

RESEARCH ARTICLE

Open Access



# Prenatal environmental exposures associated with sex differences in childhood obesity and neurodevelopment

Alejandro Cáceres<sup>1,2,3\*†</sup> , Natàlia Carreras-Gallo<sup>1†</sup>, Sandra Andrusaityte<sup>4</sup>, Mariona Bustamante<sup>1,2,5,6</sup>, Àngel Carracedo<sup>7,8</sup>, Leda Chatzi<sup>9</sup>, Varun B. Dwaraka<sup>10</sup>, Regina Grazuleviciene<sup>4</sup>, Kristine Bjerve Gutzkow<sup>11</sup>, Johanna Lepeule<sup>12</sup>, Léa Maitre<sup>1,2,5</sup>, Tavis L. Mendez<sup>10</sup>, Mark Nieuwenhuijsen<sup>1,2,5</sup>, Remy Slama<sup>12</sup>, Ryan Smith<sup>10</sup>, Nikos Stratakis<sup>1</sup>, Cathrine Thomsen<sup>11</sup>, Jose Urquiza<sup>1,2,5</sup>, Hannah Went<sup>10</sup>, John Wright<sup>13</sup>, Tiffany Yang<sup>13</sup>, Maribel Casas<sup>1,2,5</sup>, Martine Vrijheid<sup>1,2,5</sup> and Juan R. González<sup>1,2,14\*</sup>

## Abstract

**Background** Obesity and neurodevelopmental delay are complex traits that often co-occur and differ between boys and girls. Prenatal exposures are believed to influence children's obesity, but it is unknown whether exposures of pregnant mothers can confer a different risk of obesity between sexes, and whether they can affect neurodevelopment.

**Methods** We analyzed data from 1044 children from the HELIX project, comprising 93 exposures during pregnancy, and clinical, neuropsychological, and methylation data during childhood (5–11 years). Using exposome-wide interaction analyses, we identified prenatal exposures with the highest sexual dimorphism in obesity risk, which were used to create a multiexposure profile. We applied causal random forest to classify individuals into two environments: E1 and E0. E1 consists of a combination of exposure levels where girls have significantly less risk of obesity than boys, as compared to E0, which consists of the remaining combination of exposure levels. We investigated whether the association between sex and neurodevelopmental delay also differed between E0 and E1. We used methylation data to perform an epigenome-wide association study between the environments to see the effect of belonging to E1 or E0 at the molecular level.

**Results** We observed that E1 was defined by the combination of low dairy consumption, non-smokers' cotinine levels in blood, low facility richness, and the presence of green spaces during pregnancy ( $OR_{\text{interaction}} = 0.070$ ,  $P = 2.59 \times 10^{-5}$ ). E1 was also associated with a lower risk of neurodevelopmental delay in girls, based on neuropsychological tests of non-verbal intelligence ( $OR_{\text{interaction}} = 0.42$ ,  $P = 0.047$ ) and working memory ( $OR_{\text{interaction}} = 0.31$ ,  $P = 0.02$ ). In line with this, several neurodevelopmental functions were enriched in significant differentially methylated probes between E1 and E0.

<sup>†</sup>Alejandro Cáceres and Natalia Carreras-Gallo equally contributed to the paper

\*Correspondence:

Alejandro Cáceres  
alejandro.caceres@isglobal.org

Juan R. González  
juan.gonzalez@isglobal.org

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



**Conclusions** The risk of obesity can be different for boys and girls in certain prenatal environments. We identified an environment combining four exposure levels that protect girls from obesity and neurodevelopment delay. The combination of single exposures into multiexposure profiles using causal inference can help determine populations at risk.

**Keywords** Prenatal environment, Sexual dimorphism, Childhood obesity, Neurodevelopment, DNA methylation, Causal inference, Multiexposure profile

## Background

Boys and girls develop differently. For instance, their immune response to infections differs from an early age, their brains grow at different rates, and the prevalence of numerous common diseases, like obesity, is also different [1–3]. As reported by Shah et al., 65% of the countries around the world and 96% of high-income countries reported a greater prevalence of obesity for boys than girls in children aged 5–9 years old [3, 4]. Given the contrasting paths of development, it is remarkable that biomedical studies typically consider sex as a confounder rather than the main effect or an effect modifier [5]. Exposome studies, in particular, are characterized by the acquisition of massive amounts of data at individual and population levels [6]. A crucial goal of these studies is to inform the likely conditions for which a given public health intervention would be optimal, such that the best intervention is applied at the right time to the right population [7]. However, as the main difference between individuals is sex, exposome studies aiming at improving precision medicine and precision public health cannot do without considering how environmental risk factors affect sexual dimorphism in development and disease.

From a mechanistic context, studying the factors that increase sexual dimorphic outcomes of disease can offer important insights into its etiology and comorbidities, and inform of possible interventions and targeted treatments. Important advancements have been made in studying sex-related risk factors for diseases like cancer, Alzheimer's, and autoimmune diseases [8]. However, a relevant component of these age-related diseases is hormonal regulation. Studying sex differences in preteens offers not only the opportunity for identifying targeted treatments for early-age illnesses but also to explore disease mechanisms unlikely influenced by sex hormones that may also onset early in life. Previous research has, for instance, underlined that maternal factors during pregnancy can affect disease outcomes later in life [9] and, therefore, motivates the question of which pregnancy factors may promote later sexual dimorphism in disease.

Environmental exposures likely orchestrate environments that are more toxic to one sex than to the other one. However, methods to determine such multiple-exposure environments are not readily available. We have developed a method of causal modeling, based

on causal random forest, that can determine profiles of multiple exposures that are associated with high sexual dimorphism [10]. Here, we aimed to adapt our method to determine which combination of prenatal exposures can produce an environment where girls are more protected from obesity than boys during the preteen years. Furthermore, obesity in children is associated with lower cognitive function, particularly inhibitory control and working memory, critical for academic achievement [11]. Obesity often co-occurs with neurodevelopmental disorders, particularly in boys [12]. Therefore, we also evaluated whether the environment of high sexual dimorphism in obesity also shows a significant sexual dimorphism in non-verbal intelligence, working memory, attention, and ADHD.

Finally, we investigated whether the protective environment may be associated with epigenetic changes since many exposures during pregnancy are associated with specific methylation profiles [13]. This analysis may provide information about the molecular pathways that may be participating in the association between the environment and the sexual differences in obesity and neurodevelopment.

Here, we aimed to (1) combine multiple exposure levels to define an environment with high sexual dimorphism in obesity risk; (2) given the correlation between obesity and neurodevelopmental delay in children, we also enquired if the subpopulation exposed to this environment shows a significant sexual dimorphism in neurodevelopment; and (3) we then hypothesized that the individuals who belong to such an environment can be characterized by specific patterns of DNA methylation.

## Methods

### Study population

We analyzed data from The Human Early Life Exposome (HELIX). This is a multi-center study that included a total of 1301 mother–child pairs from six existing birth cohorts in Europe: BIB (Born in Bradford; the UK) [14], EDEN (Etude des Déterminants pré et postnatals du développement et de la santé de l'Enfant; France) [15], INMA-SAB (Infancia y Medio Ambiente; Spain; subcohort Sabadell) [16], KANC (Kaunas cohort; Lithuania) [17], MoBa (The Norwegian Mother, Father and Child

Cohort study; Norway) [18]), and Rhea (Greece) [19]. The pairs participated in a common, completely harmonized, follow-up examination, when children were between 5–11 years old to fully characterize the pregnancy and childhood exposome [20]. During the clinical examination, urine (pooled spot urine samples from before bedtime and first morning void) and blood samples were collected from the children. Urine and blood samples previously collected from mothers during pregnancy were also available for biomarkers of chemical exposure assessment. In our analyses, we selected the individuals who had data on prenatal exposures, performed the clinical and neurodevelopment examination, and had methylation data ( $N = 1044$ ). All studies received approval from the ethics committees of the centers involved and written informed consent was obtained from all participants. Cohort characteristics are shown in Table 1.

### Clinical outcomes

Height and weight measurements were measured during the clinical visit performed at ages 5 to 11 years. These measurements were converted to body mass index (BMI in  $\text{kg}/\text{m}^2$ ) for age-and-sex  $z$ -scores using the international WHO reference curves to allow comparison with other studies [21]. Obese children were defined as those above the age- and sex-specific 95th percentile, as recommended by WHO.

Neurodevelopmental outcomes were assessed through a battery of internationally standardized, non-linguistic, and culturally blind computer tests also at ages 5 to 11 years. We assessed working memory, attention, and general non-verbal intelligence with the N-back test [22], the attention network test (ANT) [23], and Raven's colored progressive matrices [24]; respectively. The tests were administered in

**Table 1** Characteristics of HELIX cohort. Clinical characteristics of children during pregnancy and follow-up

Children assessed at follow-up	<b>N = 1044</b>
<b>Sex, male</b>	571 (54.6%)
<b>Cohort</b>	
BIB (UK)	90 (8.6%)
EDEN (France)	135 (12.9%)
INMA (Spain)	198 (19.0%)
KANK (Lithuania)	196 (18.8%)
MOBA (Norway)	239 (22.9%)
RHEA (Greece)	136 (17.8%)
<b>Age in years, mean (range)</b>	7.9 (5.4–11.9)
<b>BMI (<math>\text{kg}/\text{m}^2</math>), median (range)</b>	16.3 (12.2–29.5)
Obesity	62 (5.9%) – F: 23 (4.9%) – M: 39 (6.8%)
<b>Raven's matrices, median (range)</b>	27 (9–36)
Affected	189 (18.2%) – F: 78 (16.6%) – M: 111 (19.5%)
<b>N-back (2-back accuracy), median (range)</b>	0.91 (0.36–1)
Affected	104 (12.9%) – F: 46 (12.6%) – M: 58 (13.0%)
<b>ANT (accuracy), median (range)</b>	0.97 (0.51–1)
Affected	206 (20.0%) – F: 72 (15.6%) – M: 134 (23.7%)
<b>ADHD</b>	104 (10.0%) – F: 27 (5.8%) – M: 77 (13.6%)
<b>Prenatal characteristics</b>	<b>N = 1044</b>
<b>Mother's age during pregnancy, mean (range)</b>	30.9 (16–34)
<b>Mother's BMI during pregnancy, median (range)</b>	23.6 (15.8–51.4)
<b>Mother's weight gain during pregnancy, mean (range)</b>	13.9 (0–40)
<b>Maternal education</b>	
Primary school	119 (11.3%)
Secondary school	359 (34.3%)
University degree or higher	566 (54.2%)
<b>Gestational age, median (range)</b>	40 (30.8–44.1)
<b>Year of birth, median (range)</b>	2006 (2003–2009)
<b>Number of parents native from the country cohort, mean (range)</b>	1.9 (02)
Parity	0.69 (0–2)

BMI body mass index, ANT attention network test, ADHD attention deficit hyperactive disorder, F female, M male

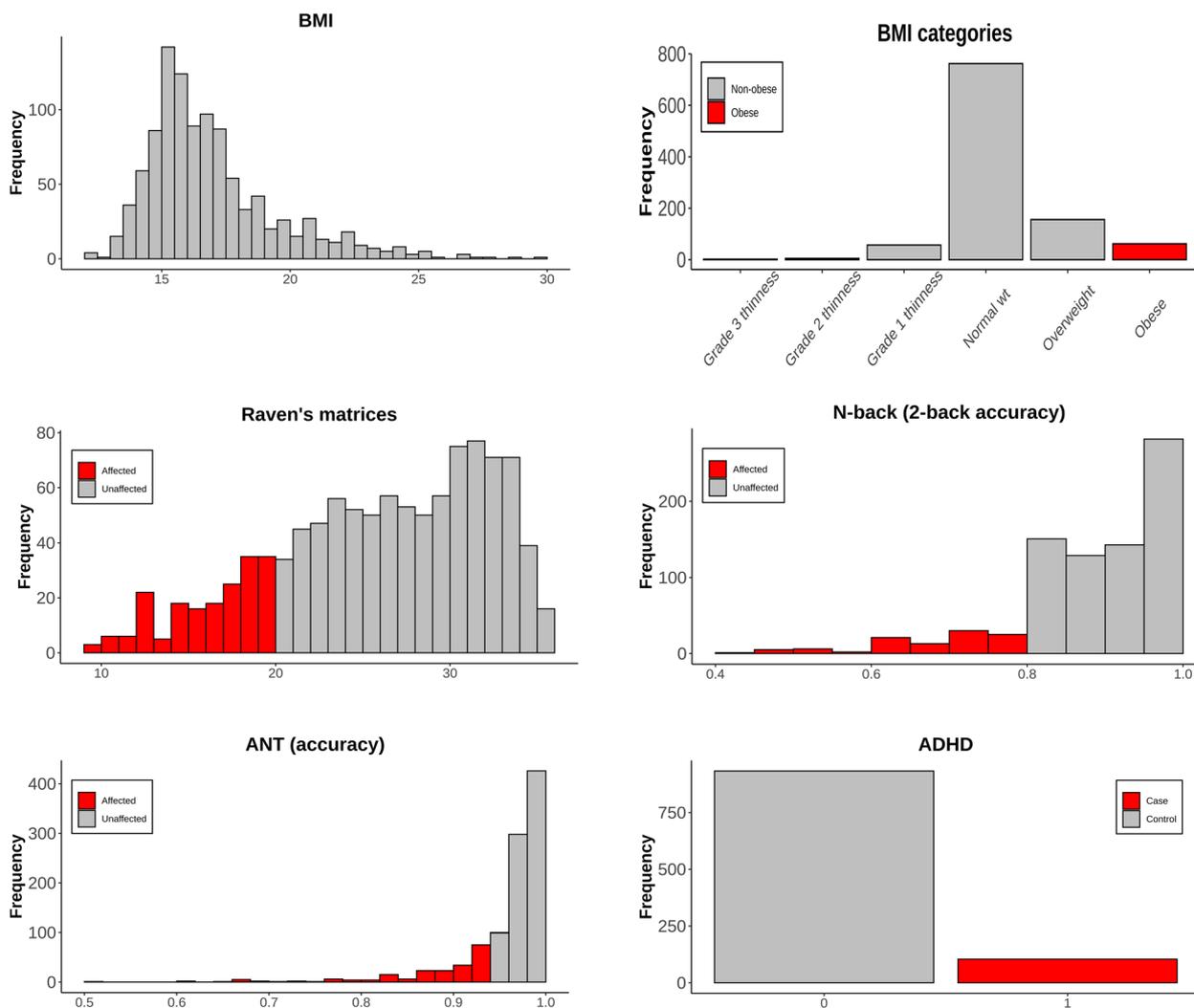
a standardized way by trained field workers through study-provided laptops. The outcomes did not distribute normally. We dichotomized them, taking as cases individuals with outcomes below the first quintiles (20%), which clearly captured the long lower tail of the outcomes' distributions (Fig. 1). These percentiles were chosen because they allowed the selection of the lowest performers in the outcomes using a single criterion and preserving a representative number of individuals within the groups. We thus studied as clinical outcomes the events of having these cognitive abilities affected. We also considered ADHD diagnosis.

We considered as common covariates (covariates used in all the analyses) 10 variables, based on Maitre et al. [25]. These covariates are cohort, year of birth,

mother's BMI, mother's weight gain during pregnancy, gestational age, mother's age during pregnancy, mother's education, whether parents were native from the country cohort, parity, and children's age at clinical assessment (Table 1).

**Pregnancy exposome**

HELIX has collected a wide range of exposures measured during two main windows: a prenatal window including the pregnancy period and a postnatal window including the exposome data of children at the same time as omics sampling (childhood). In this study, we only considered the first window (pregnancy exposome) which consists of 93 exposures distributed across 17 exposure families, including the urban environment, the chemical exposome, and social and lifestyle factors.



**Fig. 1** Distributions of clinical outcomes in the HELIX study. Analyses were performed for categorized variables shown in gray (reference) and red (affected)

The urban environment includes exposure estimates for built environment, surrounding green and blue spaces, ultraviolet (UV) radiation, road traffic noise levels, air pollution, noise, meteorology, and socioeconomic deprivation index [20, 26]. These exposures were assessed during pregnancy by environmental geographic information systems (GIS) according to their residential addresses. Tobacco smoke and diet were evaluated by questionnaires. Biomarkers of contaminant exposure, like cotinine levels, were measured in appropriate biological samples (urine or blood) collected from mothers during pregnancy. Details on the exposure assessment methods and exposure factors can be seen in the Additional file 1: Supplementary Methods [20, 27–33].

Missing values for all exposures were imputed using the method of chained equations using the *mice* package in R [32], as described in detail elsewhere [33]. When possible, multiple imputation procedure was applied (missing values are imputed stochastically several times). For the imputation process, continuous variables should have a normal distribution. Thus, skewed exposure variables were transformed to achieve normality or categorized if no transformation worked. Exposure variables with their corresponding transformation are described in Additional file 2: Table S1. Exposures with more than 70% of missing values in each cohort were excluded from the imputation process. Therefore, missing values ranged from 1.5% in traffic density to 65% in fast-food intake during pregnancy. Although none of the participants had complete data on all exposures, 95% of individuals had missing values in less than 30% of exposures.

### DNA methylation

One of the main goals of HELIX was to associate multiple environmental factors with omics biomarkers and child health outcomes. For these same children, multi-omics molecular phenotyping was performed, which included measurement of blood DNA methylation (450 K, Illumina), among others.

The DNA was obtained from buffy coat collected in EDTA tubes at 5–11 years of age. Briefly, DNA was extracted using the Chemagen kit (Perkin Elmer) in batches of 12 samples. Samples were extracted by cohort and following their position in the original boxes. DNA concentration was determined in a NanoDrop 1000 UV–Vis Spectrophotometer (ThermoScientific) and with Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies). DNA methylation was assessed using the Infinium Human Methylation 450 beadchip (Illumina), following the manufacturer's protocol. Preprocessing of methylation data has been described elsewhere [34]. After sample and probe quality control measures, the number of CpG probes analyzed was 371,533, initially available for 1192

subjects. We used the Combat algorithm to remove the batch effects supported by the slide. Methylation levels were expressed as beta values corrected by surrogate variables and CpG sites were annotated to genes by Illumina HM450 manifest file (version 1.2). We discarded the subjects without exposome data and without European ancestry based on genomic data, resulting in 993 individuals for the methylome analysis. We computed blood cell type proportions following Houseman et al. algorithm [35] and Reinius reference panel [36].

### Statistical methods

Figure 2 shows the statistical workflow.

#### Identification of prenatal exposures with sexual dimorphism in obesity risk

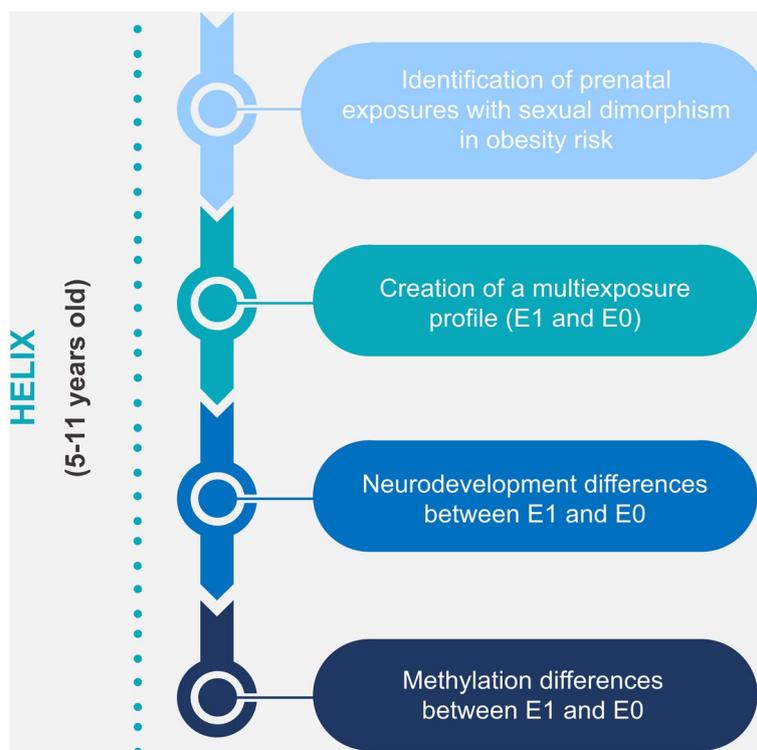
We used exposome-wide interaction analyses to determine the exposures whose association with obesity was significantly different between sexes. We assessed the associations between obesity (cases and controls) and the interactions between sex ( $S$ ) and each of the prenatal exposures ( $D_i$ ) using the logistic regression model.

$$E(Y) = \text{logit}^{-1}(\alpha_i + \beta_i(S \times D_i) + \sum_{r=1\dots k} \gamma_{ir}C_{ri})$$

where  $Y$  is the obesity status of an individual with sex  $S$  and  $i$ th exposure  $D_i$ .  $\gamma_{ir}$  are the regression coefficients of the  $k$  covariates  $C_{ri}$  that included sex, exposure  $I$ , and the 10 common covariates mentioned before (cohort, year of birth, mother's BMI, mother's weight gain during pregnancy, gestational age, mother's age during pregnancy, mother's education, whether parents were native from the country cohort, parity, and children age at clinical assessment).  $\beta_i$  were the effects of interest that measure the association between obesity and the interaction between sex and each exposure  $i$ . We adjusted  $p$ -values using false discovery rate to correct for multiple comparisons.

#### Creation of a multiexposure profile (E1 and E0)

We calculated the residuals of the exposures with nominal significant interactions adjusted by the 10 common covariates. Then, we used these residuals as covariates in causal inference modeling, using causal random forest and taking sex as the treatment variable, to determine which children in HELIX had been in personal environments with significant sexual dimorphism in obesity (female > male or female < male). We then aimed to determine whether the personal environments of the children with one of the significant dimorphisms (F > M or F < M) could be averaged into two prenatal environments, one whose female protection against obesity was stronger



**Fig. 2** Statistical workflow. The figure shows all the statistical analyses carried out along the paper

than those observed for the individual exposures and the other the opposite. We did not find enough children with negative dimorphism ( $F > M$ ). We thus created an average environment with highly significant female protection against obesity, which hereinafter we will refer to it as E1. E1 was defined as a binary vector, with one entry for each level of the exposures, indicating whether a given exposure averaged across the children with positive dimorphism was higher or lower than the average across all other children in the training set. We used the multi-exposure profile to classify all the individuals in the entire HELIX cohort depending on whether they belong to the E1 or not (E0:  $F < M$  and  $F = M$ ). To this end, we used soft targeting that tested whether they matched the environment in at least 60% of the exposures. The causal inference and the classification into the multiexposure profile associated with the E1 environment were performed with the algorithm *teff*, taking sex as the treatment variable [10] (<https://teff-package.github.io/>).

#### Neurodevelopment differences between E1 and E0

We used the classification of individuals into E1 and E0 to assess their relationship with sex differences in neurodevelopment. For this analysis, we used logistic regression models on the clinical outcomes (working memory, attention and general non-verbal intelligence with the

N-back test, ANT, Raven's colored progressive matrices, and ADHD) and we tested the interaction between the environment (E1/E0) with sex. We adjusted the model by sex, the environment, and the 10 common covariates.

#### Methylation differences between E1 and E0

We performed an epigenome-wide association study (EWAS) in the HELIX cohort between E1 and E0. As previously, we used logistic regression models to identify the probes that were differentially methylated between environments. We adjusted the analysis by the 10 common covariates used in previous analyses and counts of different immune cells in the blood. Associations were corrected for multiple comparisons using false discovery rate, as computed by *limma*. For the enrichment analysis, we used *clusterProfiler* Bioconductor package (V.3156). The commented analysis code is available in Additional file 3: Supplementary Code.

## Results

#### Sexual dimorphism of clinical outcomes

We first assessed whether obesity and the categorized neuropsychological measures were associated with differences between sexes (Fig. 1). We fitted logistic regression models adjusting by the 10 common covariates. Girls showed a lower frequency of obesity than boys, but

it was not statistically significant (OR=0.64,  $P=0.13$ , see Fig. 3A). For the neuropsychological measures, we observed that ADHD was lower in girls than boys, consistent with girls' higher protection in attention difficulty. Both associations were statistically significant (OR=0.37,  $P=2.87 \times 10^{-5}$ , OR=0.54,  $P=4.32 \times 10^{-4}$ ). For Raven's matrices and N-back, we did not see significant associations with sex (OR=0.72,  $P=0.10$ , OR=0.94,  $P=0.78$ , respectively) (Fig. 3A).

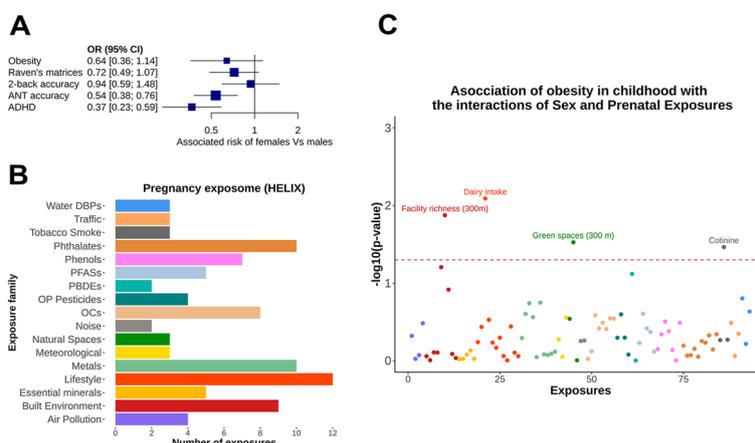
**Exposome-wide analysis of sex-exposure interactions on obesity**

We searched for prenatal exposures that could modulate the association between sex and obesity in childhood. Particularly, we searched for maternal exposure levels in which one sex would be more obese than the other at 5–11 years of age. We performed logistic regressions on obesity for all 93 sex-prenatal exposures interactions, adjusting by the common covariates, sex, and each exposure (Fig. 3B, C). We did not observe any interaction that passed multiple comparison corrections. However, at the nominal level ( $P<0.05$ ), we observed four interactions between sex (males as reference) and prenatal exposures. First, dairy consumption (OR<sub>interaction</sub> = 2.44,  $P=0.008$ ) is defined as mother's dairy consumption during pregnancy times per week and categorized as less than 18 times per week (low), between 18 and 27 (moderate), and more than 27 (high). Second, cotinine levels in mothers during pregnancy (OR<sub>interaction</sub> = 1.92,  $P=0.034$ ) are classified into three categories: non-smokers (less than 18.4 µg/L), second-hand smokers (between 18.4 and 48.4 µg/L), and smokers (more than 48.4 µg/L). Third, facility richness (OR<sub>interaction</sub> = 1.11,  $P=0.013$ ) is defined as the percentage of different facility types present compared to the maximum potential number of facility types at a 300-m

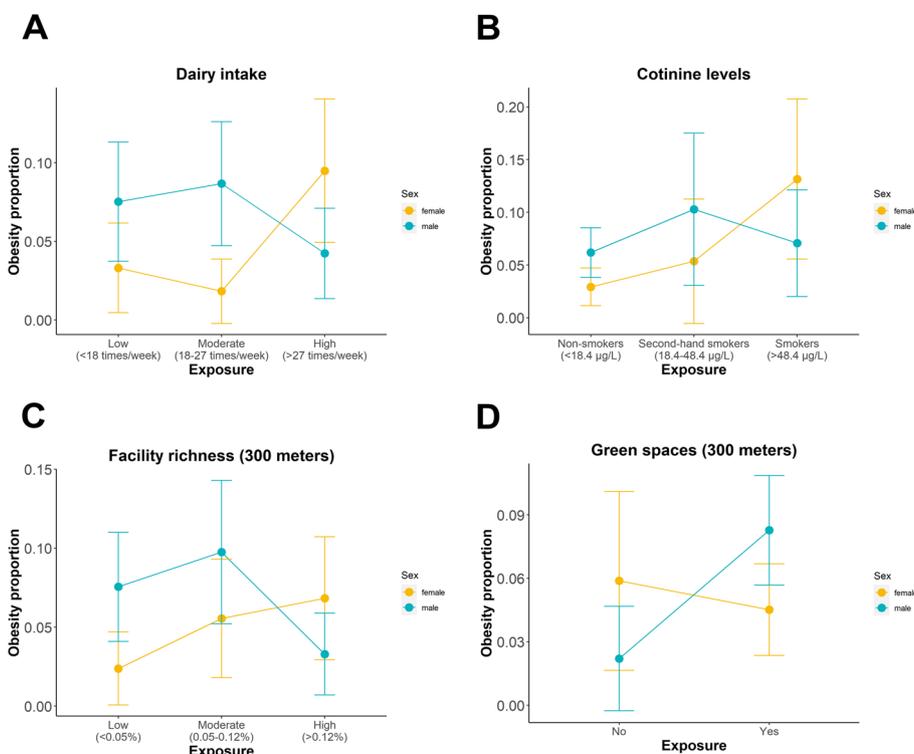
buffer during the pregnancy period. We categorized this variable into low (less than 0.05%), moderate (between 0.05 and 0.12%), and high abundance (more than 0.12%). Fourth, the presence of green spaces (OR<sub>interaction</sub> = 0.27,  $P=0.029$ ), answering the question of whether the mother lived within a distance of 300 m of green space during the pregnancy period (yes/no). A stratified analysis by sex of the association between obesity and the significant exposures revealed that dairy consumption and cotinine levels were risk factors only for girls (OR = 2.88,  $P=0.0009$ ; OR = 1.91,  $P=0.0128$ ) while facility richness and green spaces were protective and risk factors for boys, respectively (OR = 0.92,  $P=0.005$ ; OR = 5.06,  $P=0.007$ ), see Fig. 4. We finally asked the extent to which the four exposures were correlated between each other. Interestingly, we found weak but significant Pearson's correlations of facility richness with dairy intake ( $r = -0.11$ ,  $P=0.0002$ ) and cotinine levels ( $r = 0.07$ ,  $P=0.01$ ).

**Exposure environment of high differences in obesity risk between sexes**

We asked whether a combination of the four significant exposures and their levels could define specific environments where one sex is likely more obese than the other one. The exposure residuals, adjusted by common covariates, were used in causal inference modeling, with the aim to classify individuals into environments of high sexual dimorphism in obesity. We considered the multiexposure profile defined by the mother's dairy intake, cotinine levels, living richness facilities, and green spaces during pregnancy. We randomly selected a set of 208 individuals from the HELIX cohort to infer their expected sex difference in obesity risk given their personal multiexposure profiles. We thus applied the causal modeling algorithm *teff*, taking sex as the treatment variable, and observed



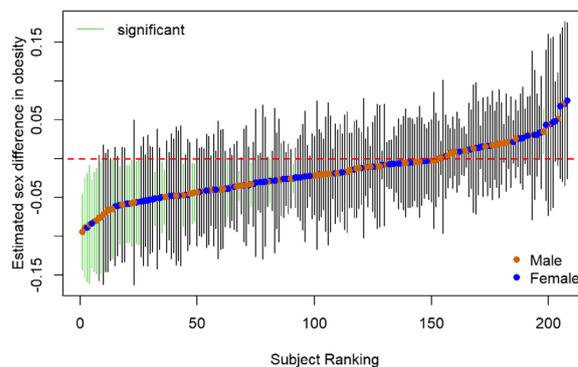
**Fig. 3** **A** Association of sex with the clinical outcomes, adjusting by covariates. **B** Number of prenatal exposures in each family measured in HELIX. **C** Exposome-wide Manhattan plot. Association of obesity with 93 sex-prenatal exposure interactions (the color follows the exposure family from panel **A**). The dotted line marks nominal significance ( $P=0.05$ )



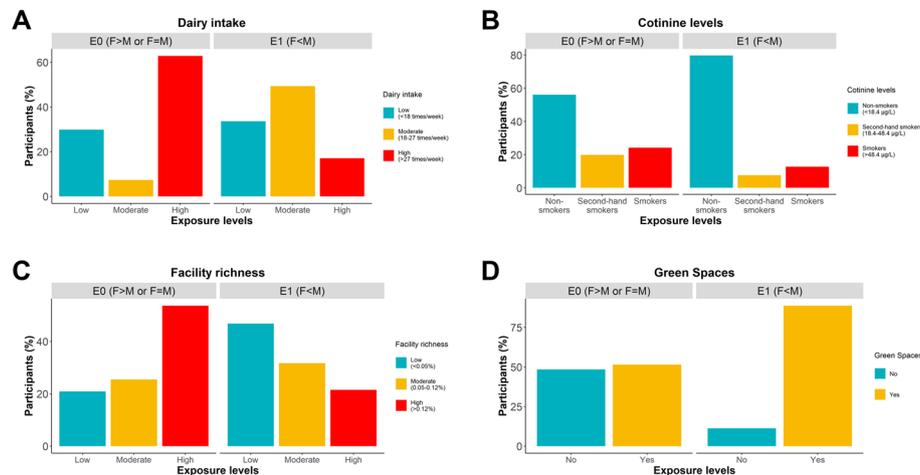
**Fig. 4** Sex-exposure interaction plots on obesity. **A** Mother’s dairy consumption during pregnancy. The figure shows the highest proportion of obesity in girls with the highest level of dairy consumption. **B** Mother’s cotinine levels in the blood. The highest levels of obesity were observed in girls with high cotinine levels. **C** Facility richness in living neighborhoods of pregnant mothers. A high abundance of facility richness is correlated with a low prevalence of obesity in boys. **D** Green spaces at 300 m from pregnant mothers’ homes. The highest prevalence of obesity was observed in boys with mothers living in the presence of green spaces. The bars represent the 95% confidence intervals for the estimated proportion of obesity

27 children (13 females, 14 males) living in personal environments where girls are less likely obese than boys. By contrast, we found only one boy living in a personal environment where girls are more likely obese than boys (Fig. 5).

We aimed to classify all individuals into two environments: E1 and E0. The first one (E1) consisted of a combination of exposure levels that protects girls against obesity ( $F < M$ ). The second one (E0) consists of the remaining combinations of exposure levels ( $F > M$  and  $F = M$ ). E1 was obtained using the personal environments of the 27 children where girls are expected to be less obese than boys. E1 was defined as a binary vector, with one entry for each level of the four exposures, indicating whether a given exposure averaged across 27 individuals was higher or lower than the average across the entire training set of 208 children. We used the multiexposure profile to classify all the individuals in HELIX and observed a total of 675 (64%) individuals classified into E1. We found that E1 was characterized by moderate



**Fig. 5** The estimated sex difference on obesity risk in personal prenatal environments. The personal prenatal exposure environments were defined by the mother’s dairy intake, cotinine levels, living richness facilities, and green spaces. The sex of the individual living in a particular prenatal environment is shown in blue (male) and orange (female). The bars show the 95%CI for the effect of a personal prenatal environment on females in relation to males. The intervals were estimated using causal modeling implemented in *teff*. Green lines are significant sex differences in obesity risk given by the prenatal environments



**Fig. 6** Characterization of the common prenatal environment where girls are more protected than boys against obesity (E1) against the reference environment (E0). Environment E1 is the common environment of individuals with personal environments where girls are significantly less obese than boys (female < male), these are the individuals with green confidence intervals in Fig. 5. E1 is defined by low mother dairy intake, non-smokers' cotinine levels, low richness facilities, and the presence of green spaces. An individual belongs to E0 if he/she does not belong to E1 (female > male or female = male)

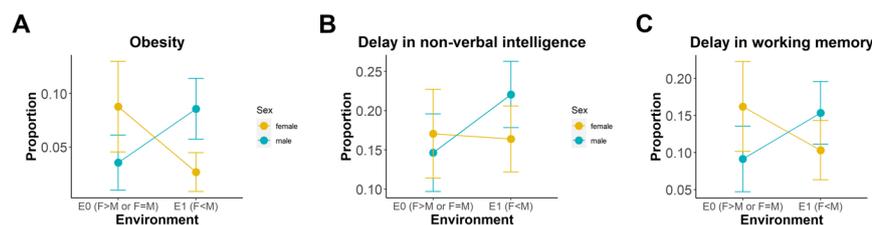
dairy consumption, non-smokers' cotinine levels, low abundance of facility richness, and the presence of green spaces (Fig. 6A–D). Therefore, the environment captured both obesity protection for girls and obesity risk for boys, as expected from the individual exposures.

We then observed a strong association of the sex-environment interaction on child obesity, adjusting by covariates ( $OR_{interaction} = 0.070, P = 2.59 \times 10^{-5}$ ). Stratified associations by sex between the environment and obesity risk were also significant (girls:  $OR = 0.18, P = 4.73 \times 10^{-4}$ ; boys:  $OR = 3.14, P = 0.012$ ), suggesting stronger environment gains in the protection for girls than in the risk for boys (Fig. 7A). These results show that E1 can be regarded as a prenatal environment of female protection against childhood obesity, with much stronger protection than those given its individual exposure components.

### Sexual dimorphism in neurodevelopment

We asked whether the environment of high differences in obesity between sexes was also an environment of high differences in neurodevelopment. First, we assessed the association between obesity and four neuropsychological outcomes, fitting logistic regression models on obesity and adjusting by common covariates, sex, and the environment (E1/E0). We observed that low values of Raven's matrices and N-back test tests were significant risk factors for obesity ( $OR = 2.42, P = 0.01$ ;  $OR = 2.65, P = 0.02$ , see Fig. 7B), as ADHD diagnosis increased the risk ( $OR = 2.15, P = 0.03$ , see Fig. 7C). However, we did not find significant associations between obesity and attention outcome.

We tested whether the subject classification into the environments E1 and E0 significantly interacted with sex on each of the neuropsychological outcomes, as it did



**Fig. 7** **A** Sex-environment interaction plot on obesity risk. The figure shows that E1 defines a prenatal environment across HELIX of strong female protection against childhood obesity (female < male), while E0 defines a prenatal environment without sexual dimorphism in obesity (F = M) or with male protection (F > M). **B** Sex-environment interaction plot on raven's matrices underperformance. Affected individuals are those with outcomes below the first quintiles. **C** Sex-environment interaction plot on N-back underperformance. Affected individuals are those with outcomes below the first quintiles

with obesity. We found that the sex-environment interaction was associated with higher outcomes of both Raven's matrices ( $OR_{interaction} = 0.42, P = 0.047$ ) and N-back test ( $OR_{interaction} = 0.31, P = 0.02$ ), suggesting a higher performance of girls with respect to boys in these two tests, within E1. Associations were fully adjusted by covariates.

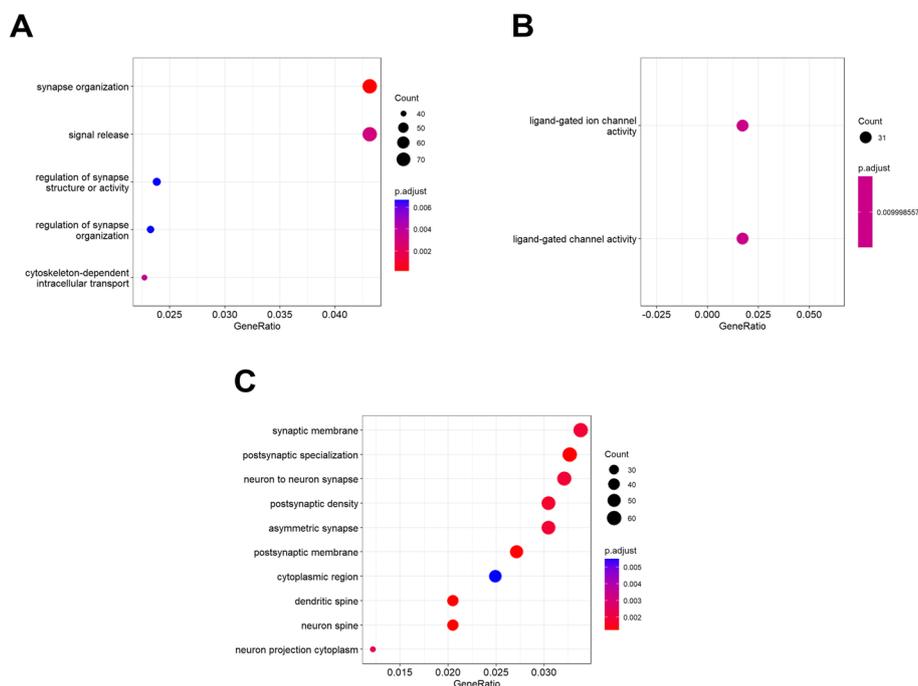
**Methylation profile associated with the prenatal environment of high sex differences in obesity**

We aimed to investigate whether the methylome captured the differences between individuals belonging to E1 or E0. We performed an EWAS of the classification of children in the prenatal environment, adjusting by common covariates and immune cell counts. Methylation data was extracted from blood samples and were previously normalized and corrected for surrogate variation. We did not observe any significant association at a genome-wide level after correcting for multiple comparisons, see top associations in Additional file 2: Table S2. We also performed an enrichment analysis for the top associations (nominal  $P < 0.01$ ). We tested different GO terms from molecular function, cellular components, and biological processes (Fig. 8), and observed several pathways related to neuronal processes. Most remarkably, *synapse organization* ( $P_{adjusted} = 0.0001$ ) and *regulation of synapse structure or activity* ( $P_{adjusted} = 0.006$ ) are two biological processes directly related to neurodevelopment.

**Discussion**

We have shown in the HELIX cohort that environments defined by a multiexposure profile with different effects on obesity for each sex can be identified with the novel use of causal inference [10]. In a previous study on the same cohort, no significant associations were observed for individual prenatal exposures with overweight and obesity status, while cotinine levels were associated with BMI only at nominal significance [37]. Although we observed only four nominally significant interactions between prenatal exposures and sex on obesity, we revealed a prenatal environment defined by specific levels of these exposures whose effect on obesity strongly changed between sexes, with a 93% reduction in obesity risk for girls in relation to boys ( $OR_{interaction} = 0.070, P = 2.59 \times 10^{-5}$ ). In the environment defined by moderate dairy consumption, non-smokers' cotinine levels, low facility richness, and the presence of green spaces, girls are more protected than boys against obesity.

Previous studies have shown conflictive findings on dairy intake during pregnancy and its relation to long-term body composition of children. Voerman et al. reported significant associations with abdominal fat in children and strong interaction with sex on the pericardial fat mass index, with a higher risk for girls [38]. However, other studies have reported no significant



**Fig. 8** Enrichment analysis for the differentially methylated sites associated with E1. **A** Gene Ontology. **B** Cellular components. **C** Biological processes. Epigenome-wide analysis for the prenatal environment E1 was performed and methylation probes with associations at  $P < 0.001$  were selected. Probes were mapped to genes that were used in enrichment analyses. The analyses mainly show pathways related to neuronal function

associations [39, 40]. Our findings suggest that part of the discrepancy could be due to the interaction with sex.

Concerning obesity and cotinine levels in the blood of pregnant mothers, previous studies have shown a 50% increase in childhood overweight for smoking during pregnancy [41], with a dose–response relationship [42]. Cotinine levels have also been associated with low birth weight but rapid gains in BMI after delivery [43]. In a Japanese population, Susuki et al. observed that boys of mothers who smoked during pregnancy had higher gains in BMI trajectories compared with girls [44]. We found, however, higher obesity frequency for girls of mothers with smoker's cotinine levels. In a large study of ~90,000 mother-children pairs, also in Japan [43], they observed that rapid gains in BMI of children were associated with urinary cotinine concentration of mothers but not with self-reported smoking status. While their results were not stratified by sex, it shows that cotinine is a more accurate assessment of pregnancy smoking.

In relation to green spaces, systematic reviews have shown weak evidence for its relationship with children's obesity [45, 46]. Associations of green spaces during pregnancy and their differential effect on sex have not been previously assessed. We found that prenatal green space is a risk factor for boys' obesity only. A recent study of the HELIX cohort showed significant associations between children's overweight and obese status with the built environment (land use mix) [47]. Children living in built environments in absence of green spaces could be at higher risk of obesity (likely due to its relationship with physical activity). However, we observed that a low abundance of facility richness and the presence of green spaces during pregnancy are risk factors for obesity in boys. Both environmental conditions of the pregnant mother are consistent with less urbanized environments where adult obesity may be more frequent [48].

In this study, we observed that the combination of the specific levels for the four exposures maximizes the differences in obesity risk between girls and boys. Previous studies have already suggested that better prediction of an outcome can be obtained from the aggregation of multiple environmental factors into risk scores [49, 50] or the use of mixture models [51]. In line with this, we used causal inference for classifying the individuals in two environments (E0 and E1) based on the combination of the four exposures.

After classifying individuals in the two environments, we further investigated whether the individuals belonging to the environment with higher sexual dimorphism in obesity presented also sexual dimorphism in neurodevelopmental delay. Based on previous studies, prenatal factors, such as maternal obesity, have been seen associated with both obesity in children and lower cognitive

abilities and ADHD [52, 53]. Animal studies have shown that mice whose mothers were on high-fat diets during pregnancy have alterations in brain methylation of dopaminergic and opioid genes [54, 55]. In addition, the neurodevelopmental delay appears to be more frequent in obese boys [12]. A longitudinal prospective study has shown that working memory and attention performance are reduced by increasing BMI in children [56]. Our study offers additional evidence of this relationship, since the environmental changes that modulate the association between sex and obesity also modulate the association between sex and neurodevelopmental delay. Furthermore, the environment is associated with methylation probes that are enriched in neurodevelopmental pathways, providing more evidence for this hypothesis.

The generalizability of these results is subject to certain limitations. First, we did not observe a significant sexual dimorphism in obesity as expected. In the HELIX population, 4.9% of girls were obese contrasting with 6.8% of boys. Since we found that there is a clear difference, we suspect that it was not significant due to the low number of children with obesity (39 boys and 23 girls). By contrast, we found significant differences between the sexes for the ADHD diagnosis and the ANT distribution. In this case, the number of children affected was higher than in obesity (77 boys and 27 girls for ADHD and 134 boys and 72 girls for ANT). Second, the interactions between prenatal exposures and sex in obesity were significant at a nominal level but not after correcting by multiple comparisons. Again, this could be because of the small sample size and the low statistical power, which is especially important when evaluating interactions.

Our study also had notable strengths. We confirmed that the combination of exposures greatly increased the significance of interactions between prenatal exposures and sex. We also confirmed the importance of the relationship between the obesogenic prenatal environment and neurodevelopment. We not only found a significant sexual dimorphism in neurodevelopment delay when comparing E1 and E0, but we also found enrichment in neurodevelopmental pathways in the methylation probes associated with the environment. This provides a possible molecular mechanism that could explain the association between the obesogenic environment and sexual dimorphism in neurodevelopment. Moreover, this study evaluates the different effects of a prenatal environment in girls and boys, which is very innovative and important to consider sex differences in the prenatal exposure guidelines.

## Conclusions

We aimed to advance a novel approach to the study of sexual dimorphism, based on high dimensional exposure data and recent methods of causal inference. The

methodological approach can also be used to determine the environmental landscape that promotes sexual dimorphisms in studies with high dimensional exposure data.

In summary, girls in childhood may be protected against obesity if their pregnant mothers had moderate dairy consumption, non-smokers cotinine levels, and lived in environments with a low abundance of rich facilities and the presence of green spaces. The environment is also protective against the neurodevelopmental delay of non-verbal intelligence and working memory. While female protection is measured against male risk, female protection outweighs the risk of obesity in boys. Our study motivates further public health efforts to raise public awareness of moderating a high-fat diet and avoiding smoking and second-hand smoking during pregnancy to protect children against obesity and neurodevelopmental delay.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-023-02815-9>.

**Additional file 1.** Supplementary Methods.

**Additional file 2: Table S1.** Exposure variables in the HELIX cohort.

**Table S2.** CpG sites with a nominal *P*-value lower than 0.001 in the epigenome-wide association study for the classification of children in the prenatal environment (E0 and E1).

**Additional file 3.** Supplementary Code.

## Acknowledgements

We are grateful to all participants and researchers who took part in this study.

## Authors' contributions

AC1 and JRG conceived the study. AC1 designed the analysis. AC1 and NC-G performed statistical analyses and wrote the first draft of the manuscript. JRG supervised the study and contributed to reviewing and editing the manuscript. MV coordinates the HELIX project, JU is the data manager, and LM is the scientific coordinator. MB, JW, RSI, MC, JRG, and MV designed the omics study in HELIX. The following authors participated in omics data acquisition and quality control: MB (DNA methylation), AC2 (DNA methylation), MJN (exposome), and CT (exposome). JW, RG, MV, RSI, LC, and CT are the PIs of the cohorts. TY, SA, MC, JL, NS, and KG participated in sample and data acquisition. AC1 and NC-G wrote the original draft of the paper and JRG, MC, KG, VBD, RS, HW, and TLM contributed to reviewing and editing the manuscript. All authors read and approved the final version of the manuscript.

## Funding

The study has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement no 308333 (HELIX project); and the H2020-EU.3.1.2.—Preventing Disease Programme under grant agreement no 874583 (ATHLETE project). BiB received core infrastructure funding from the Wellcome Trust (WT101597MA) and a joint grant from the UK Medical Research Council (MRC) and Economic and Social Science Research Council (ESRC) (MR/N024397/1). INMA-SAB data collections were supported by grants from the Instituto de Salud Carlos III, CIBERESP, and the Generalitat de Catalunya-CIRIT. KANC was funded by the grant of the Lithuanian Agency for Science Innovation and Technology (6–04-2014\_31V-66). The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. The Rhea project was financially supported by European projects (EU FP6-2003-Food-3-NewGeneris,

EU FP6. STREP Hiwate, EU FP7 ENV.2007.1.2.2.2. Project No 211250 Escape, EU FP7-2008-ENV-1.2.1.4 Envirogenomarkers, EU FP7-HEALTH-2009- single stage CHICOS, EU FP7 ENV.2008.1.2.1.6. Proposal No 226285 ENRIECO, EU- FP7-HEALTH-2012 Proposal No 308333 HELIX), and the Greek Ministry of Health (Program of Prevention of obesity and neurodevelopmental disorders in preschool children, in Heraklion district, Crete, Greece: 2011–2014; "Rhea Plus": Primary Prevention Program of Environmental Risk Factors for Reproductive Health, and Child Health: 2012–15).

This research has received funding from the Spanish Ministry of Science and Innovation through the "Centro de Excelencia Severo Ochoa 2019–2023 (CEX2018-000,806-S) program, and support from the Generalitat de Catalunya through the CERCA Program. NC and JU are supported by Spanish regional program PERIS (Ref.: SLT017/20/000061 and SLT017/20/000119, respectively), granted by Departament de Salut de la Generalitat de Catalunya. TruDiagnos- tics also provided funding for data analysis.

## Availability of data and materials

Any custom code or software used in our analysis is available at Additional file 3: Supplementary Code.

The HELIX data warehouse has been established as an accessible resource for collaborative research involving researchers external to the project. Access to HELIX data is based on approval by the HELIX Project Executive Committee and by the individual cohorts. Further details on the content of the data warehouse (data catalog) and procedures for external access are described on the project website (<http://www.projecthelix.eu/index.php/es/data-inventory>). The data used in this analysis are not available for replication because specific approvals from the HELIX Project Executive Committee and the University of Southern California Institutional Review Board must be obtained to access them.

## Declarations

### Ethics approval and consent to participate

Local ethical committees approved the studies that were conducted according to the guidelines laid down in the Declaration of Helsinki. The ethical committees for each cohort were the following: BiB: Bradford Teaching Hospitals NHS Foundation Trust, EDEN: Agence nationale de sécurité du médicament et des produits de santé, INMA: Comité Ético de Investigación Clínica Parc de Salut MAR, KANC: LIETUVOS BIOETIKOS KOMITETAS, MoBa: Regional komité for medisinsk og helsefaglig forskningsetikk, Rhea: Ethical committee of the general university hospital of Heraklion, Crete. Informed consent was obtained from a parent and/or legal guardian of all participants in the study.

### Consent for publication

Not applicable.

### Competing interests

VBD, RS, HW, and TLM are employees of TruDiagnostic. JRG has received funding from TruDiagnostic as a scientific advisor. The other authors have nothing to declare.

### Author details

<sup>1</sup>Instituto de Salud Global de Barcelona (ISGlobal), 08003 Barcelona, Spain. <sup>2</sup>Centro de Investigación Biomédica en Red en Epidemiología Y Salud Pública (CIBERESP), Madrid, Spain. <sup>3</sup>Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Universitat Politècnica de Catalunya, 08019 Barcelona, Spain. <sup>4</sup>Department of Environmental Science, Vytautas Magnus University, 44248 Kaunas, Lithuania. <sup>5</sup>Department of Health and Experimental Sciences, Universitat Pompeu Fabra (UPF), Barcelona, Spain. <sup>6</sup>Center for Genomic Regulation (CRG), Barcelona, Institute of Science and Technology (BIST), Barcelona, Spain. <sup>7</sup>Medicine Genomics Group, Centro de Investigación Biomédica en Red Enfermedades Raras (CIBERER), CIMUS, University of Santiago de Compostela, Santiago de Compostela, Spain. <sup>8</sup>Galician Foundation of Genomic Medicine, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Servicio Gallego de Salud (SERGAS), Galicia, Santiago de Compostela, Spain. <sup>9</sup>Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, USA. <sup>10</sup>TruDiagnostic, Lexington, KY, USA. <sup>11</sup>Division of Climate and Environmental Health, Norwegian Institute of Public Health, 0456 Oslo, Norway. <sup>12</sup>Institut National de La Santé Et de La Recherche Médicale (Inserm) and Université Grenoble-Alpes, Institute for Advanced

Biosciences (IAB), Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health, Grenoble, France. <sup>13</sup>Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK.

<sup>14</sup>Department of Mathematics, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain.

Received: 2 November 2022 Accepted: 6 March 2023

Published online: 12 April 2023

## References

- Muenchhoff M, Goulder PJR. Sex differences in pediatric infectious diseases. *J Infect Dis*. 2014;209(Suppl 3):S120.
- De Bellis MD, Keshavan MS, Beers SR, Hall J, Frustaci K, Masalehdan A, et al. Sex differences in brain maturation during childhood and adolescence. *Cereb Cortex*. 2001;11:552–7.
- Shah B, Tombeau Cost K, Fuller A, Birken CS, Anderson LN. Sex and gender differences in childhood obesity: contributing to the research agenda. *BMJ Nutr Prev Heal*. 2020;3:387–90.
- Lobstein T, Brinsden H. Atlas of childhood obesity. *World Obes Fed*. 2019;211. [https://s3-eu-west-1.amazonaws.com/wof-files/11996\\_Childhood\\_Obesity\\_Atlas\\_Report\\_ART\\_V2.pdf](https://s3-eu-west-1.amazonaws.com/wof-files/11996_Childhood_Obesity_Atlas_Report_ART_V2.pdf).
- Stachenfeld NS, Mazure CM. Precision medicine requires understanding how both sex and gender influence health. *Cell*. 2022;185:1619–22.
- Vrijheid M, Slama R, Robinson O, Chatzi L, Coen M, van den Hazel P, et al. The human early-life exposome (HELIX): project rationale and design. *Environ Health Perspect*. 2014;122:535–44.
- Zhang P, Carlsten C, Chaleckis R, Hanhineva K, Huang M, Isobe T, et al. Defining the scope of exposome studies and research needs from a multidisciplinary perspective. *Environ Sci Technol Lett*. 2021;8:839–52.
- Mauvais-Jarvis F, BaireyMerz N, Barnes PJ, Brinton RD, Carrero JJ, DeMeo DL, et al. Sex and gender: modifiers of health, disease, and medicine. *Lancet (London, England)*. 2020;396:565.
- Sadovsky Y, Mesiano S, Burton GJ, Lampl M, Murray JC, Freathy RM, et al. Advancing human health in the decade ahead: pregnancy as a key window for discovery: a Burroughs Wellcome Fund Pregnancy Think Tank. *Am J Obstet Gynecol*. 2020;223:312–21.
- Cáceres A, González JR. teff: estimation of Treatment Effects on transcriptomic data using causal random forest. *Bioinformatics*. 2022. <https://doi.org/10.1093/BIOINFORMATICS/BTAC269>.
- Miller AL, Lee HJ, Lumeng JC. Obesity-associated biomarkers and executive function in children. *Pediatr Res*. 2015;77:143–7.
- Wentz E, Björk A, Dahlgren J. Neurodevelopmental disorders are highly over-represented in children with obesity: a cross-sectional study. *Obesity (Silver Spring)*. 2017;25:178–84.
- Perera F, Herbstman J. Prenatal environmental exposures, epigenetics, and disease. *Reprod Toxicol*. 2011;31:363–73.
- Wright J, Small N, Raynor P, Tuffnell D, Bhopal R, Cameron N, et al. Cohort profile: the born in Bradford multi-ethnic family cohort study. *Int J Epidemiol*. 2013;42:978–91.
- Heude B, Forhan A, Slama R, Douhaud L, Bedel S, Saurel-Cubizolles MJ, et al. Cohort profile: the EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *Int J Epidemiol*. 2016;45:353–63.
- Guxens M, Ballester F, Espada M, Fernández MF, Grimalt JO, Ibarluzea J, et al. Cohort profile: the INMA-Infancia y Medio Ambiente-(environment and childhood) project. *Int J Epidemiol*. 2012;41:930–40.
- Grazuleviciene R, Danileviciute A, Dedele A, Vencloviene J, Andrusaityte S, Uždanaviciute I, et al. Surrounding greenness, proximity to city parks and pregnancy outcomes in Kaunas cohort study. *Int J Hyg Environ Health*. 2015;218:358–65.
- Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, et al. Cohort profile update: the Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol*. 2016;45:382–8.
- Chatzi L, Leventakou V, Vafeiadi M, Koutra K, Roumeliotaki T, Chalkiadaki G, et al. Cohort profile: the Mother-Child Cohort in Crete, Greece (Rhea Study). *Int J Epidemiol*. 2017;46:1392–3.
- Maitre L, De Bont J, Casas M, Robinson O, Aasvang GM, Agier L, et al. Human Early Life Exposome (HELIX) study: a European population-based exposome cohort. *BMJ Open*. 2018;8: e021311.
- De Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ*. 2007;85:660–7.
- Vuontela V, Steenari MR, Carlson S, Koivisto J, Fjällberg M, Aronen ET. Audiospatial and visuospatial working memory in 6–13 year old school children. *Learn Mem*. 2003;10:74–81.
- Rueda MR, Fan J, McCandliss BD, Halparin JD, Gruber DB, Lercari LP, et al. Development of attentional networks in childhood. *Neuropsychologia*. 2004;42:1029–40.
- Raven JC, Raven J. Progressive matrices couleur/colored progressive matrices. Paris: Centre de Psychologie Appliquée; 1998.
- Maitre L, Guimbaud J-B, Warembourg C, Güil-Oumrait N, Petrone PM, Chadeau-Hyam M, et al. State-of-the-art methods for exposure-health studies: Results from the exposome data challenge event. *Environ Int*. 2022;168: 107422.
- Wild CP. Complementing the genome with an “Exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev*. 2005;14:1847–50.
- Haug LS, Sakhi AK, Cequier E, Casas M, Maitre L, Basagana X, et al. In-utero and childhood chemical exposome in six European mother-child cohorts. *Environ Int*. 2018;121(Pt 1):751–63.
- Serra-Majem L, Ribas L, Ngo J, Ortega RM, García A, Pérez-Rodrigo C, et al. Food, youth and the Mediterranean diet in Spain. Development of KIDMED, Mediterranean Diet Quality Index in children and adolescents. *Public Health Nutr*. 2004;7:931–5.
- Boyce W, Torsheim T, Currie C, Zambon A. The family affluence scale as a measure of national wealth: validation of an adolescent self-report measure. *Soc Indic Res*. 2006;78:473–87.
- Kritsotakis G, Koutis AD, Alegakis AK, Philalithis AE. Development of the Social Capital Questionnaire in Greece. *Res Nurs Health*. 2008;31:217–25.
- White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30:377–99.
- van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45:1–67.
- Tamayo-Uria I, Maitre L, Thomsen C, Nieuwenhuijsen MJ, Chatzi L, Siroux V, et al. The early-life exposome: description and patterns in six European countries. *Environ Int*. 2019;123:189–200.
- Carreras-Gallo N, Cáceres A, Balagué-Dobón L, Ruiz-Arenas C, Andrusaityte S, Carracedo Á, et al. The early-life exposome modulates the effect of polymorphic inversions on DNA methylation. *Commun Biol*. 2022;5:1–13.
- Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinforma*. 2012;13:1–16.
- Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlén S-E, Greco D, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One*. 2012;7:e41361.
- Vrijheid M, Fossati S, Maitre L, Márquez S, Roumeliotaki T, Agier L, et al. Early-life environmental exposures and childhood obesity: an exposome-wide approach. *Environ Health Perspect*. 2020;128:1–14.
- Voerman E, Gaillard R, Geurtsen ML, Jaddoe VVW. Maternal first-trimester cow-milk intake is positively associated with childhood general and abdominal visceral fat mass and lean mass but not with other cardio-metabolic risk factors at the age of 10 years. *J Nutr*. 2021;151:1965–75.
- Hrolfsdottir L, Rytter D, Hammer Bech B, Brink Henriksen T, Danielsen I, Steingrimsdottir L, et al. Maternal milk consumption, birth size and adult height of offspring: a prospective cohort study with 20 years of follow-up. *Eur J Clin Nutr*. 2013;67:1036–41.
- Leary SD, Ness A, Emmett P, Smith GD. Maternal diet in pregnancy and offspring height, sitting height, and leg length. *J Epidemiol Community Health*. 2005;59:467–72.
- Oken E, Levitan EB, Gillman MW. Maternal smoking during pregnancy and child overweight: systematic review and meta-analysis. *Int J Obes (Lond)*. 2008;32:201–10.
- Møller SE, Ajslev TA, Andersen CS, Dalgård C, Sørensen TIA. Risk of childhood overweight after exposure to tobacco smoking in prenatal and early postnatal life. *PLoS One*. 2014;9(10):e109184.

43. Hirai H, Okamoto S, Masuzaki H, Murata T, Ogata Y, Sato A, et al. Maternal urinary cotinine concentrations during pregnancy predict infant BMI trajectory after birth: analysis of 89617 mother-infant pairs in the Japan environment and children's study. *Front Endocrinol (Lausanne)*. 2022;13:850784.
44. Suzuki K, Kondo N, Sato M, Tanaka T, Ando D, Yamagata Z. Gender differences in the association between maternal smoking during pregnancy and childhood growth trajectories: multilevel analysis. *Int J Obes (Lond)*. 2011;35:53–9.
45. Jia P, Cao X, Yang H, Dai S, He P, Huang G, et al. Green space access in the neighbourhood and childhood obesity. *Obes Rev*. 2021;22(Suppl 1 Suppl):1.
46. Lachowycz K, Jones AP. Greenspace and obesity: a systematic review of the evidence. *Obes Rev*. 2011;12(5):e183–9.
47. de Bont J, Márquez S, Fernández-Barrés S, Warembourg C, Koch S, Persavento C, et al. Urban environment and obesity and weight-related behaviours in primary school children. *Environ Int*. 2021;155:106700.
48. Okobi OE, Ajayi OO, Okobi TJ, Anaya IC, Fasehun OO, Diala CS, et al. The burden of obesity in the rural adult population of America. *Cureus*. 2021;13(6):e15770.
49. Padmanabhan JL, Shah JL, Tandon N, Keshavan MS. The "polyenviromic risk score": aggregating environmental risk factors predicts conversion to psychosis in familial high-risk subjects. *Schizophr Res*. 2017;181:17–22.
50. Jeon EJ, Kang SH, Piao YH, Kim SW, Kim JJ, Lee BJ, et al. Development of the Korea-Polyenvironmental Risk Score for Psychosis. *Psychiatry Investig*. 2022;19:197–206.
51. Güil-Oumrait N, Cano-Sancho G, Montazeri P, Stratakis N, Warembourg C, Lopez-Espinosa MJ, et al. Prenatal exposure to mixtures of phthalates and phenols and body mass index and blood pressure in Spanish preadolescents. *Environ Int*. 2022;169: 107527.
52. Edlow AG. Maternal obesity and neurodevelopmental and psychiatric disorders in offspring. *Prenat Diagn*. 2017;37:95–110.
53. Contu L, Hawkes CA. A review of the impact of maternal obesity on the cognitive function and mental health of the offspring. *Int J Mol Sci*. 2017;18(5):1093.
54. Grissom NM, Herdt CT, Desilets J, Lidsky-Everson J, Reyes TM. Dissociable deficits of executive function caused by gestational adversity are linked to specific transcriptional changes in the prefrontal cortex. *Neuropsychopharmacology*. 2015;40:1353–63.
55. Vucetic Z, Kimmel J, Totoki K, Hollenbeck E, Reyes TM. Maternal high-fat diet alters methylation and gene expression of dopamine and opioid-related genes. *Endocrinology*. 2010;151:4756–64.
56. Li N, Yolton K, Lanphear BP, Chen A, Kalkwarf HJ, Braun JM. Impact of early-life weight status on cognitive abilities in children. *Obesity (Silver Spring)*. 2018;26:1088–95.

## Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

