>PRPA</td>
<td>Propyl paraben</td>
</tr>
<tr>
<td>SDQ</td>
<td>Strengths and Difficulties Questionnaire</td>
</tr>
<tr>
<td>Se</td>
<td>Selenium</td>
</tr>
<tr>
<td>SLI</td>
<td>Specific language impairment</td>
</tr>
<tr>
<td>TRCS</td>
<td>Triclosan</td>
</tr>
<tr>
<td>Zn</td>
<td>Zinc</td>
</tr>
</tbody>
</table>
VI. Bibliography

[24] V. Somogyi et al., “Bisphenol A influences oestrogen- and thyroid hormone-regulated thyroid hormone receptor


45, no. 5, pp. 91–99, Dec. 2014.


Pages I – XII were removed