



HAL
open science

Using methylome data to inform exposome-health association studies: An application to the identification of environmental drivers of child body mass index

Solène Cadiou, Mariona Bustamante, Lydiane Agier, Sandra Andrusaityte, Xavier Basagaña, Angel Carracedo, Leda Chatzi, Regina Grazuleviciene, Juan R Gonzalez, Kristine B Gutzkow, et al.

► To cite this version:

Solène Cadiou, Mariona Bustamante, Lydiane Agier, Sandra Andrusaityte, Xavier Basagaña, et al.. Using methylome data to inform exposome-health association studies: An application to the identification of environmental drivers of child body mass index. *Environment International*, 2020, 138, pp.105622. <10.1016/j.envint.2020.105622>. <hal-03179041>

HAL Id: hal-03179041

<https://hal.science/hal-03179041v1>

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

1 Title:

2 Using methylome data to inform exposome-health association studies: an
3 application to the identification of environmental drivers of child body mass
4 index.

5 Authors: Solène Cadiou^a, Mariona Bustamante^{b,c,d}, Lydiane Agier^a, Sandra Andrusaityte^e, Xavier
6 Basagaña^{b,c,d}, Angel Carracedo^{fg}, Lida Chatzi^h, Regina Grazuleviciene^e, Juan R Gonzalez^{b,c,d},
7 Kristine B. Gutzkowⁱ, Léa Maitre^{b,c,d}, Dan Mason^j, Frédéric Millot^{k,l}, Mark Nieuwenhuijsen^{b,c,d},
8 Eleni Papadopoulouⁱ, Gillian Santorelli^j Pierre-Jean Saulnier^{k,l,m,n}, Marta Vives^d, John Wright^j,
9 Martine Vrijheid^{b,c,d}, Rémy Slama^a

10 **Affiliations:**

- 11 a. Team of environmental epidemiology, IAB, Institute for Advanced Biosciences, Inserm,
12 CNRS, CHU-Grenoble-Alpes, University Grenoble-Alpes, Grenoble, France.
- 13 b. ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain.
- 14 c. Universitat Pompeu Fabra (UPF), Barcelona, Spain.
- 15 d. CIBER Epidemiología y Salud Pública (CIBERESP), Spain.
- 16 e. Department of Environmental Sciences, Vytautas Magnus University, Kaunas, Lithuania.
- 17 f. Fundación Pública Galega de Medicina Xenómica (SERGAS), Santiago, Spain
- 18 g. Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Universidad
19 de Santiago de Compostela, Santiago, Spain
- 20 h. Department of Social Medicine, Faculty of Medicine, University of Crete, Heraklion, Crete,
21 Greece.
- 22 i. Norwegian Institute of Public Health, Oslo, Norway.
- 23 j. Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust,
24 Bradford, United Kingdom.
- 25 k. CHU Poitiers, Clinical Investigation Centre, CIC 1402, Poitiers, France.
- 26 l. Poitiers University, Clinical Investigation Centre CIC 1402, Poitiers, France.
- 27 m. INSERM, CIC 1402, F-86000 Poitiers, France.
- 28 n. CHU Poitiers, Endocrinology, Diabetology, Nutrition Service, Poitiers, France.

29 **Corresponding author:**

30 Rémy Slama, Team of environmental epidemiology, IAB, Institute for Advanced Biosciences,
31 Inserm, CNRS, CHU-Grenoble-Alpes, University Grenoble-Alpes, Allée de Alpes, Grenoble,
32 France.

33 Remy.slama@univ-grenoble-alpes.fr

34

35 **Highlights:**

- 36 • Exposome studies are promising tools to identify hazardous exposures.
- 37 • Agnostic exposome studies may suffer from low statistical power and low specificity.
- 38 • Information from DNA methylation could be used to relevantly reduce exposome dimension
39 and thus gain efficiency.
- 40 • A Meet-in-the-Middle approach using methylation data was used to relate the exposome with
41 BMI.
- 42 • Postnatal blood copper and PFOS levels were associated with child BMI (respectively
43 positively and negatively).
- 44 • An ExWAS ignoring methylation data identified more significant associations, many of which
45 possibly due to reverse causality.

46

47 **Abstract (300 words)**

48 **Background:** The exposome is defined as encompassing all environmental exposures one undergoes from
49 conception onwards. Challenges of the application of this concept to environmental-health association
50 studies include a possibly high false-positive rate.

51 **Objectives:** We aimed to reduce the dimension of the exposome using information from DNA methylation
52 to more efficiently characterize the relation between exposome and child body mass index (BMI).

53 **Methods:** Among 1,173 mother-child pairs from HELIX cohort, 216 exposures (“whole exposome”) were
54 characterized. BMI and DNA methylation from immune cells of peripheral blood were assessed in children
55 at age 6-10 years. A priori reduction of the methylome to preselect BMI-relevant CpGs was performed using
56 biological pathways. We then implemented a tailored Meet-in-the-Middle approach to identify from these
57 CpGs candidate mediators in the exposome-BMI association. Univariate linear regression models corrected
58 for multiple testing then allowed to point out exposures most likely to be associated with BMI (“reduced
59 exposome”). Associations of this reduced exposome with BMI were finally tested. The approach was
60 compared to an agnostic exposome-wide association study (ExWAS) ignoring the methylome.

61 **Results:** Among the 2284 preselected CpGs (0.6% of the assessed CpGs), 62 were associated with BMI.
62 Four factors (3 postnatal and 1 prenatal) of the exposome were associated with at least one of these CpGs,
63 among which postnatal blood level of copper and PFOS were directly associated with BMI, with respectively
64 positive and negative estimated effects. The agnostic ExWAS identified 18 additional postnatal exposures,
65 including many persistent pollutants.

66 **Discussion:** Our approach incorporating a priori information identified fewer significant associations
67 than an agnostic approach. We hypothesize that this smaller number corresponds to a higher specificity
68 (and possibly lower sensitivity), compared to the agnostic approach. Indeed, the latter cannot distinguish
69 causal relations from reverse causation, e.g. for persistent compounds stored in fat, whose circulating level
70 is influenced by BMI.

71 **Keywords:**

72 biological a priori – child body mass index - exposome –dimension reduction - DNA methylation –
73 reverse causality

74

75 Abbreviations used:

76 BMI: Body Mass Index ; BPA: bisphenol A; DDE: 4,4'dichlorodiphenyl dichloroethylene; DDT:
77 4,4' dichlorodiphenyltrichloroethane; DNA: Deoxyribonucleic acid; EDTA:
78 Ethylenediaminetetraacetic acid; ExWAS: Exposome-Wide Association Study; FDP: False
79 Discovery Proportion; FDR :False Discovery Rate; HCB: hexachlorobenzene; MWAS:
80 Methylome-Wide Association Study; PBDE: polybrominated diphenyl ether; PCB; polychlorinated
81 biphenyl; PFNA: perfluorononanoate; PFOA: perfluorooctanoate; PFOS: perfluorooctane
82 sulfonate; PFUNDA: perfluoroundecanoate; PM: Particulate Matter; POP: persistent organic
83 pollutants; zBMI : z-score of Body Mass Index.

84

86 1. Introduction

87 The exposome concept recognizes that individuals are simultaneously exposed to a multitude of
88 environmental factors from conception onwards (Wild, 2005). The exposome might explain an
89 important, yet currently not accurately quantified, part of the variability in chronic diseases risk
90 (Manrai et al., 2017). Since the 2010s, environmental epidemiology has progressively embraced the
91 exposome concept and complemented common “single exposure studies” with studies relying on
92 simultaneous measurements of several environmental factors (Agier et al., 2019; Buck Louis et al.,
93 2011; Lenters et al., 2016), which, although not including all possible environmental factors, can
94 be seen as examples of exposome studies. Amongst the many challenges faced by these exposome
95 studies (Agier et al., 2016; Siroux et al., 2016) is a possibly high false discovery rate (Agier et al.
96 2016). Specifically, a simulation study considering a realistic exposome of 237 exposures assessed
97 in 1200 individuals, with a proportion of the health outcome variability explained by the exposome
98 varying between 3% and 70%, demonstrated that regression-based methods had a suboptimal
99 sensitivity and high false discovery proportion (FDP) (Agier et al. 2016). All of the approaches
100 tested displayed a FDP well above 5% when there was correlation within the exposome. The
101 widely-used ExWAS (Exposome wide association study) approach, consisting in applying
102 independent linear regression models and correcting association p-values for multiple testing,
103 provided the highest sensitivity, at the cost of a very high FDP.

104 Dimension reduction could be a way to overcome this issue related to false positive findings.
105 Dimension reduction can be performed agnostically, i.e. without relying on external information,
106 with purely statistical techniques, such as variable selection via penalized regression (Lenters et al.,
107 2018; Zou and Hastie, 2005) or Partial-Least-Square (PLS) regressions (Chun & Keles, 2010).
108 However, as Agier et al. (2016) showed, these methods, even if they tend to perform better than
109 ExWAS, are still expected to yield relatively high FDP.

110 Dimension reduction can also be biologically-driven. Relying on a priori information, one may
111 integrate into statistical models relevant information from, for example, the toxicology and
112 fundamental biology fields. Typically, this could be done by restraining analyses to exposures
113 having biological plausibility, based on existing knowledge on associations with biological layers or
114 on pathways linking exposures and the health outcome of interest. This logic has similarities with
115 the concepts of *Mode of Action* and *Adverse Outcome Pathways* used in toxicology (OECD-
116 Organisation for Economic Co-operation and Development, 2012; Vinken, 2013). DNA
117 methylation can be regarded as such an intermediate informative layer, as it is expected to be under
118 environmental (in addition to genetic) influences (Baccarelli et al., 2009; Feil and Fraga, 2012;
119 Joubert et al., 2016; Marioni et al., 2018) and as these epigenetic alterations can result in
120 modifications of disease risk (Ho et al. 2012, Fasanelli et al., 2015). Epigenetic mechanisms, defined
121 as changes in a chromosome which result in heritable phenotype without alterations in the DNA
122 sequence (Berger et al., 2009), have a key role in regulating transcription and thus cell
123 differentiation, cell functioning, and they can ultimately influence the phenotype.

124 An option to identify biomarkers associated with both exposures and the health outcome from a
125 single intermediate DNA methylation layer is the Meet-in-the-Middle approach. It has been
126 developed to point out intermediate biomarkers by considering as putative mediators the overlap
127 between omics signals associated with an exposure and omics signals associated with the outcome
128 (Chadeau-Hyam et al., 2011; Vineis and Perera, 2007).

129 In the case of an intermediate layer with a high dimension, one would have to test the associations
130 of the intermediate putative biomarkers with exposures and health, which might entail a high false
131 positive rate, in particular in the context of correlated exposures or biomarkers. It would appear
132 relevant here to reduce the dimension of the intermediate layer, focusing on biological pathways
133 (or intermediate biomarkers) a priori relevant for the outcome of interest.

134 In this study, we aimed to identify, in an exposome context, environmental factors associated with
135 child BMI, by using information from child methylome layer to reduce the exposome dimension.
136 Childhood obesity and overweight, whose prevalence has rapidly increased over the last three
137 decades (Finucane et al., 2011), are multifactorial conditions, and the most important risk factors,
138 genetic predispositions and energy imbalance, may not suffice to fully explain the magnitude and
139 rapidity of their recent prevalence increase (Park et al., 2017). The effects on BMI of some
140 environmental factors, such as maternal smoking during pregnancy (Oken et al., 2008) or endocrine
141 and metabolic disruptors exposures in early life (Thayer et al., 2012), have already been identified
142 (Agay-Shay et al., 2015; Holtcamp, 2012). The environmental obesogenic hypothesis states that
143 these early exposures play a key role in future obesity risk by altering metabolic programming
144 (Janesick and Blumberg, 2011; Park et al., 2017). Only a few large multi-exposures (Braun, 2017;
145 Fan et al., 2017; Lauritzen et al., 2018) or methylome-wide (Fradin et al., 2017; Rzehak et al., 2017)
146 approaches to the study of child postnatal growth have been conducted.

147

148 2. Materials and methods

149

150 2.1. Overall strategy

151 We relied on the HELIX project, in which the exposome (pregnancy and childhood), the DNA
152 methylome (from peripheral blood in childhood) and BMI were assessed in 1,173 mother-child
153 pairs (Haug et al., 2018; Tamayo-Uria et al., 2019). Biological information from genetic databases
154 was used to a priori reduce the methylome dimension. We implemented a « Meet-in-the-Middle »
155 approach to identify exposures sharing differentially methylated CpGs (i.e. methylation sites) with
156 BMI, as a way to build a *reduced exposome*. The association of this reduced exposome with BMI was
157 then tested.

158 More precisely, the approach consisted in 5 steps:

- 159 a) preselection of CpGs located in genes relevant for BMI, using external databases;
160 b) test of the associations between the methylation levels of these CpGs and BMI;
161 c) test of the association between the methylation levels of the CpGs found to be associated with
162 BMI in *b)* and each exposure, using child BMI as an adjustment factor, allowing to obtain a *reduced*
163 *exposome*;
164 d) test of the association between BMI and the *reduced exposome*;
165 e) comparison with a purely agnostic ExWAS approach ignoring the methylome (i.e. without steps
166 *a)* to *c)* allowing exposome dimension reduction, sensitivity analysis I).

167

168 We also implemented mediation analyses for the significant associations of the main approaches,
169 sensitivity analyses testing the same approach without initial a priori selection of BMI-relevant
170 CpGs, and additionally correcting for cell-type, as well as a sensitivity analysis considering the cell-
171 types as the intermediate layer.

172 **2.2. Study population and outcome**

173 The study was part of the Human Early Life Exposome (HELIX) project (Maitre et al., 2018;
174 Vrijheid et al., 2014), which aimed to describe the early-life exposome and its relations with child
175 development and health.

176 In HELIX, six population-based European birth cohorts were pooled: BiB (Born in Bradford;
177 United Kingdom) (Wright et al., 2013), EDEN (Étude des Déterminants pré et postnatals du
178 développement et de la santé de l'ENfant; France) (Heude et al., 2016), INMA (INfancia y Medio
179 Ambiente; Spain) (Guxens et al., 2012), KANC (Kaunus Cohort; Lithuania) (Grazuleviciene et al.,
180 2015), MoBa (Norwegian Mother, Father and Child Cohort Study; Norway) (Magnus et al., 2016)
181 and Rhea (Greece) (Chatzi et al., 2017), summing up to 1,301 mother–child pairs from singleton
182 pregnancies for whom external exposures (Tamayo-Uria et al., 2019), health outcomes and
183 confounders were measured and harmonized.

184 Height and weight were measured according to standardized procedures between 6 and 10 years
185 of age (Maitre et al., 2018). BMI was calculated as the mass in kilograms divided by the squared
186 height in meters. We used age- and sex-standardized z-scores (named hereafter zBMI) according
187 to the international World Health Organization reference curves (de Onis et al., 2007) in order to
188 allow comparison with other studies on child BMI and to take into account the age-related shift in
189 BMI in childhood. Various lifestyle, social and anthropometric factors were additionally assessed
190 (Table 1).

191

Characteristic	Mean (SD)	n (%)
Child BMI (kg/m²)	16.9 (2.7)	
Child sex		
Female		529 (45)
Male		644 (55)
Child age (years)	7.9 (1.5)	
Cohort		
BiB		203 (17)
EDEN		146 (12)
INMA		215 (18)
KANC		198 (17)
MoBa		212 (18)
RHEA		199 (17)
Maternal education		
Low		176 (15)
Middle		402 (34)
High		595 (51)
Maternal pre-pregnancy BMI (kg/m²)	25.0 (5.0)	
Parity before index pregnancy		
0		530 (45)
1		430 (37)
2 or more		213 (18)
Trimester of conception		
January-March		368 (31)
April-June		234 (20)
July-September		260 (22)
October-December		311 (27)
Maternal tobacco smoke pregnancy exposure		
None		624 (53)
Only passive exposure		374 (32)
Smoker		175 (15)
Child postnatal tobacco smoke exposure		
Not exposed		745 (64)
Exposed		428 (36)
Maternal age (years)	30.7 (4.9)	
Birthweight		
< 2500g		40 (4)
2500 to 3500g		662 (56)
3500 to 4000g		357 (30)
≥ 4000g		114 (10)
Breastfeeding duration		
< 10.8 weeks		361 (31)
10.8 to 34.9 weeks		419 (36)
> 34.9 weeks		393 (34)
Parents born in the country of inclusion		
None		134 (11)
Only one		58 (5)
Both		981 (84)
Ethnicity		
African		7 (1)
Asian		19 (2)
European ancestry		1048 (89)
Native American		2 (0)
Pakistani		79 (7)
Other		18 (2)

192 Table 1: Characteristics of the 1,173 mother-child pairs from HELIX cohort.

193

194 **2.3. Exposome assessment**

195 Details of the exposome assessment have been published elsewhere (Haug et al., 2018; Tamayo et
196 al., 2018). Among the 234 exposures assessed in HELIX (Tamayo et al., 2018), we excluded
197 exposures with a time window of one day or one week, which were a priori considered unlikely to
198 influence BMI. This led to 216 prenatal and postnatal exposures (list available in Supplementary
199 Material 1). Metals, organochlorines, organophosphate pesticides, polybrominated diphenyl ethers
200 (PBDE), perfluorinated alkylated substances (PFAS), phenols and phthalates were assessed by
201 biomarkers in mothers during pregnancy from one urine or blood sample and in children at the
202 time of the clinical examination, from a pool of two urine samples or one blood sample (Casas et
203 al., 2018; Haug et al., 2018). Built environment exposures, indoor air exposures, lifestyle factors,
204 meteorological data, natural spaces quantification, noise, traffic, socio-economic capital and
205 concentrations of disinfection by-products in drinking water were assessed during pregnancy and
206 during the year before child examination by environmental models and questionnaires.

207 Exposures were transformed to approach normality, using a Box-Cox power transformation
208 approach that chooses among log-transforming or raising the data to the powers -2 , -1 , -0.5 ,
209 $1/3$, 0.5 , 1 , or 2 . Transformation chosen for each variable is detailed in Supplementary Material 1.
210 Exposures were standardized using their interquartile range after imputing missing data for all
211 exposures using mice R package (Buuren and Groothuis-Oudshoorn, 2011).

212 **2.4. DNA Methylation**

213 Peripheral blood was collected in EDTA tubes during the clinical examination that took place when
214 children were between 6 and 10 years old. DNA was extracted from buffy coat; DNA methylation
215 was assessed with the Infinium Human Methylation 450 beadchip (Illumina), following the
216 manufacturer's protocol. Sample locations on chips were drawn at random balancing chips for

217 cohort and infant sex. Some samples were analysed in duplicate and a control HapMap sample was
218 added in each 96-well plate.

219 DNA methylation data were pre-processed using the minfi R package (Aryee et al., 2014). A first
220 quality control of the data was done with MethylAid package (van Iterson et al., 2014); probes with
221 low call rate were then filtered following guidelines of Lehne et al. (2015). The functional
222 normalization method was further applied, including Noob background subtraction and dye-bias
223 correction (Triche et al., 2013). Several quality control checks were performed: sex consistency
224 using the shinyMethyl package (Fortin et al., 2014); consistency of duplicates; genetic consistency
225 for the samples that had genome-wide genotypic data. Finally, duplicated samples and control
226 samples were removed as well as probes to measure methylation levels at non-CpG sites (Jang et
227 al., 2017). A final filtering was performed to eliminate probes with a single-nucleotide
228 polymorphism (SNP), probes that cross-hybridize and probes on sex chromosomes, restricting to
229 386,518 CpG probes available for 1,192 subjects. The study was performed on the 1173 subjects
230 among them who also had valid exposures data.

231 We then used Combat procedure to remove the batch effects supported by the slide. Methylation
232 levels were expressed as Beta values (average methylation levels for an individual, between 0 for a
233 never methylated CpG site and 1 for an always-methylated CpG site).

234 Cell types were computed according to Houseman et al. (2012) algorithm and Reinius reference
235 panel (Reinius et al., 2012); tests of associations including methylation levels were corrected for cell
236 types only in a sensitivity analysis.

237 Correlation within the methylome was estimated by averaging the Pearson's correlation within 10
238 sets of 2284 randomly selected CpGs (same size as the restricted methylome, see next paragraph),
239 to avoid computing all pairwise correlations between the 386,518 CpGs.

240 **2.5. A priori preselection of BMI-relevant CpGs**

241 Biological pathways a priori relevant for BMI were selected using the KEGG database (Tanabe
242 and Kanehisa, 2012), searching with the key words “growth” “obesity” and “fat” in the following
243 categories: “Human energy metabolism”, “Human lipid metabolism”, “Human endocrine system”,
244 “Human digestive system”, “Human Excretory system”, “Human endocrine and metabolic
245 diseases”, “Human genetic Information Processing: ‘Transcription – Translation - Folding, sorting
246 and degradation - Replication and repair’”. We thus identified a list of 16 pathways (Supplementary
247 Material 2) and the corresponding list of genes, which were restricted to the CpGs identified as
248 enhancers, leading to a final dataset of 2284 CpGs belonging to 387 genes and 16 different
249 biological pathways (Supplementary Material 2), which we further refer to as the “restricted
250 methylome”. Correspondence between genes and CpGs as well as enhancer annotation and CpGs
251 was based on Illumina annotation (Hansen, 2016). A sensitivity analysis not restricted to enhancers
252 and these 16 pathways was performed.

253 **2.6. Meet-in-the-Middle and ExWAS approaches – statistical analyses**

254 Our Meet-in-the-Middle design itself consisted in three successive steps, as described in 2.1 (steps
255 *b*), *c*) and *d*): in step *b*), we tested the association of the methylation levels of the preselected CpGs
256 with BMI considered as the outcome; in step *c*), we tested the associations (adjusted for child BMI)
257 of each exposure with the CpGs found to be associated with BMI in *b*), leading to the identification
258 of a “reduced exposome”; step *d*) is the test of the association of this reduced exposome (the
259 exposures found to be associated to some CpGs in step *c*)) with the outcome.

260 Univariate linear regressions models were applied, and p-values were corrected for multiple testing
261 using a FDR (False-Discovery Rate) control procedure (Benjamini & Hochberg, 1995) at all steps
262 involving regression modelling. Adjustment factors coded linearly in all our regression analyses
263 were maternal pre-pregnancy BMI (additionally coded with a quadratic term in the exposome-
264 outcome associations test), maternal education, maternal age, parental country of birth, maternal
265 smoking during pregnancy, cohort (fixed effect), parity, trimester of conception, child age and child

266 sex (see Table 1 for the categories). We additionally adjusted analyses of postnatal exposures effects
267 for birth weight, breastfeeding duration and passive smoking during childhood (see Table 1) and
268 models including methylation data for ethnicity (self-reported by parents, with different questions
269 across the cohorts). At step *c*), correction for multiple testing was done considering together all
270 associations tested between exposures and CpGs, *i.e.* a number of test equal to the product of the
271 number of exposures with the number of CpGs associated with zBMI. A mediation analysis using
272 package MMA (Yu and Li, 2017) was performed for exposures found associated with the outcome
273 in step *d*), considering the CpGs both associated with the exposure and the outcome.

274 **2.7. Sensitivity analyses and test of selection relevance**

275 We compared our results to those obtained with a *totally agnostic* approach, consisting in an ExWAS
276 between the exposome and zBMI, ignoring the methylome, with exactly the same statistical
277 methods as in our step *d*) (sensitivity analysis I). Our agnostic ExWAS has some differences with
278 that performed by Vrijheid et al. in Helix data (submitted manuscript, available upon request). In
279 particular, we restricted the population of 1301 children used by Vrijheid et al. to 1173 children
280 with methylome data (see paragraph 2.4.), and we chose to additionally adjust for trimester of
281 conception, ethnicity and pre- and postnatal smoking, which could influence the methylome.

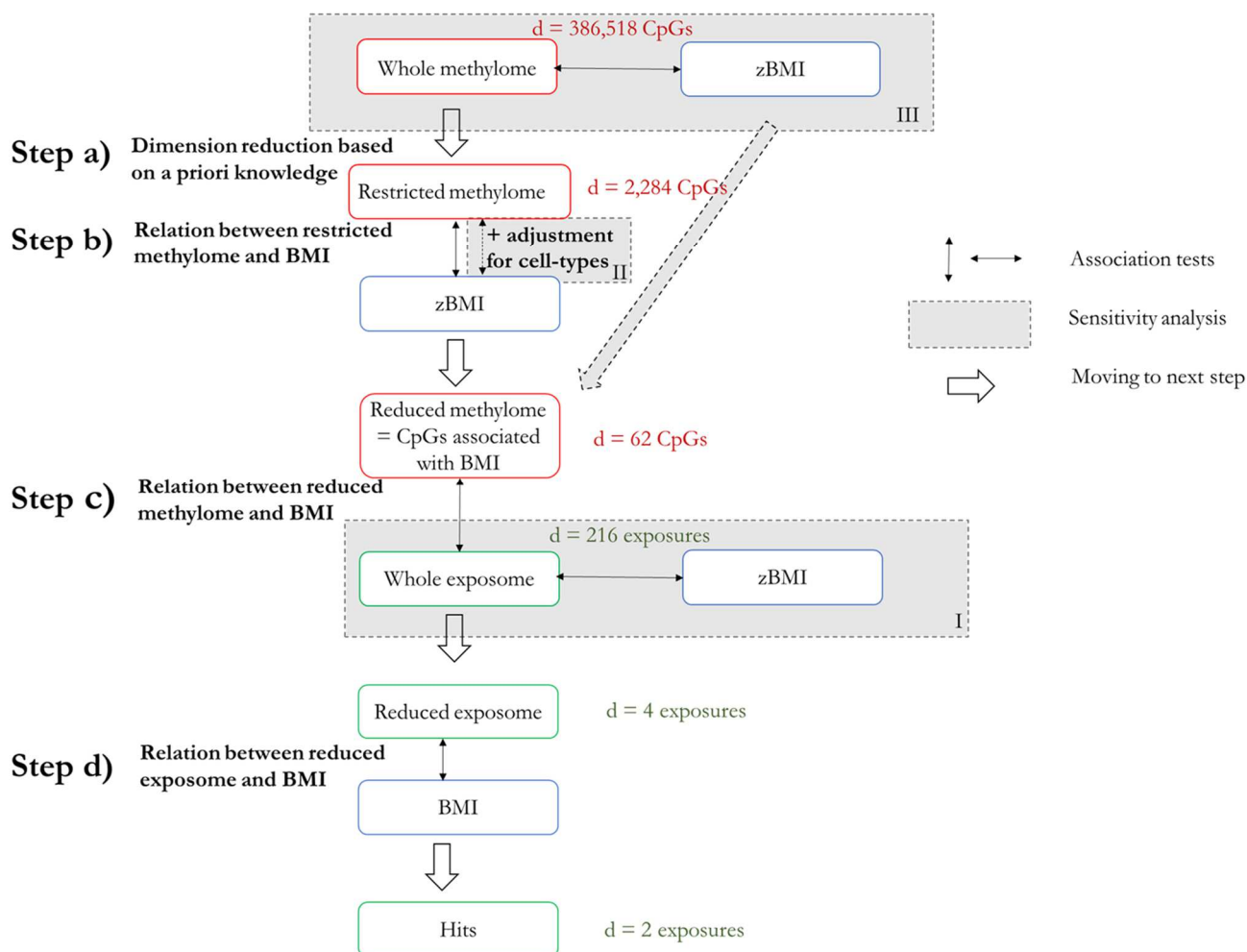
282 We performed three other sensitivity analyses. First, we repeated our approach with an additional
283 adjustment on cell-type heterogeneity for all association tests involving the methylome (sensitivity
284 analysis II). Second, we repeated our approach on the unrestricted methylome, *i.e.* without the first
285 step of a priori selection of CpGs using information from biological pathways database and
286 annotation (*i.e.* removing step *a*) and starting from step *b*) with 386,518 CpGs). This modified step
287 *b*) corresponds to a methylome-wide analysis, or MWAS (methylome-wide association study) in
288 which we tested the association between methylation levels of the whole unrestricted methylome
289 and zBMI (sensitivity analysis III).

290 To further inform the relevance of the a priori CpG selection of step a), the proportion of CpGs
 291 belonging to our candidate list of a priori BMI-relevant CpGs (i.e. our restricted methylome) among
 292 CpGs whose methylation levels was found associated with zBMI by MWAS was compared to the
 293 corresponding proportion in the whole methylome.

294 A workflow of the statistical analyses is shown Figure 1.

295 We additionally performed a fourth sensitivity analysis (sensitivity analysis IV) by repeating the
 296 whole Meet-in-the-Middle approach considering the cell-types instead of methylation data as the
 297 intermediate layer.

298



299

300 Figure 1: Workflow of the main statistical analyses. (in color; 1-column fitting image)

301

302

303 3. Results

304 3.1. Population characteristics

305 At the time of BMI measurement, mean age was 7.9 years. Mean child BMI was 16.9 kg/m² (5th
306 and 95th percentiles: 13.8; 22.4), with substantial differences between cohorts (Supplementary
307 Material 3), INMA and RHEA showing higher zBMI compared to the other cohorts. The other
308 characteristics of the study population are given Table 1 and by cohorts in Supplementary Material
309 4.

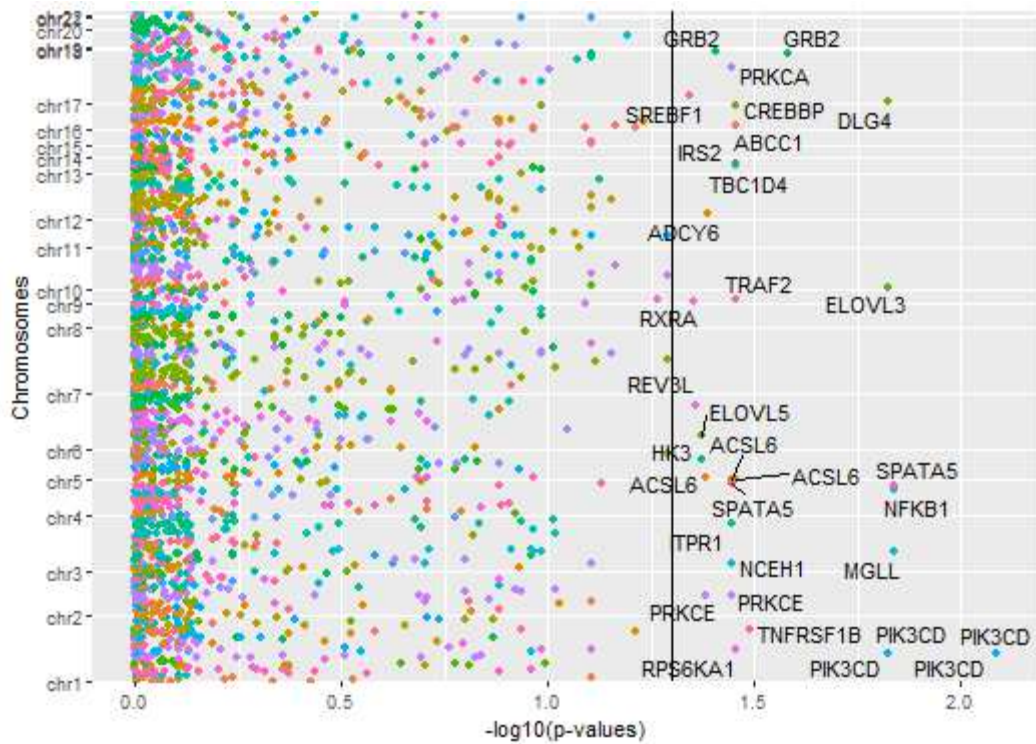
310 The mean levels of the 216 exposures considered are displayed in Supplementary Material 1. Mean
311 absolute correlation between quantitative exposures was 0.11 (5th and 95th percentiles: 0.00, 0.35);
312 the distribution of the coefficients of correlation is given in Supplementary Material 5.

313 Within the whole methylome, the estimated mean correlation was 0.09 while it was 0.12 for the
314 2284 CpGs of the reduced methylome (5th and 95th percentiles: 0.00; 0.37).

315 3.2. Meet-in-the-Middle approach

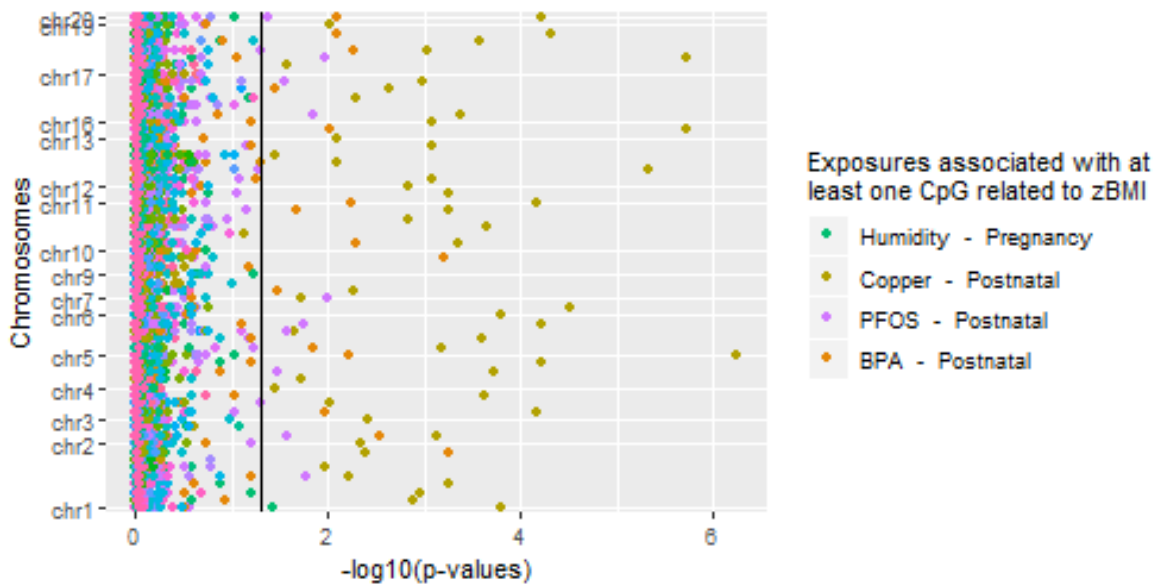
316 The analysis testing the association between the restricted methylome and zBMI (step *b*) identified
317 62 CpGs belonging to 43 different genes (FDR adjusted p-values ≤ 0.05 ; Supplementary Material
318 6). The mean correlation among these CpGs was 0.59. Figure 2A shows a Manhattan plot of the
319 FDR-adjusted p-values.

A.



320

B.



321

322 Figure 2: Manhattan plots of the FDR corrected p-values of adjusted associations obtained with
 323 the Meet-in-the-middle approach applied on the reduced methylome at steps *b*) and *c*). **A:**
 324 Associations between preselected CpG and zBMI. Each colour corresponds to a gene. The black
 325 vertical line shows the (FDR-corrected) 0.05 significance level. Lowest p-value: 3.20×10^{-3} . **B.**
 326 Associations between exposures and CpGs associated with childhood zBMI. Each color
 327 corresponds to a different exposure. The black vertical line shows the (FDR-corrected) 0.05
 328 significance level. Lowest p-value: 5.66×10^{-7} . BPA: Bisphenol A; PFOS: Perfluorooctanesulfonic
 329 acid. (in color; 2-column fitting image)

330

331
 332 The test of association of these 62 CpGs with each of the 216 environmental factors adjusted for
 333 child BMI (step *c*) identified 4 exposures associated with at least one CpG (Figure 2B, Table 2,
 334 Supplementary Material 7): copper (postnatal level), BPA (Bisphenol A, postnatal level), PFOS
 335 (Perfluorooctanesulfonic acid, postnatal level) and one meteorological variable (humidity,
 336 pregnancy average); this constituted our *reduced exposome*. In total, 53 CpGs were associated with at
 337 least one exposure.

338

Exposure	Number of CpGs associated both with the exposure and zBMI	Number of corresponding genes*
Copper (postnatal)	52	37
Bisphenol A (BPA) (postnatal)	15	14
Perfluorooctanesulfonic acid (PFOS) (postnatal)	12	12
Humidity average (pregnancy)	1	1

339 Table 2: Number of CpGs associated with both exposures and zBMI in the adjusted associations
 340 between the exposome and CpGs associated with zBMI in 1,173 children from the HELIX cohort
 341 (ExWAS model adjusted on zBMI, step *c*) of the Meet-in-the-Middle approach applied on the
 342 reduced methylome). Results are presented only for exposures associated with a (stringently
 343 corrected for multiple hypothesis testing) p-value of less than 0.05 in exposure-CpGs ExWAS, with
 344 CpGs being previously selected in a CpGs-zBMI ExWAS. *Details of the CpGs and genes are
 345 given in Supplementary Material 6.

346
 347 The last step (step *d*) identified that within the reduced exposome, postnatal blood copper and
 348 PFOS levels were directly associated with zBMI. The corresponding estimated parameters were
 349 respectively 0.22 (95% CI: 0.14; 0.30; adjusted p-value, 1.43×10^{-6}) and -0.13 (95% CI: -0.23; -0.04;
 350 adjusted p-value, 0.02) (see Table 3 for the other components of the reduced exposome). A
 351 mediation analysis quantified that for copper, the 52 CpGs mediated 29% of the total effect of
 352 postnatal blood copper level on zBMI, while for PFOS, the 12 selected CpGs mediated 28% of
 353 the total effect of postnatal blood PFOS level.

354

Group	Label	Unit	Transformation	Effect estimate*	95% CI	Uncorrected p-Value	FDR-corrected p-Value
Meteorological	Humidity (pregnancy)	%	None	0.05	-0.34; 0.44	0.81	0.81
Phenols	BPA (postnatal)	µg/g	Log2	-0.07	-0.14; 2.8x10 ⁻⁴	0.05	0.07
Metals	Copper (postnatal)	µg/L	Log2	0.22	0.14; 0.30	3.57x10 ⁻⁷	1.43x10 ⁻⁶
Perfluorinated alkylated substances (PFAS)	PFOS (postnatal)	µg/L	Log2	-0.13	-0.23; -0.04	7.69x10⁻³	0.02

356 Table 3: Adjusted associations between the reduced exposome and zBMI in 1,173 children from
 357 HELIX cohort (ExWAS model, step *d*) of the Meet-in-the-Middle approach applied to the reduced
 358 methylome). “Significant” associations are indicated in bold.

359 * Adjusted change in mean zBMI for each unit increase in transformed exposure level. Models
 360 were adjusted for maternal BMI, maternal education, maternal smoking during pregnancy, parental
 361 country, cohort, parity, trimester of conception, ethnicity, child age and child sex and additionally
 362 only for postnatal exposures birth weight, passive smoking during childhood and breastfeeding
 363 duration.

364

365 3.3. Agnostic exposome-wide approach

366 An agnostic ExWAS using FDR correction for multiple testing between the whole (not reduced)
 367 exposome and zBMI identified 20 postnatal exposures significantly associated with zBMI
 368 (Sensitivity analysis I, Table 4). These included postnatal copper and PFOS level (also identified at
 369 step *d*) of the main approach). In addition to metals and perfluorinated alkylated substances (PFAS),
 370 these exposures belonged to the organochlorines, polybrominated diphenyl ethers (PBDE),
 371 lifestyle and indoor air pollution families. Organochlorines, PBDE and PFAS compounds, as well
 372 as postnatal cobalt levels showed negative regression coefficients, corresponding to a decreased
 373 zBMI with increasing exposure levels. The most significant associations were observed for 5 of
 374 the postnatal PCB levels, which formed a group with higher correlation (mean absolute correlation,
 375 0.50) than the rest of the quantitative exposome (mean absolute correlation, 0.11).

376

Exposure group	Exposure variable	Unit	Transformation	Effect estimate*	95%CI		Uncorrected p-value	FDR-corrected p-value
Organochlorines	PCB 180 - Postnatal	ng/g lipids	Log2	-0.92	-1.05	-0.78	2.14x10 ⁻³⁸	4.62x10 ⁻³⁶
Organochlorines	HCB - Postnatal	ng/g lipids	Log2	-0.66	-0.76	-0.57	5.08x10 ⁻³⁸	5.48x10 ⁻³⁶
Organochlorines	PCB 170 - Postnatal	ng/g lipids	Log2	-0.82	-0.95	-0.70	9.56x10 ⁻³⁷	6.88x10 ⁻³⁵
Organochlorines	Sum of PCBs - Postnatal	ng/g lipids	Log2	-0.80	-0.92	-0.67	7.15x10 ⁻³⁴	3.86x10 ⁻³²
Organochlorines	PCB 138 - Postnatal	ng/g lipids	Log2	-0.67	-0.78	-0.55	5.56x10 ⁻²⁸	2.40x10 ⁻²⁶
Organochlorines	PCB 153 - Postnatal	ng/g lipids	Log2	-0.70	-0.82	-0.58	1.29x10 ⁻²⁷	4.66x10 ⁻²⁶
Organochlorines	DDE - Postnatal	ng/g lipids	Log2	-0.54	-0.64	-0.43	3.42x10 ⁻²²	1.06x10 ⁻²⁰
Polybrominated diphenyl ethers	PBDE 153 - Postnatal	ng/g lipids	Log2	-0.40	-0.52	-0.28	8.70x10 ⁻¹¹	2.35x10 ⁻⁹
Organochlorines	PCB 118 - Postnatal	ng/g lipids	Log2	-0.28	-0.38	-0.19	1.39x10 ⁻⁹	3.33x10 ⁻⁸
Metals	Copper - Postnatal	µg/L	Log2	0.21	0.13	0.30	5.67x10 ⁻⁷	1.11x10 ⁻⁵
Perfluorinated alkylated substances	PFOA - Postnatal	µg/L	Log2	-0.24	-0.33	-0.15	5.29x10 ⁻⁷	1.11x10 ⁻⁵
Organochlorines	DDT - Postnatal	ng/g lipids	Log2	-0.27	-0.38	-0.16	1.11x10 ⁻⁶	2.01x10 ⁻⁵
Perfluorinated alkylated substances	PFNA - Postnatal	µg/L	Log2	-0.19	-0.28	-0.10	3.65x10 ⁻⁵	6.07x10 ⁻⁴
Perfluorinated alkylated substances	PFUNDA - Postnatal	µg/L	Log2	-0.21	-0.32	-0.09	4.06x10 ⁻⁴	6.27x10 ⁻³
Metals	Cesium - Postnatal	µg/L	Log2	0.20	0.09	0.31	4.89x10 ⁻⁴	7.04x10 ⁻³
Metals	Cobalt - Postnatal	µg/L	Log2	-0.12	-0.19	-0.05	7.52x10 ⁻⁴	0.01
Tobacco Smoke	Active smoking - Pregnancy	-	-	0.34	0.13	0.55	1.45x10 ⁻³	0.02
Indoor air	Indoor PM absorbance - Postnatal	10 ⁻⁵ m ⁻¹	Log	0.13	0.04	0.21	2.95x10 ⁻³	0.03
Indoor air	Indoor PM _{2.5} - Postnatal	µg / m ³	Log	0.11	0.04	0.19	2.94x10 ⁻³	0.03
Perfluorinated alkylated substances	PFOS - Postnatal	µg/L	Log2	-0.14	-0.24	-0.04	4.59x10 ⁻³	0.05

377 Table 4: Sensitivity analysis I: adjusted associations between the whole exposome and zBMI in
378 1,173 children from the HELIX cohort (ExWAS agnostic approach, ignoring the methylome).
379 Results are presented only for exposures with a (FDR-corrected for multiple hypothesis testing) p-
380 value lower than 0.05.

381 * Adjusted change in mean zBMI for each unit increase in transformed exposure level. Models
382 were adjusted for maternal BMI, maternal education, maternal smoking during pregnancy, parental
383 country, cohort, parity, trimester of conception, ethnicity, child age and child sex and additionally
384 only for postnatal exposures birth weight, passive smoking during childhood and breastfeeding
385 duration.
386

387 **3.4. Other sensitivity analyses**

388 In order to determine if our a priori CpGs selection led to a concentration of information, we
389 quantified the overrepresentation of our preselected BMI-relevant CpGs among the discoveries of
390 a methylome-wide analysis linking the whole methylome to zBMI. As expected, the CpGs
391 associated with zBMI in an MWAS considering the whole exposome were enriched in enhancers
392 CpGs selected as being relevant for BMI from KEGG database (1.22%) compared to the other
393 CpGs (0.46%, a ratio of 2.6 to 1). However, most significant associations found by the MWAS
394 were not part of the a priori selected list of CpGs (1760 out of 1788, Supplementary Material 8).

395 When the whole Meet-in-the-Middle was repeated without the step of CpGs preselection using
396 external biological information (sensitivity analysis III), final results differed from those obtained
397 using the restricted methylome: additionally to blood postnatal copper and PFOS levels (which
398 were also found significant in this analysis), blood postnatal hexachlorobenzene (HCB),
399 Pentabromodiphenyl ether (PBDE) 153 and dichlorodiphenyltrichloroethane (DDT) levels (with
400 negative associations with zBMI) and blood postnatal caesium level (with a positive association
401 with zBMI) were identified in step *d*) (Table 5). These 5 exposures were also associated with zBMI
402 in the agnostic ExWAS approach (sensitivity analysis I). To give more details, 1788 out of 386,518
403 CpGs were associated with zBMI in step *b*) (Figure 3A and Supplementary Material 9). In step *c*)
404 of the analysis, 28 exposures were significantly associated with at least one of these 1788 CpGs.
405 Among them, postnatal blood levels of copper, BPA and PFOS (Supplementary Material 10 and
406 Figure 3B) and prenatal humidity exposure, which had all been previously found in the main

407 analysis, were associated with respectively 110, 449, 180 and 47 CpGs. All the other exposures were
 408 associated with less than 10 CpGs.

409

Exposure group	Exposure variable	Unit	Transformation	Effect estimate*	95%CI	Unadjusted p-value	FDR adjusted p-value
Organochlorines	HCB - Postnatal	ng/g lipids	Log2	-0.66	-0.76; -0.56	7.87x10 ⁻³⁸	2.20x10 ⁻³⁶
Polybrominated diphenyl ethers	PBDE 153 - Postnatal	ng/g lipids	Log2	-0.40	-0.52; -0.28	1.62x10 ⁻¹⁰	2.26x10 ⁻⁹
Metals	Copper - Postnatal	µg/L	Log2	0.22	0.14; 0.30	3.57x10 ⁻⁷	3.33x10 ⁻⁶
Organochlorines	DDT - Postnatal	ng/g lipids	Log2	-0.28	-0.39; -0.17	6.06x10 ⁻⁷	4.24x10 ⁻⁶
Metals	Caesium - Postnatal	µg/L	Log2	0.19	0.08; 0.30	8.50x10 ⁻⁴	4.76x10 ⁻³
Perfluorinated alkylated substances	PFOS - Postnatal	µg/L	Log2	-0.13	-0.23; -0.04	7.69x10 ⁻³	0.04
Phenols	BPA - Postnatal	µg/g	Log2	-0.07	-0.14; 2.83x10 ⁻⁴	0.05	0.20
Phthalates	OH-MiNP - Pregnancy	µg/g	Log2	-0.07	-0.14; 4.04x10 ⁻³	0.06	0.23
Phthalates	MEHP - Postnatal	µg/g	Log2	-0.07	-0.16; 0.01	0.10	0.31
Perfluorinated alkylated substances	PFHXS - Postnatal	µg/L	Log2	-0.10	-0.22; 0.02	0.12	0.33
Built Environment	Population density - Postnatal	people / km ²	Square root	0.06	-0.02; 0.15	0.15	0.38
Lifestyle	Soda intake - Postnatal	Times/ week	Tertiles	-0.10	-0.26; 0.05	0.18	0.42
Socio-eco capital	Social participation - Postnatal	-	None	-0.08	-0.24; 0.07	0.28	0.60
Lifestyle	Fastfood intake - Pregnancy	Times/ week	Tertiles	0.13	-0.14; 0.39	0.34	0.66
Phthalates	MiBP - Postnatal	µg/g	Log2	-0.05	-0.16; 0.06	0.35	0.66
Phthalates	MBzP - Postnatal	µg/g	Log2	-0.03	-0.11; 0.04	0.40	0.70
Organochlorines	PCB 138 - Pregnancy	ng/g lipids	Log2	-0.05	-0.18; 0.08	0.48	0.78
Organochlorines	Sum of PCBs - Pregnancy	ng/g lipids	Log2	-0.05	-0.20; 0.11	0.55	0.80
Air Pollution	PM2.5 - Pregnancy	µg / m ³	None	-0.04	-0.17; 0.10	0.57	0.80
Noise	Traffic noise (24h) - Postnatal	dB(A)	None	-0.06	-0.26; 0.13	0.54	0.80
Meteorological	Humidity - Pregnancy	%	None	0.05	-0.34; 0.44	0.81	0.85

Exposure group	Exposure variable	Unit	Transformation	Effect estimate*	95%CI	Unadjusted p-value	FDR adjusted p-value
Organochlorines	PCB 170 - Pregnancy	ng/g lipids	Log2	-0.03	-0.16; 0.10	0.69	0.85
OP Pesticides	DETP - Pregnancy	µg/g	Log2	0.02	-0.09; 0.12	0.77	0.85
Lifestyle	Vegetables intake - Pregnancy	Times/ week	Tertiles	-0.02	-0.18; 0.15	0.85	0.85
Natural Spaces	Green spaces (300 m) - Pregnancy	-	None	-0.02	-0.17; 0.14	0.84	0.85
Phenols	PRPA - Pregnancy	µg/g	Log2	0.01	-0.09; 0.12	0.82	0.85
Metals	Thallium - Postnatal		None	-0.04	-0.29; 0.21	0.78	0.85
Lifestyle	Yogurt intake - Postnatal	Times/ week	Tertiles	0.04	-0.14; 0.21	0.69	0.85

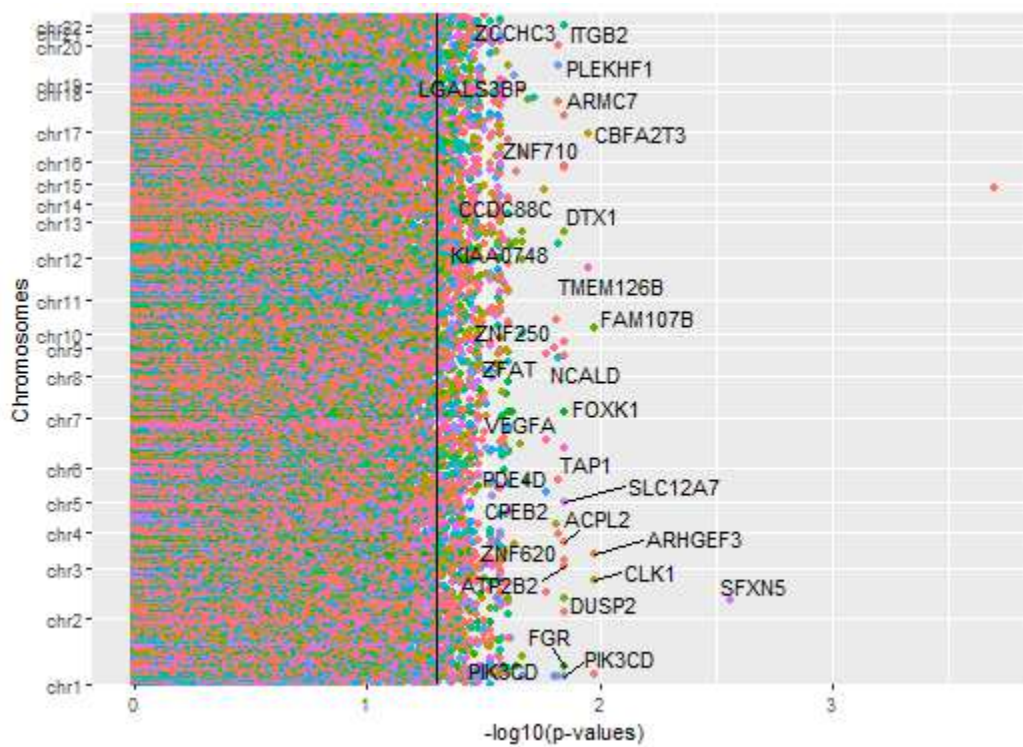
410 Table 5: Sensitivity analysis III - Meet-in-the-Middle without CpGs preselection based on KEGG
411 database: adjusted association between the reduced exposome and zBMI in 1,173 children from
412 the HELIX cohort (ExWAS model, step *d*) of the Meet-in-the-Middle approach applied on the
413 reduced methylome). Bold lines indicate significant associations.

414 * Adjusted change in mean zBMI for each increase by 1 in transformed exposure level. Models
415 were adjusted for maternal BMI, maternal education, maternal smoking during pregnancy, parental
416 country, cohort, parity, trimester of conception, ethnicity, child age and child sex and additionally
417 only for postnatal exposures birth weight, passive smoking during childhood, breastfeeding
418 duration.

419

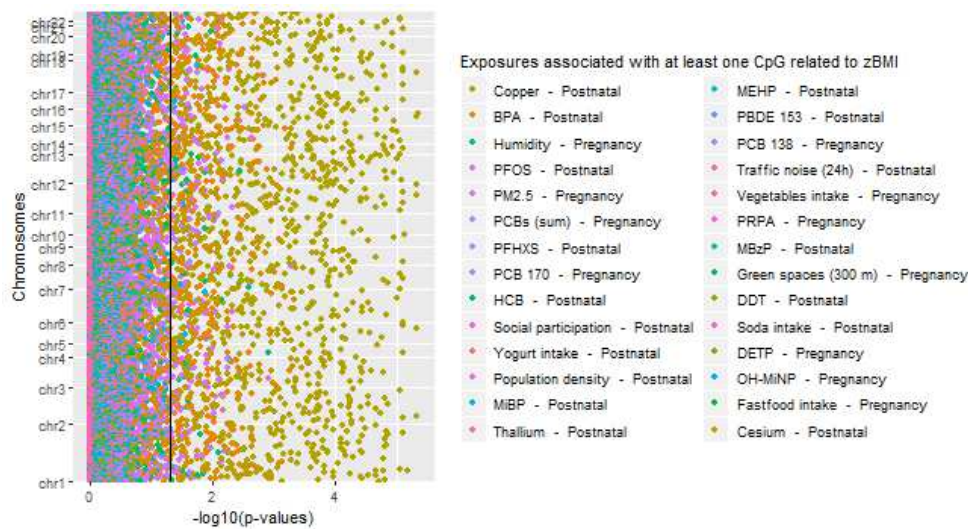
420

421 A.



422

423 B.



424

425 Figure 3: Sensitivity analysis III - Meet-in-the-Middle approach without a priori preselection of
 426 CpGs: Manhattan plots of the FDR corrected p-values of adjusted associations obtained with the
 427 Meet-in-the-middle approach applied on the reduced methylome at steps b and c. **A:** Associations
 428 between all CpG and zBMI. Each colour corresponds to genes. The black line is the (FDR-
 429 corrected) 0.05 significance threshold. Lowest p-value: 2.04×10^{-4} . **B.** Associations between
 430 exposures and CpGs associated with childhood zBMI. Each colour corresponds to a different
 431 exposure. Lowest p-value: 4.51×10^{-6} . (in color; 2-column fitting image)

432

433
434 When we repeated our analysis adding a correction for blood cell-types (sensitivity analysis II), no
435 association was significant at step *b*) (lowest p-value with Benjamini-Hochberg correction: 0.72).
436 When we repeated the analysis corrected for blood cell-types without the preselection step, we
437 found one association between the whole methylome and zBMI, but the corresponding CpG
438 (cg02032125) was not associated with any exposure at step *c*) so that no exposure was eventually
439 selected as associated with BMI.

440 When we considered the cell-types instead of the methylation data as our intermediate layer,
441 (sensitivity analysis IV), results were very similar to those of the main analysis: the reduced
442 exposome consisted in three exposures, postnatal blood copper and BPA levels and average
443 pregnancy humidity exposure. In the last step, only copper level was associated with zBMI. The
444 three cell-types associated with both copper and zBMI mediated 13% of the effect of copper on
445 zBMI. Detailed results of sensitivity analysis IV are available in Supplementary Material 11.

446

447 4. Discussion:

448 We implemented a modified Meet-in-the-Middle approach among 1,173 mother-child pairs to
449 identify components of the exposome influencing child BMI through DNA methylation changes.
450 The analysis highlighted postnatal copper blood level as being positively associated with zBMI, an
451 association supported by changes with copper levels in the methylation levels of 52 CpGs from
452 genes that are relevant for BMI based on a priori knowledge. Blood perfluorooctanesulfonic acid
453 postnatal level was also found related to zBMI in our Meet-in-the-Middle approach, an association
454 likely due to reverse causality.

455 Our work is one of the first studying the link of an exposome including both chemical and
456 nonchemical stressors during the prenatal and postnatal time windows with child BMI. Beside
457 Helix studies, the largest studies considering multiple chemical exposures and child BMI considered
458 up to 27 components (Agay-Shay et al., 2015; Fan et al., 2017; Lauritzen et al., 2018)).

459 The efficiency of our Meet-in-the-Middle approach to detect true predictors of zBMI within the
460 exposome relies on three main assumptions: 1) that part of the effects of the exposome on child
461 BMI are mediated by changes in methylation levels that can be observed from peripheral blood; 2)
462 that these methylation changes are strong enough to be detectable and that they can be used to
463 select plausible exposures and thus reduce the exposome dimension; 3) that existing databases of
464 biological pathways and regulatory regions (enhancers) allow to relevantly reduce the dimension of
465 the methylome a priori to study its association with BMI. We discuss here the relevance of these
466 three assumptions, as well as our choice not to correct for cell-types heterogeneity.

467

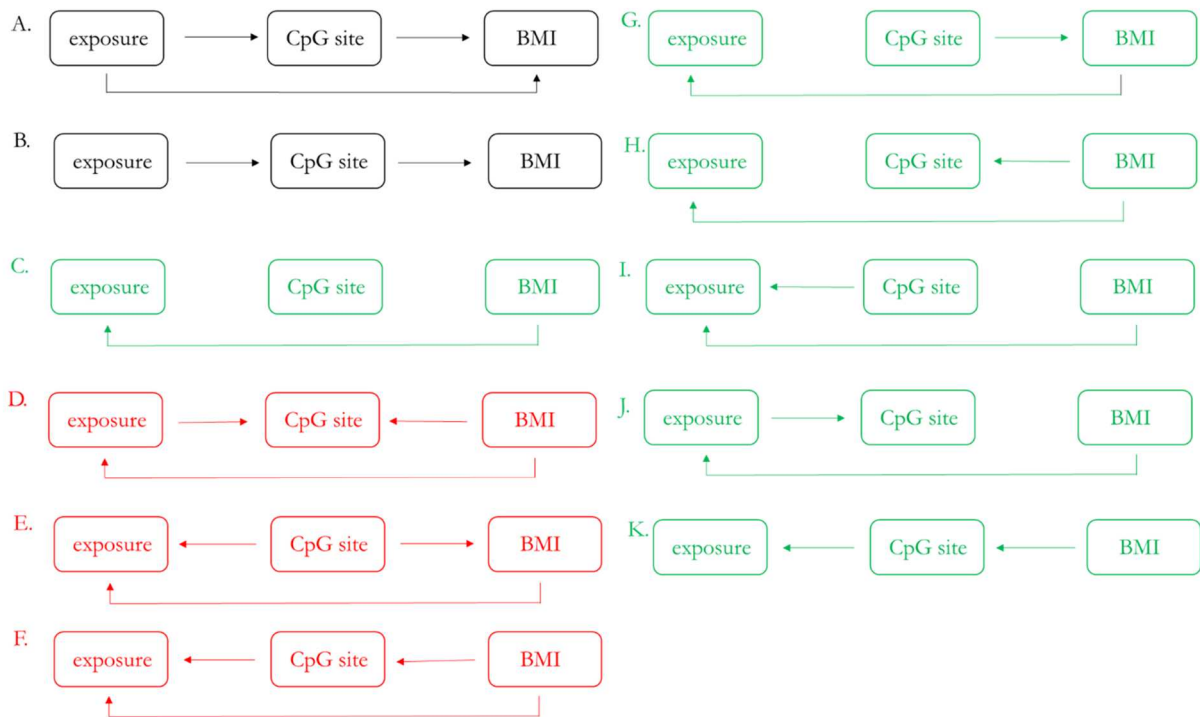
468 **4.1. Are some effects of the exposome on child BMI likely to be mediated by the**
469 **methylome?**

470 Methylation has been reported to mediate part of the effect of specific exposures on health. This
471 has been suggested for example for smoking effects on health: Fasanelli et al. (2015) pointed
472 methylation mediation for smoking effects on lung cancer and Wahl et al. (2018) showed that site-
473 specific methylation can mediate the effect of smoking on the expression of inflammatory proteins.
474 For BMI, evidence of mediation arises from animal toxicological studies, which showed that long-
475 term obesity risk can result from effects of early overfeeding/underfeeding mediated by
476 methylation changes on specific regulatory CpG sites in different tissues (Carone et al., 2010;
477 Lillycrop et al., 2008, 2005; Plagemann et al., 2009). In humans, studies based on a subjects who
478 experienced famine during intra-uterine life suggested that blood cells methylation could mediate
479 effects of prenatal undernutrition on later overweight: early-life exposures to famine had an effect
480 on CpGs regulating growth and metabolic mechanisms involved in obesity (Heijmans et al., 2008;
481 Tobi et al., 2009). In addition, such a mediation between prenatal exposure to famine and adult
482 metabolic traits has been statistically demonstrated; indeed, studies (Tobi, Sliker, et al., 2018; Tobi
483 et al., 2018) pointed out that even if, from a biological point of view, DNA methylation measured
484 in peripheral blood was not likely to be a causal mediator of BMI change, it could be a proxy of
485 epigenetic regulation changes in specific tissues, and thus allow to detect mediation. For early
486 exposures other than nutrition, little evidence is currently available regarding an effect on BMI or
487 growth mediated by epigenetic changes (Richmond et al., 2015): however, Cao-Lei et al. (2015)
488 suggested that some gene-specific methylation could mediate part of the effect of prenatal maternal
489 stress child BMI and central adiposity. Less directly, several exposures were identified as possibly
490 influencing epigenetic marks, and some of these alterations may occur on genes involved in
491 signaling pathways controlling growth and adipose tissue development (Richmond et al., 2015).
492 Importantly, effects in the opposite causal directions are also likely, in that changes in BMI could
493 influence methylation levels on specific loci, as suggested by Dekkers et al. (2016) and Richmond
494 et al. (2016).

495 Our approach, as well as the agnostic ExWAS, identified postnatal blood copper level as positively
496 associated with zBMI and additionally with changes in BMI-relevant CpGs. The mediating effect
497 that we estimated was of 29% of the estimated total copper effect. Copper is an essential trace
498 element involved via oxidoreduction reactions in a broad range of processes, including energy
499 expenditure, mitochondrial respiration, antioxidant defences and inflammation (Tisato et al., 2010).
500 Human copper intake is most often due to presence of copper in drinking water, food or vitamin
501 supplement (Brewer, 2010; Pal et al., 2014) and is known to influence blood copper level (Silverio
502 Amancio et al., 2003; Uauy et al., 1998). Elevated copper concentrations have been observed in
503 many diseases, including cancer, Alzheimer and metabolic diseases (Brewer, 2010; Salustri et al.,
504 2010; Squitti et al., 2009; Tisato et al., 2010). Specifically, a positive link between copper level and
505 high BMI or obesity in children has been previously described in the same data from HELIX
506 (Vrijheid et al., submitted, available upon request) and elsewhere (Fan et al., 2017; Lima et al., 2006;
507 Yakinci et al., 1997). These studies in children are cross-sectional and an important question relates
508 to the direction of any causal link between copper level and obesity. Overweight might disrupt
509 copper level, for example due to a higher food intake linked to an increased appetite, as
510 hypothesized by Yakinci et al. (1997), or due to metabolic changes. Other arguments exist in favour
511 of copper being a proximal cause of overweight. Nutrition studies in human showed that changes
512 in copper intake (depletion or supplementation) can have adverse health effects such as metabolic
513 and cardiovascular abnormalities (Klevay, 2018; Milne and Weswig, 2018). The toxicity of copper
514 (Brewer, 2010) and its ability to induce oxidative stress are well-known in humans (Brewer, 2010;
515 Uriu-Adams and Keen, 2005) and from animal models (Galhardi et al., 2004; Pereira et al., 2016).
516 Part of this process can occur via methylation changes, as shown in zebrafish, in which stress-
517 related gene expression can be modified by early-life copper exposure (Dorts et al., 2016). Although
518 we cannot formally exclude a situation in which copper levels are influenced by the child's
519 overweight status (e.g. as in Figure 4D), in particular due to the cross-sectional design of our study
520 of DNA methylation-BMI links, the above-mentioned experimental and prospective studies make

521 copper a plausible causal biomarker or predictor of BMI , with clues for effects possibly mediated
 522 by methylation changes (as in Figure 4A).

523



524

525 Figure 4: Different causal models involving one exposure, one CpG site and the outcome (BMI).
 526 Among the causal models corresponding to the exposure-BMI link corresponding to reverse
 527 causality (C to K), the models in which the Meet-in-the-Middle approach is expected to be able to
 528 provide a truly negative result are displayed in green, and causal models in which our Meet-in-the-
 529 Middle approach can be expected to provide a false-positive result are displayed in red. (in color;
 530 1.5-column fitting image)

531

532

533 On the contrary, the negative and less strong association of PFOS level with zBMI may correspond
 534 to reverse causality (see below). An influence of PFOS levels on blood or serum methylation
 535 change has some plausibility (Ruiz-Hernandez et al., 2015; Watkins et al., 2014).

536 **4.2. Can information be borrowed from the blood methylome to reduce the dimension**
 537 **of the exposome?**

538 Dimension reduction is one of the possible cures of the curse of dimensionality (Jimenez and
 539 Landgrebe, 1998). Assuming that part of the effects of the exposome on BMI are mediated by the

540 methylome, and that blood methylome constitutes a proxy of methylation levels in other target
541 organs, identifying exposures associated with methylation changes on CpGs relevant for BMI is a
542 way to restrict the analysis to a subgroup of exposures with higher likelihood of having an effect
543 on BMI. If the a priori CpG selection is accurate, one can expect the reduced exposome to contain
544 a higher proportion of true predictors of BMI than the full exposome. In a situation of expected
545 limited power, testing only the association of this reduced exposome with BMI could lead to a
546 better specificity (fewer false positives) than an agnostic ExWAS performed on the whole
547 exposome, which may suffer from high FDP because of the correlation within the exposome (Agier
548 et al., 2016). This approach however comes at the cost of possibly excluding true BMI predictors
549 within the exposome whose effect on BMI are not identifiable from the blood methylome. Yet, we
550 considered a high FDP to be of greater concern for exposome studies than low sensitivity.

551 Compared to the classical Meet-in-the-Middle framework (Vineis et al., 2013), we additionally
552 adjusted for the outcome when testing for associations between exposures and CpGs. This
553 adjustment was meant to exclude some cases of spurious association between the exposome and
554 the methylome. Thus, we used this tailored Meet-in-the-Middle approach with a different goal than
555 mediators identification, to focus on a subset of the exposome relevant for the considered
556 outcome, with the ultimate goal to increase specificity.

557 We used ExWAS-type methods to identify our reduced exposome, which has a possibly high
558 sensitivity and an expected high false positive rate (Agier et al. 2016). This could make our reduced
559 exposome possibly inaccurate, containing exposures selected by chance even if the test is adapted
560 to the underlying causal structure. However, our reduced exposome was considerably smaller than
561 the full exposome (4 exposures *vs.* 216) and our results provided far fewer discoveries (2 *vs.* 20)
562 than an agnostic ExWAS ignoring the methylome data. This lower number of discoveries is not
563 consistent with our approach having a higher rate of spuriously mediated exposures. Rather, it
564 could be explained by our approach having a lower FDP and/or a lower sensitivity. To discuss

565 these hypotheses, we compare the plausibility of the results of the Meet-in-the-Middle approach
566 and of the agnostic ExWAS obtained in the same population.

567 As discussed before, copper, identified by both approaches, is a plausible causal predictor of BMI
568 mediated by methylation. Among the 19 other exposures significantly associated with zBMI in the
569 agnostic approach, four associations corresponded to positive slopes: postnatal blood caesium
570 level, prenatal maternal active smoking and postnatal indoor particulate matter (PM_{2.5})
571 concentration and absorbance. The last three associations may be (at least partly) due to the well-
572 known effect of prenatal and postnatal smoking on the obesity (Oken et al., 2008; Vázquez Nava
573 et al., 2006): indeed, postnatal smoking variables were used to compute indoor particulate matter
574 levels. An odd ratio greater than 1 was also reported for the influence of high urinary caesium levels
575 on diabetes, an obesity-related outcome (Menke et al., 2016). The remaining 15 associations
576 corresponded to negative slopes (that is, a lower BMI with increasing exposure levels assessed in
577 child blood): it was the case for postnatal blood levels of perfluorinated compounds, cobalt and of
578 persistent organochlorine compounds (some PCBs, DDT (an insecticide), its metabolite DDE, and
579 HCB). For some of these exposures, associations of prenatal levels with overweight or obesity have
580 been reported in the literature. PFAS, including PFOS, prenatal exposures have been associated
581 with *higher* BMI in childhood and adulthood (Braun, 2017; Lauritzen et al., 2018; Saikat et al., 2013);
582 obesogenic effects of early-life PCBs levels have also been reported (Heindel & vom Saal, 2009;
583 Thayer et al., 2012). However, for childhood exposure, several studies found *negative* associations
584 of PCBs (Rönn et al., 2011) and PFAS (Nelson et al., 2010) with BMI, similarly to our results.
585 These negative associations may be indicative of reverse causality. This is supported by the facts
586 that 1) the postnatal exposome and outcome were assessed simultaneously; 2) lipophilic
587 compounds such as PFAS, PCBs and DDT are stored in fat, which makes the blood level a possibly
588 inaccurate marker of exposure: studies on seals showed that, for identical levels of exposures,
589 higher persistent organic pollutants (POP) levels in blood are found in thin compared to fatter
590 animals (Debier et al., 2006, Lydersen et al., 2002). In humans, Rönn et al. (2011) found positive

591 associations of fat mass with blood levels of lightly chlorinated PCBs with fat mass and negative
592 association for highly chlorinated PCBs, which are more lipophilic and therefore more stored in
593 fat. This is in favour of the negative associations between POPs (persistent organic pollutants)
594 levels and BMI being explained by fat levels influencing the blood POPs levels (causal models C
595 to K, Figure 4), rather than by POPs influencing adiposity (causal models A or B). Unfortunately,
596 we could not access fat tissue POP concentrations, which may have been a better exposure
597 biomarker.

598 The Meet-in-the-Middle approach did not select PCBs, DDT, DDE, HCB and other POPs
599 highlighted by the agnostic ExWAS. This may be due to our approach coping more efficiently with
600 reverse causality. A recent simulation study showed that Sobel's test of mediation, under specific
601 hypotheses, has far better detection rates for true mediation effects than in a reverse causality
602 situation (Tobi et al., 2018). Our approach, which has similarities to a mediation test, is moreover
603 specifically designed to efficiently handle some situations of reverse causality.

604 Figure 4C to K shows all possible causal situations corresponding to reverse causality between an
605 exposure (E) and zBMI. An ExWAS approach, which cannot distinguish between E influencing
606 zBMI and zBMI influencing E, is expected to detect all of them (assuming perfect power), whereas
607 our Meet-in-the-Middle approach is not expected to conclude that E and zBMI are linked in cases
608 C and I. Moreover, since we adjusted our second test (the exposome-methylome test) for zBMI,
609 cases G and H, which seem very plausible from a biological point of view, should not be
610 erroneously detected by our approach either (indeed, there is no exposure-CpG association
611 conditionally on BMI in models G and H). Finally, as methylation is not likely to influence exposure
612 levels in another way than via its link with the level of fat and as we preselected CpGs in pathways
613 likely to link exposures and BMI, cases E, F and K may not be very frequent situations. Case D
614 however cannot be excluded and may be erroneously detected by our approach, which may be the
615 case for PFOS in our study. We can thus hypothesize that our Meet-in-the-Middle approach is

616 likely to (erroneously) identify a link between E and Y in fewer causal models corresponding to
617 reverse causation than the ExWAS approach. This may limit the number of false-positive findings
618 compared to the agnostic ExWAS.

619 Regarding now the sensitivity of our approach, it strongly depends on the proportion of exposures
620 truly affecting BMI for which one of the underlying mechanisms relates to changes in blood DNA
621 methylation. If the effect of most exposures is (at least partly) reflected in the blood methylome,
622 then we can expect a high sensitivity, empowered by the smaller dimension of the reduced
623 exposome compared to the whole exposome, which makes the correction for multiple testing less
624 penalizing than in the agnostic ExWAS. If on the contrary most exposures affecting BMI do so by
625 pathways unrelated to the methylome, then our approach is expected to have a low sensitivity. We
626 did not find any association with prenatal exposures to endocrine disruptors (some of which are
627 possible obesogens (Braun, 2017)) or with dietary factors such as soda intake, whose effects on
628 BMI are well documented (Heo et al., 2017; Hu and Malik, 2010; Murakami and Livingstone, 2016)
629 but may not be visible from the blood methylome. Overall, the sensitivity of our approach will
630 depend on the (unknown) proportion of exposures whose effect is mediated by the methylome.
631 Whether the likely lower sensitivity is considered to be compensated by the expected decrease in
632 FDP would depend on the specific nature of the exposome study, with e.g. confirmatory studies
633 putting the emphasis on the limitation of false positive signals.

634 **4.3. Is a preliminary dimension reduction of the methylome necessary to efficiently**
635 **borrow information from it?**

636 We tried to identify a relevantly reduced exposome with ExWAS-type methods applied to
637 methylome data. ExWAS is known to have a high-false positive rate, particularly if correlation
638 exists among predictors (Agier et al. 2016). To cope with this problem, we modified the original
639 Meet-in-the-Middle framework by testing the exposome-CpGs associations only for CpGs which
640 we had found to be associated with zBMI. This drastically reduced the number of tests done at

641 step *c*) (10,152 tests versus 83,487,888, a division by 8200) and thus increased power for this step.
642 Moreover, we also performed a preliminary reduction of the size of the methylome, relying on
643 external biological information from the KEGG pathways database (Tanabe and Kanehisa, 2012).
644 Our preselection of CpGs was, again, drastic, as it reduced the methylome from 386,518 to 2284
645 CpGs, possibly allowing a gain of power to test associations between methylation levels and zBMI,
646 under strong assumptions. These assumptions relate to the quality and completeness of the KEGG
647 database (and of our query) to identify BMI-relevant genes and to the quality of the ILLUMINA
648 annotation about enhancers CpGs and on the link between genes and CpGs. These can be
649 questioned: in particular the KEGG database is based on publications of various quality, and
650 pathways selected may not be relevant for 6-10-year-old children. Moreover, ILLUMINA
651 annotation is not tissue-specific whereas enhancer characteristic is; it is consequently unclear
652 whether the ILLUMINA enhancer tag is relevant for blood immune cells, on which methylation
653 was assessed. Finally, changes in the methylation level of the enhancer CpGs of one gene may not
654 be linked with the protein level of this same gene, but of a remote gene (Jang et al., 2017).

655 To try to quantify the impact of our CpGs preselection, we used two approaches. First, we
656 performed a sensitivity analysis, in which the Meet-in-the-Middle approach was performed without
657 step *a*) of CpGs preselection. This led to quite different results. Although, in the last step, postnatal
658 levels of copper and PFOS were again found to be associated with zBMI, 4 other exposures were
659 additionally significantly negatively associated with zBMI: postnatal blood levels of HCB, PBDE
660 153, DDT and caesium. Contrarily to copper and PFOS levels, which were associated with
661 respectively 1110 and 180 CpGs at step *c*), these four additional exposures were each associated
662 with not more than 2 CpGs. Moreover, as discussed above, for the organochlorine compounds,
663 their association with BMI may be due to reverse causality. Thus, the discoveries of this analysis
664 without CpGs preselection contained more compounds little likely to be causative predictors of
665 the zBMI than the discoveries of the main analysis. This might be explained by the high number
666 of false positives expected from an ExWAS-type method in high dimension, at steps *b*) and *c*). The

667 dimension reduction of the methylome in our case seemed to help to cope with this problem.
668 Further studies are needed to determine if this impact of CpGs preselection is generally expected
669 or not.

670 To further test the quality of our a priori CpGs preselection, we tried to establish if our a priori
671 reduction led or not to concentration of information, by quantifying the overrepresentation of our
672 preselected CpGs in the significant associations found by a methylome-wide analysis relating the
673 whole methylome to zBMI. The information was indeed concentrated (apparently higher specificity
674 of the selected CpGs), but at the cost of a considerable loss of information: giving up at least 1760
675 CpGs associated with zBMI implies a risk not to identify some exposures truly associated with
676 CpGs which are themselves associated with zBMI, and consequently may decrease the sensitivity
677 of our approach (Supplementary Material 8). This illustrates, again, that our approach, which was
678 built to gain in specificity, may in principle have a cost in terms of loss in sensitivity compared to
679 the agnostic approach.

680 **4.4. Correction for cell-type heterogeneity**

681 Correction for the proportion of the cell-types in which DNA methylation is assessed is now
682 applied in most methylome studies, although its relevance is debated (Holbrook et al., 2017).
683 Between-subject differences in DNA methylation may or may not depend on cellular
684 heterogeneity. Generally, it is assumed that differences in DNA methylation not due to cell-types
685 mixture are more likely to be causative of disease, while methylation differences caused by
686 differences in cell-types proportion are considered a likely consequence of disease or at least of the
687 disease process (i.e. to correspond to reverse causality, for example in the case of obesity-induced
688 inflammation leading to changes in leukocytes proportion (Zeyda and Stulnig, 2009)), or to be a
689 confounder whose effect needs to be controlled for. However, diseases are also often associated
690 with the distribution of cell types, and cell type proportion can also in some situations be a cause
691 of disease or a marker for e.g. inflammatory or immune-related diseases (Holbrook et al., 2017).

692 This is particularly relevant for obesity development, which is known to involve inflammatory
693 pathways (Hotamisligil, 2003). The differential methylation driven by cell-type heterogeneity could
694 therefore mediate exposure effects rather than be a consequence of the overweight. Thus, cell types
695 proportion could be 1) a consequence or very close marker of our outcome; in such case, adjusting
696 for it in the model linking methylation and zBMI is irrelevant, as adjustment for consequences of
697 the diseases, which are by definition not confounders, could have harmful consequences (Barton
698 et al., 2019); 2) a cause of the health outcome. In this case, if (option 2a) it is a mediator of an effect
699 of an exposure on zBMI or a proxy of such a mediator, we should not adjust for it: as we want to
700 identify exposures whose effect on zBMI is biologically mediated by the methylome layer,
701 correcting for cell-types could prevent us from selecting potential exposures of interest associated
702 with zBMI and whose effect is mediated by cell-type-dependent methylation. If (2b) it is not a
703 mediator, i.e. if it is a cause of outcome but not a consequence of exposure, it is a potential
704 confounder and as such should be corrected for in the test of association between methylation and
705 zBMI. We chose not to adjust the DNA methylation-BMI model of the main analysis for cell-type
706 heterogeneity because we a priori consider hypotheses 1) and 2a) as more likely than 2b), and
707 because the consequence of erroneously adjusting may, in our study in which identifying the
708 intermediate causal factors was not the main aim, be more harmful (Barton et al., 2019) than
709 adjusting.

710 In our sensitivity analysis IV, we repeated the whole Meet-in-the-Middle design considering the
711 cell-types instead of the methylation data as the intermediate layer. Final results were very similar
712 to those of the main analysis, making it possible that differential methylation driven by cell-type
713 heterogeneity explained the association between methylome layer and both copper and zBMI, and
714 which therefore mediate copper effects (case 2a). However, the reduced exposome was slightly
715 smaller when considering cell-types and the proportion mediated was lower (13% *vs.* 29%). This
716 may mean that the cell-type did not convey as much information on the biological effect of
717 exposures than DNA methylation on the path between selected exposures and zBMI.

718 We acknowledge several limitations to this original work; first, our results are dependent on the
719 quality of a priori information; second, the methylome and the outcome (and a part of the
720 exposome) were assessed simultaneously, a design which makes difficult to exclude reverse
721 causation with an agnostic test of association. Measurement error is also expected for both the
722 methylome and the exposome; finally, only monotonous associations were tested to limit the
723 number of tests. The strengths of our study include a multilayer analysis of the exposome,
724 methylome and BMI in a large and well-characterized population, the consideration of multiple
725 testing in the composite tests and of a priori information on the methylome-BMI relation to restrict
726 the number of tests done. This composite design may allow improving the specificity of exposome-
727 health studies and limit associations due to reverse causality, with a possible cost on sensitivity.

728

729 Conclusion

730 This work is to our knowledge the first epidemiological study relying on an intermediate blood
731 methylome layer to try to better characterize the exposome-health association. Purely agnostic
732 exposome studies are expected to suffer from a high false positives rate, and possibly a low
733 sensitivity (due to the correlation within the exposome) (Agier et al. 2016). Our approach may allow
734 reducing the false positive rate by using a modified Meet-in-the-Middle design that permits a
735 biologically driven reduction of the exposome, which may in particular allow discarding some of
736 the associations of the outcome with the exposome due to reverse causality. It comes at a cost of
737 being insensitive to exposures acting on the outcome via pathways not causing changes in the
738 methylome of peripheral blood. Extensions of this approach to other biologically relevant layers
739 might allow avoiding this limitation in the future.

740

741 Funding:

742 The study has received funding from the European Community's Seventh Framework Programme
743 (FP7/2007-2013) under grant agreement n° 308333 – the HELIX project for data collection and
744 analyses. The Norwegian Mother and Child Cohort Study is supported by the Norwegian Ministry
745 of Health and Care Services and the Ministry of Education and Research, NIH/NIEHS (contract
746 no N01-ES-75558), NIH/NINDS (grant n°.1 UO1 NS 047537-01 and grant no.2 UO1 NS
747 047537-06A1). We also received support from Région Auvergne-Rhône-Alpes for collaborations
748 with Catalunya.

749

750 Acknowledgments:

751 We acknowledge the input of the HELIX consortium. We are grateful to all the participating
752 families in the six countries who took part in this cohort study (BiB, EDEN, INMA, KANC, MoBa

753 and RHEA cohorts). We acknowledge the work of all fieldworkers and health professionals who
754 make the data collection possible. Regarding the EDEN cohort, we further thank Sonia Brishoual,
755 Angélique Serre and Michele Grosdenier (Poitiers Biobank, CRB BB-0033-00068, Poitiers, France)
756 for biological sample management, Elodie Migault, Manuela Boue and Sandy Bertin (Clinical
757 Investigation Centre, Inserm CIC1402, CHU de Poitiers, Poitiers, France) for planification and
758 investigational actions. We are also grateful to Veronique Ferrand-Rigalleau, Céline Leger and
759 Noella Gorry (CHU de Poitiers, Poitiers, France) for administrative assistance (EDEN). We
760 acknowledge the support of Région Auvergne Rhône Alpes for scientific collaboration with
761 Catalogna. Regarding the MoBa study, we like to thank Ingvild Essen and Jorunn Evandt for
762 thorough field work, Heidi Marie Nordheim for biological sample management and the MoBa
763 administrative unit (MoBa). The Norwegian Mother, Father and Child Cohort Study is supported
764 by the Norwegian Ministry of Health and Care Services and the Ministry of Education and
765 Research. We are grateful to all the participating families in Norway who take part in this on-going
766 cohort study.

767 **Contributions:**

768 SC and RS designed the analytical and statistical methods. SC analysed the data. SC, MB, LA, MV
769 and RS interpreted the results and wrote the paper. MV, MB and LM coordinated the HELIX data
770 collection. All authors contributed to the data collection and to the manuscript, and approved the
771 manuscript.

772

773 **Declaration of interest:**

774 We declare no competing interests.

References

- Agay-Shay, K., Martinez, D., Valvi, D., Garcia-Esteban, R., Basagaña, X., Robinson, O., Casas, M., Sunyer, J., Vrijheid, M., 2015. Exposure to endocrine-disrupting chemicals during pregnancy and weight at 7 years of age: A multi-pollutant approach. *Environ. Health Perspect.* 123, 1030–1037. <https://doi.org/10.1289/ehp.1409049>
- Agier, L., Basagaña, X., Maitre, L., Granum, B., Bird, P.K., Casas, M., Oftedal, B., Wright, J., Andrusaityte, S., de Castro, M., Cequier, E., Chatzi, L., Donaire-Gonzalez, D., Grazuleviciene, R., Haug, L.S., Sakhi, A.K., Leventakou, V., McEachan, R., Nieuwenhuijsen, M., Petraviciene, I., Robinson, O., Roumeliotaki, T., Sunyer, J., Tamayo-Uria, I., Thomsen, C., Urquiza, J., Valentin, A., Slama, R., Vrijheid, M., Siroux, V., 2019. Early-life exposome and lung function in children in Europe: an analysis of data from the longitudinal, population-based HELIX cohort. *Lancet Planet. Heal.* 3, e81–e92. [https://doi.org/10.1016/S2542-5196\(19\)30010-5](https://doi.org/10.1016/S2542-5196(19)30010-5)
- Agier, L., Portengen, L., Chadeau-Hyam, M., Basagaña, X., Giorgis-Allemand, L., Siroux, V., Robinson, O., Vlaanderen, J., González, J.R., Nieuwenhuijsen, M.J., Vineis, P., Vrijheid, M., Slama, R., Vermeulen, R., 2016. A Systematic Comparison of Linear Regression-Based Statistical Methods to Assess Exposome-Health Associations. *Environ. Health Perspect.* 124, 1848–1856. <https://doi.org/10.1289/EHP172>
- Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., Irizarry, R.A., 2014. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 30, 1363–9. <https://doi.org/10.1093/bioinformatics/btu049>
- Baccarelli, A., Wright, R.O., Bollati, V., Tarantini, L., Litonjua, A.A., Suh, H.H., Zanobetti, A., Sparrow, D., Vokonas, P.S., Schwartz, J., 2009. Rapid DNA Methylation Changes after Exposure to Traffic Particles. *Am. J. Respir. Crit. Care Med.* 179, 572–578. <https://doi.org/10.1164/rccm.200807-1097OC>
- Barton, S.J., Melton, P.E., Titcombe, P., Murray, R., Rauschert, S., Lillycrop, K.A., Huang, R.C., Holbrook, J.D., Godfrey, K.M., 2019. In Epigenomic Studies, Including Cell-Type Adjustments in Regression Models Can Introduce Multicollinearity, Resulting in Apparent Reversal of Direction of Association. *Front. Genet.* 10. <https://doi.org/10.3389/fgene.2019.00816>
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B.* <https://doi.org/10.2307/2346101>
- Berger, S.L., Kouzarides, T., Shiekhatar, R., Shilatifard, A., 2009. An operational definition of epigenetics. *Genes Dev.* 23, 781–783. <https://doi.org/10.1101/gad.1787609>
- Braun, J.M., 2017. Early-life exposure to EDCs: Role in childhood obesity and neurodevelopment. *Nat. Rev. Endocrinol.* 13, 161–173. <https://doi.org/10.1038/nrendo.2016.186>
- Brewer, G.J., 2010. Copper toxicity in the general population. *Clin. Neurophysiol.* 121, 459–460. <https://doi.org/10.1016/j.clinph.2009.12.015>
- Buck Louis, G.M., Schisterman, E.F., Sweeney, A.M., Wilcosky, T.C., Gore-Langton, R.E., Lynch, C.D., Boyd Barr, D., Schrader, S.M., Kim, S., Chen, Z., Sundaram, R., 2011. Designing prospective cohort studies for assessing reproductive and developmental toxicity

- during sensitive windows of human reproduction and development - the LIFE Study. *Paediatr. Perinat. Epidemiol.* 25, 413–424. <https://doi.org/10.1111/j.1365-3016.2011.01205.x>
- Buuren, S. van, Groothuis-Oudshoorn, K., 2011. **mice** : Multivariate Imputation by Chained Equations in R. *J. Stat. Softw.* 45, 1–67. <https://doi.org/10.18637/jss.v045.i03>
- Cao-Lei, L., Dancause, K.N., Elgbeili, G., Massart, R., Szyf, M., Liu, A., Laplante, D.P., King, S., 2015. DNA methylation mediates the impact of exposure to prenatal maternal stress on BMI and central adiposity in children at age 13½ years: Project Ice Storm. *Epigenetics* 10, 749–761. <https://doi.org/10.1080/15592294.2015.1063771>
- Carone, B.R., Fauquier, L., Habib, N., Shea, J.M., Hart, C.E., Li, R., Bock, C., Li, C., Gu, H., Zamore, P.D., Meissner, A., Weng, Z., Hofmann, H.A., Friedman, N., Rando, O.J., 2010. Paternally Induced Transgenerational Environmental Reprogramming of Metabolic Gene Expression in Mammals. *Cell* 143, 1084–1096. <https://doi.org/10.1016/J.CELL.2010.12.008>
- Casas, M., Basagaña, X., Sakhi, A.K., Haug, L.S., Philippat, C., Granum, B., Manzano-Salgado, C.B., Brochot, C., Zeman, F., de Bont, J., Andrusaityte, S., Chatzi, L., Donaire-Gonzalez, D., Giorgis-Allemand, L., Gonzalez, J.R., Gracia-Lavedan, E., Grazuleviciene, R., Kampouri, M., Lyon-Caen, S., Pañella, P., Petraviciene, I., Robinson, O., Urquiza, J., Vafeiadi, M., Vernet, C., Waiblinger, D., Wright, J., Thomsen, C., Slama, R., Vrijheid, M., 2018. Variability of urinary concentrations of non-persistent chemicals in pregnant women and school-aged children. *Environ. Int.* 121, 561–573. <https://doi.org/10.1016/J.ENVINT.2018.09.046>
- Chadeau-Hyam, M., Athersuch, T.J., Keun, H.C., De Iorio, M., Ebbels, T.M.D., Jenab, M., Sacerdote, C., Bruce, S.J., Holmes, E., Vineis, P., 2011. Meeting-in-the-middle using metabolic profiling – a strategy for the identification of intermediate biomarkers in cohort studies. *Biomarkers* 16, 83–88. <https://doi.org/10.3109/1354750X.2010.533285>
- Chatzi, L., Leventakou, V., Vafeiadi, M., Koutra, K., Roumeliotaki, T., Chalkiadaki, G., Karachaliou, M., Daraki, V., Kyriklaki, A., Kampouri, M., Fthenou, E., Sarri, K., Vassilaki, M., Fasoulaki, M., Bitsios, P., Koutis, A., Stephanou, E.G., Kogevinas, M., 2017. Cohort Profile: The Mother-Child Cohort in Crete, Greece (Rhea Study). *Int. J. Epidemiol.* 46, 1392-1393k. <https://doi.org/10.1093/ije/dyx084>
- Chun, H., Keleş, S., 2010. Sparse partial least squares regression for simultaneous dimension reduction and variable selection. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 72, 3–25. <https://doi.org/10.1111/j.1467-9868.2009.00723.x>
- de Onis, M., Onyango, A.W., Borghi, E., Siyam, A., Nishida, C., Siekmann, J., 2007. Development of a WHO growth reference for school-aged children and adolescents. *Bull. World Health Organ.* 85, 660–7.
- Dekkers, K.F., van Ijzerman, M., Sliker, R.C., Moed, M.H., Bonder, M.J., van Galen, M., Mei, H., Zhernakova, D. V., van den Berg, L.H., Deelen, J., van Dongen, J., van Heemst, D., Hofman, A., Hottenga, J.J., van der Kallen, C.J.H., Schalkwijk, C.G., Stehouwer, C.D.A., Tigchelaar, E.F., Uitterlinden, A.G., Willemsen, G., Zhernakova, A., Franke, L., 't Hoen, P.A.C., Jansen, R., van Meurs, J., Boomsma, D.I., van Duijn, C.M., van Greevenbroek, M.M.J., Veldink, J.H., Wijmenga, C., van Zwet, E.W., Slagboom, P.E., Jukema, J.W., Heijmans, B.T., 2016. Blood lipids influence DNA methylation in circulating cells. *Genome Biol.* 17, 138. <https://doi.org/10.1186/s13059-016-1000-6>

- Dorts, J., Falisse, E., Schoofs, E., Flamion, E., Kestemont, P., Silvestre, F., 2016. DNA methyltransferases and stress-related genes expression in zebrafish larvae after exposure to heat and copper during reprogramming of DNA methylation. *Sci. Rep.* 6, 34254. <https://doi.org/10.1038/srep34254>
- Fan, Y., Zhang, C., Bu, J., 2017. Relationship between selected serum metallic elements and obesity in children and adolescent in the U.S. *Nutrients* 9, 1–12. <https://doi.org/10.3390/nu9020104>
- Fasanelli, F., Baglietto, L., Ponzi, E., Guida, F., Campanella, G., Johansson, Mattias, Grankvist, K., Johansson, Mikael, Assumma, M.B., Naccarati, A., Chadeau-Hyam, M., Ala, U., Faltus, C., Kaaks, R., Risch, A., De Stavola, B., Hodge, A., Giles, G.G., Southey, M.C., Relton, C.L., Haycock, P.C., Lund, E., Polidoro, S., Sandanger, T.M., Severi, G., Vineis, P., 2015. Hypomethylation of smoking-related genes is associated with future lung cancer in four prospective cohorts. *Nat. Commun.* 6, 10192. <https://doi.org/10.1038/ncomms10192>
- Feil, R., Fraga, M.F., 2012. Epigenetics and the environment: emerging patterns and implications. *Nat. Rev. Genet.* 13, 97–109. <https://doi.org/10.1038/nrg3142>
- Finucane, M.M., Stevens, G.A., Cowan, M.J., Danaei, G., Lin, J.K., Paciorek, C.J., Singh, G.M., Gutierrez, H.R., Lu, Y., Bahalim, A.N., Farzadfar, F., Riley, L.M., Ezzati, M., 2011. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9·1 million participants. *Lancet* 377, 557–567. [https://doi.org/10.1016/S0140-6736\(10\)62037-5](https://doi.org/10.1016/S0140-6736(10)62037-5)
- Fortin, J.-P., Fertig, E., Hansen, K., 2014. shinyMethyl: interactive quality control of Illumina 450k DNA methylation arrays in R. *F1000Research* 3, 175. <https://doi.org/10.12688/f1000research.4680.2>
- Fradin, D., Boëlle, P.Y., Belot, M.P., Lachaux, F., Tost, J., Besse, C., Deleuze, J.F., De Filippo, G., Bougnères, P., 2017. Genome-Wide Methylation Analysis Identifies Specific Epigenetic Marks in Severely Obese Children. *Sci. Rep.* 7. <https://doi.org/10.1038/srep46311>
- Galhardi, C.M., Diniz, Y.S., Faine, L.A., Rodrigues, H.G., Burneiko, R.C.M., Ribas, B.O., Novelli, E.L.B., 2004. Toxicity of copper intake: Lipid profile, oxidative stress and susceptibility to renal dysfunction. *Food Chem. Toxicol.* <https://doi.org/10.1016/j.fct.2004.07.020>
- Grazuleviciene, R., Danileviciute, A., Dedele, A., Vencloviene, J., Andrusaityte, S., Uždanaviciute, I., Nieuwenhuijsen, M.J., 2015. Surrounding greenness, proximity to city parks and pregnancy outcomes in Kaunas cohort study. *Int. J. Hyg. Environ. Health* 218, 358–365. <https://doi.org/10.1016/j.ijheh.2015.02.004>
- Guxens, M., Ballester, F., Espada, M., Fernández, M.F., Grimalt, J.O., Ibarluzea, J., Olea, N., Rebagliato, M., Tardón, A., Torrent, M., Vioque, J., Vrijheid, M., Sunyer, J., 2012. Cohort Profile: The INMA—Infancia y Medio Ambiente—(Environment and Childhood) Project. *Int. J. Epidemiol.* 41, 930–940. <https://doi.org/10.1093/ije/dyr054>
- Hansen, K.D., 2016. IlluminaHumanMethylation450kanno.ilmn12.hg19: Annotation for Illumina's 450k methylation arrays. R Packag. version 0.6.0.
- Haug, L.S., Sakhi, A.K., Cequier, E., Casas, M., Maitre, L., Basagana, X., Andrusaityte, S., Chalkiadaki, G., Chatzi, L., Coen, M., de Bont, J., Dedele, A., Ferrand, J., Grazuleviciene, R., Gonzalez, J.R., Gutzkow, K.B., Keun, H., McEachan, R., Meltzer, H.M., Petraviciene, I., Robinson, O., Saulnier, P.J., Slama, R., Sunyer, J., Urquiza, J., Vafeiadi, M., Wright, J., Vrijheid, M., Thomsen, C., 2018. In-utero and childhood chemical exposome in six European mother-child cohorts. *Environ. Int.* 121, 751–763.

<https://doi.org/10.1016/j.envint.2018.09.056>

- Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G.J., Susser, E.S., Slagboom, P.E., Lumey, L.H., 2008. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci.* 105, 17046–17049. <https://doi.org/10.1073/pnas.0806560105>
- Heo, E.J., Shim, J.E., Yoon, E.Y., 2017. Systematic Review on the Study of the Childhood and Adolescent Obesity in Korea: Dietary Risk Factors. *Korean J. Community Nutr.* 22, 191. <https://doi.org/10.5720/kjcn.2017.22.3.191>
- Heude, Barbara, Forhan, Anne, Slama, Rémy, Douhaud, L., Bedel, S., Saurel-Cubizolles, M.-J., Hankard, Régis, Thiebaugeorges, Olivier, De Agostini, Maria, Annesi-Maesano, Isabella, Kaminski, Monique, Charles, Marie-Aline, Annesi-Maesano, I, Bernard, J., Botton, J., Charles, M-A, Dargent-Molina, P., de Lauzon-Guillain, B., Ducimetière, P., de Agostini, M, Foliguet, B., Forhan, A, Fritel, X., Germa, A., Goua, V., Hankard, R, Heude, B, Kaminski, M, Larroque, B., Lelong, N., Lepeule, J., Magnin, G., Marchand, L., Nabet, C., Pierre, F., Slama, R, Saurel-Cubizolles, M., Schweitzer, M., Thiebaugeorges, O, 2016. Cohort Profile: The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *Int. J. Epidemiol.* 45, 353–363. <https://doi.org/10.1093/ije/dyv151>
- Ho, S.-M., Johnson, A., Tarapore, P., Janakiram, V., Zhang, X., Leung, Y.-K., 2012. Environmental Epigenetics and Its Implication on Disease Risk and Health Outcomes. *ILAR J.* 53, 289–305. <https://doi.org/10.1093/ilar.53.3-4.289>
- Holbrook, J.D., Huang, R.-C., Barton, S.J., Saffery, R., Lillycrop, K.A., 2017. Is cellular heterogeneity merely a confounder to be removed from epigenome-wide association studies? *Epigenomics* 9, 1143–1150. <https://doi.org/10.2217/epi-2017-0032>
- Holtcamp, W., 2012. Obesogens: an environmental link to obesity. *Environ. Health Perspect.* 120, a62-8. <https://doi.org/10.1289/ehp.120-a62>
- Hotamisligil, G.S., 2003. Inflammatory pathways and insulin action. *Int. J. Obes.* 27, S53–S55. <https://doi.org/10.1038/sj.ijo.0802502>
- Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K., Kelsey, K.T., 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 13, 86. <https://doi.org/10.1186/1471-2105-13-86>
- Hu, F.B., Malik, V.S., 2010. Sugar-sweetened beverages and risk of obesity and type 2 diabetes: Epidemiologic evidence. *Physiol. Behav.* 100, 47–54. <https://doi.org/10.1016/j.physbeh.2010.01.036>
- Janesick, A., Blumberg, B., 2011. Endocrine disrupting chemicals and the developmental programming of adipogenesis and obesity. *Birth Defects Res. Part C Embryo Today Rev.* 93, 34–50. <https://doi.org/10.1002/bdrc.20197>
- Jang, H.S., Shin, W.J., Lee, J.E., Do, J.T., 2017. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes (Basel)*. 8. <https://doi.org/10.3390/GENES8060148>
- Jimenez, L.O., Landgrebe, D.A., 1998. Supervised classification in high-dimensional space: Geometrical, statistical, and asymptotical properties of multivariate data, *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*. Institute of Electrical

and Electronics Engineers. <https://doi.org/10.1109/5326.661089>

- Joubert, B.R., Felix, J.F., Yousefi, P., Bakulski, K.M., Just, A.C., Breton, C., Reese, S.E., Markunas, C.A., Richmond, R.C., Xu, C.J., Küpers, L.K., Oh, S.S., Hoyo, C., Gruziova, O., Söderhäll, C., Salas, L.A., Baiz, N., Zhang, H., Lepeule, J., Ruiz, C., Ligthart, S., Wang, T., Taylor, J.A., Duijts, L., Sharp, G.C., Jankipersadsing, S.A., Nilsen, R.M., Vaez, A., Fallin, M.D., Hu, D., Litonjua, A.A., Fuemmeler, B.F., Huen, K., Kere, J., Kull, I., Munthe-Kaas, M.C., Gehring, U., Bustamante, M., Saurel-Coubizolles, M.J., Quraishi, B.M., Ren, J., Tost, J., Gonzalez, J.R., Peters, M.J., Håberg, S.E., Xu, Z., Van Meurs, J.B., Gaunt, T.R., Kerkhof, M., Corpeleijn, E., Feinberg, A.P., Eng, C., Baccarelli, A.A., Benjamin Neelon, S.E., Bradman, A., Merid, S.K., Bergström, A., Herceg, Z., Hernandez-Vargas, H., Brunekreef, B., Pinart, M., Heude, B., Ewart, S., Yao, J., Lemonnier, N., Franco, O.H., Wu, M.C., Hofman, A., McArdle, W., Van Der Vlies, P., Falahi, F., Gillman, M.W., Barcellos, L.F., Kumar, A., Wickman, M., Guerra, S., Charles, M.A., Holloway, J., Auffray, C., Tiemeier, H.W., Smith, G.D., Postma, D., Hivert, M.F., Eskenazi, B., Vrijheid, M., Arshad, H., Antó, J.M., Dehghan, A., Karmaus, W., Annesi-Maesano, I., Sunyer, J., Ghantous, A., Pershagen, G., Holland, N., Murphy, S.K., Demeo, D.L., Burchard, E.G., Ladd-Acosta, C., Snieider, H., Nystad, W., Koppelman, G.H., Relton, C.L., Jaddoe, V.W.V., Wilcox, A., Melén, E., London, S.J., 2016. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am. J. Hum. Genet.* 98, 680–696. <https://doi.org/10.1016/j.ajhg.2016.02.019>
- Klevay, L.M., 2018. Cardiovascular Disease from Copper Deficiency—A History. *J. Nutr.* 130, 489S-492S. <https://doi.org/10.1093/jn/130.2.489s>
- Kolatorova, L., Sramkova, M., Vitku, J., Vcelak, J., Lischkova, O., Starka, L., Duskova, M., 2018. Parabens and their relation to obesity. *Physiol. Res.* 67, S465–S472.
- Lauritzen, H.B., Larose, T.L., Øien, T., Sandanger, T.M., Odland, J.O., Van De Bor, M., Jacobsen, G.W., 2018. Prenatal exposure to persistent organic pollutants and child overweight/obesity at 5-year follow-up: A prospective cohort study. *Environ. Heal. A Glob. Access Sci. Source* 17, 9. <https://doi.org/10.1186/s12940-017-0338-x>
- Lehne, B., Drong, A.W., Loh, M., Zhang, W., Scott, W.R., Tan, S.T., Afzal, U., Scott, J., Jarvelin, M.R., Elliott, P., McCarthy, M.I., Kooner, J.S., Chambers, J.C., 2015. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol.* 16, 37. <https://doi.org/10.1186/s13059-015-0600-x>
- Lenters, V., Portengen, L., Rignell-Hydbom, A., Jönsson, B.A.G., Lindh, C.H., Piersma, A.H., Toft, G., Bonde, J.P., Heederik, D., Rylander, L., Vermeulen, R., 2016. 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.* 124, 365–372. <https://doi.org/10.1289/ehp.1408933>
- Lenters, V., Vermeulen, R., Portengen, L., 2018. Performance of variable selection methods for assessing the health effects of correlated exposures in case-control studies. *Occup. Environ. Med.* 75, 522–529. <https://doi.org/10.1136/oemed-2016-104231>
- Lillicrop, K.A., Phillips, E.S., Jackson, A.A., Hanson, M.A., Burdge, G.C., 2005. Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. *J. Nutr.* 135, 1382–1386. <https://doi.org/10.1093/nj/135/6/1382> [pii]
- Lillicrop, K.A., Phillips, E.S., Torrens, C., Hanson, M.A., Jackson, A.A., Burdge, G.C., 2008.

- Feeding pregnant rats a protein-restricted diet persistently alters the methylation of specific cytosines in the hepatic PPAR α promoter of the offspring. *Br. J. Nutr.* 100, 278–282. <https://doi.org/10.1017/S0007114507894438>
- Lima, S.C.V.C., Arrais, R.F., Sales, C.H., Almeida, M.G., De Sena, K.C.M., Oliveira, V.T.L., De Andrade, A.S., Pedrosa, L.F.C., 2006. Assessment of copper and lipid profile in obese children and adolescents. *Biol. Trace Elem. Res.* 114, 19–30. <https://doi.org/10.1385/BTER:114:1:19>
- Magnus, P., Birke, C., Vejrup, K., Haugan, A., Alsaker, E., Daltveit, A.K., Handal, M., Haugen, M., Høiseth, G., Knudsen, G.P., Paltiel, L., Schreuder, P., Tambs, K., Vold, L., Stoltenberg, C., 2016. Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). *Int. J. Epidemiol.* 45, 382–388. <https://doi.org/10.1093/ije/dyw029>
- Maitre, L., de Bont, J., Casas, M., Robinson, O., Aasvang, G.M., Agier, L., Andrušaitytė, S., Ballester, F., Basagaña, X., Borràs, E., Brochot, C., Bustamante, M., Carracedo, A., de Castro, M., Dedele, A., Donaire-Gonzalez, D., Estivill, X., Evandt, J., Fossati, S., Giorgis-Allemand, L., R Gonzalez, J., Granum, B., Grazuleviciene, R., Bjerve Gützkow, K., Småstuen Haug, L., Hernandez-Ferrer, C., Heude, B., Ibarluzea, J., Julvez, J., Karachaliou, M., Keun, H.C., Hjertager Krog, N., Lau, C.-H.E., Leventakou, V., Lyon-Caen, S., Manzano, C., Mason, D., McEachan, R., Meltzer, H.M., Petraviciene, I., Quentin, J., Roumeliotaki, T., Sabido, E., Saulnier, P.-J., Siskos, A.P., Siroux, V., Sunyer, J., Tamayo, I., Urquiza, J., Vafeiadi, M., van Gent, D., Vives-Usano, M., Waiblinger, D., Warembourg, C., Chatzi, L., Coen, M., van den Hazel, P., Nieuwenhuijsen, M.J., Slama, R., Thomsen, C., Wright, J., Vrijheid, M., 2018. Human Early Life Exposome (HELIX) study: a European population-based exposome cohort. *BMJ Open* 8, e021311. <https://doi.org/10.1136/bmjopen-2017-021311>
- Manrai, A.K., Cui, Y., Bushel, P.R., Hall, M., Karakitsios, S., Mattingly, C.J., Ritchie, M., Schmitt, C., Sarigiannis, D.A., Thomas, D.C., Wishart, D., Balshaw, D.M., Patel, C.J., 2017. Informatics and Data Analytics to Support Exposome-Based Discovery for Public Health. *Annu. Rev. Public Heal.* 38, 279–94. <https://doi.org/10.1146/annurev-publhealth>
- Marioni, R.E., McRae, A.F., Bressler, J., Colicino, E., Hannon, E., Li, S., Prada, D., Smith, J.A., Trevisi, L., Tsai, P.-C., Vojinovic, D., Simino, J., Levy, D., Liu, C., Mendelson, M., Satizabal, C.L., Yang, Q., Jhun, M.A., Kardina, S.L.R., Zhao, W., Bandinelli, S., Ferrucci, L., Hernandez, D.G., Singleton, A.B., Harris, S.E., Starr, J.M., Kiel, D.P., McLean, R.R., Just, A.C., Schwartz, J., Spiro, A., Vokonas, P., Amin, N., Ikram, M.A., Uitterlinden, A.G., van Meurs, J.B.J., Spector, T.D., Steves, C., Baccarelli, A.A., Bell, J.T., van Duijn, C.M., Fornage, M., Hsu, Y.-H., Mill, J., Mosley, T.H., Seshadri, S., Deary, I.J., 2018. Meta-analysis of epigenome-wide association studies of cognitive abilities. *Mol. Psychiatry* 1. <https://doi.org/10.1038/s41380-017-0008-y>
- Menke, A., Guallar, E., Cowie, C.C., 2016. Metals in urine and diabetes in U.S. adults. *Diabetes* 65, 164–171. <https://doi.org/10.2337/db15-0316>
- Milne, D.B., Weswig, P.H., 2018. Effect of Supplementary Copper on Blood and Liver Copper-containing Fractions in Rats. *J. Nutr.* <https://doi.org/10.1093/jn/95.3.429>
- Murakami, K., Livingstone, M.B.E., 2016. Associations between meal and snack frequency and overweight and abdominal obesity in US children and adolescents from National Health and Nutrition Examination Survey (NHANES) 2003–2012. *Br. J. Nutr.* 115, 1819–1829. <https://doi.org/10.1017/S0007114516000854>
- Nelson, J.W., Hatch, E.E., Webster, T.F., 2010. Exposure to polyfluoroalkyl chemicals and

- cholesterol, body weight, and insulin resistance in the general U.S. population. *Environ. Health Perspect.* 118, 197–202. <https://doi.org/10.1289/ehp.0901165>
- OECD-- Organisation for Economic Co-operation and Development, 2012. Proposal for a template, and guidance on developing and assessing the completeness of adverse outcome pathways 1–17.
- Oken, E., Levitan, E.B., Gillman, M.W., 2008. Maternal smoking during pregnancy and child overweight: Systematic review and meta-analysis. *Int. J. Obes.* <https://doi.org/10.1038/sj.ijo.0803760>
- Pal, A., Jayamani, J., Prasad, R., 2014. An urgent need to reassess the safe levels of copper in the drinking water: Lessons from studies on healthy animals harboring no genetic deficits. *Neurotoxicology.* <https://doi.org/10.1016/j.neuro.2014.05.005>
- Park, S.S., Skaar, D.A., Jirtle, R.L., Hoyo, C., 2017. Epigenetics, obesity and early-life cadmium or lead exposure. *Epigenomics* 9, 57–75. <https://doi.org/10.2217/epi-2016-0047>
- Pereira, T.C.B., Campos, M.M., Bogo, M.R., 2016. Copper toxicology, oxidative stress and inflammation using zebrafish as experimental model. *J. Appl. Toxicol.* 36, 876–885. <https://doi.org/10.1002/jat.3303>
- Plagemann, A., Harder, T., Brunn, M., Harder, A., Roepke, K., Wittrock-Staar, M., Ziska, T., Schellong, K., Rodekamp, E., Melchior, K., Dudenhausen, J.W., 2009. Hypothalamic proopiomelanocortin promoter methylation becomes altered by early overfeeding: an epigenetic model of obesity and the metabolic syndrome. *J. Physiol.* 587, 4963–4976. <https://doi.org/10.1113/jphysiol.2009.176156>
- Reinius, L.E., Acevedo, N., Joerink, M., Pershagen, G., Dahlén, S.E., Greco, D., Söderhäll, C., Scheynius, A., Kere, J., 2012. Differential DNA methylation in purified human blood cells: Implications for cell lineage and studies on disease susceptibility. *PLoS One* 7, e41361. <https://doi.org/10.1371/journal.pone.0041361>
- Richmond, R.C., Sharp, G.C., Ward, M.E., Fraser, A., Lyttleton, O., McArdle, W.L., Ring, S.M., Gaunt, T.R., Lawlor, D.A., Smith, G.D., Relton, C.L., 2016. DNA Methylation and BMI: Investigating Identified Methylation Sites at HIF3A in a Causal Framework. *Diabetes* 65, 1231–1244. <https://doi.org/10.2337/DB15-0996>
- Richmond, R.C., Timpson, N.J., Sørensen, T.I., 2015. Exploring possible epigenetic mediation of early-life environmental exposures on adiposity and obesity development: Figure 1. *Int. J. Epidemiol.* 44, 1191–1198. <https://doi.org/10.1093/ije/dyv066>
- Ruiz-Hernandez, A., Kuo, C.C., Rentero-Garrido, P., Tang, W.Y., Redon, J., Ordoñas, J.M., Navas-Acien, A., Tellez-Plaza, M., 2015. Environmental chemicals and DNA methylation in adults: A systematic review of the epidemiologic evidence. *Clin. Epigenetics* 7. <https://doi.org/10.1186/s13148-015-0055-7>
- Rzehak, P., Covic, M., Saffery, R., Reischl, E., Wahl, S., Grote, V., Weber, M., Xhonneux, A., Langhendries, J.P., Ferre, N., Closa-Monasterolo, R., Escribano, J., Verduci, E., Riva, E., Socha, P., Gruszfeld, D., Koletzko, B., 2017. DNA-Methylation and Body Composition in Preschool Children: Epigenome-Wide-Analysis in the European Childhood Obesity Project (CHOP)-Study. *Sci. Rep.* 7. <https://doi.org/10.1038/s41598-017-13099-4>
- Saikat, S., Kreis, I., Davies, B., Bridgman, S., Kamanyire, R., 2013. The impact of PFOS on health in the general population: A review. *Environ. Sci. Process. Impacts* 15, 329–335. <https://doi.org/10.1039/c2em30698k>

- Salustri, C., Barbati, G., Ghidoni, R., Quintiliani, L., Ciappina, S., Binetti, G., Squitti, R., 2010. Is cognitive function linked to serum free copper levels? A cohort study in a normal population. *Clin. Neurophysiol.* 121, 502–7. <https://doi.org/10.1016/j.clinph.2009.11.090>
- Silverio Amancio, O.M., Alves Chaud, D.M., Yanaguibashi, G., Esteves Hilário, M.O., 2003. Copper and zinc intake and serum levels in patients with juvenile rheumatoid arthritis. *Eur. J. Clin. Nutr.* 57, 706–712. <https://doi.org/10.1038/sj.ejcn.1601601>
- Siroux, V., Agier, L., Slama, R., 2016. The exposome concept: a challenge and a potential driver for environmental health research. *Eur. Respir. Rev.* 25, 124–9. <https://doi.org/10.1183/16000617.0034-2016>
- Squitti, R., Bressi, F., Pasqualetti, P., Bonomini, C., Ghidoni, R., Binetti, G., Cassetta, E., Moffa, F., Ventriglia, M., Vernieri, F., Rossini, P.M., 2009. Longitudinal prognostic value of serum "free" copper in patients with Alzheimer disease. *Neurology* 72, 50–55. <https://doi.org/10.1212/01.wnl.0000338568.28960.3f>
- Tamayo-Uria, I., Maitre, L., Thomsen, C., Nieuwenhuijsen, M.J., Chatzi, L., Siroux, V., Aasvang, G.M., Agier, L., Andrusaityte, S., Casas, M., de Castro, M., Dedele, A., Haug, L.S., Heude, B., Grazuleviciene, R., Gutzkow, K.B., Krog, N.H., Mason, D., McEachan, R.R.C., Meltzer, H.M., Petraviciene, I., Robinson, O., Roumeliotaki, T., Sakhi, A.K., Urquiza, J., Vafeiadi, M., Waiblinger, D., Warembourg, C., Wright, J., Slama, R., Vrijheid, M., Basagaña, X., 2019. The early-life exposome: Description and patterns in six European countries. *Environ. Int.* 123, 189–200. <https://doi.org/10.1016/j.envint.2018.11.067>
- Tamayo, I., Maitre, L., Thomsen, C., Nieuwenhuijsen, M.J., Chatzi, L., 2018. The early-life exposome: description and patterns in six European countries. *Rev.*
- Tanabe, M., Kanehisa, M., 2012. Using the KEGG database resource. *Curr. Protoc. Bioinforma.* 11, 1.12.1-1.12.54. <https://doi.org/10.1002/0471250953.bi0112s38>
- Thayer, K.A., Heindel, J.J., Bucher, J.R., Gallo, M.A., 2012. Role of Environmental Chemicals in Diabetes and Obesity: A National Toxicology Program Workshop Review. *Environ. Health Perspect.* 120, 779. <https://doi.org/10.1289/EHP.1104597>
- Tisato, F., Marzano, C., Porchia, M., Pellei, M., Santini, C., 2010. Copper in diseases and treatments, and copper-based anticancer strategies. *Med. Res. Rev.* 30, 708–749. <https://doi.org/10.1002/med.20174>
- Tobi, E.W., Lumey, L.H., Talens, R.P., Kremer, D., Putter, H., Stein, A.D., Slagboom, P.E., Heijmans, B.T., 2009. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum. Mol. Genet.* 18, 4046–4053. <https://doi.org/10.1093/hmg/ddp353>
- Tobi, E.W., Sliker, R.C., Luijk, R., Dekkers, K.F., Stein, A.D., Xu, K.M., Biobank-based Integrative Omics Studies Consortium, B.I.O.S., Slagboom, P.E., van Zwet, E.W., Lumey, L.H., Heijmans, B.T., 2018a. DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Sci. Adv.* 4, eaao4364. <https://doi.org/10.1126/sciadv.aao4364>
- Tobi, E.W., Zwet, E.W. van, Lumey, L., Heijmans, B.T., 2018b. Why mediation analysis trumps Mendelian randomization in population epigenomics studies of the Dutch Famine. *bioRxiv* 362392. <https://doi.org/10.1101/362392>
- Triche, T.J., Weisenberger, D.J., Van Den Berg, D., Laird, P.W., Siegmund, K.D., 2013. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res.* 41.

- Uauy, R., Olivares, M., Gonzalez, M., 1998. Essentiality of copper in humans, in: *American Journal of Clinical Nutrition*. <https://doi.org/10.1093/ajcn/67.5.952S>
- Uriu-Adams, J.Y., Keen, C.L., 2005. Copper, oxidative stress, and human health. *Mol. Aspects Med.* 26, 268–98. <https://doi.org/10.1016/j.mam.2005.07.015>
- van Iterson, M., Tobi, E.W., Sliker, R.C., den Hollander, W., Luijk, R., Slagboom, P.E., Heijmans, B.T., 2014. MethylAid: Visual and interactive quality control of large Illumina 450k data sets. *Bioinformatics* 30, 3435–3437. <https://doi.org/10.1093/bioinformatics/btu566>
- Vázquez Nava, F., Saldívar González, A.H., Perales, G.M., Ochoa, D.L., del Carmen Barrientos Gómez, M., Rodríguez, E.M.V., Rodríguez, C.F.V., Guzmán, F.J.B., 2006. Associations Between Family History of Allergy, Exposure to Tobacco Smoke, Active Smoking, Obesity, and Asthma in Adolescents. *Arch. Bronconeumol.* ((English Ed. 42, 621–626. [https://doi.org/10.1016/s1579-2129\(07\)60003-2](https://doi.org/10.1016/s1579-2129(07)60003-2)
- Vineis, P., Perera, F., 2007. Molecular epidemiology and biomarkers in etiologic cancer research: The new in light of the old. *Cancer Epidemiol. Biomarkers Prev.* <https://doi.org/10.1158/1055-9965.EPI-07-0457>
- Vineis, P., van Veldhoven, K., Chadeau-Hyam, M., Athersuch, T.J., 2013. Advancing the application of omics-based biomarkers in environmental epidemiology. *Environ. Mol. Mutagen.* 54, 461–467. <https://doi.org/10.1002/em.21764>
- Vinken, M., 2013. The adverse outcome pathway concept: A pragmatic tool in toxicology. *Toxicology*. <https://doi.org/10.1016/j.tox.2013.08.011>
- Vrijheid, M., Slama, R., Robinson, O., Chatzi, L., Coen, M., van den Hazel, P., Thomsen, C., Wright, J., Athersuch, T.J., Avellana, N., Basagaña, X., Brochet, C., Bucchini, L., Bustamante, M., Carracedo, A., Casas, M., Estivill, X., Fairley, L., van Gent, D., Gonzalez, J.R., Granum, B., Gražulevičienė, R., Gutzkow, K.B., Julvez, J., Keun, H.C., Kogevinas, M., McEachan, R.R.C., Meltzer, H.M., Sabidó, E., Schwarze, P.E., Siroux, V., Sunyer, J., Want, E.J., Zeman, F., Nieuwenhuijsen, M.J., 2014. The human early-life exposome (HELIX): project rationale and design. *Environ. Health Perspect.* 122, 535–44. <https://doi.org/10.1289/ehp.1307204>
- Wahl, A., Kasela, S., Carnero-Montoro, E., van Iterson, M., Štambuk, J., Sharma, S., van den Akker, E., Klaric, L., Benedetti, E., Razdorov, G., Trbojević-Akmačić, I., Vučković, F., Ugrina, I., Beekman, M., Deelen, J., van Heemst, D., Heijmans, B.T., B.I.O.S. Consortium, Wuhler, M., Plomp, R., Keser, T., Šimurina, M., Pavić, T., Gudelj, I., Krištić, J., Grallert, H., Kunze, S., Peters, A., Bell, J.T., Spector, T.D., Milani, L., Slagboom, P.E., Lauc, G., Gieger, C., 2018. IgG glycosylation and DNA methylation are interconnected with smoking. *Biochim. Biophys. Acta - Gen. Subj.* 1862, 637–648. <https://doi.org/10.1016/J.BBAGEN.2017.10.012>
- Watkins, D.J., Wellenius, G.A., Butler, R.A., Bartell, S.M., Fletcher, T., Kelsey, K.T., 2014. Associations between serum perfluoroalkyl acids and LINE-1 DNA methylation. *Environ. Int.* 63, 71–76. <https://doi.org/10.1016/j.envint.2013.10.018>
- Wild, C.P., 2005. Complementing the Genome with an “Exposome”: The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. *Cancer Epidemiol. Prev. Biomarkers* 14.
- Wright, J., Small, N., Raynor, P., Tuffnell, D., Bhopal, R., Cameron, N., Fairley, L., Lawlor, D.A., Parslow, R., Petherick, E.S., Pickett, K.E., Waiblinger, D., West, J., 2013. Cohort Profile:

- The Born in Bradford multi-ethnic family cohort study. *Int. J. Epidemiol.* 42, 978–991. <https://doi.org/10.1093/ije/dys112>
- Yakinci, G., Paç, A., Küçükbay, F.Z., Tayfun, M., Gül, A., 1997. Serum zinc, copper, and magnesium levels in obese children. *Pediatr. Int.* 39, 339–341. <https://doi.org/10.1111/j.1442-200X.1997.tb03748.x>
- Yu, Q., Li, B., 2017. mma: An R Package for Mediation Analysis with Multiple Mediators. *J. Open Res. Softw.* 5. <https://doi.org/10.5334/jors.160>
- Zeyda, M., Stulnig, T.M., 2009. Obesity, inflammation, and insulin resistance - A mini-review. *Gerontology* 55, 379–386. <https://doi.org/10.1159/000212758>
- Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67, 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

Supplementary materials

Supplementary Material 1: Exposure levels in 1,173 mother-child pairs from the HELIX cohort.

Exposure name	Exposure window	Group	Unit - transformation	Modality	Mean +- SE	n (%)	Assessment	Detailed exposure name
Arsenic	Postnatal	Metals	$\mu\text{g/L} - \text{Log2}$		0.6 ± 0		By biomarkers	
Arsenic	Pregnancy	Metals	$\mu\text{g/L} - \text{Log2}$		0.5 ± 0		By biomarkers	
Cadmium	Postnatal	Metals	$\mu\text{g/L} - \text{Log2}$		-3.4 ± 0		By biomarkers	
Cadmium	Pregnancy	Metals	$\mu\text{g/L} - \text{Log2}$		-1.3 ± 0		By biomarkers	
Cobalt	Postnatal	Metals	$\mu\text{g/L} - \text{Log2}$		-1.5 ± 0		By biomarkers	
Cobalt	Pregnancy	Metals	$\mu\text{g/L} - \text{Log2}$		-1.4 ± 0		By biomarkers	
Cesium	Postnatal	Metals	$\mu\text{g/L} - \text{Log2}$		1.3 ± 0		By biomarkers	
Cesium	Pregnancy	Metals	$\mu\text{g/L} - \text{Log2}$		1.5 ± 0		By biomarkers	
Copper	Postnatal	Metals	$\mu\text{g/L} - \text{Log2}$		10.8 ± 0		By biomarkers	
Copper	Pregnancy	Metals	$\mu\text{g/L} - \text{Log2}$		11.4 ± 0		By biomarkers	
Mercury	Postnatal	Metals	$\mu\text{g/L} - \text{Log2}$		0.4 ± 0		By biomarkers	
Mercury	Pregnancy	Metals	$\mu\text{g/L} - \text{Log2}$		1.6 ± 0		By biomarkers	
Manganese	Postnatal	Metals	$\mu\text{g/L} - \text{Log2}$		4 ± 0		By biomarkers	
Manganese	Pregnancy	Metals	$\mu\text{g/L} - \text{Log2}$		4.3 ± 0		By biomarkers	
Molybdenum	Postnatal	Metals	$\mu\text{g/L} - \text{Log2}$		0.5 ± 0		By biomarkers	
Molybdenum	Pregnancy	Metals	$\mu\text{g/L} - \text{Log2}$		0.2 ± 0		By biomarkers	
Lead	Postnatal	Metals	$\mu\text{g/L} - \text{Log2}$		3.9 ± 0		By biomarkers	
Lead	Pregnancy	Metals	$\mu\text{g/L} - \text{Log2}$		4.1 ± 0		By biomarkers	
Thallium	Postnatal	Metals		Undetected		1087 (0.93)	By biomarkers	
Thallium	Postnatal	Metals		Detected		86 (0.07)	By biomarkers	
Thallium	Pregnancy	Metals		Undetected		1134 (0.97)	By biomarkers	
Thallium	Pregnancy	Metals		Detected		39 (0.03)	By biomarkers	
DDE	Postnatal	OCs	$\text{ng/g lipids} - \text{Log2}$		5.4 ± 0		By biomarkers	Dichlorodiphenyldichloroethylene

DDE	Pregnancy	OCs	ng/g lipids – Log2	6.6 ± 0		By biomarkers	Dichlorodiphenyldichloroethylene
DDT	Postnatal	OCs	ng/g lipids – Log2	0.1 ± 0		By biomarkers	Dichlorodiphenyltrichloroethane
DDT	Pregnancy	OCs	ng/g lipids – Log2	1.5 ± 0		By biomarkers	Dichlorodiphenyltrichloroethane
HCB	Postnatal	OCs	ng/g lipids – Log2	4 ± 0		By biomarkers	Hexachlorobenzene
HCB	Pregnancy	OCs	ng/g lipids – Log2	3.9 ± 0		By biomarkers	Hexachlorobenzene
PCB 118	Postnatal	OCs	ng/g lipids – Log2	1.9 ± 0		By biomarkers	Polybrominated diphenyl ether-47,Dichlorodiphenyldic
PCB 118	Pregnancy	OCs	ng/g lipids – Log2	2.1 ± 0		By biomarkers	Polybrominated diphenyl ether-47,Dichlorodiphenyldic
PCB 138	Postnatal	OCs	ng/g lipids – Log2	3.2 ± 0		By biomarkers	Polybrominated diphenyl ether-153,Dichlorodiphenyld
PCB 138	Pregnancy	OCs	ng/g lipids – Log2	4 ± 0		By biomarkers	Polybrominated diphenyl ether-153,Dichlorodiphenyld
PCB 153	Postnatal	OCs	ng/g lipids – Log2	4.3 ± 0		By biomarkers	Polybrominated diphenyl ether-153,Dichlorodiphenyld
PCB 153	Pregnancy	OCs	ng/g lipids – Log2	4.8 ± 0		By biomarkers	Polybrominated diphenyl ether-153,Dichlorodiphenyld
PCB 170	Postnatal	OCs	ng/g lipids – Log2	0.9 ± 0		By biomarkers	Polybrominated diphenyl ether-153,Dichlorodiphenyld
PCB 170	Pregnancy	OCs	ng/g lipids – Log2	2.7 ± 0		By biomarkers	Polybrominated diphenyl ether-153,Dichlorodiphenyld
PCB 180	Postnatal	OCs	ng/g lipids – Log2	2.5 ± 0		By biomarkers	Polybrominated diphenyl ether-153,Dichlorodiphenyld
PCB 180	Pregnancy	OCs	ng/g lipids – Log2	4 ± 0		By biomarkers	Polybrominated diphenyl ether-153,Dichlorodiphenyld
PCBs (sum)	Postnatal	OCs	ng/g lipids – Log2	6.4 ± 0		By biomarkers	
PCBs (sum)	Pregnancy	OCs	ng/g lipids – Log2	7.3 ± 0		By biomarkers	
DEP	Postnatal	OP Pesticides	µg/L – Log2	1.2 ± 0.1		By biomarkers	Diethyl phosphate (DEP)
DEP	Pregnancy	OP Pesticides	µg/L – Log2	2.5 ± 0		By biomarkers	Diethyl phosphate (DEP)
DETP	Postnatal	OP Pesticides	µg/L – Log2	-0.7 ± 0.1		By biomarkers	Diethyl thiophosphate (DETP)
DETP	Pregnancy	OP Pesticides	µg/L – Log2	-0.1 ± 0.1		By biomarkers	Diethyl thiophosphate (DETP)
DMDTP	Postnatal	OP Pesticides		Undetected	960 (0.82)	By biomarkers	Dimethyldithiophosphate
DMDTP	Postnatal	OP Pesticides		Detected	213 (0.18)	By biomarkers	Dimethyldithiophosphate
DMP	Postnatal	OP Pesticides	µg/L – Log2	0.7 ± 0.1		By biomarkers	Dimethyl phosphate
DMP	Pregnancy	OP Pesticides	µg/L – Log2	3.7 ± 0		By biomarkers	Dimethyl phosphate

DMTP	Postnatal	OP Pesticides	$\mu\text{g/L} - \text{Log2}$	2.2 ± 0	By biomarkers	Dimethyl thiophosphate
DMTP	Pregnancy	OP Pesticides	$\mu\text{g/L} - \text{Log2}$	3 ± 0	By biomarkers	Dimethyl thiophosphate
PBDE 153	Postnatal	PBDEs	$\mu\text{g/L} - \text{Log2}$	-2.6 ± 0	By biomarkers	Polybrominated diphenyl ether 153
PBDE 153	Pregnancy	PBDEs	$\mu\text{g/L} - \text{Log2}$	-0.8 ± 0.1	By biomarkers	Polybrominated diphenyl ether 153
PBDE 47	Postnatal	PBDEs	$\mu\text{g/L} - \text{Log2}$	-1.3 ± 0	By biomarkers	Polybrominated diphenyl ether 47
PBDE 47	Pregnancy	PBDEs	$\mu\text{g/L} - \text{Log2}$	0.1 ± 0	By biomarkers	Polybrominated diphenyl ether 47
PFHXS	Postnatal	PFASs	$\mu\text{g/L} - \text{Log2}$	-1.6 ± 0	By biomarkers	Perfluorohexane sulfonate
PFHXS	Pregnancy	PFASs	$\mu\text{g/L} - \text{Log2}$	-0.2 ± 0	By biomarkers	Perfluorohexane sulfonate
PFNA	Postnatal	PFASs	$\mu\text{g/L} - \text{Log2}$	-1.1 ± 0	By biomarkers	Perfluorononanoate
PFNA	Pregnancy	PFASs	$\mu\text{g/L} - \text{Log2}$	0.2 ± 0	By biomarkers	Perfluorononanoate
PFOA	Postnatal	PFASs	$\mu\text{g/L} - \text{Log2}$	0.6 ± 0	By biomarkers	Perfluorooctanoate
PFOA	Pregnancy	PFASs	$\mu\text{g/L} - \text{Log2}$	1.9 ± 0	By biomarkers	Perfluorooctanoate
PFOS	Postnatal	PFASs	$\mu\text{g/L} - \text{Log2}$	1.1 ± 0	By biomarkers	Perfluorooctanoate
PFOS	Pregnancy	PFASs	$\mu\text{g/L} - \text{Log2}$	3.4 ± 0	By biomarkers	Perfluorooctanoate
PFUNDA	Postnatal	PFASs	$\mu\text{g/L} - \text{Log2}$	-4.6 ± 0	By biomarkers	Perfluoroundecanoate
PFUNDA	Pregnancy	PFASs	$\mu\text{g/L} - \text{Log2}$	-1.8 ± 0	By biomarkers	Perfluoroundecanoate
BPA	Postnatal	Phenols	$\mu\text{g/g} - \text{Log2}$	2.9 ± 0	By biomarkers	Bisphenol A
BPA	Pregnancy	Phenols	$\mu\text{g/g} - \text{Log2}$	2.4 ± 0	By biomarkers	Bisphenol A
BUPA	Postnatal	Phenols	$\mu\text{g/g} - \text{Log2}$	-2.7 ± 0	By biomarkers	N-Butyl paraben
BUPA	Pregnancy	Phenols	$\mu\text{g/g} - \text{Log2}$	1.7 ± 0.1	By biomarkers	N-Butyl paraben
ETPA	Postnatal	Phenols	$\mu\text{g/g} - \text{Log2}$	0.5 ± 0	By biomarkers	Ethyl paraben
ETPA	Pregnancy	Phenols	$\mu\text{g/g} - \text{Log2}$	3.3 ± 0.1	By biomarkers	Ethyl paraben
MEPA	Postnatal	Phenols	$\mu\text{g/g} - \text{Log2}$	4.1 ± 0.1	By biomarkers	Methyl paraben
MEPA	Pregnancy	Phenols	$\mu\text{g/g} - \text{Log2}$	7.8 ± 0.1	By biomarkers	Methyl paraben
OXBE	Postnatal	Phenols	$\mu\text{g/g} - \text{Log2}$	1.5 ± 0.1	By biomarkers	Oxybenzone
OXBE	Pregnancy	Phenols	$\mu\text{g/g} - \text{Log2}$	3.4 ± 0.1	By biomarkers	Oxybenzone
PRPA	Postnatal	Phenols	$\mu\text{g/g} - \text{Log2}$	-1.3 ± 0.1	By biomarkers	Propyl paraben

PRPA	Pregnancy	Phenols	$\mu\text{g/g} - \text{Log2}$	5.9 ± 0.1	By biomarkers	Propyl paraben
TRCS	Postnatal	Phenols	$\mu\text{g/g} - \text{Log2}$	-0.3 ± 0.1	By biomarkers	Triclosan
TRCS	Pregnancy	Phenols	$\mu\text{g/g} - \text{Log2}$	3.8 ± 0.1	By biomarkers	Triclosan
MBzP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	3.2 ± 0	By biomarkers	Mono benzyl phthalate
MBzP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	3.8 ± 0	By biomarkers	Mono benzyl phthalate
MECPP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	6 ± 0	By biomarkers	Mono-2-ethyl 5-carboxypentyl phthalate
MECPP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	6 ± 0	By biomarkers	Mono-2-ethyl 5-carboxypentyl phthalate
MEHHP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	5.2 ± 0	By biomarkers	Mono-2-ethyl-5-hydroxyhexyl phthalate
MEHHP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	5 ± 0	By biomarkers	Mono-2-ethyl-5-hydroxyhexyl phthalate
MEHP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	2.4 ± 0	By biomarkers	Mono-2-ethylhexyl phthalate
MEHP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	3.7 ± 0	By biomarkers	Mono-2-ethylhexyl phthalate
MEOHP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	4.5 ± 0	By biomarkers	Mono-2-ethyl-5-oxohexyl phthalate
MEOHP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	4.6 ± 0	By biomarkers	Mono-2-ethyl-5-oxohexyl phthalate
MEP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	5.3 ± 0	By biomarkers	Mono-2-ethyl-5-oxohexyl phthalate
MEP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	8.2 ± 0	By biomarkers	Mono-2-ethyl-5-oxohexyl phthalate
MiBP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	5.4 ± 0	By biomarkers	Mono-iso-butyl phthalate
MiBP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	6.1 ± 0	By biomarkers	Mono-iso-butyl phthalate
MnBP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	4.6 ± 0	By biomarkers	Mono-n-butyl phthalate
MnBP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	5.7 ± 0	By biomarkers	Mono-n-butyl phthalate
OH-MiNP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	2.6 ± 0	By biomarkers	Mono- ⁴ -methyl-7-hydroxyoctyl phthalate
OH-MiNP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	0.7 ± 0	By biomarkers	Mono- ⁴ -methyl-7-hydroxyoctyl phthalate
oxo-MiNP	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	1.7 ± 0	By biomarkers	Mono- ⁴ -methyl-7-oxooctyl phthalate
oxo-MiNP	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	0.9 ± 0	By biomarkers	Mono- ⁴ -methyl-7-oxooctyl phthalate
DEHP (sum of metabolites)	Postnatal	Phthalates	$\mu\text{g/g} - \text{Log2}$	7 ± 0	By biomarkers	Sum of DEHP metabolites
DEHP (sum of metabolites)	Pregnancy	Phthalates	$\mu\text{g/g} - \text{Log2}$	7.8 ± 0	By biomarkers	Sum of DEHP metabolites
Cotinine	Postnatal	Tobacco Smoke		Undetected	957 (0.82)	By biomarkers

Cotinine	Postnatal	Tobacco Smoke	Detected	216 (0.18)	By biomarkers
Cotinine	Pregnancy	Tobacco Smoke	$\mu\text{g/L} - \text{None}$	1.5 ± 0	By biomarkers
PMabsorbance (preg)	Pregnancy	Air Pollution	$10^{-5} \text{ m}^{-1} - \text{Log}$	1 ± 0	By environmental models or questionnaires
NO2 (preg)	Pregnancy	Air Pollution	$\mu\text{g}/\text{m}^3 - \text{Log}$	3.5 ± 0	By environmental models or questionnaires
PM10 (preg)	Pregnancy	Air Pollution	$\mu\text{g}/\text{m}^3 - \text{None}$	23.9 ± 0.2	By environmental models or questionnaires
PM2.5 (preg)	Pregnancy	Air Pollution	$\mu\text{g}/\text{m}^3 - \text{None}$	15.1 ± 0.1	By environmental models or questionnaires
NO2 (year)	Postnatal	Air Pollution	$\mu\text{g}/\text{m}^3 - \text{Log}$	3.5 ± 0	By environmental models or questionnaires
PM10 (year)	Postnatal	Air Pollution	$\mu\text{g}/\text{m}^3 - \text{None}$	25.9 ± 0.2	By environmental models or questionnaires
PM2.5 (year)	Postnatal	Air Pollution	$\mu\text{g}/\text{m}^3 - \text{None}$	13.7 ± 0.1	By environmental models or questionnaires
PMabsorbance (year)	Postnatal	Air Pollution	$10^{-5} \text{ m}^{-1} - \text{Log}$	0.8 ± 0	By environmental models or questionnaires
Accessibility (bus lines 300m)	Pregnancy	Built Environmen t	$\text{m} / \text{km}^2 - \text{Dic}$	0.7 ± 0	By environmental models or questionnaires
Accessibility (bus stops 300m)	Pregnancy	Built Environmen t	$\text{m} / \text{km}^2 - \text{Dic}$	2.5 ± 0	By environmental models or questionnaires
Built density (300m)	Pregnancy	Built Environmen t	$\text{m}^2 \text{ built} / \text{km}^2$	418.9 ± 4.3	By environmental models or questionnaires
Connectivity (300m)	Pregnancy	Built Environmen t	number of intersections / km^2	12.2 ± 0.1	By environmental models or questionnaires
Facility richness (300m)	Pregnancy	Built Environmen t	-	0.1 ± 0	By environmental models or questionnaires
Land use (300m)	Pregnancy	Built Environmen t	-	0.4 ± 0	By environmental models or questionnaires
Population density	Pregnancy	Built Environmen t	people / km^2	77.2 ± 1.3	By environmental models or questionnaires
Walkability	Pregnancy	Built Environmen t	-	0.3 ± 0	By environmental models or questionnaires
Accessibility (bus lines 300m)	Postnatal	Built Environmen t	$\text{m} / \text{km}^2 - \text{Dic}$	0.7 ± 0	By environmental models or questionnaires

Accessibility (bus lines 300m - school)	Postnatal	Built Environment	m / km ² - Dic	0.7 ± 0	By environmental models or questionnaires
Accessibility (bus stops 300m - school)	Postnatal	Built Environment	m / km ² - Log	2.5 ± 0	By environmental models or questionnaires
Built density (300m)	Postnatal	Built Environment	m ² built / km ² – Square root	390.3 ± 4.6	By environmental models or questionnaires
Built density (300m - school)	Postnatal	Built Environment	m ² built / km ² – Square root	409.8 ± 4.3	By environmental models or questionnaires
Connectivity density (300 m)	Postnatal	Built Environment	number of intersections / km ² - Log	5.3 ± 0	By environmental models or questionnaires
Connectivity density (300m - school)	Postnatal	Built Environment	number of intersections / km ² - Log	5.4 ± 0	By environmental models or questionnaires
Facility density (300m - school)	Postnatal	Built Environment	n / km ² - Log	3.4 ± 0	By environmental models or questionnaires
Facility richness	Postnatal	Built Environment	-- None	0.1 ± 0	By environmental models or questionnaires
Facility richness (300m - school)	Postnatal	Built Environment	-- None	0.1 ± 0	By environmental models or questionnaires
Land use (300m)	Postnatal	Built Environment	-- None	0.4 ± 0	By environmental models or questionnaires
Land use (300m - school)	Postnatal	Built Environment	-- None	0.4 ± 0	By environmental models or questionnaires
Population density	Postnatal	Built Environment	people / km ² – Square root	69.4 ± 1.3	By environmental models or questionnaires
Population density (school)	Postnatal	Built Environment	people / km ² – Square root	69.8 ± 1.1	By environmental models or questionnaires
Walkability index	Postnatal	Built Environment	-- None	0.3 ± 0	By environmental models or questionnaires

Facility density (300m)	Pregnancy	Built Environment	n / km ² - Log	3.4 ± 0		By environmental models or questionnaires
Facility density (300m)	Postnatal	Built Environment	n / km ² - Log	3.3 ± 0		By environmental models or questionnaires
Accessibility (bus stops 300m)	Postnatal	Built Environment	m / km ² - Log	2.7 ± 0		By environmental models or questionnaires
Indoor PMabsorbance	Postnatal	Indoor air	10 ⁻⁵ m ⁻¹ – Log	0.4 ± 0		By environmental models or questionnaires
Indoor benzene	Postnatal	Indoor air	µg/m ³ – Log	1.2 ± 0		By environmental models or questionnaires
Indoor NO2	Postnatal	Indoor air	µg/m ³ – Log	4.3 ± 0		By environmental models or questionnaires
Indoor PM2.5	Postnatal	Indoor air	µg/m ³ – Log	3 ± 0		By environmental models or questionnaires
Indoor BTEX	Postnatal	Indoor air	µg/m ³ – Log	3.6 ± 0		By environmental models or questionnaires
Alcohol intake	Pregnancy	Lifestyle		0	811 (0.69)	By environmental models or questionnaires
Alcohol intake	Pregnancy	Lifestyle		1	362 (0.31)	By environmental models or questionnaires
Cereals intake	Pregnancy	Lifestyle		1	459 (0.39)	By environmental models or questionnaires
Cereals intake	Pregnancy	Lifestyle		2	393 (0.34)	By environmental models or questionnaires
Cereals intake	Pregnancy	Lifestyle		3	321 (0.27)	By environmental models or questionnaires
Dairy intake	Pregnancy	Lifestyle		1	399 (0.34)	By environmental models or questionnaires
Dairy intake	Pregnancy	Lifestyle		2	404 (0.34)	By environmental models or questionnaires
Dairy intake	Pregnancy	Lifestyle		3	370 (0.32)	By environmental models or questionnaires
Fastfood intake	Pregnancy	Lifestyle		1	257 (0.22)	By environmental models or questionnaires
Fastfood intake	Pregnancy	Lifestyle		2	127 (0.11)	By environmental models or questionnaires
Fastfood intake	Pregnancy	Lifestyle		3	789 (0.67)	By environmental models or questionnaires
Fish and seafood intake	Pregnancy	Lifestyle		1	482 (0.41)	By environmental models or questionnaires
Fish and seafood intake	Pregnancy	Lifestyle		2	358 (0.31)	By environmental models or questionnaires
Fish and seafood intake	Pregnancy	Lifestyle		3	333 (0.28)	By environmental models or questionnaires
Folic acid supplementation	Pregnancy	Lifestyle		0	524 (0.45)	By environmental models or questionnaires
Folic acid supplementation	Pregnancy	Lifestyle		1	649 (0.55)	By environmental models or questionnaires

Fruits intake	Pregnancy	Lifestyle	1	447 (0.38)	By environmental models or questionnaires
Fruits intake	Pregnancy	Lifestyle	2	357 (0.3)	By environmental models or questionnaires
Fruits intake	Pregnancy	Lifestyle	3	369 (0.31)	By environmental models or questionnaires
Legumes intake	Pregnancy	Lifestyle	1	277 (0.24)	By environmental models or questionnaires
Legumes intake	Pregnancy	Lifestyle	2	509 (0.43)	By environmental models or questionnaires
Legumes intake	Pregnancy	Lifestyle	3	387 (0.33)	By environmental models or questionnaires
Meat intake	Pregnancy	Lifestyle	1	381 (0.32)	By environmental models or questionnaires
Meat intake	Pregnancy	Lifestyle	2	384 (0.33)	By environmental models or questionnaires
Meat intake	Pregnancy	Lifestyle	3	408 (0.35)	By environmental models or questionnaires
Moderate physical activity (t3)	Pregnancy	Lifestyle	None or sometimes	565 (0.48)	By environmental models or questionnaires
Moderate physical activity (t3)	Pregnancy	Lifestyle	Often	283 (0.24)	By environmental models or questionnaires
Moderate physical activity (t3)	Pregnancy	Lifestyle	Very Often	325 (0.28)	By environmental models or questionnaires
Vigorous physical activity (t3)	Pregnancy	Lifestyle	Low	514 (0.44)	By environmental models or questionnaires
Vigorous physical activity (t3)	Pregnancy	Lifestyle	Medium-High	659 (0.56)	By environmental models or questionnaires
Vegetables intake	Pregnancy	Lifestyle	1	375 (0.32)	By environmental models or questionnaires
Vegetables intake	Pregnancy	Lifestyle	2	411 (0.35)	By environmental models or questionnaires
Vegetables intake	Pregnancy	Lifestyle	3	387 (0.33)	By environmental models or questionnaires
Bakery products intake	Postnatal	Lifestyle	1	409 (0.35)	By environmental models or questionnaires
Bakery products intake	Postnatal	Lifestyle	2	451 (0.38)	By environmental models or questionnaires
Bakery products intake	Postnatal	Lifestyle	3	313 (0.27)	By environmental models or questionnaires
Soda intake	Postnatal	Lifestyle	1	460 (0.39)	By environmental models or questionnaires
Soda intake	Postnatal	Lifestyle	2	418 (0.36)	By environmental models or questionnaires
Soda intake	Postnatal	Lifestyle	3	295 (0.25)	By environmental models or questionnaires
Breakfast cereals intake	Postnatal	Lifestyle	1	408 (0.35)	By environmental models or questionnaires
Breakfast cereals intake	Postnatal	Lifestyle	2	384 (0.33)	By environmental models or questionnaires

Breakfast cereals intake	Postnatal	Lifestyle		3	381 (0.32)	By environmental models or questionnaires
Caffeinated drinks	Postnatal	Lifestyle		1	746 (0.64)	By environmental models or questionnaires
Caffeinated drinks	Postnatal	Lifestyle		2	163 (0.14)	By environmental models or questionnaires
Caffeinated drinks	Postnatal	Lifestyle		3	264 (0.23)	By environmental models or questionnaires
Dairy products intake	Postnatal	Lifestyle		1	396 (0.34)	By environmental models or questionnaires
Dairy products intake	Postnatal	Lifestyle		2	402 (0.34)	By environmental models or questionnaires
Dairy products intake	Postnatal	Lifestyle		3	375 (0.32)	By environmental models or questionnaires
Sleep duration	Postnatal	Lifestyle		10.3 ± 0		By environmental models or questionnaires
Fastfood intake	Postnatal	Lifestyle		1	651 (0.55)	By environmental models or questionnaires
Fastfood intake	Postnatal	Lifestyle		2	382 (0.33)	By environmental models or questionnaires
Fastfood intake	Postnatal	Lifestyle		3	140 (0.12)	By environmental models or questionnaires
KIDMED score	Postnatal	Lifestyle		2.8 ± 0.1		By environmental models or questionnaires
Moderate and vigorous PA	Postnatal	Lifestyle		40.3 ± 0.7		By environmental models or questionnaires
Organic food intake	Postnatal	Lifestyle		1	409 (0.35)	By environmental models or questionnaires
Organic food intake	Postnatal	Lifestyle		2	468 (0.4)	By environmental models or questionnaires
Organic food intake	Postnatal	Lifestyle		3	296 (0.25)	By environmental models or questionnaires
Cat at home	Postnatal	Lifestyle		0	967 (0.82)	By environmental models or questionnaires
Cat at home	Postnatal	Lifestyle		1	206 (0.18)	By environmental models or questionnaires
Dog at home	Postnatal	Lifestyle		0	1006 (0.86)	By environmental models or questionnaires
Dog at home	Postnatal	Lifestyle		1	167 (0.14)	By environmental models or questionnaires
Other pets at home	Postnatal	Lifestyle	No		728 (0.62)	By environmental models or questionnaires
Other pets at home	Postnatal	Lifestyle	Yes		445 (0.38)	By environmental models or questionnaires
Processed meat intake	Postnatal	Lifestyle		1	427 (0.36)	By environmental models or questionnaires
Processed meat intake	Postnatal	Lifestyle		2	484 (0.41)	By environmental models or questionnaires
Processed meat intake	Postnatal	Lifestyle		3	262 (0.22)	By environmental models or questionnaires
Ready made food intake	Postnatal	Lifestyle		1	581 (0.5)	By environmental models or questionnaires

Ready made food intake	Postnatal	Lifestyle	2	293 (0.25)	By environmental models or questionnaires
Ready made food intake	Postnatal	Lifestyle	3	299 (0.25)	By environmental models or questionnaires
Sedentary behaviour	Postnatal	Lifestyle	239.6 ± 3.8		By environmental models or questionnaires
Bread intake	Postnatal	Lifestyle	1	531 (0.45)	By environmental models or questionnaires
Bread intake	Postnatal	Lifestyle	2	360 (0.31)	By environmental models or questionnaires
Bread intake	Postnatal	Lifestyle	3	282 (0.24)	By environmental models or questionnaires
Cereals intake	Postnatal	Lifestyle	1	395 (0.34)	By environmental models or questionnaires
Cereals intake	Postnatal	Lifestyle	2	394 (0.34)	By environmental models or questionnaires
Cereals intake	Postnatal	Lifestyle	3	384 (0.33)	By environmental models or questionnaires
Fish and seafood intake	Postnatal	Lifestyle	1	468 (0.4)	By environmental models or questionnaires
Fish and seafood intake	Postnatal	Lifestyle	2	356 (0.3)	By environmental models or questionnaires
Fish and seafood intake	Postnatal	Lifestyle	3	349 (0.3)	By environmental models or questionnaires
Fruits intake	Postnatal	Lifestyle	1	386 (0.33)	By environmental models or questionnaires
Fruits intake	Postnatal	Lifestyle	2	396 (0.34)	By environmental models or questionnaires
Fruits intake	Postnatal	Lifestyle	3	391 (0.33)	By environmental models or questionnaires
Total fat intake	Postnatal	Lifestyle	1	467 (0.4)	By environmental models or questionnaires
Total fat intake	Postnatal	Lifestyle	2	323 (0.28)	By environmental models or questionnaires
Total fat intake	Postnatal	Lifestyle	3	383 (0.33)	By environmental models or questionnaires
Meat intake	Postnatal	Lifestyle	1	444 (0.38)	By environmental models or questionnaires
Meat intake	Postnatal	Lifestyle	2	329 (0.28)	By environmental models or questionnaires
Meat intake	Postnatal	Lifestyle	3	400 (0.34)	By environmental models or questionnaires
Potatoes intake	Postnatal	Lifestyle	1	417 (0.36)	By environmental models or questionnaires
Potatoes intake	Postnatal	Lifestyle	2	506 (0.43)	By environmental models or questionnaires
Potatoes intake	Postnatal	Lifestyle	3	250 (0.21)	By environmental models or questionnaires
Sweets intake	Postnatal	Lifestyle	1	382 (0.33)	By environmental models or questionnaires
Sweets intake	Postnatal	Lifestyle	2	394 (0.34)	By environmental models or questionnaires

Sweets intake	Postnatal	Lifestyle		3	397 (0.34)	By environmental models or questionnaires
Vegetables intake	Postnatal	Lifestyle		1	594 (0.51)	By environmental models or questionnaires
Vegetables intake	Postnatal	Lifestyle		2	224 (0.19)	By environmental models or questionnaires
Vegetables intake	Postnatal	Lifestyle		3	355 (0.3)	By environmental models or questionnaires
Yogurt intake	Postnatal	Lifestyle		1	525 (0.45)	By environmental models or questionnaires
Yogurt intake	Postnatal	Lifestyle		2	264 (0.23)	By environmental models or questionnaires
Yogurt intake	Postnatal	Lifestyle		3	384 (0.33)	By environmental models or questionnaires
Humidity (preg)	Pregnancy	Meteorological	% - None	76 ± 0.3		By environmental models or questionnaires
Humidity (t1)	Pregnancy	Meteorological	% - None	76 ± 0.3		By environmental models or questionnaires
Pressure (preg)	Pregnancy	Meteorological	Bar - None	995.1 ± 0.4		By environmental models or questionnaires
Pressure (t1)	Pregnancy	Meteorological	Bar - None	994.7 ± 0.4		By environmental models or questionnaires
Temperature (preg)	Pregnancy	Meteorological	°C - None	11.4 ± 0.1		By environmental models or questionnaires
Temperature (t1)	Pregnancy	Meteorological	°C - None	10.8 ± 0.2		By environmental models or questionnaires
Blue spaces (300 m)	Pregnancy	Natural Spaces	m - Log	0	1085 (0.92)	By environmental models or questionnaires
Blue spaces (300 m)	Pregnancy	Natural Spaces	m - Log	1	88 (0.08)	By environmental models or questionnaires
Green spaces (300 m)	Pregnancy	Natural Spaces	m - Log	0	292 (0.25)	By environmental models or questionnaires
Green spaces (300 m)	Pregnancy	Natural Spaces	m - Log	1	881 (0.75)	By environmental models or questionnaires
NDVI (100 m)	Pregnancy	Natural Spaces	NDVI - None	0.4 ± 0		By environmental models or questionnaires
Blue spaces (300 m)	Postnatal	Natural Spaces	m - Log	0	1077 (0.92)	By environmental models or questionnaires
Blue spaces (300 m)	Postnatal	Natural Spaces	m - Log	1	96 (0.08)	By environmental models or questionnaires
Blue spaces (300m - school)	Postnatal	Natural Spaces	m - Log	0	1092 (0.93)	By environmental models or questionnaires
Blue spaces (300m - school)	Postnatal	Natural Spaces	m - Log	1	81 (0.07)	By environmental models or questionnaires
Green spaces (300 m)	Postnatal	Natural Spaces	m - Log	0	245 (0.21)	By environmental models or questionnaires

Green spaces (300 m)	Postnatal	Natural Spaces	m - Log	1	928 (0.79)	By environmental models or questionnaires
Green spaces (300m - school)	Postnatal	Natural Spaces	m - Log	0	263 (0.22)	By environmental models or questionnaires
Green spaces (300m - school)	Postnatal	Natural Spaces	m - Log	1	910 (0.78)	By environmental models or questionnaires
NDVI (100 m)	Postnatal	Natural Spaces	NDVI - None	0.4 ± 0		By environmental models or questionnaires
NDVI (100 m - school)	Postnatal	Natural Spaces	NDVI - None	0.4 ± 0		By environmental models or questionnaires
Traffic noise (24h)	Pregnancy	Noise		1	551 (0.47)	By environmental models or questionnaires
Traffic noise (24h)	Pregnancy	Noise		2	207 (0.18)	By environmental models or questionnaires
Traffic noise (24h)	Pregnancy	Noise		3	259 (0.22)	By environmental models or questionnaires
Traffic noise (24h)	Pregnancy	Noise		4	156 (0.13)	By environmental models or questionnaires
Traffic noise (night)	Pregnancy	Noise		1	911 (0.78)	By environmental models or questionnaires
Traffic noise (night)	Pregnancy	Noise		2	152 (0.13)	By environmental models or questionnaires
Traffic noise (night)	Pregnancy	Noise		3	72 (0.06)	By environmental models or questionnaires
Traffic noise (night)	Pregnancy	Noise		4	24 (0.02)	By environmental models or questionnaires
Traffic noise (night)	Pregnancy	Noise		5	14 (0.01)	By environmental models or questionnaires
Traffic noise (24h)	Postnatal	Noise		1	616 (0.53)	By environmental models or questionnaires
Traffic noise (24h)	Postnatal	Noise		2	190 (0.16)	By environmental models or questionnaires
Traffic noise (24h)	Postnatal	Noise		3	250 (0.21)	By environmental models or questionnaires
Traffic noise (24h)	Postnatal	Noise		4	117 (0.1)	By environmental models or questionnaires
Traffic noise (24h - school)	Postnatal	Noise	dB - None	54.1 ± 0.2		By environmental models or questionnaires
Traffic noise (night)	Postnatal	Noise		1	934 (0.8)	By environmental models or questionnaires
Traffic noise (night)	Postnatal	Noise		2	94 (0.08)	By environmental models or questionnaires
Traffic noise (night)	Postnatal	Noise		3	86 (0.07)	By environmental models or questionnaires
Traffic noise (night)	Postnatal	Noise		4	59 (0.05)	By environmental models or questionnaires
Family affluence score	Postnatal	Socio-eco capital		Low	134 (0.11)	By environmental models or questionnaires
Family affluence score	Postnatal	Socio-eco capital		Middle	469 (0.4)	By environmental models or questionnaires

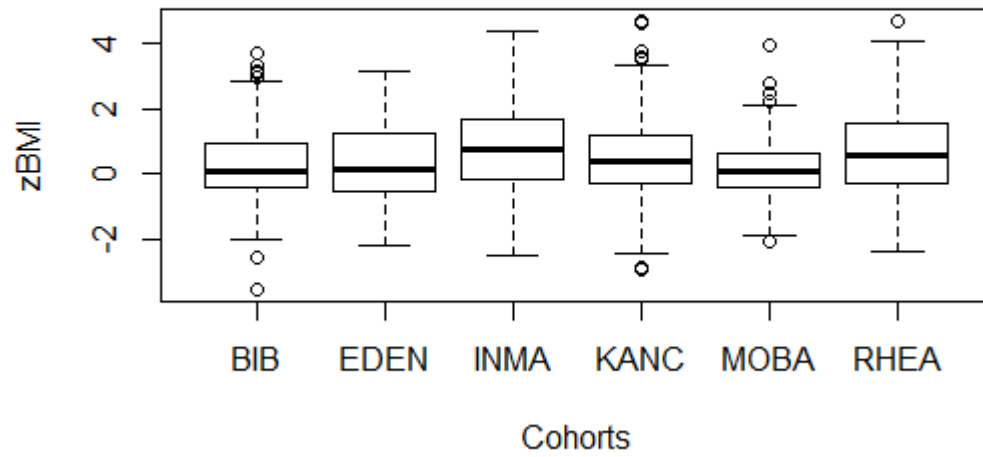
Family affluence score	Postnatal	Socio-eco capital		High	570 (0.49)	By environmental models or questionnaires
Contact with family and friends	Postnatal	Socio-eco capital		Less than once a week	50 (0.04)	By environmental models or questionnaires
Contact with family and friends	Postnatal	Socio-eco capital		Once a week	338 (0.29)	By environmental models or questionnaires
Contact with family and friends	Postnatal	Socio-eco capital		(almost) Daily	785 (0.67)	By environmental models or questionnaires
House crowding	Postnatal	Socio-eco capital	- None		4.3 ± 0	By environmental models or questionnaires
Social participation	Postnatal	Socio-eco capital		None	667 (0.57)	By environmental models or questionnaires
Social participation	Postnatal	Socio-eco capital		1 organisation	332 (0.28)	By environmental models or questionnaires
Social participation	Postnatal	Socio-eco capital		2 or more organisations	174 (0.15)	By environmental models or questionnaires
Cigarette	Pregnancy	Tobacco Smoke	- None		0.6 ± 0.1	By environmental models or questionnaires
Active smoking (preg)	Pregnancy	Tobacco Smoke		no	998 (0.85)	By environmental models or questionnaires
Active smoking (preg)	Pregnancy	Tobacco Smoke		yes	175 (0.15)	By environmental models or questionnaires
ETS	Postnatal	Tobacco Smoke		no exposure	745 (0.64)	By environmental models or questionnaires
ETS	Postnatal	Tobacco Smoke		exposure	428 (0.36)	By environmental models or questionnaires
Parental smoking	Postnatal	Tobacco Smoke		neither	722 (0.62)	By environmental models or questionnaires
Parental smoking	Postnatal	Tobacco Smoke		one	325 (0.28)	By environmental models or questionnaires
Parental smoking	Postnatal	Tobacco Smoke		both	126 (0.11)	By environmental models or questionnaires
Maternal smoking (active and ETS)	Pregnancy	Tobacco Smoke		no exposure	624 (0.53)	By environmental models or questionnaires
Maternal smoking (active and ETS)	Pregnancy	Tobacco Smoke		only passive exposure	374 (0.32)	By environmental models or questionnaires
Maternal smoking (active and ETS)	Pregnancy	Tobacco Smoke		smoker	175 (0.15)	By environmental models or questionnaires
Inverse distance to nearest road	Pregnancy	Traffic	m ⁻¹ - None		-2.5 ± 0	By environmental models or questionnaires
Road traffic load (100 m)	Pregnancy	Traffic	Veh./day.m -cube root		71.9 ± 1.9	By environmental models or questionnaires

Traffic density on nearest road	Pregnancy	Traffic	veh/day.m -cube root	12.5 ± 0.2		By environmental models or questionnaires
Inverse distance to nearest road	Postnatal	Traffic	m ⁻¹ - None	-3.5 ± 0		By environmental models or questionnaires
Road traffic load (100 m)	Postnatal	Traffic	veh/day.m -cube root	71.7 ± 2		By environmental models or questionnaires
Inverse distance to nearest road (school)	Pregnancy	Traffic	m ⁻¹ - None	16.2 ± 0.3		By environmental models or questionnaires
Traffic load of major roads (100 m)	Pregnancy	Traffic		0	915 (0.78)	By environmental models or questionnaires
Traffic load of major roads (100 m)	Pregnancy	Traffic		1	258 (0.22)	By environmental models or questionnaires
Traffic load of major roads (100 m - school)	Pregnancy	Traffic		0	884 (0.75)	By environmental models or questionnaires
Traffic load of major roads (100 m - school)	Pregnancy	Traffic		1	289 (0.25)	By environmental models or questionnaires
Water Brominated THMs	Pregnancy	Water disinfection by-products	ng/L - None	1.6 ± 0.1		By environmental models or questionnaires
Water Chloroform	Pregnancy	Water disinfection by-products	ng/L - None	1.5 ± 0		By environmental models or questionnaires
Water THMs	Pregnancy	Water disinfection by-products	ng/L - None	3 ± 0		By environmental models or questionnaires

OCs : Organochlorine compounds Brominated ; PFAS: Perfluorinated alkylated substances; PBDEs: Brominated compounds Ops: Organophosphate pesticide metabolites

Pathway	Number of genes	Number of enhancer CpGs
Fat digestion and absorption	41	84
Fatty acid elongation	30	93
Fatty acid degradation	44	79
Biosynthesis of unsaturated fatty acids	23	66
Vitamin digestion and absorption	24	80
Bile secretion	71	304
PPAR signalling pathway	74	202
Insulin resistance	107	478
Regulation of lipolysis in adipocytes	54	299
Adipocytokine signalling pathway	69	194
Type II diabetes mellitus	46	269
Ribosome biogenesis in eukaryotes	105	52
Thyroid hormone signalling pathway	116	524
Fanconi anemia pathway	54	51
GnRH signalling pathway	93	518
Prolactin signalling pathway	70	271

Supplementary Material 2: Pathways identified as relevant for zBMI relying on KEGG database, and corresponding numbers of genes and enhancer CpGs



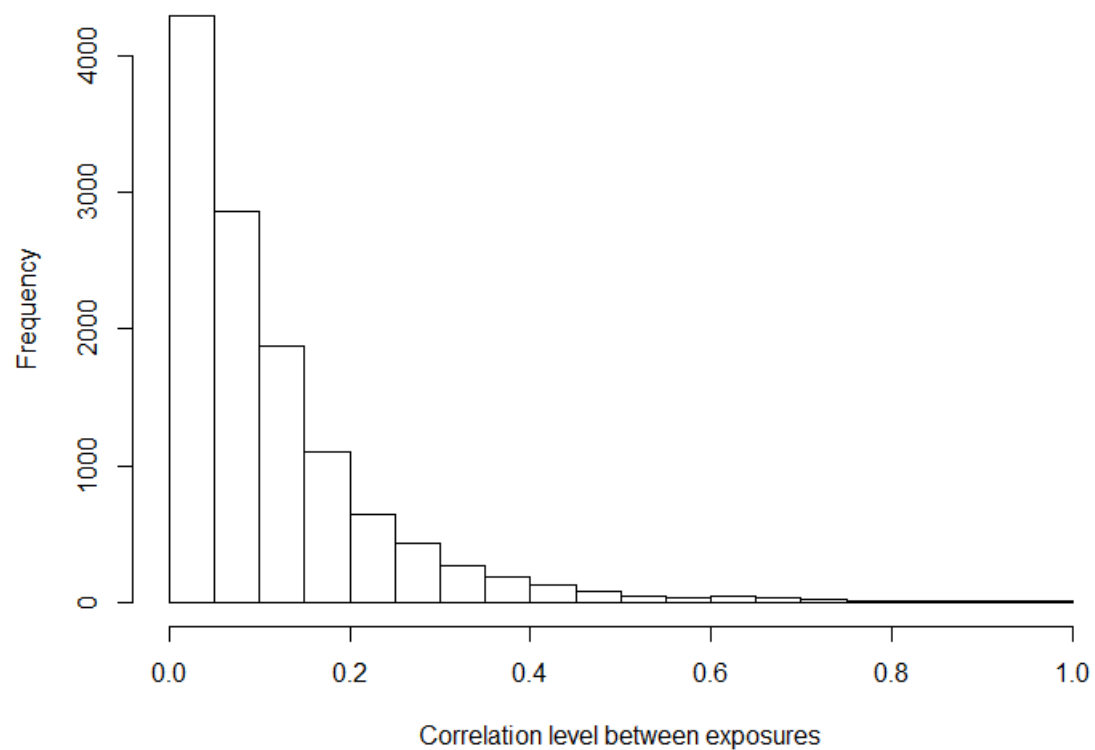
Supplementary Material 3: Boxplot of child zBMI in HELIX data, by cohorts.

Cohorts (number of individuals)	BIB		EDEN		INMA		KANC		MOBA		RHEA	
Characteristic	Mean (SD)	n (%)	Mean (SD)	n (%)	Mean (SD)	n (%)	Mean (SD)	n (%)	Mean (SD)	n (%)	Mean (SD)	n (%)
Child BMI	16.0 (2.0)		17.9 (2.7)		18.1 (3.1)		16.4 (2.3)		16.3 (1.9)		16.8 (2.6)	
Child sex												
Female		91 (45)		63 (43)		98 (46)		91 (46)		98 (46)		88 (44)
Male		112 (55)		83 (57)		117 (54)		107 (54)		114 (54)		111 (56)
Child age (year)	6.6 (0.2)		10.7 (0.5)		8.82 (0.6)		6.5 (0.5)		8.5 (0.5)		6.5 (0.3)	
Maternal education												
low		95 (47)		10 (7)		50 (23)		12 (6)		0 (0)		9 (5)
middle		36 (18)		53 (36)		91 (42)		69 (35)		43 (20)		110 (55)
high		72 (35)		83 (57)		74 (34)		117 (59)		169 (80)		80 (40)
Maternal pre- pregnancy BMI (kg/m ²)	28.2 (5.3)		23.3 (4.2)		24 (4.7)		27.6 (5)		22.7 (3.3)		24.1 (4.3)	
Parity before index pregnancy												
0		86 (42)		70 (48)		116 (54)		85 (43)		97 (46)		76 (38)
1		57 (28)		51 (35)		90 (42)		57 (29)		88 (42)		87 (44)
2 or more		60 (30)		25 (17)		9 (4)		56 (28)		27 (13)		36 (18)
Trimester of conception												
January- March		73 (36)		51 (35)		46 (21)		63 (32)		75 (35)		60 (30)
April-June		20 (10)		30 (21)		48 (22)		28 (14)		43 (20)		65 (33)
July- September		34 (17)		24 (16)		60 (28)		64 (32)		35 (17)		43 (22)
October- December		76 (37)		41 (28)		61 (28)		43 (22)		59 (28)		31 (16)
Maternal smoking status during pregnancy												
no exposure		137 (67)		99 (68)		78 (36)		102 (52)		190 (90)		18 (9)
only passive exposure		39 (19)		16 (11)		84 (39)		84 (42)		12 (6)		139 (70)
smoker		27 (13)		31 (21)		53 (25)		12 (6)		10 (5)		42 (21)
Postnatal tobacco												

smoke exposure							
<i>not exposed</i>	145 (71)	108 (74)	146 (68)	115 (58)	171 (81)	60 (30)	
<i>exposed</i>	58 (29)	38 (26)	69 (32)	83 (42)	41 (19)	139 (70)	
Maternal age (years)	28.7 (5.8)	30.8 (4.9)	31.9 (4)	29.1 (4.9)	32.7 (3.7)	30.9 (4.8)	
Birthweight							
< 2500g	13 (6)	5 (3)	5 (2)	5 (3)	6 (3)	6 (3)	
2500 to 3500g	114 (56)	92 (63)	141 (66)	75 (38)	97 (46)	143 (72)	
3500 to 4000g	56 (28)	36 (25)	62 (29)	84 (42)	76 (36)	43 (22)	
≥ 4000g	20 (10)	13 (9)	7 (3)	34 (17)	33 (16)	7 (4)	
Breastfeeding duration							
< 10.8 weeks	68 (33)	81 (55)	60 (28)	32 (16)	32 (15)	88 (44)	
10.8 to 34.9 weeks	87 (43)	52 (36)	98 (46)	61 (31)	42 (20)	79 (40)	
> 34.9 weeks	48 (24)	13 (9)	57 (27)	105 (53)	138 (65)	32 (16)	
Parents born in the country of inclusion							
None	81 (40)	0 (0)	1 (0)	8 (4)	42 (20)	2 (1)	
Only one	31 (15)	10 (7)	10 (5)	0 (0)	1 (0)	6 (3)	
Both	91 (45)	136 (93)	204 (95)	190 (96)	169 (80)	191 (96)	
Ethnicity							
African	7 (3)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
Asian	13 (6)	0 (0)	0 (0)	0 (0)	6 (3)	0 (0)	
European	87 (43)	146 (100)	215 (100)	198 (100)	203 (96)	199 (100)	
ancestry							
Native	0 (0)	0 (0)	0 (0)	0 (0)	2 (1)	0 (0)	
American							
Pakistani	79 (39)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
Other	17 (8)	0 (0)	0 (0)	0 (0)	1 (0)	0 (0)	

Supplementary Material 4: Population characteristics by cohort

Supplementary Material 5: Distribution of pairwise coefficients of correlation within quantitative variables of the full exposome assessed in 1,173 mother-child pairs from HELIX cohort



Supplementary Material 6: Adjusted associations between the reduced methylome (2284 CpGs) and zBMI in 1,173 children from the HELIX cohort (ExWAS model, step b) of the Meet-in-the-Middle approach). Results are presented only for CpGs with a (FDR-corrected for multiple hypothesis testing) p-value below 0.05 in ExWAS.

CpG site	Gene	Effect estimate	95% CI	Uncorrected p-Value	FDR-corrected p-Value
cg23098018	PIK3CD	3.64	2.12 5.16	2.80x10 ⁻⁶	3.20x10 ⁻³
cg01943221	PIK3CD	4.16	2.47 5.85	1.51x10 ⁻⁶	3.20x10 ⁻³
cg06695691	SPATA5	-3.99	-5.70 -2.27	5.65x10 ⁻⁶	4.30x10 ⁻³
cg03781224	MGLL	-5.79	-8.38 -3.20	1.29x10 ⁻⁵	5.96x10 ⁻³
cg25905215	NFKB1	5.23	2.88 7.57	1.31x10 ⁻⁵	5.96x10 ⁻³
cg09706586	PIK3CD	3.37	1.81 4.93	2.46x10 ⁻⁵	8.19x10 ⁻³
cg11947782	PIK3CD	3.05	1.63 4.48	2.80x10 ⁻⁵	8.19x10 ⁻³
cg18288462	ELOVL3	-3.70	-5.43 -1.96	3.14x10 ⁻⁵	8.19x10 ⁻³
cg12228229	DLG4	-9.94	-14.63 -5.24	3.58x10 ⁻⁵	8.19x10 ⁻³
cg27110374	GRB2	2.68	1.41 3.94	3.48x10 ⁻⁵	8.19x10 ⁻³
cg15526535	TNFRSF1B	2.31	1.18 3.44	6.48x10 ⁻⁵	0.01
cg09900893	RPS6KA1	-4.54	-6.78 -2.31	7.10x10 ⁻⁵	0.01
cg05805445	SPATA5	2.64	1.29 4.00	1.31x10 ⁻⁴	0.02
cg09035699	ACSL6	2.55	1.23 3.86	1.56x10 ⁻⁴	0.02
cg14841483	ACSL6	2.80	1.37 4.24	1.35x10 ⁻⁴	0.02
cg14003265	TRAF2	2.74	1.33 4.15	1.51x10 ⁻⁴	0.02
cg02423534	ADCY6	-2.69	-4.08 -1.31	1.37x10 ⁻⁴	0.02
cg00810292	TBC1D4	1.66	0.81 2.50	1.28x10 ⁻⁴	0.02
cg12085119	IRS2	3.31	1.63 4.99	1.12x10 ⁻⁴	0.02
cg16702014	ABCC1	-2.13	-3.23 -1.04	1.38x10 ⁻⁴	0.02
cg01312837	CREBBP	-4.14	-6.27 -2.01	1.44x10 ⁻⁴	0.02
cg22435313	PRKCA	2.00	0.97 3.04	1.52x10 ⁻⁴	0.02
cg22284398	PRKCE	-2.98	-4.53 -1.43	1.73x10 ⁻⁴	0.02
cg02105211	ITPR1	-1.57	-2.39 -0.75	1.74x10 ⁻⁴	0.02
cg02099877	NCEH1	-3.07	-4.68 -1.46	1.89x10 ⁻⁴	0.02
cg11010552	ACSL6	-2.44	-3.72 -1.15	2.08x10 ⁻⁴	0.02

CpG site	Gene	Effect estimate	95% CI	Uncorrected p-Value	FDR-corrected p-Value
cg05004855	CAMK2A	-3.84	-5.86 -1.82	2.03x10 ⁻⁴	0.02
cg09365147	REV3L	-1.18	-1.80 -0.56	2.09x10 ⁻⁴	0.02
cg19791262	HK3	-3.22	-4.94 -1.49	2.68x10 ⁻⁴	0.02
cg18681426	ELOVL5	-4.29	-6.61 -1.98	2.78x10 ⁻⁴	0.02
cg13941235	RXRA	-2.13	-3.27 -0.98	2.72x10 ⁻⁴	0.02
cg04056757	GRB2	3.12	1.45 4.78	2.53x10 ⁻⁴	0.02
cg07298473	NR1H3	-2.34	-3.61 -1.06	3.28x10 ⁻⁴	0.02
cg09877009	PRKCQ	-3.72	-5.76 -1.69	3.44x10 ⁻⁴	0.02
cg13832670	CREB3L2	-1.98	-3.08 -0.88	4.18x10 ⁻⁴	0.03
cg23875758	SREBF1	-5.80	-9.02 -2.59	4.14x10 ⁻⁴	0.03
cg00793946	PRKCE	-9.05	-14.07 -4.02	4.35x10 ⁻⁴	0.03
cg04730825	ABCC1	-2.47	-3.85 -1.10	4.41x10 ⁻⁴	0.03
cg07217499	CACNA1C	-3.87	-6.07 -1.67	5.87x10 ⁻⁴	0.03
cg19542445	CACNA1C	-2.98	-4.68 -1.28	5.96x10 ⁻⁴	0.03
cg07382687	CREB3L2	-2.32	-3.65 -0.99	6.30x10 ⁻⁴	0.03
cg12978800	PRKAG2	-2.49	-3.93 -1.06	6.75x10 ⁻⁴	0.03
cg13487983	RXRA	-4.39	-6.91 -1.86	6.75x10 ⁻⁴	0.03
cg16401207	PRKG1	-3.91	-6.15 -1.66	6.67x10 ⁻⁴	0.03
cg25338454	ITPR2	-1.43	-2.26 -0.61	6.45x10 ⁻⁴	0.03
cg27340723	ADCY9	-1.91	-3.02 -0.81	7.21x10 ⁻⁴	0.04
cg09265397	NOTCH3	-2.55	-4.02 -1.07	7.54x10 ⁻⁴	0.04
cg23257225	ADORA1	-3.82	-6.04 -1.60	7.72x10 ⁻⁴	0.04
cg19782686	SPATA5	-1.43	-2.26 -0.59	8.33x10 ⁻⁴	0.04
cg18912768	ABCB11	-2.14	-3.41 -0.87	9.64x10 ⁻⁴	0.04
cg18390025	ELOVL3	-1.57	-2.50 -0.64	9.75x10 ⁻⁴	0.04
cg00536939	NR1H3	-1.62	-2.59 -0.66	9.59x10 ⁻⁴	0.04
cg16942632	CAMKK2	-1.19	-1.89 -0.49	8.96x10 ⁻⁴	0.04
cg19982668	MAP3K3	-2.32	-3.70 -0.95	9.56x10 ⁻⁴	0.04
cg24961795	PLCG1	1.67	0.68 2.66	9.29x10 ⁻⁴	0.04
cg13092108	RPS6KA1	-2.97	-4.74 -1.20	1.03x10 ⁻³	0.04

CpG site	Gene	Effect estimate	95% CI	Uncorrected p-Value	FDR-corrected p-Value
cg18537222	PPARG	-1.08	-1.73 -0.43	1.08x10 ⁻³	0.04
cg05379597	CAMK2G	-2.04	-3.26 -0.81	1.12x10 ⁻³	0.04
cg09473249	ABCC1	-6.15	-9.84 -2.46	1.11x10 ⁻³	0.04
cg12005026	PTEN	-2.37	-3.80 -0.94	1.22x10 ⁻³	0.05
cg04885396	ABCC1	-1.71	-2.76 -0.67	1.28x10 ⁻³	0.05
cg23369234	CACNA1C	-5.80	-9.34 -2.26	1.35x10 ⁻³	0.05

Supplementary Material 7: Adjusted associations between exposures and CpGs associated with childhood zBMI in 1,173 children from HELIX cohort (ExWAS model, step c) of the Meet-in-the-Middle approach). Results are presented only for CpGs with a (FDR-corrected for multiple hypothesis testing) p-value below 0.05 in ExWAS.

Supplementary Material 8: Characteristics of the CpG selected by a methylome wide analysis on the whole methylome (row percentages)

CpGs	Number of CpGs	Number (%) of CpGs selected by MWAS on the whole methylome - Benjamini Hochberg correction
All	386,518	1788 (0.46%)
Belonging to a pathway AND enhancer	2,284	28 (1.22%)
Neither belonging to a pathway nor being an enhancer	384,234	1760 (0.46%)

Supplementary Material 9: Sensitivity analysis III - adjusted association between the whole methylome and zBMI in 1,173 children from the HELIX cohort (ExWAS model, step b of the Meet-in-the-Middle approach applied to the whole methylome). Results are presented only for CpGs with a (FDR - corrected for multiple hypothesis testing) p-value below 0.05 in ExWAS.

This supplementary is provided as a separated file due to its large size.

Exposure	Number of CpGs associated with the exposure and with zBMI	Number of corresponding genes
Copper - postnatal	1110	677
Bisphenol A (BPA) -postnatal	449	293
Perfluorooctanesulfonic acid (PFOS) - postnatal	180	133
Humidity - Pregnancy	47	38
PCBs (sum) - Pregnancy	10	10
PCB 170 - Pregnancy	9	7
MEHP - Postnatal	8	7
MiBP - Postnatal	7	7
DETP - Pregnancy	3	3
DDT - Postnatal	2	1
PBDE 153 - Postnatal	2	2
PCB 138 - Pregnancy	2	2
PFHXS - Postnatal	2	2
Thallium - Postnatal	2	2
Traffic noise (24h) - Postnatal	2	2
Cesium - Postnatal	1	1
Fastfood intake - Pregnancy	1	1
Green spaces (300 m) - Pregnancy	1	1
HCB - Postnatal	1	1
MBzP - Postnatal	1	1
OH-MiNP - Pregnancy	1	1
PM2.5 - Pregnancy	1	1

Population density - Postnatal	1	1
PRPA - Pregnancy	1	1
Social participation - Postnatal	1	1
Soda intake - Postnatal	1	1
Vegetables intake - Pregnancy	1	1
Yogurt intake - Postnatal	1	1

Supplementary Material 10: Sensitivity analysis III, Meet-in-the-Middle without CpGs preselection: adjusted associations between the exposome and CpGs associated with zBMI in 1,173 children from the HELIX cohort (ExWAS model adjusted on zBMI, step c of the Meet-in-the-Middle approach applied on the whole methylome). Results are presented only for exposures associated with a (stringently corrected for multiple hypothesis testing) p-value of less than 0.05 in exposure-CpGs ExWAS, with CpGs being previously selected in a CpGs-zBMI ExWAS.

Supplementary Material 11: Sensitivity analysis IV: Meet-in-the-Middle approach considering the cell-types as the intermediate layer. Adjusted associations at steps *b*),*c*) and *d*).

A.

Cell types	Effect estimate*	95%CI		Unadjusted p-Value	FDR adjusted p -Value
NK-cells	0.87	-0.82	2.56	0.31	0.38
B-cells	-0.33	-2.19	1.53	0.72	0.72
CD4+ T-cells	-2.22	-3.31	-1.13	7.00x10⁻⁵	4.20x10⁻⁴
CD8+ T-cells	-1.91	-3.35	-0.47	9.20x10⁻³	0.02
Granulocytes	0.87	0.23	1.51	7.86x10⁻³	0.02
Monocytes	3.11	0.70	5.53	0.01	0.02

* Adjusted change in mean zBMI for each increase by 1 in cell-type level. Models were adjusted for maternal BMI, maternal education, maternal smoking during pregnancy, parental country, cohort, parity, trimester of conception, ethnicity, child age and child sex and additionally only for postnatal exposures birth weight, passive smoking during childhood, breastfeeding duration.

B.

Cell types	Exposure variable	Effect estimate*	95%CI	Unadjusted p-Value	FDR adjusted p –Value
CD4+ T-cells	Copper - Postnatal	-0.01	-0.02 -8.76x10 ⁻³	6.56x10 ⁻⁹	5.67x10 ⁻⁶
Granulocytes	Copper - Postnatal	0.02	0.01 0.03	8.50x10 ⁻⁸	3.67x10 ⁻⁵
CD8+ T-cells	Copper - Postnatal	-7.68x10 ⁻³	-0.01 -4.27x10 ⁻³	1.09x10 ⁻⁵	3.13x10 ⁻³
Granulocytes	BPA - Postnatal	-0.01	-0.02 -6.96x10 ⁻³	3.36x10 ⁻⁵	7.26x10 ⁻³
CD4+ T-cells	Humidity - Pregnancy	-0.04	-0.06 -0.02	2.36x10 ⁻⁴	0.04

* Adjusted change in transformed exposure for each increase by 1 in cell-type level. Models were adjusted for maternal BMI, maternal education, maternal smoking during pregnancy, parental country, cohort, parity, trimester of conception, ethnicity, child age and child sex and additionally only for postnatal exposures birth weight, passive smoking during childhood, breastfeeding duration.

C.

Exposure group	Exposure variable	Unit	Transformation	Effect estimate*	95%CI	Unadjusted p-Value	FDR adjusted p –Value
Metals	Copper - Postnatal	µg/L	Log2	0.22	0.14 0.30	3.57x10⁻⁷	1.07x10⁻⁶
Phenols	BPA - Postnatal	µg/g	Log2	-0.07	-0.14 2.83x10 ⁻⁴	0.05	0.08
Meteorological	Humidity - Pregnancy	%	None	0.05	-0.34 0.44	0.81	0.81

* Adjusted change in mean zBMI for each increase by 1 in transformed exposure level. Models were adjusted for maternal BMI, maternal education, maternal smoking during pregnancy, parental country, cohort, parity, trimester of conception, ethnicity, child age and child sex and additionally only for postnatal exposures birth weight, passive smoking during childhood, breastfeeding duration.