Human Early Life Exposome (HELIX) study: a European population-based exposome cohort

Léa Maitre, Jeroen de Bont, Maribel Casas, Olivier Robinson, Gunn Marit Aasvang, Lydiane Agier, Sandra Andrusaityte, Ferran Ballester, Xavier Basagana, Eva Borras, et al.

BMJ Open

Human Early Life Exposome (HELIX) study: a European population-based exposome cohort

Léa Maitre,[1,2,3] Jeroen de Bont,[1,2,3] Maribel Casas,[1,2,3] Oliver Robinson,[1,2,3,4] Gunn Marit Aasvang,[5] Lydiane Agier,[6] Sandra Andrušaitytė,[7] Ferran Ballester,[3,8,9] Xavier Basagaña,[1,2,3] Eva Borràs,[2,10] Céline Brochot,[11] Mariona Bustamante,[1,2,3,10] Angel Carracedo,[12,13] Montserrat de Castro,[1,2,3] Audrius Dedele,[7] David Donaire-Gonzalez,[1,2,3] Xavier Estivill,[14,15] Jorunn Evandt,[5] Serena Fossati,[1,2,3] Lise Giorgis-Allemand,[6] Juan R Gonzalez,[1,2,3] Berit Granum,[5] Regina Grazuleviciene,[7] Kristine Bjerve Gützkow,[5] Line Småstuen Haug,[5] Carles Hernandez-Ferrer,[1,2,3] Barbara Heude,[16] Jesus Ibarluzea,[3,17,18,19] Jordi Julvez,[1,2,3,4] Marianna Karachaliou,[20] Hector C Keun,[21] Norun Hjertager Krog,[5] Chung-Ho E Lau,[21,22] Vasiliki Leventakou,[20] Sarah Lyon-Caen,[6] Cyntia Manzano,[1,2,3] Dan Mason,[23] Rosemary McEachan,[23] Helle Margrete Meltzer,[5] Inga Petraviciene,[7] Joane Quentin,[6] Theano Roumeliotaki,[20] Eduard Sabido,[2] Pierre-Jean Saulnier,[24] Alexandros P Siskos,[21] Valérie Siroux,[6] Jordi Sunyer,[1,2,3,4] Ibon Tamayo,[1,3,25] Jose Urquiza,[1,2,3] Marina Vafeiadi,[20] Diana van Gent,[1,2,3] Marta Vives-Usano,[1,2,3,10] Dagmar Waiblinger,[23] Charline Warembourg,[1,2,3] Leda Chatzi,[26,27] Muireann Coen,[22] Peter van den Hazel,[28] Mark J Nieuwenhuijsen,[1,2,3] Rémy Slama,[6] Cathrine Thomsen,[5] John Wright,[23] Martine Vrijheid[1,2,3]

To cite: Maitre L, de Bont J, Casas M, *et al*. Human Early Life Exposome (HELIX) study: a European population-based exposome cohort. *BMJ Open* 2018;**8**:e021311. doi:10.1136/bmjopen-2017-021311

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-021311).

Received 9 January 2018
Revised 2 July 2018
Accepted 17 July 2018

For numbered affiliations see end of article.

Correspondence to
Professor Martine Vrijheid;
martine.vrijheid@isglobal.org

ABSTRACT

Purpose Essential to exposome research is the collection of data on many environmental exposures from different domains in the same subjects. The aim of the Human Early Life Exposome (HELIX) study was to measure and describe multiple environmental exposures during early life (pregnancy and childhood) in a prospective cohort and associate these exposures with molecular omics signatures and child health outcomes. Here, we describe recruitment, measurements available and baseline data of the HELIX study populations.

Participants The HELIX study represents a collaborative project across six established and ongoing longitudinal population-based birth cohort studies in six European countries (France, Greece, Lithuania, Norway, Spain and the UK). HELIX used a multilevel study design with the entire study population totalling 31 472 mother-child pairs, recruited during pregnancy, in the six existing cohorts (first level); a subcohort of 1301 mother-child pairs where biomarkers, omics signatures and child health outcomes were measured at age 6–11 years (second level) and repeat-sampling panel studies with around 150 children and 150 pregnant women aimed at collecting personal exposure data (third level).

Findings to date Cohort data include urban environment, hazardous substances and lifestyle-related exposures for women during pregnancy and their offspring from birth until 6–11 years. Common, standardised protocols were used to collect biological samples, measure exposure biomarkers and omics signatures and assess child health across the six cohorts. Baseline data of the cohort show substantial variation in health outcomes and determinants between the six countries, for example, in family affluence levels, tobacco smoking, physical activity, dietary habits and prevalence of childhood obesity, asthma, allergies and attention deficit hyperactivity disorder.

Future plans HELIX study results will inform on the early life exposome and its association with molecular omics signatures and child health outcomes. Cohort data are accessible for future research involving researchers external to the project.

INTRODUCTION

The 'exposome' concept encompasses the totality of non-genetic exposures from conception throughout the life course, complementing the genome.[1] The exposome concept carries the expectation that the use of holistic and data-driven approaches, similar to those pioneered in the genomics fields, can result in advances in our understanding of the complex environmental component of disease aetiology. The exposome has been delineated to include three overlapping and complementary domains[2]: (1) a general external domain including macrolevel factors such as climate, urban environment and societal factors; (2) an individual external domain

including agents such as environmental pollutants, tobacco smoke, diet and physical activity and (3) a specific internal domain including gene expression, inflammation and metabolism, often assessed through high-throughput molecular omics methodologies such as transcriptomics, proteomics and metabolomics.

The HELIX project aims to measure and describe multiple environmental exposures from the different exposome domains during early life (pregnancy and childhood) and associate these with omics markers and child health outcomes. The background, rationale and detailed objectives of the HELIX project have been described elsewhere at the start of the project.[3] HELIX takes early life as a key starting point for defining the exposome because, as recognised in the Developmental Origins of Health and Disease research, it is well recognised that the periods of organ development during prenatal life and infancy are especially vulnerable to the effects of environmental risk factors, which may manifest themselves throughout the lifetime in adult diseases.[4] Essential to exposome research is the collection of data on many environmental exposures from the different exposome domains in the same subjects. Here, we describe recruitment, study population, measurements available and baseline data of the HELIX nested study populations, with an aim to provide a detailed description of the cohort and of the data available for future collaborative research.

COHORTS PARTICIPATING IN HELIX

The HELIX study represents a collaborative project across six established and ongoing longitudinal population-based birth cohort studies in Europe: the Born in Bradford (BiB) study in the UK,[5] the Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant (EDEN) study in France,[6] the INfancia y Medio Ambiente (INMA) cohort in Spain,[7] the Kaunus cohort (KANC) in Lithuania,[8] the Norwegian Mother and Child Cohort Study (MoBa)[9] and the RHEA Mother Child Cohort study in Crete, Greece[10] (table 1). These cohorts were selected for participation in the HELIX project because: (a) they could provide substantial existing longitudinal data from early pregnancy through childhood, (b) they could follow-up children at similar ages, (c) they could integrate questionnaires,

Table 1 Characteristics of the cohorts contributing to the HELIX cohort

Cohort	Recruitment in original cohort	Exclusions made during recruitment	Years of birth	Region covered by HELIX	No. of births in HELIX entire cohort
BiB, UK[5]	All pregnant women who attended the oral glucose tolerance test clinic at Bradford Royal Infirmary in weeks 26–28 of pregnancy.	Women who planned to move away from Bradford before birth were excluded.	2007–2010	Bradford	10849
EDEN, France[6]	Pregnant women who attended prenatal care at the University hospitals of Nancy and Poitiers recruited before 24 weeks of amenorrhoea.	Twin pregnancies, women with known diabetes before pregnancy, insufficient French language skills and intention to move away from the recruitment area were excluded.	2003–2006	Nancy and Poitiers, urban areas	1900
INMA, Spain[7]	Pregnant women who attended a prenatal care centre in the study region during weeks 6–10 of pregnancy.	Women who resided or intended to deliver outside the study area, who were aged under 16 years, who had twin or multiple pregnancies, who had assisted reproduction or who had communication problems were excluded.	2003–2008	Gipuzkoa Sabadell Valencia	2063
KANC, Lithuania[8]	Pregnant women who attended one of four prenatal care clinics affiliated to the hospitals of the Kaunas University of Medicine during first trimester of pregnancy.	Women who lived outside Kaunas municipality, had medical records of pregnancy induced hypertension and/or diabetes were excluded.	2007–2008	Kaunas	4107
MoBa, Norway[9]	Recruitment at the first ultrasound (US) scan, ie, during the 17–18 weeks of gestation. All women who gave singleton births in the participating maternity units.	None	1999–2008	Oslo	11095
RHEA, Greece[10]	Pregnant women who attended US examination before 15 week of pregnancy with residence in and near Heraklion at Crete.	Women who were aged under 16 years or who had communication problems were excluded.	2007–2008	Heraklion	1458
Total					31472

BiB, Born in Bradford; EDEN, Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant; INMA, INfancia y Medio Ambiente; KANC, Kaunus cohort; MoBa, Norwegian Mother and Child Cohort Study.

biosampling and clinical examinations using common HELIX protocols and (d) they offered heterogeneity in terms of exposure and population characteristics.

Pregnant women in the original cohorts were recruited between 1999 and 2010. Three cohorts (INMA, KANC, RHEA) recruited during the first trimester of pregnancy, two through the first and second trimesters (EDEN, MoBa), while in the BiB cohort women were recruited between weeks 26 and 28 of gestation (second/third trimesters). Inclusion and exclusion criteria varied between cohorts, as described in table 1. All cohorts included at least one follow-up point during pregnancy, one at birth and several after delivery.

Based on these six existing cohorts, HELIX used a multilevel study design, drawing on nested study populations for data collection of different intensities (figure 1): (1) the *entire cohort* in which factors arising primarily from outdoor exposures were assessed through geospatial models and linked to existing health outcome data; (2) a *subcohort* in which one new follow-up examination of the children between ages 6 and 11 years was carried out in order to assess child health outcomes and to fully characterise different areas of the exposome through questionnaires, biological sample collection and biomarker and omics measurements and (3) two *panel studies* in children and pregnant women to characterise in depth the variability in exposure biomarkers and omics biomarkers, individual exposure-related behaviours and personal exposures.

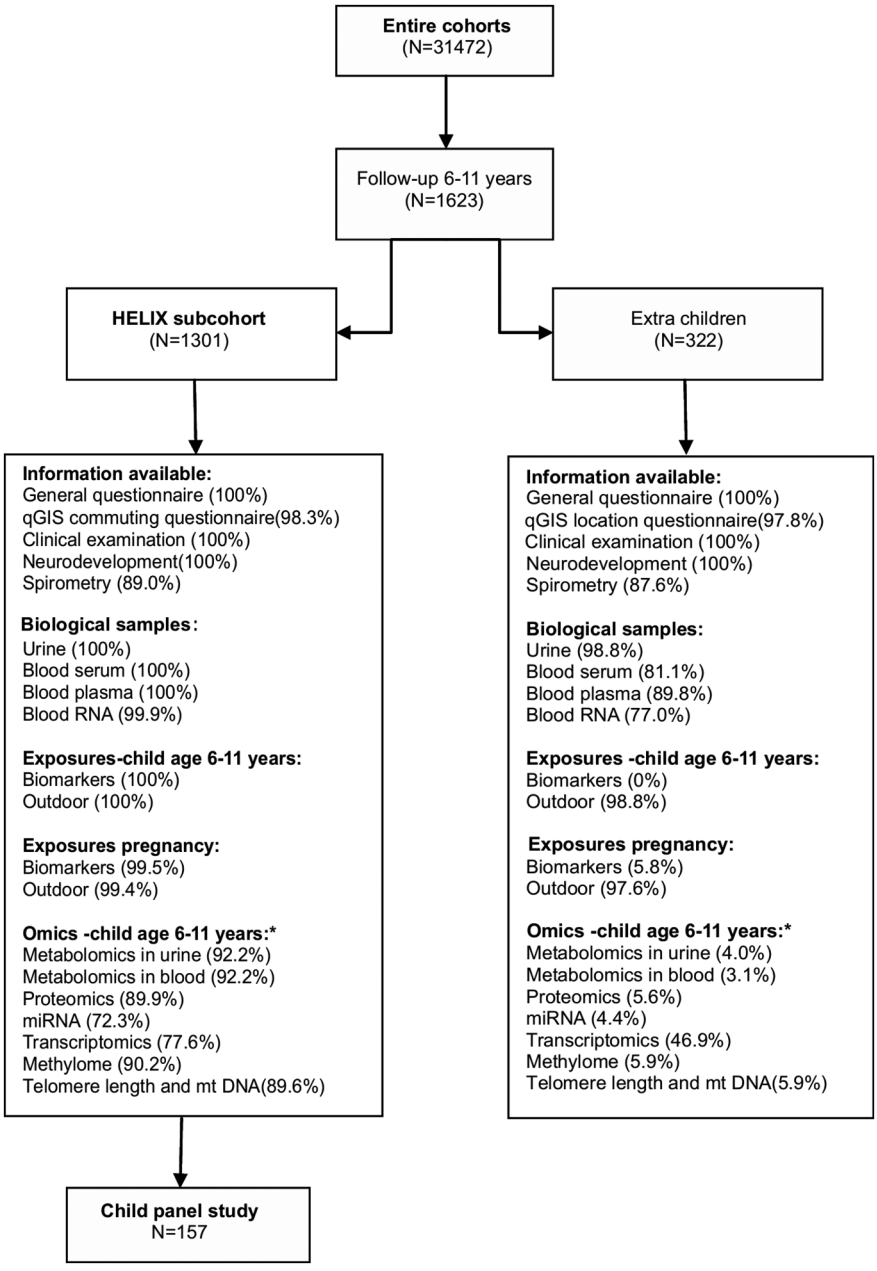

Figure 1 Flow chart describing design and available data. GIS, Geographic Intelligent Software; HELIX, Human Early Life Exposome; miRNA, microRNA; mtDNA, mitochondrial DNA. *Omics data available after quality control

HELIX ENTIRE COHORT

The study population for the entire HELIX cohort includes 31 472 women who had singleton deliveries between 1999 and 2010, and for whom exposure to ambient air pollution during pregnancy had been estimated as part of the European Study of Cohorts for Air Pollution Effects (ESCAPE) project.[11] The entire cohort includes nine regions from the six cohorts; we included only regions where geographic data were available to calculate air pollution levels and built environment indicators (table 1). This meant, for example, that the city of Oslo and not the whole of the national MoBa cohort was included, and that only the Gipuzkoa, Sabadell and Valencia regions of the INMA study were included. In the other cohorts, women residing outside the main urban areas were excluded for the same reason.

In this study population, data on many variables had been collected in the individual cohorts during previous data collection points (during pregnancy and between birth and 5 years of age). Existing data included information on certain exposures (eg, maternal tobacco smoking during pregnancy, environmental tobacco smoke), key covariates (eg, pregnancy complications, maternal and child diet, maternal and child physical activity, child sleep, breast feeding, other health-related behaviours, indicators of socioeconomic status) and health and development outcomes. As part of HELIX, relevant datasets from all 31 472 mother-child pairs were transferred from the six cohorts to the central HELIX data warehouse located at the Barcelona Institute for Global Health (ISGlobal) (see below). Through data harmonisation, these cohort-specific variables were converted to harmonised variables. This process involved summarising, checking and matching the specific variable cohort-by-cohort and deciding a common coding system appropriate to each variable. Specific expert working groups throughout the HELIX consortium advised on the harmonisation rules for each variable. The child health and developmental outcomes harmonised as part of HELIX include birth outcomes, growth-related and obesity-related outcomes, blood pressure, neurodevelopment and respiratory health between birth and 5 years of age (table 2).

HELIX SUBCOHORT

From the entire cohort, a subcohort of mother-child pairs was selected to be fully characterised for a broad

Table 2 Health outcomes harmonised across the entire cohort between birth and 5 years of age

Health/development outcomes	Methods	BiB	EDEN	INMA	KANC	MoBa	RHEA	Total number of subjects in the harmonised dataset
Birth								
Birth weight	Measurements	√	√	√	√	√	√	31 472
Gestational duration	Medical records/ ultrasound	√	√	√	√	√	√	31 472
0–5 years								
Repeated weight, height, BMI	Measurements and records	√	√	√	√	√	√	28 305
Waist circumference	Measurements							
1–2 years		√	√	√			√	4598
4–5 years		√	√	√			√	4275
Skinfolds	Measurements							
1–2 years		√	√				√	3364
4–5 years		√	√				√	2774
Blood pressure (4–5 years)	Measurements	√	√	√			√	5182
Cognition	Psychologist-administered tests and parental questionnaires		√	√			√	3470
Motor skills, language	Psychologist-administered tests and parental questionnaires	√	√	√		√	√	10 245
Behaviour	Questionnaires	√	√	√	√	√	√	12 644
Asthma, wheeze	Questionnaires	√	√	√	√	√	√	12 068
Lung function (4–5 years)	Spirometry	√	√	√			√	2719

BiB, Born in Bradford; BMI, body mass index; EDEN, Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant; INMA, INfancia y Medio Ambiente; KANC, Kaunus cohort; MoBa, Norwegian Mother and Child Cohort Study.

suite of environmental exposures and 'omics' data, to be clinically examined and to have biological samples collected. A new follow-up visit was organised for these mother-child pairs between December 2013 and February 2016. Subcohort subjects were recruited from within the entire cohorts such that there were approximately 200 mother-child pairs from each of the six cohorts. Subcohort recruitment in the EDEN cohort was restricted to the Poitiers area and in the INMA cohort to the city of Sabadell.

Eligibility criteria for inclusion in the subcohort were: (a) age 6–11 years at the time of the visit, with a preference for ages 7–9 years if possible; (b) sufficient stored pregnancy blood and urine samples available for analysis of prenatal exposure biomarkers; (c) complete address history available from first to last follow-up point; (d) no serious health problems that may affect the performance of the clinical testing or impact the volunteer's safety (eg, acute respiratory infection). In addition, the selection considered whether data on important covariates (diet, socioeconomic factors) were available. Each cohort selected participants at random from the eligible pool in the entire cohort and invited them to participate in this subcohort until the required number of participants was reached. In total, 1301 mother-child pairs with complete questionnaire and clinical examination data, and urine and blood samples, were included in the HELIX subcohort (figure 1).

Several cohorts then invited and examined further subjects (n=322) following the same protocols for clinical examination and sample collection, and the same questionnaires, but these were not included in the measurement of exposure biomarkers for the HELIX study (figure 1). Among the 322 extra individuals, 266 came from INMA-Sabadell, 26 from BiB, 7 from EDEN, 3 from KANC, 19 from MoBa and 1 from RHEA. For some of these individuals omics data were collected. These individuals may be included in studies with another focus than exposure biomarkers, which is why they are shown in figure 1.

The new follow-up visits for the subcohort took place in the six study centres at a local hospital, a primary care centre or at the National Institute for Public Health (NIPH) in Oslo. During the follow-up examination, trained nurses interviewed the mothers, carried out health examinations of the children and collected biological samples using standardised operating procedures.

Questionnaire information

Interviews with the mothers during the visit used a computer-aided version of a common standardised questionnaire developed for HELIX. Questionnaires were translated and back-translated in each of the country languages. If it was not possible for the mother to attend (although the mother's attendance was greatly encouraged on recruitment), then the father or legal guardian completed the questionnaire and the mother checked it at a later date at home. This happened for 4% of the children. The full questionnaire can be accessed online (https://tinyurl.com/yat9hao4).

The questionnaire collected information on child's diet including an internationally agreed food frequency questionnaire with portion size examples appropriate to each cohort and the reliability of the Mediterranean diet quality index questionnaires to assess Mediterranean diet,[12] physical activity of the child, sleeping patterns of child, socioeconomic status (family affluence scale (FAS II),[13] ie, subjective wealth), social capital of the family,[14] stress of the mother,[15] exposure to environmental tobacco smoke, water consumption habits, cooking and heating methods at the home, cleaning products, bedroom location, noise perception, child's use of mobile phones and other electronic devices, use of green spaces, commuting behaviour, holidays and sun exposure and puberty development of the child.

Questions on exposures during the day and week before the visit were asked separately in a short questionnaire and repeated during the second period of nested child panel study (see below). Additionally, questions related to addresses, places visited and travel were collected using a custom-made Google maps-aided commuting questionnaire based on the free Geographic Intelligent Software (GIS), qGIS software[16] that allowed mothers to trace their child's commuting routes directly on the computer. This information was used to enhance the accuracy of location data collected through the main questionnaires and integrated with the outdoor exposure estimates to provide exposure estimates at different locations and for different time-activity patterns (eg, home, school, commuting exposures). A total of 98.3% of mothers in the subcohort completed the qGIS commuting questionnaire.

Anthropometry and body composition

During the subcohort follow-up examination, anthropometric data were collected using regularly calibrated instruments: height was measured with a stadiometer and weight with a digital weight scale, both without shoes and with light clothing. Height and weight measurements were converted to body mass index (BMI in kg/m^2) for age-and-sex z-scores using the international WHO reference curves in order to allow comparison with other studies.[17] Overweight and obese children were defined as those above the age-and-sex-specific 85th and 95th percentiles, respectively, as recommended by WHO (http://www.who.int/mediacentre/factsheets/fs311/en/). Circumferences (arm, waist, head) were measured with a metric tape and recorded in duplicate.

Four skinfolds were measured (triceps, subscapular, suprailiac, thigh), following the protocols as described in the report from National Health and Nutrition Examination Survey III Body Measurements (anthropometry).[18] Three complete sets of each skinfold measurement were taken consecutively, and the mean

was used as the representative value for each site. A skinfold is the thickness of a double fold of skin and subcutaneous fat, excluding the underlying muscle. Skinfolds are highly correlated with total body fat and a way to assess the distribution of fat tissue. Specific training workshops (one before and one during field work) were organised to standardise skinfold and other anthropometric measurements between the cohorts. In these workshops, all field workers participated and were trained to obtain measurements that were comparable to those measured by an expert anthropometrist.

Bioelectric impedance analyses readings were performed with the Bodystat 1500 (Bodystat, Douglas, Isle of Man) equipment after 5 min of lying down. Bioelectric impedance provides an objective measure of body composition when standard protocols are followed and population-specific equations are available and used. Fat free mass and fat mass (in grams and as proportion of total body mass) were calculated based on values of impedance, using published age-specific and race-specific equations validated for use in children.[19] The multiracial equations developed recently for children based on impedance values obtained by a single frequency tetra-polar Bodystat device[19] fit well the measures obtained from children in HELIX using the same device.

Blood pressure

Blood pressure was taken in sitting position after 5 min of rest using the OMRON 705-CPII automated oscillometric device. The mean of three consecutive measurements that were taken with 1 min intervals was used.[20] Blood pressure was measured towards the end of the visit to ensure that children had not consumed anything that may affect the results (chocolate, cola drinks) in the previous hour. Systolic and diastolic blood pressures and pulse rate from each measurement were recorded.

Respiratory health

Lung function measures were obtained through forced spirometry test, using the EasyOne spirometer in children by trained field workers using a standardised protocol. During the measurements, the child was sitting straight and equipped with a nose clip and asked to breathe in as deeply possible until his/her lungs were totally full, and then to quickly position the mouthpiece and blast out the air as hard and as fast as possible. The child was asked to perform at least six of these manoeuvres to achieve the three acceptable and reproducible manoeuvres needed for a valid test. While acceptability and reproducibility criteria for spirometry have been well defined for adults, the criteria to be used for children lack clarity and consistency.[21–23] Based on international standards,[21 24] a manoeuvre was defined as acceptable if there was no hesitation of false starts (ie, if the back-extrapolated volume (BEV) was <5% of the forced vital capacity (FVC) or if BEV <100 mL if FVC <1000 mL) and if the forced expiratory time (FET) was in an acceptable range (1.5 s<FET<10 s). The highest values for forced expiratory volume at one second (FEV_1) taken from acceptable forced expiratory manoeuvres should not vary >150 mL or 5% (or <100 mL if FVC <1000 mL) from the second highest FEV_1. Then, the per cent predicted values FEV_1 were computed using the Global Lung Initiative equations, and any best FEV_1% predicted value in the (60%; 140%) range was retained. Following these criteria, 79.4% of the HELIX children performed a valid test.

Information on occurrence of wheeze, asthma, eczema, allergic rhinitis and food allergy in the children was obtained through questions adapted from the International Study on Asthma and Allergy in Childhood by the Mechanisms of the Development of Allergy project (MeDALL).[25]

Neurodevelopment

Neurodevelopmental outcomes were assessed through a battery of internationally standardised, non-linguistic, and culturally blind computer tests. The tests included N-back[26] and the attention network test[27] to assess working memory and sustained attention, the trail making test[28] to assess speed of processing and executive functions, the finger tapping test to assess motor speed and lateralisation[29] and Raven's coloured progressive matrices[30] to assess general non-verbal intelligence. The tests were administered through standardised study-provided laptops and lasted a maximum of 1 hour. Rooms for testing were ensured to be quiet and the tests were done with minimal interference. Field workers were trained to instruct children in a standardised way.

A proxy of maternal IQ or cognitive functioning is an important cofactor to be assessed in any study where the neurodevelopment is the main outcome. The mothers (or father if the mother was not available) completed a short version of N-back (no more than 6 min) adapted to adults.

Parents completed the Conner rating scale's and child behaviour checklist (CBCL) before the visit to assess child behavioural problems. The Conner rating scale's of 27 items provides information on child behaviour, particularly in relation to inattention and hyperactivity.[31] CBCL is one of the most recognised and extended tools to fully assess a child behavioural functioning and contains several subscales, including aggressive behaviour, anxiety/depression, attention problems, internalising problems, externalising problems, etc.[32]

Biological sample collection

During the subcohort follow-up examination, new biological samples suitable for all planned exposure biomarker and omics analyses were collected using the same standardised protocols across all six cohorts as shown in figure 1. Two spot urine samples (one before bedtime and one first morning void) were collected in high-quality polypropylene tubes (Sarstedt: 75.9922.744). The two urine samples were brought by

the participants to the centre in cool packs and stored at −4°C until processing. After aliquoting, the urine samples were frozen at −80°C under optimised and standardised procedures. If the families did not bring urine samples with them, a new sample was collected on arrival at the centre. This occurred in 6.6% of the subcohort children; 18 mL of blood was collected during the follow-up visit at the end of the clinical examination of the child to ensure an approximate 3 hour (median 3.5 hours, SD 1.1 hour) fasting time since the last meal. Blood samples were collected using a 'butterfly' vacuum clip and local anaesthetic and processed into a variety of sample matrices. It included EDTA Vacutainers designed for trace element testing, used for plasma proteomics, microRNA (miRNA) and perfluorinated alkylated substance analyses, blood smears, whole blood heavy metal and DNA isolation (BD: 368381, K2EDTA coated), tempus tubes for RNA isolation (Life Technologies Cat. No.: 4342792), plastic silica Vacutainers for serum metabolomics and clinical parameter analyses (BD: 368813 silica coated, clot activator), glass silica Vacutainer for serum polychlorinated biphenyl, dichlorodiphenyldichloroethylene, hexachlorobenzene, polybrominated diphenyl ether analyses (BD: 367614, silica coated, no activator). After processing, these samples were frozen at −80°C under optimised and standardised procedures. After performing the relevant assays, blood and urine samples, hair samples, RNA and DNA samples remain in storage for the subcohort children.

EXPOSURE ASSESSMENT

To construct the exposome, HELIX has estimated exposure to a wide range of environmental contaminant exposures and indicators of the built environment. In the entire cohort and subcohort, a GIS environment for the nine study areas was constructed, and, based on residential address histories, exposure estimates were assigned for ambient air pollutants, road traffic noise levels, surrounding (natural spaces green and blue spaces), built environment, ultraviolet (UV) radiation and meteorological variables during pregnancy and childhood (table 3). These estimates build on existing land-use regression air pollution models (ESCAPE project[33 34]), city noise maps, land use maps ('Urban Atlas' by European Environmental Protection Agency), raster maps of the normalised difference vegetation index (NDVI),[35 36] raster maps of land surface temperature, building density, population density, connectivity, walkability and public bus transport map information for the built environment and meteorological data, as described in more detail elsewhere (Robinson and colleagues[37]). Data from existing regulatory monitors were used to back extrapolate ambient air pollution exposure models. The estimates for these outdoor exposures were calculated for the prenatal period and several postnatal periods up to the HELIX subcohort follow-up time point (table 3).

Furthermore, in the subcohort, biomarkers of contaminant exposure were measured in appropriate biological samples collected from the children at age 6–11 years and in samples previously collected from mothers during pregnancy or from the neonates during delivery (cord blood) and stored in cohort biobanks (table 3). Chemical assays were conducted in the laboratory at the Department of Environmental Exposure and Epidemiology at the NIPH, apart from analyses of metals/elements and cotinine, creatinine and blood lipids, which were subcontracted to ALS Laboratory Group Norway AS and Dr Fürst Medisinsk Laboratorium AS, respectively. Biomarkers include: organochlorine compounds and brominated compounds, perfluoroalkyl substances and metals in blood, and non-persistent chemicals (phthalate metabolites, phenols, organophosphate pesticide metabolites and cotinine) in urine samples (table 3). Concentrations of OCs and PBDEs were adjusted for to total lipid percentage and expressed in ng/g of lipids. Urinary concentrations were adjusted for creatinine and expressed in µg/g of creatinine. Urine samples of the night before the visit and the first morning void on the day of the visit were combined to provide a slightly long-term exposure assessment than can be achieved with one spot urine sample (Haug *et al*, s).

Concentrations of drinking water disinfection by-products (DBPs) during pregnancy were estimated from water company concentration and distribution data as part of the water contaminants and stillbirth, congenital anomalies, birth weight, preterm delivery (HiWate) project in four of the cohorts (BiB, KANC, INMA, RHEA).[38] For EDEN and MoBa, we followed the same methodology to obtain estimates during pregnancy. Data were not sufficiently complete to estimate child exposure to DBPs. Indoor air concentrations of nitrogen dioxide (NO_2), particulate matter <2.5 µm ($PM_{2.5}$), particulate matter absorbance, benzene and toluene, ethylbenzene, xylene were estimated by combining measurements in the homes of a subgroup of children during the two periods of the nested panel studies (see below) with questionnaire data from the subcohort.

MEASUREMENT OF MOLECULAR SIGNATURES

In the subcohort, we also obtained the following measurements of molecular omics signatures at the age of 6–11 years: blood leucocyte DNA methylation (450K, Illumina), whole blood transcription (HTA V.2.0, Affymetrix and SurePrint Human miRNA rel 21, Agilent), serum metabolites (AbsoluteIDQ p180 kit, Biocrates), urine metabolites (proton nuclear magnetic resonance (^{1}H NMR) spectroscopy) and plasma proteins (Luminex, cytokines 30-plex, apoliprotein 5-plex and adipokine 15-plex). Among the samples available for omics analyses, some were excluded because of the absence of genetic analysis consent (n=1), because the

Table 3 Exposure estimates available in the HELIX entire cohort and subcohort

Exposure group	Description*	Pregnancy (and specific trimesters)*	Postnatal 0–5 years	Subcohort 6–11 years
Outdoor and urban exposure estimates available in the entire cohort and in the subcohort				
Atmospheric pollutants	NO_2, $PM_{2.5}$, PM_{10}, $PM_{2.5}$ (absorbance ratio)	√ *	√	√
Ultraviolet (UV)	Ambient UV radiation levels	√	√	√
Surrounding natural space	Average normalised difference vegetation index within buffers of 100, 300 and 500 m Size of and distance to nearest major green or blue space (>5000 m^2) Presence of a major green or blue space in a distance of 300 m	√	√	√
Meteorology	Land surface temperature average in a buffer of 50 m Temperature from meteorological stations (mean, minimum and maximum) Humidity percentage from meteorological stations Atmospheric pressure data from the ESCAPE project	√ *	√	√
Built environment	Population density: inhabitants per km^2 Building density: built area in m^2 per km^2 within 100 and 300 m buffers Street connectivity: number of road intersections per km^2 within 100 and 300 m buffers Accessibility: metres of bus public transport lines and number of bus public transport stops per km^2 within 100, 300 and 500 m buffers Facilities: facility richness index and facility density index in a 300 m buffer Land use evenness index Walkability index in 300 m buffer*	√	√	√
Traffic	Total traffic load of major roads in a 100 m buffer, total traffic load in a 100 m buffer, traffic density on nearest road and inverse distance to nearest road	√	√	√
Road traffic noise	Day and night time road noise levels	√	√	√
Contaminant exposure estimates available in the HELIX subcohort				
Organochlorine compounds	Blood concentrations of dichlorodiphenyldichloroethylene, dichlorodiphenyltrichloroethane, hexachlorobenzene and polychlorinated biphenyl—118, 68, 153, 170, 180. With and without lipid adjustment.	√	–	√
Brominated compounds	Blood concentrations of polybrominated diphenyl ether—47, 153. With and without lipid adjustment.	√	–	√
Perfluorinated alkylated substances	Blood concentrations of perfluorooctanoate, perfluorononanoate, perfluoroundecanoate, perfluorohexane sulfonate, perfluorooctane sulfonate	√	–	√
Metals and essential elements	Whole blood concentrations of arsenic, cadmium, cesium, cobalt, copper, lead, manganese, mercury, molybdenum, thallium, potassium, magnesium, sodium, selenium and zinc	√	–	√
Phthalate metabolites	Urine concentrations of monoethyl phthalate, mono-iso-butyl phthalate, mono-n-butyl phthalate, mono benzyl phthalate, mono-2-ethylhexyl phthalate, mono-2-ethyl-5-hydroxyhexyl phthalate, mono-2-ethyl-5-oxohexyl phthalate, mono-2-ethyl 5-carboxypentyl phthalate, mono-4-methyl-7-hydroxyoctyl phthalate, mono-4-methyl-7-oxooctyl phthalate. With and without creatinine adjustment.	√	–	√

Continued

Table 3 Continued

Exposure group	Description*	Pregnancy (and specific trimesters)*	Postnatal 0–5 years	Subcohort 6–11 years
Phenols	Urine concentrations of methyl paraben, ethyl paraben , bisphenol A, propyl paraben, N-butyl paraben, oxybenzone, triclosan. With and without creatinine adjustment.	√	–	√
Organophosphate pesticide metabolites	Urine concentrations of dimethyl phosphate, dimethyl thiophosphate, dimethyl dithiophosphate, diethyl phosphate, diethyl thiophosphate, diethyl dithiophosphate. With and without creatinine adjustment.	√	–	√
Tobacco smoking	Urine levels of cotinine. Questionnaire on active and passive smoking.	√	–	√
Water disinfection by-products	Total concentration of total trihalomethanes (THMs), chloroform and total brominated THMs estimated in tap water from water company concentration and distribution data.	√	–	–
Indoor air	Prediction models for indoor air concentrations of NO_2, $PM_{2.5}$, PM_{abs}, benzene and toluene, ethylbenzene, xylene using panel study data from indoor air samplers.	–	–	√

*Walkability indicator adapted from the previous walkability indexes[45 46]: calculated as the mean of the deciles of population density, street connectivity, facility richness index and land use Shannon's Evenness Index within 300 m buffers, giving a walkability score ranging from 0 to 1.
ESCAPE, European Study of Cohorts for Air Pollution Effects; HELIX, Human Early Life Exposome; NO_2, nitrogen dioxide; $PM_{2.5}$, mass concentration of particles <2.5 µm in aerodynamical diameter; PM_{10}, mass concentration of particles <10 µm in aerodynamical diameter; PM_{abs}, absorbance of $PM_{2.5}$ filters; a proxy for elemental carbon, which is the dominant light absorbing substance.

blood DNA/RNA extraction failed (n=386, 22.8%, for RNA; n=22, 1.5%, for DNA), because omics data were of low quality (n=10 for transcriptomics), because of technical outliers (serum haemolysed, n=1), or because of failed sample identity checks for the methylome (n=4 based on sex mismatch, and n=6 based on genotype mismatch between longitudinal HELIX samples from the same child or from existing genome-wide genetic data of the child). Telomere length and mitochondrial DNA content were measured by quantitative PCR as part of a separate project. Genome-wide genotyping will be completed as part of a separate project using the Infinium Global Screening Array from Illumina.

The number of omics markers varies greatly across the omics platforms: from 36 for proteomics to 480 071 for the methylome (table 4). The platforms and data processing procedures selected for the proteins and serum metabolome were in fact targeted assays (<200

Table 4 Omics features available in the HELIX subcohort and repeat child panel study

Omics	Sample	Platform	Features	Subcohort (n=1300)*	Second period child panel study
Proteomics	Plasma	Luminex kits: cytokines 30-plex, apoliprotein 5-plex and adipokine 15-plex	36	1170	154
Methylation	Buffy coat	450K, Illumina	386 518	1173	153
Transcriptomics	Whole blood	HTA V.2.0, Affymetrix	35 841	1010	127
MicroRNA (miRNA)	Whole blood	SurePrint Human miRNA rel 21, Agilent	330	941	123
Urinary metabolomics	Urine	^{1}H NMR spectroscopy	44	1198	153
Serum metabolomics	Serum	AbsoluteIDQ p180 kit, Biocrates	177	1198	154
Telomere length Mitochondrial DNA content	Buffy coat	Quantitative real-time PCR	2	1166	153

*One child's parents in the subcohort did not give genetic consent, therefore this child was excluded from all omics analyses.
HELIX, Human Early Life Exposome.

features) in order to obtain the best quality data for a large number of samples with fully annotated proteins and metabolites.[39] Further data filtering was applied to decrease the apparent complexity in the omics data. For example, in the urine metabolome, generated from untargeted NMR spectroscopic analysis from 128K spectral data points, 44 metabolite integrals were calculated only for resonances with high abundance and limited overlap with other metabolite signals. Urine metabolites were normalised using the median fold change.[40] Proteins were filtered out if 30% of samples were outside of the linear range of quantification. After an initial quality control, the number of CpGs, transcript clusters and miRNAs were: 480 071, 67 528 and 2549, respectively. Additional filtering (ie, probes in sexual chromosomes, cross-hybridisation probes, etc) might be applied during data analysis.

Within the HELIX subcohort of 1301 mother-child pairs, we obtained the following final numbers of children with omics data: n=941 for miRNA, 1010 for transcripts, 1170 for proteins, 1173 for methylation and 1198 for urine and serum metabolites; a total of 874 children (67% of the subcohort) had complete exposure and omics data (table 4, figure 1). Among these children, between 123 and 154, depending on the omic platform, also had a sample analysed for the second visit approximately 6 months later (table 4).

PANEL STUDIES

Intensive repeat panel studies collected data on short-term temporal variability in exposure biomarkers and omics biomarkers, individual behaviours (physical activity, mobility) and personal and indoor exposures

Table 5 Measurements performed in the child and pregnancy panel studies

Measurement	No. of subjects in child panel study*	No. of subjects in pregnancy panel study*	Description	Measurement point/ period
Geolocation and mobility	146	126	Smartphone GPS with ExpoApp application installed	7 days in each study period
Physical activity	145	148	Smartphone and Actigraph accelerometer	7 days in each study period
NO_2	154	158	Passive samplers for NO_2 installed in the home	7 days in each study period
BTEX	154	158	Passive samplers for BTEX installed in the home	7 days in each study period
$PM_{2.5}$	92	90	Active $PM_{2.5}$ Cyclone pumps (BGI-400–4), carried by participants in backpack and installed in the home	Last 24 hours of each of the two study periods
Black carbon	89	66	MicroAthelometer (AE51) for continuous monitoring	Last 24 hours of each of the two study periods
UV	69	141	Electronic wrist band UV dosimeters[47]	7 days in each study period
Phthalates, phenols, organophosphate pesticides	152	–	Pool of bedtime and first morning urine	4 separate days in one study period
Phthalates, phenols, organophosphate pesticides, cotinine	152	154	Pool of daily urine samples (2 or 3 per day) during 1 week	One pool in each of the two study periods
Phthalates	–	44	All morning and bed time urines during 1 week	7 days in one study period
^1H NMR metabolomics	22	–	All morning and bed time urines during 1 week	7 days in one study period
Lung function	62	–	Spirometry	Last day of period 1 and 2
Blood pressure	157	154	OMRON 705-CPII automated oscillometric device	Last day of period 1 and 2
Height and weight	157	145		Last day of period 1 and 2

*With data in both periods.
BTEX, benzene, toluene, ethylene and meta-xylene, para-xylene and ortho-xylene; ^1H NMR, proton nuclear magnetic resonance; NO_2, nitrogen dioxide; $PM_{2.5}$, mass concentration of particles <2.5 µm in aerodynamical diameter; UV, ultraviolet.

(table 5). The child panel study included children from the HELIX subcohort (n=157, from all cohorts except MoBa) who lived in a first floor apartment or private house and were sampled following a maximum variation sampling strategy to high traffic-density exposure at home address. The pregnancy panel study included pregnant women from outside the cohorts in three cities, Barcelona, Grenoble and Oslo (n=154). The inclusion criteria for these pregnant women were to be 18 years or older at the start of pregnancy, to have a singleton pregnancy, to be living in the study area until delivery and to have the first visit before the end of gestational week 20. Participants in the child panel study were followed for 1 week in two seasons, whereas in the pregnancy panel study the participants were followed for 1 week in two trimesters. In the child panel, the last day of the first week coincided with the subcohort examination, detailed above.

Participants carried smartphones for measurement of physical activity and to collect geolocalisation data through the ExpoApp, a smartphone-based application[41] specifically developed for the project (table 5). Indoor air pollution exposure to NO_2 and to volatile organic compounds benzene, toluene, ethylbenzene, meta-xylene, para-xylene and ortho-xylene, was measured through passive samplers installed in the homes. For the last 24 hours of the panel study periods, participants carried backpacks containing Active $PM_{2.5}$ Cyclone pumps and black carbon MicroAthelometer monitors (Model AE51, AethLabs, California, USA). Electronic wrist bands measured UV exposure (Scienterra, New Zealand).

Urine samples were collected twice daily (first morning void and bedtime sample) in the child panel study and three times per day (morning, afternoon, bed time) in the pregnancy panel. Urine samples were used to measure repeat biomarkers for non-persistent exposures (phthalates, phenols, organophosphate pesticides and cotinine) and they were used to assess the variance in NMR metabolomics measured in the first morning void, bedtime and pooled urine[42].

At the end of each monitoring week, blood samples were collected following the same procedures as for the subcohort, indeed the collection in week 1 was part of subcohort examination. Blood samples in the child panel study were also used to measure repeat omics signals (table 4). Lung function, blood pressure and anthropometric data were measured at the end of the panel study week following the same protocol as the subcohort clinical examination.

PATIENT AND PUBLIC INVOLVEMENT

There was no patient involvement in this study. The six cohorts participating in the HELIX project recruited healthy pregnant mothers and followed their children up to 6–11 years. The cohort studies kept the families involved throughout these years through regular clinical and questionnaire follow-ups and disseminated study results to them through newsletters, family meetings and open days. The results are also regularly disseminated to local, national and international stakeholders.

FINDINGS TO DATE

Baseline characteristics of the entire cohort

Main characteristics of the entire cohort are shown in table 6. Fifty-one per cent of the children in the entire cohort are boys; the average birth weight was 3372 g and the average gestational age 39.7 weeks; maternal age at delivery was 29.6 years on average; the majority of participants were from the highest educational level (51.6%); maternal BMI at the beginning of pregnancy showed a high percentage of overweight (25.6%) or mothers with obesity (15.8%) and 12.1% of mothers smoked during the entire pregnancy.

Baseline characteristics of the subcohort

Basic characteristics of the subcohort were somewhat different to those of the entire cohort, probably reflecting selective participation of families in the intensive subcohort follow-up visit and data completeness requirements (table 6). Compared with the entire cohort, the subcohort contained a greater percentage of boys, fewer children whose parents were born abroad (in particular in INMA and RHEA), a lower percentage of mothers with low education (in particular in BiB), a lower percentage of primiparous mothers (mainly in MoBa) and older mothers. The higher percentage of active smoking observed in the subcohort compared with the entire cohort was due to the fact that there were less missing values for smoking in the subcohort.

The age of the children in the subcohort at the time of the examination was 8.1 years on average, and this varied substantially between cohorts with the youngest ages being observed in KANC, RHEA and BiB (median age 6.4, 6.5 and 6.6 years, respectively), followed by MoBa and INMA (8.4 and 8.8 years, respectively), and the oldest ages in EDEN (11.0 years) (table 7). On average, 45.4% of the subcohort participants were girls, ranging from 42.9% in EDEN to 47.8% in MoBa. Most of the subcohort children were of white European origin, although the subcohort within BiB comprised 43% white British and 44.9% of South Asian origin families with 12.1% of other ethnicities. The family's economic situation as measured by the family affluence scale, showed marked differences between the cohorts with the majority of families in EDEN (78%), MoBa (72%) and INMA (54%) scoring high affluence, while lower affluence scores were observed in BiB, KANC and RHEA with 29.3%, 32.7% and 33.7% in the highest affluence category in those cohorts, respectively. The percentage of children classified with low family affluence was highest in BiB with 27.8%, while only around 1% of children in EDEN and MoBa were from low family affluence.

Table 6 Comparison of basic characteristics between the HELIX entire cohort (n=31 472), the subcohort (n=1301) and the child panel study (n=157)

	Entire cohort (n=31 472)	Subcohort (n=1301)	Panel study* (n=157)
Child characteristics			
Sex (%)			
Male	51.2	54.7†	56.1
Female	48.8	45.4†	43.9
Birth weight, g (SD)	3372 (547)	3379 (508)	3346 (479)
Gestational age, weeks (SD)	39.7 (1.8)	39.6 (1.7)	39.5 (1.6)
Family characteristics			
Maternal age, years (SD)	29.6 (5.2)	30.8 (4.9)†	31.0 (4.9)‡
Maternal education (%)			
Low	22.7	6.8†	6.8
Middle	25.7	34.5†	34.5
High	51.6	51.8†	51.8
Maternal BMI* (%)			
Underweight/normal weight	58.6	60.9	58.1
Overweight	25.6	24.4	25.8
Obesity	15.8	14.7	16.1
Country of origin of parents (%)			
Both from other country	21	11.1†	5.2‡
One parent from other country	6.3	7.5†	4.5‡
Both parents from country of cohort	65.7	81.4†	90.3‡
Parity (%)			
0 previous pregnancies	50.7	45.9†	42.4
≥1 previous pregnancies	49.3	54.1†	57.6
Maternal smoking during pregnancy (%)			
No	87.9	83.1†	85.3
Yes	12.1	14.8†	14.7

*BMI grouped according to WHO categories for underweight (<18.5 kg/m^2), normal (18.5–24.9 kg/m^2), overweight (25–29.9 kg/m^2) and obese (≥30 kg/m^2).
†P<0.05 comparing the subcohort (n=1301) with the entire cohort.
‡P<0.10 comparing the child panel study with non-panel subcohort children in the same cohorts (excluding MoBa).
HELIX, Human Early Life Exposome; MoBa, Norwegian Mother and Child Cohort Study.

Maternal smoking during pregnancy was most prevalent in INMA (24.7%), EDEN (23.7%) and RHEA (21.2%) (table 7). Mothers' replies to questions on environmental tobacco smoke exposure of the child showed that 34% of the children were exposed to environmental tobacco smoke in at least one place (outdoor or indoor), ranging from 19% in MoBa to 69% in RHEA. Consumption of fruits was highest in BiB and MoBa, and of vegetables in EDEN and MoBa. Visits to fast food restaurants/takeaways were most frequent in BiB. Physical activity levels were constructed based on a self-reported questionnaire where we asked about the frequency, intensity and duration of performing physical activities at school, out of school, during weekends and during summer. Over-reporting and abnormal data were corrected based on predictive models built from the panel population accelerometer (Actigraph) data. Estimates in minutes per day of moderate to vigorous activity, that is, activities with intensity above three metabolic equivalents, were low in EDEN (17 min) and INMA (26 min) and high in KANC (42 min), BiB (42 min) and RHEA (49 min).

Food allergy questionnaires showed that overall 21% of children were reported to have at least one food allergy (ever experienced), ranging from 15.6% in RHEA to 35% in INMA (table 7, figure 2). The percentage of children who had ever had asthma was low in INMA (3.6%) and high in BiB (18.5%) and EDEN (20.2%). Overall, 18.8% of children were overweight and 9.9% were obese (total 27.7%). The percentage of overweight and obese

Table 7 HELIX subcohort characteristics

	Total	BiB	EDEN	INMA	KANC	MoBa	RHEA
	(n=1301)	(n=205)	(n=198)	(n=223)	(n=204)	(n=272)	(n=199)
Age, years, median (IQR)	8.1 (6.5; 8.9)	6.6 (6.4; 6.8)	11 (10; 11)	8.8 (8.4; 9.2)	6.4 (6.1; 6.8)	8.4 (8.2; 8.8)	6.5 (6.4; 6.6)
Sex, female (%)	45.4	44.9	42.9	46.2	45.6	47.8	44.2
Ethnicity, Caucasian (%)	90	42.9*	99.5	100	100	95.6	100
Family affluence (%)							
Low	10.7	27.8	1	6.8	15.3	1.5	15.1
Middle	38.5	42.9	21.2	39.6	52.7	26.8	51.3
High	50.8	29.3	77.8	53.6	32	71.7	33.7
Maternal smoking during pregnancy, yes (%)	14.8	12.2	23.7	24.7	6	3.4	21.2
Environmental tobacco smoke, yes (%)	34.3	27.3	24.8	31.5	40.1	18.8	69.4
Fruit intake, times/week, median (IQR)	9 (6; 18)	16 (10; 21)	6.6 (3.3; 14)	7.5 (3.6; 12)	7.3 (3.8; 9.6)	8.5 (6.2; 14)	14 (8.6; 21)
Vegetable intake, times/week, median (IQR)	6.5 (4; 10)	6 (4; 10)	8.2 (4.4; 11)	6 (3; 8.5)	6 (3.5; 8.5)	6.5 (4; 10)	8.5 (6; 14)
Visits to fast food restaurant, take away times/week, median (IQR)	0.13 (0.13; 0.5)	0.5 (0.13; 1)	0.13 (0.13; 0.5)	0.13 (0.13; 0.5)	0.13 (0; 0.13)	0.5 (0.13; 0.5)	0.13 (0.13; 0.5)
Moderate to vigorous physical activity, min/day, median (IQR)	36 (22; 55)	42 (32; 62)	17 (7.8; 29)	26 (17; 44)	42 (35; 57)	35 (23; 63)	49 (36; 61)
Food allergy, yes (%)	20.7	21	16.7	35	16.2	18.8	15.6
Asthma, yes (%)	11.6	18.5	20.2	3.6	7.8	11.4	9.1
Child zBMI (%)							
Underweight/normal weight	71.4	77.1	73.7	57.7	70	84.1	62.8
Overweight	18.8	15.1	21.2	23.6	20.2	6.2	20.6
Obesity	9.9	7.8	5.1	18.6	9.9	2.6	16.6
Conners ADHD symptoms, yes (%)	10.1	9.3	8.6	6.6	15.2	4.4	10.8
CBCL, total problems, median (IQR)	21 (10; 33)	15 (8; 28)	26 (16; 37)	22 (12; 40)	27 (17; 42)	9 (5; 17)	26 (16; 36)

*44.9% of the subcohort BiB population is from South Asian ethnic origin and 12.1% of other ethnicity.
ADHD, Attention Deficit Hyperactivity Disorder; BiB, Born in Bradford; CBCL, child behaviour checklist; EDEN, Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant; HELIX, Human Early Life Exposome; INMA, INfancia y Medio Ambiente; KANC, Kaunus cohort; MoBa, Norwegian Mother and Child Cohort Study; zBMI, age standardized z-score for body mass index.

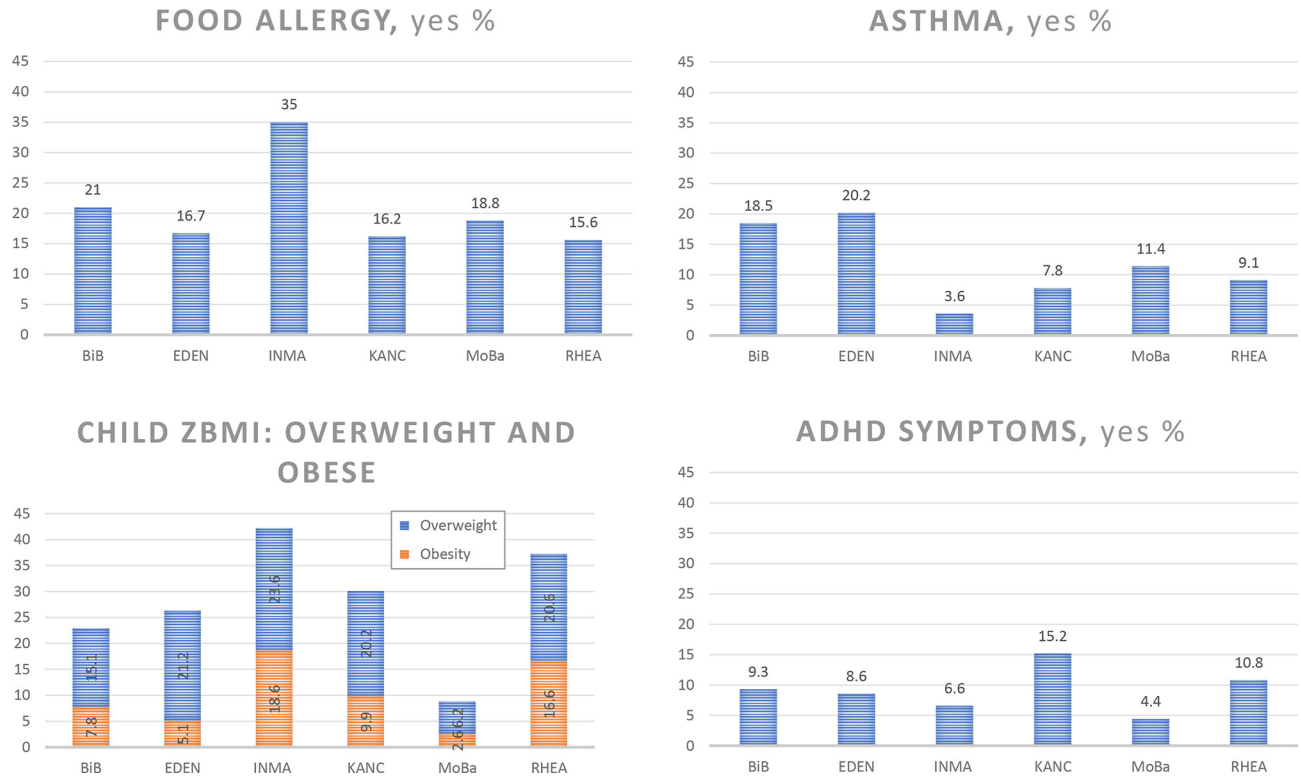

Figure 2 Prevalence of children with food allergy, asthma, overweight/obesity and ADHD symptoms in the HELIX subcohort at 6–11 years. BiB, Born in Bradford; EDEN, Étude des Déterminants pré et postnatals du développement et de la santé de l'Enfant; HELIX, Human Early Life Exposome; INMA, INfancia y Medio Ambiente; KANC, Kaunus cohort; MoBa, Norwegian Mother and Child Cohort Study; zBMI , age standardized z-score for body mass index .

children (using the age-and-sex-standardised z-scores) was highest in RHEA (37.2%) and INMA (42.3%) and lowest in MoBa (15.8%) (table 7, figure 2).

ADHD symptoms assessed through the Conner's rating scale were classified using the cut-off score of the 80th percentile.[42] Using this classification, 10.1% of children in the subcohort were classified as having ADHD symptoms, ranging from 4.4% in MoBa to 15.2% in KANC (table 7, figure 2). The total problems score of the CBCL, which consists of the sum of ratings on all 120 behavioural and emotional items of the CBCL, also showed that mothers in MoBa reported the lowest total score (median score 9) and mothers in KANC the highest (median score 27).

Baseline characteristics of the child panel study

Participants in the child panel study (n=157, 28 from BiB, 28 from EDEN, 42 from INMA, 29 from KANC and 30 from RHEA) were similar to the HELIX non-panel subcohort children of the same cohorts in terms of sociodemographic characteristics (table 6). The panel study included children whose mothers had similar ages, weight status and education than children not included in the panel. Birth weights and gestational ages were also similar between panel and non-panel children.

Through the child panel study, we showed that the pooled urine sample (before bedtime and first morning void) provided more coverage of the stable metabolome than would be achieved with either morning or bedtime urine sample alone.[43] Through the repeated analysis of non-persistent exposures, we provided variability indicators for each chemicals that can be used to correct dose-response relations and optimise sampling designs in future biomonitoring and exposome studies, and thus limit exposure misclassification (Casas, manuscript under revision).

Strengths and limitations

The HELIX project has constructed a unique large exposome cohort, which included the prospective collection of objective data from different sources (biomonitoring data, geospatial data, sensor data, child health outcomes and omics signatures). These data can facilitate cross-validation of repeated information across different sources (eg, tobacco exposure estimated from questionnaire and cotinine biomarker) and the use of standardised tools and objective measures can allow international comparisons with other studies. The pluridisciplinary aspect of the HELIX study means that a wide range of environmental factors were measured including detailed information of socioeconomic factors, which will help unravelling the influences of pregnancy risk factors, the chemical and physical environment, early family life and that of the school-age exposures on child development. Weaknesses include the loss to follow-up over time, a typical issue in most

prospective longitudinal studies and lack of statistical power to study rare outcomes. Our sample size does not allow the investigation of rare diseases or extreme values for continuous traits unless data are pooled with those of other cohorts. In addition, those living outside urban areas were not included in the study due to the lack of outdoor environment data.

ETHICS AND DATA PROTECTION

Prior to the start of HELIX, all six cohorts on which HELIX is based had been in existence for some years, had undergone the required evaluation by national ethics committees and had obtained all the required permissions for their cohort recruitment and follow-up visits. Each cohort also confirmed that relevant informed consent and approval were in place for secondary use of data from pre-existing data. The work in HELIX was covered by new ethics approvals in each country, and at enrolment in the HELIX subcohort and panel studies participants were asked to sign an informed consent form for the specific HELIX work including clinical examination and biospecimen collection and analysis. An Ethics Task Force was established to support the HELIX project on ethical issues, for advice on the project's ethical compliance, identification and alerting to changes in legislation where applicable.

Specific procedures are in place within HELIX to safeguard the privacy of study subjects and confidentiality of data. First, any reported study results pertain to analyses of aggregate data; no variables or combination of variables that can identify an individual will be associated with any published or unpublished report of this study. Primary databases with personal information (such as geocodes, dates, questionnaires or health outcomes) have been stored on separate computers with personal identifiers removed. Subjects are identified by a unique study number, linking all basic data required for the study. The master key file linking the study numbers with personal identifiers is maintained in each cohort. For the dataset analysis, all information that enables identification of an individual (dates, geocodes, etc) is removed before distribution of datasets to the researchers. All data exchanges will adhere to the most up-to-date EU and national data protection regulations.

DATA WAREHOUSE

Relevant datasets from all 31 472 mother-child pairs were transferred from the six cohorts to the central HELIX data warehouse located at ISGlobal. The HELIX data warehouse consists of several schemas, which are linked by means of common identifiers in a relational database created in MySQL.[44] New data, collected through the common protocols during the subcohort and panel study fieldwork, were entered directly into an electronic database and then uploaded into the data warehouse. Questionnaires were computer-based with a direct entry to the database. All data were locally and centrally checked by examination of the ranges, distributions, means, SD, outliers and logical checks. Data outliers and missing values were checked with the local cohort field workers and, where possible and relevant, replaced by correct values. All new measurements of exposure biomarkers and omics from the labs, and all exposure variables estimated through geospatial models and other methods, were added to the data warehouse as they became available.

Author affiliations
[1]ISGlobal, Institute for Global Health, Barcelona, Spain
[2]Universitat Pompeu Fabra (UPF), Barcelona, Spain
[3]CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
[4]Municipal Institute of Medical Research (IMIM-Hospital del Mar), Barcelona, Spain
[5]Norwegian Institute of Public Health, Oslo, Norway
[6]Team of Environmental Epidemiology, IAB, Institute for Advanced Biosciences, Inserm, CNRS, CHU-Grenoble-Alpes, University Grenoble-Alpes, CNRS, Grenoble, France
[7]Department of Environmental Sciences, Vytautas Magnus University, Kaunas, Lithuania
[8]Nursing School, Universitat de València, Valencia, Spain
[9]FISABIO–Universitat Jaume I–Universitat de València Joint Research Unit of Epidemiology and Environmental Health, Valencia, Spain
[10]Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, Spain
[11]Unité Modèles pour l'Ecotoxicologie et la Toxicologie (METO), Institut National de l'Environnement Industriel et des Risques (INERIS), Verneuil en Halatte, France
[12]Fundación Pública Galega de Medicina Xenómica (SERGAS), Santiago, Spain
[13]Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Universidad de Santiago de Compostela, Santiago, Spain
[14]Research Department, Sidra Medicine, Doha, Qatar
[15]Genomics Unit, Dexeus Woman's Health, Barcelona, Spain
[16]Inserm UMR 1153—Centre de Recherche Epidémiologie et Biostatistique Sorbonne Paris Cité (CRESS), Equipe de recherche sur les origines précoces de la santé et du développement de l'enfant (ORCHAD), Villejuif, France
[17]School of Psychology, University of the Basque Country UPV/EHU, San Sebastian, Spain
[18]Biodonostia Health Research Institute, San Sebastian, Spain
[19]Department of Health, Public Health of Gipuzkoa, Government of the Basque Country, San Sebastian, Spain
[20]Department of Social Medicine, Faculty of Medicine, University of Crete, Heraklion, Greece
[21]Division of Cancer, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK
[22]Integrative Systems Medicine and Digestive Disease, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK
[23]Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK
[24]Centre d'Investigation Clinique CIC1402, Inserm, Université de Poitiers, CHU Poitiers, Poitiers, France
[25]Department of Statistics, Faculty of Arts and Sciences, Harvard University, Cambridge, Massachusetts, USA
[26]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA
[27]Department of Genetics and Cell Biology, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands
[28]Veiligheids- en Gezondheidsregio Gelderland Midden (VGGM), Arnhem, The Netherlands

Acknowledgements The authors would like to thank all the participating children, parents, practitioners and researchers in the six countries who took part in this study. The authors would like to thank Sonia Brishoual, Angelique Serre and Michele Grosdenier (Poitiers Biobank, CRB BB-0033-00068, Poitiers, France) for biological sample management and Professor Frederic Millot (Principal Investigator), Elodie

Migault, Manuela Boue and Sandy Bertin (Clinical Investigation Center, Inserm CIC1402, CHU de Poitiers, Poitiers, France) for planning and investigational actions. The authors would like to thank Veronique Ferrand-Rigalleau, Céline Leger and Noella Gorry (CHU de Poitiers, Poitiers, France) for administrative assistance (EDEN). The authors would like to thank Silvia Fochs, Nuria Pey, Cecilia Persavente and Susana Gross for field work, sample management and overall management in INMA. The authors would like to thank Georgia Chalkiadaki and Danai Feida for biological sample management, to Eirini Michalaki, Mariza Kampouri, Anny Kyriklaki and Minas Iakovidis for field study performance and to Maria Fasoulaki for administrative assistance (RHEA). The authors would also like to thank Ingvild Essen for thorough field work, Heidi Marie Nordheim for biological sample management and the MoBa administrative unit (MoBa).

Contributors LM coordinated the collection and harmonisation of the data as the HELIX project scientific coordinator (2016–2018) and drafted the first draft of the manuscript. JdB performed the panel study fieldwork, data harmonisation and data description and assisted in drafting the manuscript. MCasas coordinated field work in the INMA Sabadell cohort, designed the biomarker database, coordinated sample collection and assisted in drafting the manuscript. OR prepared fieldwork protocols and questionnaires and supervised the fieldwork across all cohorts as the HELIX project coordinator (2013–2016). The following authors contributed to the collection of data on chemical contaminants: CT led the workpackage and oversaw all aspects of the work on the biomarker measurements of chemical contaminants; LSH performed the biological sample management and biomarker analysis; CB conducted the pharmacokinetics models data collection and protocols preparation. The following authors contributed to the collection of data on outdoor exposures: MJN led the workpackage and oversaw all aspects of the work on outdoor exposures; MdC conducted the outdoor exposure calculations; DDG conducted the exposure monitoring of panel children and physical activity; IT conducted the outdoor exposome data harmonisation and modelled indoor air pollution and water contamination exposures. The following authors contributed to the omics data collection, analysis and interpretation: MCoen led the omics workpackage, designed the study and oversaw the metabolomics data collection; HCK designed the study and oversaw the metabolomics data collection; CEL performed the NMR metabolite quantification; APS performed the MS metabolite quantification; EB conducted the proteomics analysis; ES designed the proteomics study and oversaw the proteomics data collection; MB designed the study and conducted the analysis for the DNA methylation and transcriptomics (gene expression and miRNAs); MVives conducted the gene expression and miRNA analysis; AC facilitated the analysis and oversaw research for DNA methylation data; XE coordinated the analysis and oversaw research for transcriptomics (gene expression and miRNAs) data collection; JRG designed the omics and exposome bioinformatics and statistical analyses; CHF programmed the R package and contributed to the design of the omics and exposome bioinformatics and statistical analyses. The following authors contributed to data analysis and interpretation: RS led the workpackage and oversaw the preparation of statistical analysis protocols; XB led the statistical analysis working group and prepared statistical analysis protocols; LC led a workpackage, prepared clinical examination protocols and contributed to the clinical data harmonisation and interpretation; SF prepared clinical examination protocols and contributed to the clinical data harmonisation and interpretation; JJ prepared the neurodevelopment protocols and coordinated the neurodevelopment data preparation and interpretation. VS and BG led the allergy and respiratory health data collection, harmonisation and interpretation. VS also assisted in the preparation of statistical analysis protocols. LA conducted the spirometry data harmonisation and contributed to the statistical protocol preparation. CW checked pooled data for accuracy of information and revised the manuscript critically. The following authors contributed to the cohort data collection. MoBa cohort: HMM designed the study and oversaw all aspects of subcohort and panel study data collection. KBG coordinated the subcohort fieldwork; BG coordinated the pregnancy panel fieldwork; GMA constructed and harmonised the MoBa existing database; JE was responsible for the neurological testing in the subcohort, NHK collected GIS input data and prepared routine monitoring data;. KANC cohort: RG (PI of the KANC cohort) designed the study and oversaw all aspects of KANC data collection. SA coordinated the fieldwork for subcohort and panel study and checked pooled data for accuracy of information; AD conducted fieldwork and GIS work; IP revised KANC data and revised the manuscript critically. INMA cohort: MVrijheid designed the study and oversaw all aspects of INMA subcohort and panel study data collection. FB (PI of the INMA-Valencia cohort) oversaw data collection in Valencia; JI (PI of the INMA-Gipuzkoa cohort) oversaw data collection in Gipuzkoa; JS (PI of the INMA-Sabadell cohort and of the entire INMA study) oversaw all previous INMA data collections; CM coordinated the Barcelona pregnant woman panel fieldwork and data preparation. EDEN cohort: RS designed the study and oversaw all aspects of EDEN subcohort and panel study data collection and critically reviewed the manuscript. BH (PI of the EDEN cohort) oversaw previous follow-ups of EDEN population; SLC coordinated the pregnant women panel fieldwork; JQ co-coordinated the children subcohort fieldwork and database integration; PJS was responsible for the subcohort fieldwork in Poitiers; LGA co-coordinated the children panel follow-up, checked pooled data for accuracy of information, conducted the ESCAPE data harmonisation and prepared GIS data. RHEA cohort: LC (PI of the RHEA cohort) designed the study and oversaw all aspects of RHEA subcohort and panel study data collection. JMK carried out the field work and helped design the clinical examination protocols, VL coordinated and carried out the fieldwork; TR checked pooled data for accuracy of information, prepared GIS data and conducted the clinical data harmonisation; MVafeiadi coordinated fieldwork and sample management. BiB cohort: JW designed and oversaw all aspects of BiB subcohort and panel study data collection data; DM constructed the database; RMc designed and oversaw all aspects of BiB subcohort and panel study data collection data; DW coordinated the fieldwork. PvH was responsible for dissemination aspects of the HELIX project. JU constructed and managed the HELIX database and performed data harmonisation, cleaning and validation. DvG is the HELIX project coordinator; she drafted the ethical and data protection and sharing proposal. Finally, the following authors designed the HELIX study and supervised all aspects of the work as members of the HELIX Project Executive Committee: LC, MCoen, PvdH, MJN, RS, CT and JW. MVrijheid coordinated the HELIX project, supervised all data collection, supervised all work related to the manuscript and drafted the manuscript. All authors read and approved the final manuscript. ISGlobal is a member of the Agency for the Research Centres of Catalonia (CERCA) Programme, Generalitat de Catalunya. MC is a member of the MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, UK. ISGlobal is a member of the Agency for the Research Centres of Catalonia (CERCA) Programme, Generalitat de Catalunya. MC is a member of the MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, UK.

Funding The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-206) under grant agreement no 308333—the HELIX project. Dr Maribel Casas and Dr Jordi Julvez received funding from Instituto de Salud Carlos III (Ministry of Economy and Competitiveness) (MS16/00128, MS14/00108). INMA data collections were supported by grants from the Instituto de Salud Carlos III, CIBERESP, the Conselleria de Sanitat, Generalitat Valenciana, Department of Health of the Basque Government; the Provincial Government of Gipuzkoa, and the Generalitat de Catalunya-CIRIT. KANC was funded by the grant of the Lithuanian Agency for Science Innovation and Technology (6-04-2014_31V-66). The Norwegian Mother and Child Cohort Study (MoBa) is supported by the Norwegian Ministry of Health and the Ministry of Education and Research, NIH/NIEHS (contract no. N01-ES-75558), and NIH/NINDS (grant no. 1 UO1 NS 047537-01 and grant no. 2 UO1 NS 047537-06A1). The Rhea project was financially supported by European projects, 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: 2012–2015). The work was also supported by MICINN (MTM2015-68140-R) and Centro Nacional de Genotipado-CEGEN-PRB2-ISCIII. CW received funding from the Fondation de France.

Competing interests None declared.

Patient consent Parental/guardian consent obtained.

Ethics approval Comité Ético de investigación Clínica Parc de Salut MAR.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The data warehouse has been established in a format that allows future use beyond the project lifespan (2013–2017) 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, who will evaluate potential overlap with ongoing work, adequacy of data protection plans, logistic and financial consequences and adequacy of authorship and acknowledgement plans. Further details on the content of the data warehouse (data catalogue) and procedures for external access are described on the project website (http://www.projecthelix.eu/index.php/es/data-inventory). The authors encourage interested researchers to contact them to set up collaborations.

REFERENCES

1. Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 2005;14:1847–50.
2. Wild CP. The exposome: from concept to utility. *Int J Epidemiol* 2012;41:24–32.
3. Vrijheid M, Slama R, Robinson O, *et al*. The human early-life exposome (HELIX): project rationale and design. *Environ Health Perspect* 2014;122:535–44.
4. Robinson O, Vrijheid M. The pregnancy exposome. *Curr Environ Health Rep* 2015;2:204–13.
5. Wright J, Small N, Raynor P, *et al*. Cohort profile: the born in Bradford multi-ethnic family cohort study. *Int J Epidemiol* 2013;42:978–91.
6. Heude B, Forhan A, Slama R, *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.
7. Guxens M, Ballester F, Espada M, *et al*. Cohort profile: the INMA-INfancia y Medio Ambiente-(Environment and Childhood) Project. *Int J Epidemiol* 2011;054.
8. Grazuleviciene R, Danileviciute A, Nadisauskiene R, *et al*. Maternal smoking, GSTM1 and GSTT1 polymorphism and susceptibility to adverse pregnancy outcomes. *Int J Environ Res Public Health* 2009;6:1282–97.
9. Magnus P, Birke C, Vejrup K, *et al*. Cohort profile update: the Norwegian Mother and child cohort study (MoBa). *Int J Epidemiol* 2016;45:382–8.
10. Chatzi L, Leventakou V, Vafeiadi M, *et al*. Cohort profile: the mother-child cohort in crete, Greece (Rhea Study). *Int J Epidemiol* 2017;46:1392–3.
11. Pedersen M, Giorgis-Allemand L, Bernard C, *et al*. Ambient air pollution and low birthweight: a European cohort study (ESCAPE). *Lancet Respir Med* 2013;1:695–704.
12. Serra-Majem L, Ribas L, Ngo J, *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.
13. Boyce W, Torsheim T, Currie C, *et al*. The family affluence scale as a measure of national wealth: validation of an adolescent self-report measure. *Soc Indic Res* 2006;78:473–87.
14. Kritsotakis G, Koutis AD, Alegakis AK, *et al*. Development of the social capital questionnaire in Greece. *Res Nurs Health* 2008;31:217–25.
15. Cohen S, Williamson G, Spacapam S, *et al*. Perceived stress in a probability sample of the United States. *Soc Psychol Heal Claremont Symp Appl Soc Psychol* 1988.
16. QGIS Geographic Information System. Open source geospatial foundation project. 2016.
17. de Onis M, Onyango AW, Borghi E, *et al*. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007;85:660–7.
18. Nhanes I. *Body measurements (Anthropometry)*. 20850. Rockville, MD, 1988.
19. Clasey JL, Bradley KD, Bradley JW, *et al*. A new BIA equation estimating the body composition of young children. *Obesity* 2011;19:1813–7.
20. Gillman MW, Cook NR. Blood pressure measurement in childhood epidemiological studies. *Circulation* 1995;92:1049–57.
21. Standardization of Spirometry. American thoracic society. *Am J Respir Crit Care Med* 1995;152:1107–36.
22. Le Souef P, Lafortune B, Landau L. Spirometric assessment ofasthmatic children aged 2 to 6 years. *Aust NZ Med* 1986;16.
23. Aurora P, Stocks J, Oliver C, *et al*. Quality control for spirometry in preschool children with and without lung disease. *Am J Respir Crit Care Med* 2004;169:1152–9.
24. Quanjer PH, Tammeling GJ, Cotes JE, *et al*. Lung volumes and forced ventilatory flows. Report working party standardization of lung function tests, European community for steel and coal. Official statement of the European Respiratory society. *Eur Respir J Suppl* 1993;16:5–40.
25. Bousquet J, Anto J, Auffray C, *et al*. MeDALL (Mechanisms of the Development of ALLergy): an integrated approach from phenotypes to systems medicine. *Allergy* 2011;66:596–604.
26. Vuontela V, Steenari MR, Carlson S, *et al*. Audiospatial and visuospatial working memory in 6-13 year old school children. *Learn Mem* 2003;10:74–81.
27. Rueda MR, Fan J, McCandliss BD, *et al*. Development of attentional networks in childhood. *Neuropsychologia* 2004;42:1029–40.
28. Lezak M, Howieson D, Loring D. Neuropsychological assessment 4. 2004.
29. Lezak MD, Diane BH, Bigler ED, *et al*. *Neuropsychological assessment 3*: Oxford University Press, 1995.
30. Raven JC, Court JH, Raven J. *Progressive matrices couleur/ colored progressive matrices*. Paris: Les Editions du Centre de Psychologie Appliquée, 1998.
31. Conners C. *Conners' Rating Scales - Revised. User's manual. Conners 3r*. North Tonawanda, New York, 1997.
32. Achenbach T, Rescorla L. *Manual for the ASEBA School-age forms and profiles. An integrated system of multi-informant assessment*. Burlington, VT: University of Vermont, Research Center for Children, Youth and Families, 2001.
33. Beelen R, Hoek G, Vienneau D, *et al*. Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. *Atmos Environ* 2013;72:10–23.
34. Eeftens M, Beelen R, de Hoogh K, *et al*. Development of Land Use Regression models for $PM_{(2.5)}$, $PM_{(2.5)}$absorbance, $PM_{(10)}$ and $PM_{(coarse)}$ in 20 European study areas; results of the ESCAPE project. *Environ Sci Technol* 2012;46:11195–205.
35. Nieuwenhuijsen MJ, Kruize H, Gidlow C, *et al*. Positive health effects of the natural outdoor environment in typical populations in different regions in Europe (PHENOTYPE): a study programme protocol. *BMJ Open* 2014;4:e004951.
36. Smith G, Cirach M, Swart W, *et al*. Characterisation of the natural environment: quantitative indicators across Europe. *Int J Health Geogr* 2017;16:16:16.
37. Robinson O, Tamayo I, de Castro M, *et al*. The Urban Exposome during Pregnancy and Its Socioeconomic Determinants. *Environ Health Perspect* 2018;126:077005.
38. Nieuwenhuijsen MJ, Smith R, Golfinopoulos S, *et al*. Health impacts of long-term exposure to disinfection by-products in drinking water in Europe: HIWATE. *J Water Health* 2009;7:185.
39. Siskos AP, Jain P, Römisch-Margl W, *et al*. Interlaboratory reproducibility of a targeted metabolomics platform for analysis of human serum and plasma. *Anal Chem* 2017;89:656–65.
40. Dieterle F, Ross A, Schlotterbeck G, *et al*. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal Chem* 2006;78:4281–90.
41. Nieuwenhuijsen MJ, Donaire-Gonzalez D, Foraster M, *et al*. Using personal sensors to assess the exposome and acute health effects. *Int J Environ Res Public Health* 2014;11:7805–19.
42. Julvez J, Ribas-Fitó N, Forns M, *et al*. Attention behaviour and hyperactivity at age 4 and duration of breast-feeding. *Acta Paediatr* 2007;96:842–7.
43. Maitre L, Lau CE, Vizcaino E, *et al*. Assessment of metabolic phenotypic variability in children's urine using H NMR spectroscopy. *Sci Rep* 2017;7:46082.
44. Axmark D, Widenius M. MySQL 5.1 Reference manual. 1997. https://docs.oracle.com/cd/E19078-01/mysql/mysql-refman-5.1/ (accessed 7 Dec 2017).
45. Walk Score®. Live Where You Love. 2016. https://www.walkscore.com/
46. Frank LD, Sallis JF, Conway TL, *et al*. Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass index, and air quality. *J Am Plann Assoc* 2006;72:75–87.
47. Seckmeyer G, Klingebiel M, Riechelmann S, *et al*. A critical assessment of two types of personal UV dosimeters. *Photochem Photobiol* 2012;88:215–22.