Review Article

ENVIRONMENTAL EPIDEMIOLOGY

OPEN

Advancing tools for human early lifecourse exposome research and translation (ATHLETE)

Project overview

Martine Vrijheid[a–c*], Xavier Basagaña[a–c], Juan R. Gonzalez[a–c], Vincent W. V. Jaddoe[d,e], Genon Jensen[f], Hector C. Keun[g], Rosemary R. C. McEachan[h], Joana Porcel[a–c], Valerie Siroux[i], Morris A. Swertz[j,k], Cathrine Thomsen[l], Gunn Marit Aasvang[l], Sandra Andrušaitytė[m], Karine Angeli[n], Demetris Avraam[o], Ferran Ballester[b,p,q], Paul Burton[o], Mariona Bustamante[a–c], Maribel Casas[a–c], Leda Chatzi[r], Cécile Chevrier[s], Natacha Cingotti[f], David Conti[r], Amélie Crépet[n], Payam Dadvand[a–c], Liesbeth Duijts[d,e], Esther van Enckevort[j,k], Ana Esplugues[b,p,q], Serena Fossati[a–c], Ronan Garlantezec[t], María Dolores Gómez Roig[u–w], Regina Grazuleviciene[m], Kristine B. Gützkow[l], Mònica Guxens[a,b,c,x], Sido Haakma[j,k], Ellen V. S. Hessel[y], Lesley Hoyles[z], Eleanor Hyde[j,k], Jana Klanova[aa], Jacob D. van Klaveren[y], Andreas Kortenkamp[bb], Laurent Le Brusquet[cc], Ivonne Leenen[f], Aitana Lertxundi[b,dd,ee], Nerea Lertxundi[dd,ee], Christos Lionis[ff], Sabrina Llop[b,p], Maria-Jose Lopez-Espinosa[b,p,q], Sarah Lyon-Caen[i], Lea Maitre[a–c], Dan Mason[h], Sandrine Mathy[gg], Edurne Mazarico[u–w], Tim Nawrot[hh,ii], Mark Nieuwenhuijsen[a–c], Rodney Ortiz[a–c], Marie Pedersen[jj], Josep Perelló[kk], Míriam Pérez-Cruz[u–w], Claire Philippat[i], Pavel Piler[aa], Costanza Pizzi[ll], Joane Quentin[i], Lorenzo Richiardi[ll], Adrian Rodriguez[kk], Theano Roumeliotaki[ff], José Manuel Sabin Capote[kk], Leonardo Santiago[kk], Susana Santos[d,e], Alexandros P. Siskos[g], Katrine Strandberg-Larsen[jj], Nikos Stratakis[a,r], Jordi Sunyer[a–c], Arthur Tenenhaus[cc], Marina Vafeiadi[ff], Rebecca C. Wilson[mm], John Wright[h], Tiffany Yang[h], Remy Slama[i]

Abstract. Early life stages are vulnerable to environmental hazards and present important windows of opportunity for lifelong disease prevention. This makes early life a relevant starting point for exposome studies. The Advancing Tools for Human Early Lifecourse Exposome Research and Translation (ATHLETE) project aims to develop a toolbox of exposome tools and a Europe-wide exposome cohort that will be used to systematically quantify the effects of a wide range of community- and individual-level environmental risk factors on mental, cardiometabolic, and respiratory health outcomes and associated biological pathways, longitudinally from early pregnancy through to adolescence. Exposome tool and data development include as follows: (1) a findable, accessible, interoperable, reusable (FAIR) data infrastructure for early life exposome cohort data, including 16 prospective birth cohorts in 11 European countries; (2) targeted and nontargeted approaches to measure a wide range of environmental exposures (urban, chemical, physical, behavioral, social); (3) advanced statistical and toxicological strategies to analyze complex multidimensional exposome data; (4) estimation of associations between the exposome and early organ development, health trajectories, and biological (metagenomic, metabolomic, epigenetic, aging, and stress) pathways; (5) intervention strategies to improve early life urban and chemical exposomes, co-produced with local communities; and (6) child health impacts and associated costs related to the exposome. Data, tools, and results will be assembled in an openly accessible toolbox, which will provide great opportunities for researchers, policymakers, and other stakeholders, beyond the duration of the project. ATHLETE's results will help to better understand and prevent health damage from environmental exposures and their mixtures from the earliest parts of the life course onward.

Keywords: Exposome; Early life; Exposure assessment; Child health; Adolescent health

Editors' note: *Related articles appear on pages XXX and XXX.*

[a]ISGlobal, Barcelona, Spain; [b]Spanish Consortium for Research on Epidemiology and Public Health (CIBERESP), Madrid, Spain; [c]Universitat Pompeu Fabra, Barcelona, Spain; [d]The Generation R Study Group, Erasmus University Medical Center, Rotterdam, The Netherlands; [e]Department of Pediatrics, Erasmus University Medical Center, Rotterdam, The Netherlands; [f]Health & Environment Alliance (HEAL), Brussels, Belgium; [g]Department of Surgery & Cancer and Department of Metabolism, Digestion & Reproduction, Imperial College London, London, United Kingdom; [h]Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford,United Kingdom; [i]University Grenoble Alpes, Inserm, CNRS, IAB (Institute for Advanced Biosciences) Joint Research Center, Team of Environmental Epidemiology Applied to Development and Respiratory Health, Grenoble, France; [j]University of Groningen, University Medical Center Groningen, Genomics Coordination Center, Groningen, The Netherlands; [k]University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands; [l]Department of Environmental Health, Norwegian Institute of Public Health, Oslo, Norway; [m]Department of Environmental Sciences, Vytautas Magnus University, Kaunas, Lithuania; [n]French Agency for Food, Environmental and Occupational Health and Safety (ANSES), Risk Assessment Department, Maisons-Alfort, France;

Introduction

Our lifetime health trajectories contain a so-called "build-up" stage, from conception and early intrauterine life to late adolescence, characterized by rapid successions of environmentally and socially sensitive periods that strongly determine subsequent later disease and aging trajectories and thereby influence the maximum attained level of health.[1,2] Starting prevention in early life is a particularly efficient way to shift or improve these trajectories.[3]

Environmental exposures during early life stages are associated with risks of impaired cognitive development, and cardiometabolic and respiratory diseases in childhood. Examples include smoking,[4,5] diet,[6,7] socioeconomic position,[8] air pollution,[9,10] noise,[11] lack of green spaces,[12–15] persistent organic pollutants,[16–18] bisphenol A, and phthalates.[19,20] Epidemiological studies on the impacts of early life environmental chemical and nonchemical stressors have, up to very recently, almost exclusively assessed the risks of single exposures or exposure

Downloaded from http://journals.lww.com/environepidem by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 12/14/2021

groups. More recently, exposome-wide discovery approaches have pioneered the simultaneous assessment of associations between many environmental risk factors and pregnancy and child health outcomes (e.g., blood pressure, lung function, birth weight, obesity, communication impairments).[21–26] First early life exposome studies have also made progress in understanding how multiple exposures correlate and co-exist,[27–30] how multiple exposures vary geographically and temporally,[27,30–32] which social and dietary factors determine parts of the early life exposome,[33–35] and how we may explore associations between multiple exposures and child health.[36–39]

Likewise, the first exposome projects have moved forward in the use of high-throughput omics techniques to characterize the internal part of the exposome and to identify biological signatures and pathways that respond to and interact with environmental exposures.[40–46] Such information may be used to develop novel exposure biomarkers, improve biological plausibility of associations, understand how different exposures may act on common or diverse pathways, and, ultimately, predict environmental health-related disease before its clinical manifestation. We hypothesize that the early part of the life course is a particularly important period to study the preclinical triggers of disease: exposures during vulnerable periods may have effects at the molecular level that may remain clinically undetectable until adulthood.

[o]Population Health Sciences Institute, Newcastle University, Newcastle, United Kingdom; [p]Epidemiology and Environmental Health Joint Research Unit, FISABIO-Universitat Jaume I-Universitat de València, València, Spain; [q]Faculty of Nursing and Chiropody, Universitat de València, Valencia, Spain; [r]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California; [s]University Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail)—UMR_S 1085, Rennes, France; [t]CHU de Rennes, University Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail)—UMR_S 1085, Rennes, France; [u]Institut de Recerca Sant Joan de Déu (IR-SJD), Barcelona, Spain; [v]Maternal and Child Health and Development Network II (SAMID II), Instituto de Salud Carlos III (ISCIII), Madrid, Spain; [w]BCNatal—Barcelona Center for Maternal Fetal and Neonatal Medicine, Hospital Sant Joan de Déu, Barcelona, Spain; [x]Department of Child and Adolescence Psychiatry, Erasmus MC, University Medical Center, Rotterdam, The Netherlands; [y]National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands; [z]Department of Biosciences, Nottingham Trent University, Nottingham, United Kingdom; [aa]RECETOX Centre, Faculty of Science, Masaryk University, Brno, Czech Republic; [bb]Brunel University London, College of Health, Medicine and Life Sciences, Uxbridge, United Kingdom; [cc]University Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes, Gif-sur-Yvette, France; [dd]University of Basque Country UPV/EHU, Basque Country, Bilbao, Spain; [ee]Biodonostia, Research Health Institute, Donostia-San Sebastian, Spain; [ff]Department of Social Medicine, School of Medicine, University of Crete, Heraklion, Crete, Greece; [gg]University Grenoble Alpes, CNRS, INRAE, Grenoble INP, GAEL, Grenoble, France; [hh]Centre for Environmental Sciences, Hasselt University, Hasselt, Belgium; [ii]Centre for Health and Environment, Leuven University, Leuven, Belgium; [jj]Department of Public Health, University of Copenhagen, Copenhagen, Denmark; [kk]Bettair Cities SL, Barcelona, Spain; [ll]Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin, Turin, Italy; and [mm]Department of Public Health, Policy and Systems, University of Liverpool, Liverpool, United Kingdom

Sponsorships or competing interests that may be relevant to content are disclosed at the end of the article.

*Corresponding Author. Address: ISGlobal, Institute for Global Health, C. Doctor Aiguader 88, 08003 Barcelona, Spain. E-mail: martine.vrijheid@isglobal.org (M. Vrijheid).

Copyright © 2021 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of The Environmental Epidemiology. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Environmental Epidemiology (2021) 5:e166

Received: 26 June 2020; Accepted 28 June 2021

Published online 1 October 2021

DOI: 10.1097/EE9.0000000000000166

Altogether, these first early life exposome studies also highlighted many challenges related to temporal exposure variability, differential measurement errors (i.e., different errors for different exposures), mixture effects, cross-sectional designs, false-positive findings, statistical power, and absence of causal structure in untargeted analyses. The Advancing Tools for Human Early Lifecourse Exposome Research and Translation (ATHLETE) project was designed to advance some of these challenges through improved tools, data, and translation of knowledge from exposome research into practice. Here, we provide an overview of the project's design, study populations, planned measurements, and tools and infrastructure development.

Project description

Aim

The general objective of ATHLETE (http://www.athleteproject.eu/) is to develop a toolbox of exposome tools and a Europe-wide exposome cohort that will be used to systematically quantify the effects of a wide range of community-level and individual-level environmental risk factors on mental, cardiometabolic, and respiratory health outcomes and associated biological pathways during the first 2 decades of life, to develop intervention strategies to improve early life urban and chemical exposomes, and to translate the resulting evidence to policy recommendations and prevention strategies. ATHLETE forms part of the European Human Exposome Network (https://www.humanexposome.eu/). The project consists of three interlinked components, containing nine research areas or work packages (WPs), focusing on data and tools, evidence, and translation (Figure 1), described in detail below.

Study population

Study populations include general population cohorts and exposome intervention studies. The intervention studies are described below. Here we detail the ATHLETE Europe-wide exposome cohort, which consists of 16 existing longitudinal population-based birth cohort studies in 11 European countries (Figure 2). Each cohort recruited mothers before or during pregnancy, or at delivery, and actively follows its participants through childhood and adolescence. Together, these cohorts include around 80,000 mother-child pairs with a wealth of already collected exposome data (Figure 3). Our rationale for this selection of cohorts is three-fold:

1. Prospective follow-up of the Human Early Life Exposome (HELIX) subcohort. The HELIX project previously generated a completely harmonized dataset with biomonitoring data (chemical exposome), geospatial data (urban exposome), questionnaire data (behavioral/lifestyle/social exposome), multiomics signatures (genome, deoxyribonucleic acid [DNA] methylome, transcriptome, proteins, metabolome), and child health data (neurodevelopment, growth, cardiometabolic health, respiratory health, allergies) up to 6–11 years, in around 1,300 mother-child pairs from six existing European cohorts (Born in Bradford [BiB],[47] Etude des Determinants pre et postnatals du developpement et de la sante de l'Enfant [EDEN],[48] Kaunas Cohort [KANC],[49] Infancia y Medio Ambiente [INMA],[50] Norwegian Mother and Child Cohort [MoBa],[51] Crete Mother Child Cohort [RHEA]),[52] as extensively documented.[53,54] ATHLETE will follow-up this cohort into adolescence (at 12–18 years, 7 years after the HELIX visit, with 1,100 adolescents expected to participate), to add a prospective data collection time point for exposure, omics and health outcome data to allow evaluation of longitudinal associations into adolescence. It will also allow inclusion of exposures of particular relevance for adolescents, such as screen time, sleep, mental health, and, topically, of questions related to the impact of

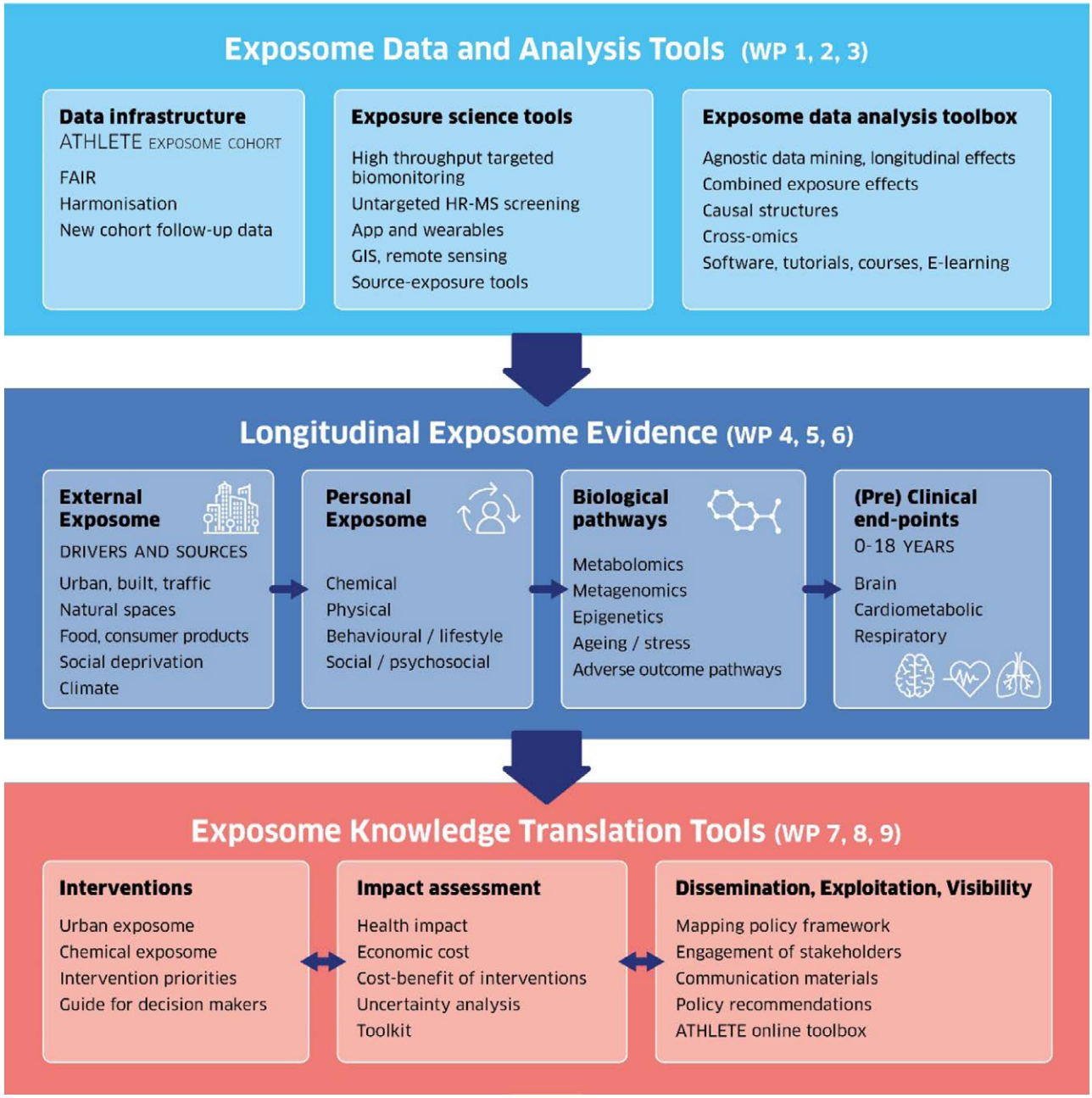


Figure 1. ATHLETE project components. GIS indicates geographic information system.

coronavirus disease 2019 (COVID-19) lockdown and social distancing measures on mental and physical health and well-being of adolescents. Standardized protocols across the six cohorts will largely repeat the common HELIX protocols,[53] and collect data as needed for the subsequent work in the project (see below): biological samples (blood, urine, stool, hair), questionnaires, smartphone app and wearable sensors, address history, and clinical examinations.

2. Enlarging the adolescent exposome cohort by including new populations. In the exposome context, testing multiple exposures and applying untargeted analysis approaches, large sample sizes and replication studies are required to improve power and causal inference. ATHLETE will build on the European Union (EU) Child Cohort Network, established as part of the EC-H2020 LifeCycle project (https://lifecycle-project.eu), which brings together many European pregnancy and child cohort studies into one harmonized and findable, accessible, interoperable, reusable (FAIR) data sharing platform. ATHLETE includes those cohorts from the network for which we have already characterized and harmonized important parts of the exposome, including the external, physical, lifestyle, and social exposome: Generation R in the Netherlands,[55] Danish National Birth Cohort (DNBC) in Denmark,[56] Nascita e Infanzia: gli Effetti dell'Ambiente (NINFEA) in Italy,[57] and Perturbateurs Endocriniens: Étude Longitudinale sur les Anomalies de la Grossesse, l'Infertilité et l'Enfance (PELAGIE) in France,[58] as well as the six HELIX cohorts. These cohorts have entered adolescence (Figure 3) and allow the investigation of repeat measurements of the exposome in association with repeated omics and outcome data up to 18 years of age.
3. Integrating newly established birth cohorts with improved in-depth exposome data. ATHLETE integrates "new,"

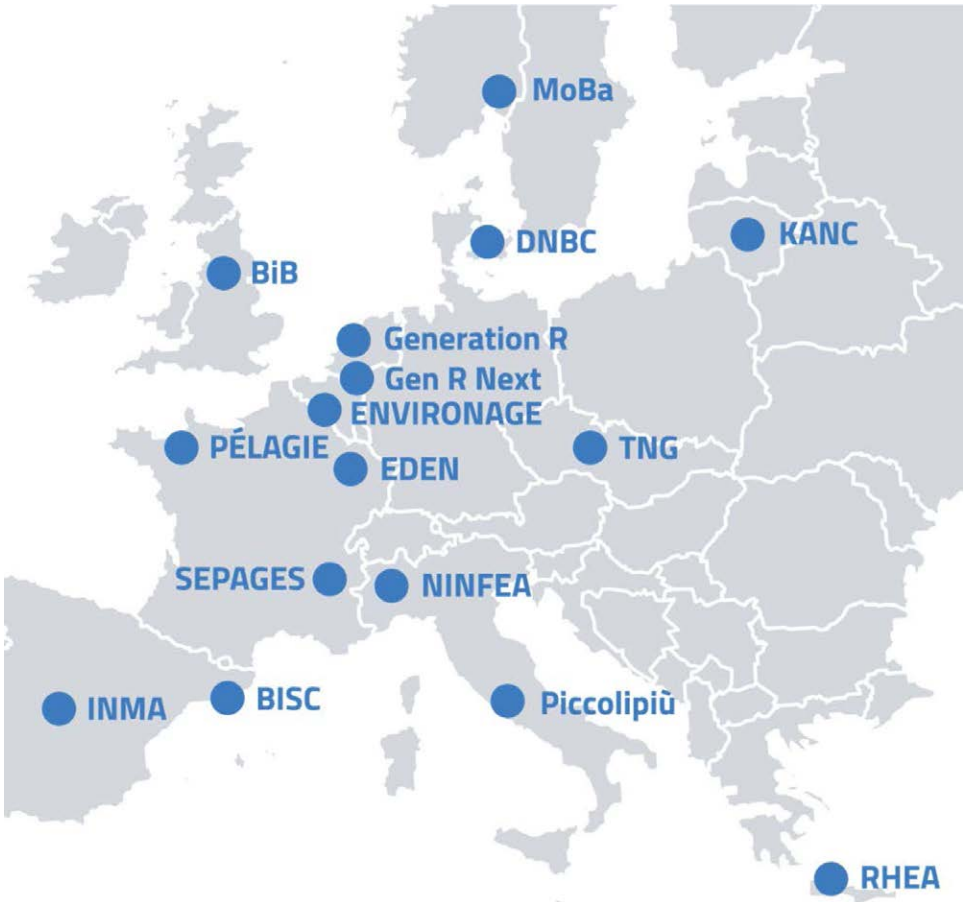


Figure 2. Cohorts participating in ATHLETE. BiSC indicates Barcelona Life Study Cohort; DNBC, Danish National Birth Cohort; EDEN, Etude des Determinants pre et postnatals du developpement et de la sante de l'Enfant; ENVIRONAGE, environmental influence on early ageing; INMA, INfancia y Medio Ambiente; KANC, Kaunas Cohort; MoBa, Norwegian Mother and Child Cohort Study; NINFEA, Nascita e Infanzia: gli Effetti dell'Ambiente; PELAGIE, Perturbateurs Endocriniens: Étude Longitudinale sur les Anomalies de la Grossesse, l'Infertilité et l'Enfance; RHEA, Crete Mother Child Cohort; SEPAGES, Suivi de l'Exposition à la Pollution Atmosphérique durant la Grossesse et Effets sur la Santé; TNG, CELSPAC The Next Generation cohort.

recently established, state-of-the-art birth cohorts that are highly suitable for exposome research: Suivi de l'Exposition à la Pollution Atmosphérique durant la Grossesse et Effets sur la Santé (SEPAGES) in France,[59] environmental influence on early ageing (ENVIRONAGE) in Belgium,[60] Generation R Next in the Netherlands, Barcelona Life Study Cohort (BiSC) in Spain (https://www.projectebisc.org), Piccolipiù in Italy,[61] and CELSPAC-The Next Generation (TNG) in the Czech Republic. The inclusion of new cohorts is important for (1) their improved sampling strategies for exposure assessment, in particular the collection of many repeated urine samples during pregnancy (BiSC and SEPAGES), personal monitoring (BiSC and SEPAGES), and placenta sampling (BiSC, SEPAGES, ENVIRONAGE) and (2) their cutting-edge outcome assessments, including imaging techniques, to study organ and placenta development (BiSC, Generation R Next). The inclusion of new exposome cohorts also allows the evaluation of new chemicals that are now produced in high volumes but that are not detectable in biosamples collected during pregnancy in older cohorts even 10 years ago (e.g., new bisphenols).

FAIR data infrastructure for the ATHLETE Exposome cohort (WP1)

At present, exposome data are scattered across hard-to-find and hard-to-access databases. A prerequisite for exposome research into the future is to bring data together in openly accessible data platforms that will allow pooling of data for larger sample size and replication of findings. ATHLETE will implement an early life exposome data infrastructure by building on the data sharing platform that has already been developed as part of LifeCycle for European birth cohorts (https://lifecycle-project.eu/for-scientists/variable-catalogue/) and that implements FAIR principles to enable findability, accessibility, interoperability, and reusability of cohort data.[62] This infrastructure makes cohorts and datasets findable for project partners and outside researchers in an easy-to-use open access web-based catalog, enabling quick assessment of available data suitable to answer specific research questions. No actual data are given in the online catalog. ATHLETE will add to the existing data catalog by proposing a new set of exposome modules with harmonized data for the participating cohorts, including, among others, data on the chemical exposome that is not currently available in the catalog. Importantly also, the richest exposome database within this project, the HELIX subcohort, will be transferred in to the FAIR infrastructure as a separate entity to make it easily accessible. The catalog structure will be based on international standards, most notably the Minimum Information About BIobank Data Sharing (MIABIS) standard,[63] and those defined in http://fairsharing.org. The catalog software will build upon the open source Molecular Genetics Information System (MOLGENIS) project,[64] which has been proven for many catalogs including the EU catalog of biobanks (http://directory.bbmri-eric.eu). ATHLETE will implement harmonization protocols to make the exposome data interoperable. Syntax files for harmonization of exposome variables will be developed, tested, and applied to

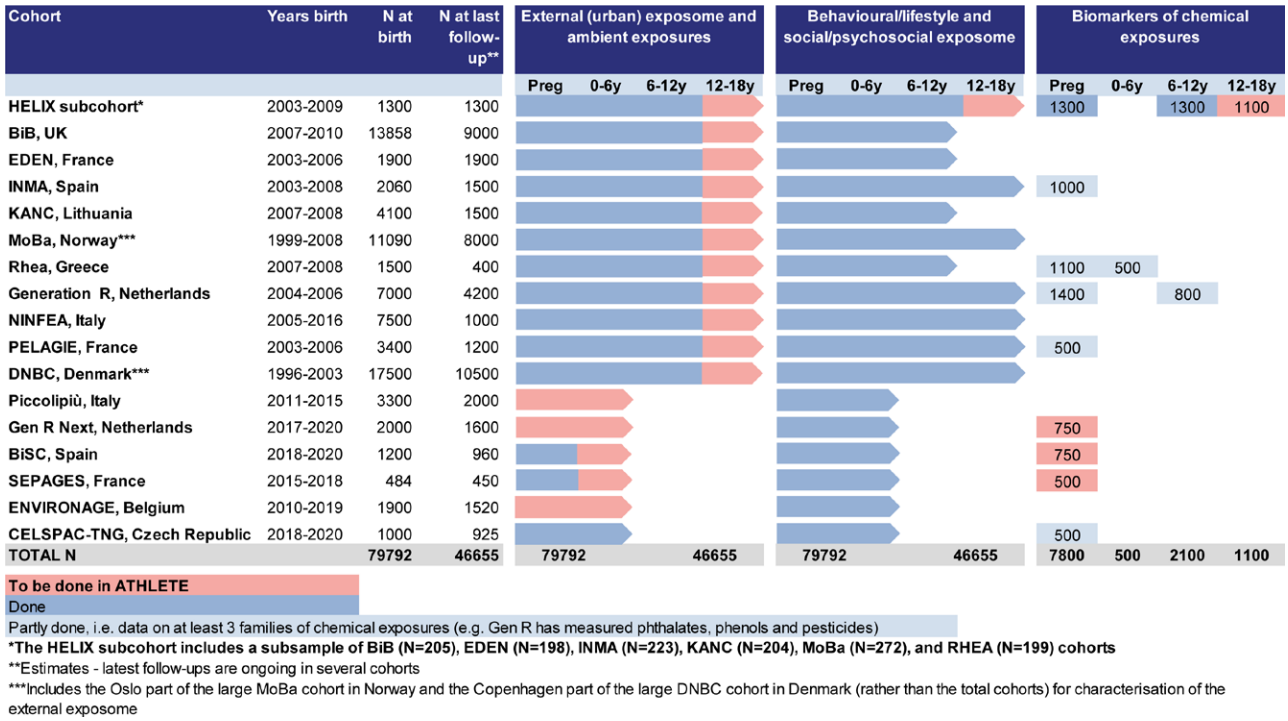


Figure 3. Timeline of available data on exposome domains in the ATHLETE cohorts. BiSC indicates Barcelona Life Study Cohort; CELSPAC-TNG, CELSPAC The Next Generation cohort; DNBC, Danish National Birth Cohort; EDEN, Etude des Determinants pre et postnatals du developpement et de la sante de l'Enfant; ENVIRONAGE, environmental influence on early ageing; HELIX, Human Early Life Exposome; INMA, INfancia y Medio Ambiente; KANC, Kaunas Cohort; MoBa, Norwegian Mother and Child Cohort Study; NINFEA, Nascita e Infanzia: gli Effetti dell'Ambiente; PELAGIE, Perturbateurs Endocriniens: Étude Longitudinale sur les Anomalies de la Grossesse, l'Infertilité et l'Enfance; RHEA, Crete Mother Child Cohort; SEPAGES, Suivi de l'Exposition à la Pollution Atmosphérique durant la Grossesse et Effets sur la Santé.

cohort data. Each cohort will harmonize and store their own harmonized data on secure local servers and make the metadata findable through the data catalog.

ATHLETE will implement "federated" (data stays on local servers and is analyzed remotely) and "centralized" (data analyzed centrally by analyst) systems for cohort owners to make their datasets accessible to project partners and outside researchers in a secure and controlled manner. For some of the exposome analyses, we expect to deploy "DataSHIELD"[65,66] as one of the federated access protocols, which enables access from the "R" statistical environment using MOLGENIS[64] or Opal software.[67] The federated system overcomes governance restrictions that prohibit the release or sharing of some of the required data, or render data access slow. Because we do not expect that all exposome analyses can be done through DataSHIELD or similar protocols, the local servers will also enable cohorts and database owners to submit their data centrally, where data are then analyzed centrally on a trusted facility with strict data access policies (managed by the project steering committee). In all cases, the cohort and data owners will be in full control of data access.

Exposure assessment tools (WP2)

An individual's exposome is made up of a great number of exposures, many of which are correlated, and which vary over time and across geographical locations.[27,30,31,33] Accuracy of exposure estimates is crucial in exposome studies. When many exposures are analyzed together, differential measurement errors (with some exposures more accurately measured than others) may lead to false negative findings and can greatly reduce our ability to compare risk estimates from these exposures. For example, we have previously established that for highly variable nonpersistent chemical exposures (which comprise most chemicals of current regulatory concern), measurements in single spot urine samples entail attenuation bias, which can amount to 80% in the case of compounds with very high within-subject variability such as bisphenol A.[31,68,69] Bias can be mitigated by within-subject pooling of many biospecimens.[39] Similarly, improved accuracy of exposure estimates in the external and urban environment can be achieved by integrating information on how people move through their environment and on their personal exposome levels.[32] Exposome measurements thus require complementary approaches to achieve both wide and accurate exposome coverage.

For the generation of new exposure data, ATHLETE will use complementary targeted and nontargeted exposure assessment approaches aimed, respectively, at obtaining solid exposure-response relationships for more established or suspected (chemical, physical, behavioral, social, urban) risk factors, and at the exploration of the "unknown" part of the chemical exposome. Exposures included are summarized in Table 1. Our choice of exposures was based on their widespread occurrence in the general population, and on their relevance for at least one of the health outcome areas under study. Our choice of chemical pollutant groups was further based on recent or current production, plausibility of frequent exposure in European pregnant women and children, and alignment with the chemicals prioritized by the European human biomonitoring project (https://www.hbm4eu.eu/). ATHLETE will have access to a very wide range of existing, and already harmonized exposure data at repeated time points in the cohorts, and will generate new exposure data to complement this (Table 1).

Targeted biomonitoring assays will be developed for the measurement of around 100 known chemical pollutants (per- and perfluoroalkyl substances [PFASs], heavy metals, phenols, phthalates, pesticides, polycyclic aromatic hydrocarbons—Table 1); these comprise highly sensitive multiassay techniques to determine different persistent and nonpersistent chemicals simultaneously in small volumes of biosample, thereby making

Table 1.

Exposure assessment—new and existing data in the ATHLETE cohorts

Exposure group	Exposure variables	Methods for new data generation	New data to be generated	Existing data in ATHLETE cohorts
Personal exposome				
Chemical exposures (traditional biomarkers)				
PFASs	19 PFASs, incl PFOS, PFOA, PFNA, PFUnDA, PFHxS	Plasma measurement	BiSC, SEPAGES, Gen R next (pregnancy N = 2,000) + HELIX subcohort new follow-up (12–18 yr, N = 1,100)	HELIX subcohort, INMA, PELAGIE, DNBC, TNG
Metals and elements	15 metals and elements including cadmium, arsenic, mercury, copper, cobalt, lead	Whole blood measurement		HELIX subcohort, INMA, TNG
Phthalate metabolites	15 phthalate metabolites including DINCH metabolites	Pools of repeat urine samples[a]		HELIX subcohort, INMA, Rhea, Gen R, PELAGIE, SEPAGES, TNG
Phenols	4 parabens, 5 bisphenols including bisphenol A, oxybenzone, triclosan	Pools of repeat urine samples[a]		HELIX subcohort, INMA, Rhea, Gen R, PELAGIE, SEPAGES, TNG
Organophosphate pesticides	6 dialkyl phosphate metabolites	Pools of repeat urine samples[a]		HELIX subcohort, INMA, Gen R, PELAGIE, TNG
Other pesticides	Metabolites of pyrethroids, 2,4-dichlorophenoxyacid, boscalid, and imazalil	Pools of repeat urine samples[a]		PELAGIE
Glycol ethers	Metabolite of phenoxyethanol	Pools of repeat urine samples[a]		PELAGIE
PAHs	18 metabolites, including: 3-hydroxybenzo[a]pyrene, 1-hydroxypyrene	Pools of repeat urine samples[a]		TNG
Tobacco smoking	Cotinine, self-reported smoking habits	Questionnaires, urine samples		All cohorts
Persistent organic pollutants	Organochlorine compounds (PCBs, DDE, HCB), brominated flame retardants	—	—	HELIX subcohort, INMA, Rhea, PELAGIE, DNBC, TNG
Chemical exposures (nontargeted)				
Unknown and emerging chemicals	Nontargeted screening	High-resolution mass spectrometry platforms	HELIX subcohort follow-up (12–18 yr, N = 1,100)	TNG
Chemical exposures (ambient)				
Outdoor air pollution	NO_2, $PM_{2.5}$, PM_{10}, $PM_{2.5abs}$ composition	ELAPSE and ESCAPE air pollution models	New cohorts and new follow-ups	All cohorts
Personal and indoor air pollution	NO_2, $PM_{2.5}$	Diffusion tubes	HELIX follow-up (12–18 yr, N = 1,100)	BiSC, SEPAGES, HELIX panel studies, TNG
Physical exposures (ambient)				
Road traffic noise	Noise levels	Regulatory noise maps combined with questionnaire data (location of bedrooms, etc.)	New cohorts and new follow-ups	All cohorts
Meteorological factors	Temperature, relative humidity	Daily average from city monitoring stations		HELIX, INMA, Gen R, Gen R Next, DNBC
UV	Ambient UV radiation levels	Remote sensing and questionnaires		HELIX
Light	Nighttime light exposure	Remote sensing combined with questionnaire data		INMA, Gen R, Gen R Next
Behavioral/lifestyle exposures				
Physical activity	Physical activity duration and intensity	Actigraphs, ExpoApp3, questionnaires	HELIX subcohort follow-up (12–18 yr, N = 1,100)	All cohorts (questionnaires)
Sleep	Sleep duration, sleep quality, sleep onset latency	Wrist watches and questionnaires		All cohorts (questionnaires)
Commuting routes	Commuting routes and modes	qGIS		HELIX subcohort
Mobility	Time spent in different environments	ExpoApp3		HELIX panel studies, BiSC
Mobile technology use/screen time	Use of mobile phones, laptops, tablets, gaming, etc.	Questionnaires		INMA, Gen R, Gen R Next
Diet	Food frequency, diet quality	Food frequency questionnaire		All cohorts
Social/psychosocial exposures				
Psychosocial	Stress	Questionnaires, hair cortisol	HELIX subcohort follow-up (12–18 yr, N = 1,100)	BiSC, INMA, Gen R (hair cortisol), TNG
Social and economic capital	Family affluence score, social contact, social participation, house crowding	Questionnaires		All cohorts
External (or urban) exposome				
Built environment	Population and building density, street connectivity, facility density, land use, walkability	Land cover/use maps	New cohorts and new follow-ups	All cohorts
Natural spaces including green space	Residential surrounding greenness, distance to nearest green and blue spaces	Remote sensing and land cover/use maps		All cohorts
Traffic and transport	Traffic load, distance to roads, public transport network	Land use maps		All cohorts
Social deprivation	Area level indicators	Local deprivation data		All cohorts
Food environment	Fast food restaurants, healthy food places	Facilities maps		—

— denotes no data.

[a]Within-subject pools of many urine samples in the new cohorts, 2–3 samples daily during 2 pregnancy weeks; pools of 5–10 urine samples in the HELIX subcohort follow-up.

BiSC indicates Barcelona Life Study Cohort; DDE, 4,4′dichlorodiphenyldichloroethylene; DINCH, 1,2-Cyclohexane dicarboxylic acid diisononyl ester; DNBC, Danish National Birth Cohort; ELAPSE, Effects of Low-Level Air Pollution: A Study in Europe; ESCAPE, European Study of Cohorts for Air Pollution Effects; Gen R, Generation R cohort; Gen R next, Generation R Next cohort; HCB, hexachlorobenzene; INMA, INfancia y Medio Ambiente cohort; NO2, nitrogen dioxide; PAH, polycyclic aromatic hydrocarbon; PCB, polychlorinated biphenyl; PELAGIE, Perturbateurs Endocriniens: Étude Longitudinale sur les Anomalies de la Grossesse, l'Infertilité et l'Enfance; PFHxS, perfluorohexanesulfonate; PFNA, perfluorononanoate; PFOA, perfluorooctanoate; PFOS, perfluorohexanesulfonate; PFUnDA, perfluoroundecanoate; PM10, particulate matter with an aerodynamic diameter of less than 10 μm; PM2.5, particulate matter with an aerodynamic diameter of less than 2.5 μm; PM2.5abs, absorbance of PM2.5 filters; RHEA, Crete Mother Child Cohort; SEPAGES, Suivi de l'Exposition à la Pollution Atmosphérique durant la Grossesse et Effets sur la Santé; UV, ultraviolet.

such targeted biomonitoring methods more suitable for the exposome era. ATHLETE will use within-subject biospecimen pooling of many urine samples to achieve greater accuracy in the measurement of nonpersistent chemical pollutants.[31,69] New data on these chemical pollutants will be generated in the newly collected HELIX subcohort samples (standardized biosample collection at 12–18 years of age) and in stored pregnancy samples in the new cohorts (BiSC, Generation R Next, SEPAGES).

In addition, we will explore high-resolution mass spectrometry (HR-MS) techniques combined with liquid and gas chromatography (LC, GC) for their capability to detect relevant, but thus far "unknown" or emerging chemical exposures. Nontargeted and suspect screening approaches to detect chemicals of emerging concern are currently in a stage of rapid development, and although the analytical technologies still face many challenges related to their ability to identify and accurately quantify exposures,[70] they hold important promises for the discovery and prioritization of chemical exposures.[70,71] We will apply HR-MS to the HELIX subcohort samples (12–18 years) for which the targeted chemical exposome is also available, allowing comparison and evaluation of both approaches.

Lastly, ATHLETE will build on the already established geospatial modeling platform, developed by the HELIX and LifeCycle projects,[33] for the characterization of the external and urban exposome in birth cohorts. We will expand this platform by including new cohorts and new exposures such as food environment and nighttime light exposure (Table 1). To improve the accuracy of estimation of external and lifestyle exposures, we will combine geospatial and questionnaire-based methods with personal monitoring approaches (Table 1). These include wearable sensors for air pollution, physical activity and sleep, and a new smartphone application for location and physical activity data (Android mobile app ExpoApp3 and associated web dashboard ExpoHub, developed by Bettair; Bettair Cities SL, Barcelona, Spain). The application will make it possible to estimate the external exposome not just at the participants' residential address but in different microenvironments (home, school, commuting routes). The wearable sensors will be deployed in the follow-up of the HELIX subcohort at 12–18 years in all participating adolescents and in the urban exposome intervention study detailed below.

To intervene on (parts of) the personal exposome, it is important to understand what the drivers and sources of exposures are. ATHLETE will tackle this question through two distinct approaches: (1) by evaluating socioeconomic position, deprivation, and urbanization as drivers of the exposome using data from all ATHLETE cohorts on the personal and external exposome. This will help to identify vulnerable subgroups of the population, that is, those that are exposed to multiple environmental hazards and (2) by identifying dietary sources of the chemical exposome. Diet is one of the main sources of exposure for a range of chemicals, including persistent organic pollutants, PFAS, bisphenol A, phthalates, and pesticides.[72] As a novel approach, we will use the Monte Carlo Risk Assessment (MCRA) models for dietary risk assessment of single chemicals and mixtures as developed in the EU-funded EuroMix project (www.euromixproject.eu). This EU project has resulted in a web-based toolbox for exposure assessment and can link dietary risk assessment using European residue monitoring and food consumption data to biomarkers of exposure.[73] The MCRA models have been applied to mixtures of residue of pesticides, food additives, and contaminants,[74] and are used to discuss regulatory implementation of mixtures of pesticides.[75] The models will be integrated with biomonitoring data from ATHLETE cohorts and intervention studies.

Exposome data analysis tools (WP3)

An important challenge in associating the exposome with health outcomes is the simultaneous consideration of many correlated exposures.[27] Our previous methodological work established that, in an exposome context, some statistical techniques are limited in their ability to efficiently differentiate true predictors from correlated covariates, so that false-positive findings are a concern.[36,37] ATHLETE will leverage these early proof-of-principle studies to develop strategies and tools to tackle the next set of analytical challenges in the context of exposome research.

1. Evaluation of longitudinal exposome and health associations

Longitudinal exposome studies introduce further complexity to already complex exposome data. Questions include how to relate the exposome to longitudinal trajectories of a health outcome, how to relate repeated assessments of the exposome to an outcome measured at a single time point, and combinations of the two. Strategies to be considered include those aimed at risk prediction, including machine-learning (black box) techniques, and those aimed at estimating dose-response functions for relevant exposures, support vector classifier for longitudinal high-dimensional data or penalized generalized estimating equations.[76–78]

2. Estimation of combined effects of exposures

Simultaneous exposure to several harmful exposures can confer extra risk for a health outcome compared with the sum of effects of isolated exposures. Such potentially complex interactions increase exponentially the dimensionality of the exposome and are difficult to capture by purely statistical methods. Simulation studies and real data will be used to assess the properties of agnostic statistical methods that have been proposed to analyze combined effects of exposures related to health risk (e.g., Bayesian Kernel Regression, Bayesian Profile Regression). This task will also develop ways to incorporate a priori information from toxicology on synergistic effects of exposure combinations.

3. Integration of exposome and cross-omics data to uncover exposome-health relationships

The availability of multilayer omics data (e.g., metabolomics, metagenomics, epigenomics, transcriptomics) in exposome studies allows the integration of data on biological pathways in exposome-health associations. This will involve extending previous work on the Regularized Generalized Canonical Correlation Analysis (RGCCA) framework, a method to integrate data from multiple sources,[79–81] and comparing other suggested approaches to data integration of multilayer omics data, for example, joint and individual variance explained, single cell analysis, joint Matrix/Tensor Factorization approaches, Latent Unknown clustering (LUCID), network analysis, and sparse penalized least squares (sPLS).[82]

4. The incorporation of a priori knowledge on causal structures and mediators to improve causal inference

To complement the agnostic methods above, we will develop strategies to incorporate a priori information on the temporal ordering of exposures, the hypothesized causal structures, or the biological pathways (from omics or toxicological data) into the exposome-health associations.[83] This will include developing penalized extensions of Structural Equation Models to the high-dimensional case, expanding methods for mediation analysis that incorporate penalized approaches for variable selection (including multiple exposures and multiple mediators), and applying analyses that incorporate hypothesized causal

structure through causal diagrams, and handle high-dimensional confounding with super-learner.[84]

In all strategies, we will take account of issues inherent to exposome data, such as correlation between exposures, missing data, cohort effects, and exposure measurement errors that differ between exposures.

In addition to developing these analytical tools, ATHLETE will develop open-access software, front-end applications, tutorials, e-learning material and courses, targeted at varying levels of expertise. These are being made available through our online toolbox (https://athleteproject.eu/toolbox/). As part of this, we aim to extend DataSHIELD tools for remote and nondisclosive data analysis (http://www.datashield.ac.uk/) by incorporating new functionalities to deal with exposome data visualization and analysis. To this end, our recent development of the "resources" architecture in DataSHIELD will facilitate handling complex big data, including omics, within DataSHIELD through the Opal data warehouse.[85]

We will create R packages that will be available through open source repositories such as Comprehensive R Archive Network (CRAN) and Bioconductor, along with Shiny apps that will facilitate their use for less experienced users. Developments will include adding functionalities to our existing R-exposome package (https://isglobal-brge.github.io/rexposome/),[86] and the RGCCA Package.[87]

Biological pathways from the exposome to health (WP4)

Omics technologies are promising tools to shed light on early, preclinical, perturbations of biological pathways in response to environmental exposures. For example, first exposome projects have shown that early life exposures (including arsenic, tobacco smoke, air pollution, polychlorinated biphenyls, PFASs, diet) may have a detectable imprint on the metabolomic and epigenetic profiles of pregnant women and children.[40–46] These early studies require replication in larger populations, prospective follow-up, and repeated time points of measurement. ATHLETE cohorts have measured multiple omics layers (genetics, epigenetics, transcriptomics, metabolomics) in a relatively large sample size (Table 2), allowing us to address these challenges. ATHLETE will focus on the prospective evaluation of biological pathways from the exposome to adverse health, including metagenomics (gut microbiome), metabolomics, epigenetics (placenta and blood), and specific aging (telomere length, epigenetic aging) and stress (allostatic load) related pathways, as well as on the evaluation of cross-omics responses. ATHLETE will include the gut microbiome as a new omics layer in exposome research. Using these datasets together with existing toxicological databases and models, we aim to discover key molecular events and biological pathways associated to the exposome that are of specific interest in early life.

New data to be generated in all participants in the new follow-up of the HELIX subcohort (12–18 years, N = 1,100) include shotgun whole metagenomic sequencing (~5 gigabases/sample) of fecal DNA and untargeted serum and urine metabolomics profiling data generated by HR-MS and nuclear magnetic resonance spectroscopy (NMR). This will allow us to define gut metagenomic signatures at the functional level that associate with different exposures, with our health outcomes, and with metabolic mediators.

Further, we will use the ATHLETE omics data resource to build a poly-environmental score for risk prediction, aggregating environmental risk factors and multiomics data, and including genetic background. The uses of such a score would be multiple and include improved risk prediction for prevention or early detection strategies aimed at individuals or at-risk groups; improved identification of vulnerable or susceptible subpopulations in epidemiological studies, which would allow risk stratification and evaluation of interactions with lesser known environmental risk factors (similar to polygenic risk scores); and improved ease of use of omics data in disease risk prediction (by developing a single index). We will evaluate a range of attributes of the score and evaluate how the addition of biological and omics markers may improve prediction models based on more easily available variables.

Finally, we will explore the use of the adverse outcome pathway (AOP) concept from toxicology to build hypotheses about toxicologically relevant mixtures in the study of specific health outcomes in our epidemiological exposome studies. The AOP concept links the exposure of chemicals to their molecular initiating events, through network/pathway disturbances and key events to responses at the cellular, organ, organism, and population levels.[88,89] Although of recognized relevance in toxicology, AOPs have seen little use in epidemiology. This work aims to provide indices of combined exposures to the exposome-health association studies.

Exposome-health associations (WP5 and 6)

The systematic evaluation of health risks related to multiple exposures will inform public health strategies or decisions, by identifying chemical agents or urban and lifestyle exposures, or combinations of these exposures, that are most likely to pose a hazard. ATHLETE focuses on health outcome areas that are known to be linked to noncommunicable disease risk in later life,[90–94] and that represent prevalent health end-points in European children.

1. Brain development: ATHLETE cohorts have assessed brain structure in embryonic, fetal and infant life through cutting-edge imaging techniques (neurosonography, brain magnetic resonance imaging [MRI]), and brain development by repeat neuropsychological and neurobehavioral assessments during childhood and adolescence.
2. The cardiometabolic system: ATHLETE cohorts have assessed early (embryonic, fetal, and infant) organ development using cutting-edge measurements of advanced cardiac and great vessel imaging (anatomical and functional echocardiography and MRI), as well as trajectories of cardiometabolic health (e.g., blood pressure, macrovascular and microvascular phenotypes, weight gain, lipid profiles) into adolescence.
3. The respiratory system: ATHLETE cohorts have assessed early lung structure, repeated lung function measurements throughout childhood (spirometry), respiratory symptoms (e.g., wheeze), clinical outcomes (e.g., doctor diagnoses of asthma), and immunological or allergy-related outcomes (e.g., eczema, rhinitis).

Exposome-health associations will be examined in two parts: (1) focusing on associations between the in utero exposome and outcomes during embryonic, fetal, and neonatal life, including novel outcomes based on imaging techniques (neonatal MRI, fetal neurosonography, echocardiography) and the use of placental function measurements and (2) focusing on longitudinal exposome-health trajectories into adolescence. For both parts, we will distinguish different populations for hypotheses related to different exposome domains; in practice, this means that for external, lifestyle and psychosocial domains, we will base our analyses on the larger cohort populations (N up to 80,000). For analyses including the chemical exposome, we will restrict our analyses to those with biomarker data (N up to 7,300). This large sample size allows us to look at the new questions not yet tackled in exposome research, for example, on interactions between exposures and between exposures and other risk factors. This work will follow the statistical strategies to be developed in WP3 as described above and follow both an agnostic approach and an approach including a priori knowledge from biological pathways and causal structures and incorporating longitudinal trajectories where relevant.

Table 2.

Omics and molecular markers—new and existing data in the ATHLETE cohorts

Omics	Data source	Age/matrix	Cohorts
Genotypic variation (genome-wide)	Existing data	Any	BiB, INMA, Gen R, HELIX
DNA methylation (genome-wide)[a]	Existing data	Placenta	INMA, EDEN, SEPAGES, PELAGIE, BiSC
		Cord blood	BiB, INMA, Gen R, Gen R Next, ENVIRONAGE, Piccolipiù
		Childhood blood	HELIX subcohort, Gen R
		Infant saliva	NINFEA
Transcriptomics (genome-wide)[b]	Existing data	Placenta	ENVIRONAGE
		Cord blood	ENVIRONAGE
		Childhood blood	HELIX subcohort, Gen R
miRNA expression (genome-wide)	Existing data	Childhood blood	HELIX subcohort
Metabolomics[c]	Existing data	Cord blood	BiB, INMA, Rhea, Gen R Next, Piccolipiù, ENVIRONAGE, PELAGIE
		Infant blood	BiB
		Childhood blood	HELIX subcohort, Gen R
		Childhood urine	HELIX subcohort
	New measurements ATHLETE	Adolescent blood	HELIX subcohort
Microbiome[d]	Existing data	Birth meconium	SEPAGES, TNG
		Infancy stool (at repeated times)	SEPAGES, Gen R Next
		Childhood stool	Gen R, INMA
	New measurements ATHLETE	Adolescent stool	HELIX subcohort
Candidate proteins	Existing data	Childhood plasma	HELIX subcohort
Telomere length	Existing data	Placenta	ENVIRONAGE
		Cord blood	ENVIRONAGE, Piccolipiù
		Childhood blood	HELIX subcohort
	New measurements ATHLETE	Adolescent blood	HELIX subcohort

[a]Methylation platforms include Illumina Infinium HumanMethylation450 (450K) BeadChip array and Infinium MethylationEPIC array.

[b]Transcriptomics platforms include whole human genome 8 × 60 K and HTAv2 microarrays, and mRNAseq.

[c]Metabolomics platforms (existing data) include ^{1}H NMR, targeted LC-MS/MS with AbsoluteIDQ p180 kit (Biocrates Life Sciences AG, Innsbruck, Austria), targeted LC-MS/MS with Helmuth C 2012 method, untargeted LC-MS with metabolon.

[d]Microbiome platforms include 16S rRNA gene sequencing (existing data) and shotgun metagenomics (new ATHLETE data).

BiSC indicates Barcelona Life Study Cohort; DNA, deoxyribonucleic acid; EDEN, Etude des Determinants pre et postnatals du developpement et de la sante de l'Enfant; ENVIRONAGE, environmental influence on early ageing; Gen R, Generation R cohort; Gen R Next, Generation R Next cohort; INMA, INfancia y Medio Ambiente; LC-MS/MS, liquid chromatography–mass spectrometry; miRNA, micro ribonucleic acid; mRNAseq, messenger ribonucleic acid sequencing; NINFEA, Nascita e Infanzia: gli Effetti dell'Ambiente; PELAGIE, Perturbateurs Endocriniens: Étude Longitudinale sur les Anomalies de la Grossesse, l'Infertilité et l'Enfance; RHEA, Crete Mother Child Cohort; rRNA, ribosomal ribonucleic acid; SEPAGES, Suivi de l'Exposition à la Pollution Atmosphérique durant la Grossesse et Effets sur la Santé; TNG, CELSPAC The Next Generation cohort.

Interventions to improve the personal exposome (WP7)

Effective preventive actions are needed to reduce the health and economic burden of the harmful environmental exposures. ATHLETE will demonstrate the development and evaluation of effective and scalable interventions to improve the urban and the chemical exposome. By developing interventions in close partnership with communities and key stakeholders, we will ensure that these interventions are both acceptable and feasible, thus increasing the likelihood of rapid translation into practice.

For the urban exposome, we will focus on primary school-aged children. The urban environment is a source of physical, chemical, and behavioral exposures (e.g., pollution, lack of green space, noise, physical activity), all of which have been associated with a variety of child health outcomes.[12,95–98] Schools are often urban exposome "hotspots" located in areas of high pollution or noise,[98,99] compounded by high levels of car use during the "school run."[100] We will work with schools in Barcelona and Bradford to co-produce interventions using citizen science and established co-production activities. A co-produced design phase will evaluate options for intervention, for example, co-designing alternative routes to school to reduce the harmful urban exposures (e.g., noise, traffic) and increase beneficial exposures (e.g., green space), working with schools and communities to implement activities designed to reduce traffic around the school (e.g., restricting traffic), reduce emissions (e.g., anti-idling campaigns), and promotion of active travel (e.g., walking buses). We will then conduct an evaluation of the co-produced interventions. Pre- and post-intervention outcomes (6 months follow-up) will be evaluated based on quantitative measures in 100 participants and will include exposure (e.g., indoor and outdoor air quality, noise), health-related behaviors (e.g., physical activity), and wellbeing (mental health, self-perceived health). Child-completed questionnaires will assess mental and physical health, physical activity, sleep quality, and school travel patterns. In a subsample (N = 40), we will conduct real-time personal exposure assessments using the ExpoApp3 (see above), combined with mobile air quality assessment. Detailed intervention logs and qualitative interviews will record activities and barriers/enablers to implementation.

The intervention on the chemical exposome will focus on women of reproductive age, aiming to modify their chemical exposome before pregnancy. Thousands of chemical ingredients, including some with known endocrine-disrupting properties, are used in cosmetics and personal care products (PCP). Early life exposure to chemicals found in PCPs (e.g., parabens and triclosan) are associated with deleterious effects on child respiratory health,[21] child growth,[101] and behavior.[102] Determining effective approaches to reduce chemical burden associated with PCPs is thus highly relevant. After a co-produced design phase, 80 nonpregnant women of reproductive age will be followed for 2 days before and after the 4-day intervention. The intervention consists of stopping the use of PCPs. For the PCPs that cannot be removed, we will provide alternative PCPs that do not contain the chemicals of interest. We will assess: (1) biomarkers of exposure to several phenols, phthalates, and glycol ethers in pools of repeated urine samples collected pre, during, and post intervention; (2) nontargeted markers of effect (e.g., nontargeted metabolomics) in blood samples pre and post intervention; (3) fidelity to the intervention and barriers to implementation in a structured post-intervention questionnaire; and (4) sustainability of behavior change related to PCP use in a questionnaire 2 months post intervention.

Health and economic impact of the exposome (WP8)

Health impact assessment (HIA) is a crucial tool to translate the knowledge generated from environmental health research into information relevant for policy making. So far, several approaches have been used in HIA, including those developed in the context of the Global Burden of Disease project (e.g., [103]) and in our own assessment of the environmental burden of childhood disease.[104] These assessments, however, are limited to the consideration of environmental factors with a strong level of evidence such as particulate matter, lead, and radon. ATHLETE will employ a weight of evidence approach to calculate health and economic impact of a wider set of key chemical and urban exposures possibly or more certainly related to child health. To achieve this, a plausibility database will synthesize the overall level of evidence regarding the effect of many environmental factors (urban exposome and chemicals) on child health, incorporating all mechanistic, animal- and human-based evidence. After classifying the overall level of evidence (from unlikely to very likely), our health impact estimation will then consider associations classified as "likely" or "very likely," weighting each impact estimate by the corresponding level of evidence (e.g., [105]). Exposure-response functions will be taken from the existing evidence, prioritizing meta-analyses done in children, if available. This will lead to an estimation of the impact of several components of the exposome, including urban exposures (particulate matter, noise, green space) and chemicals (lead, mercury, organophosphate pesticides, polybrominated flame retardants, and, depending on the estimated level of evidence, PFASs, bisphenol A, phthalates, triclosan). Health impacts will be calculated based on biomonitoring data collected in ATHLETE and on representative national consumption and chemical concentration surveys whenever available. This impact will be formulated in terms of attributable disease cases, Disability-Adjusted Life Years (DALYs), and Euros. Economic costs will take into account both direct and indirect tangible as well as intangible costs. Since, for many diseases, costs are expected to be country-specific, we will attempt deriving such country-specific estimates.

Dissemination and exploitation towards policy intervention (WP9)

Efforts to translate evidence into practice often fail because researchers have not understood, nor taken into account, complex contextual factors, because they are lacking capacities to engage relevant stakeholders, or because effect estimates (relative risks) remain an abstract notion without direct public health meaning. Rapid translation of evidence into practice will require engaging communities, regulators and decision-makers across many components of the exposome from the earliest stages of the project, and effective tailoring of dissemination strategies and key messages for different audiences in multiple languages. ATHLETE contains a WP dedicated to dissemination, including engagement channels and activities tailored to the specificities of stakeholders, policymakers, or the general public. Particular emphasis will be placed on translating the project developments and findings from all other WPs into accessible knowledge on the long-term health impacts of chronic exposure to environmental factors during the critical early life stages and on the specific contributions of the exposome compared with more traditional environmental health studies. An intervention toolkit for communities (i.e., schools, clinicians) and policymakers will be developed and promoted together with the use of the HIA estimates in the design of the environmental health, chemical safety, and urban and transport policies. Stakeholder engagement will particularly focus on noncommunicable disease and health-affected groups, as well as the environmental health community, to understand and use the ATHLETE online toolbox.

ATHLETE online toolbox

All parts of the work described above will provide input to the ATHLETE online toolbox that will ensure that exposome data and tools are not only developed and used within the project but will be available to researchers and policymakers long after the project has finished. This toolbox will include the FAIR data infrastructure, searchable results catalogs, approved analysis pipelines and protocols for different research areas, the EXPOapp3, HIA and intervention toolkits, e-learning tutorials, and policy recommendations. Data sharing and access procedures will be developed during the project and will form part of the toolbox. The online toolbox will be implemented in compliance with the General Data Protection Regulation (EU 2016/679).

Strengths and limitations

ATHLETE incorporates existing exposome data resources, the existing European network of birth cohort studies and harmonized data platform, and pilot work in exposome methodology (e.g., statistical methods,[36,37] exposome variability[31,32,106]), which provide the project with a base of data, knowledge, and solid collaborations. By focusing on the early part of the life course and on the early signs of health damage before the onset of disease, the project is of high relevance for prevention. Also, ATHLETE focuses on pollutants and risk factors that are widespread in the general population and of regulatory relevance: air pollution, noise, lack of green space, heavy metals, pesticides, endocrine disruptors. This broad range of environmental stressors, together with information on living and social environments and on personal habits and behaviors, offers unique data to study the complexities of the exposome. The main advances that ATHLETE will make to the application of the exposome concept can be summarized as follows:

- The assembly of a large, harmonized, prospective exposome cohort and FAIR data infrastructure for the early part of the life course will be a major step forward compared with the relatively small and scattered current data sets and provide a sustainable platform for future early life exposome research. Of specific relevance here is our expansion of the rich HELIX subcohort exposome database with a new follow-up data point and new measurements of exposures, omics, and health outcomes.
- The combination of targeted and nontargeted biomonitoring techniques for the measurement of chemical pollutants in the same subjects will be a powerful approach to evaluate their relevance for exposome research.
- Within-subject biospecimen pooling will provide greater accuracy in the estimates of long-term exposure to nonpersistent chemicals than has previously been achieved.
- The deployment of personal monitoring approaches in entire cohorts will improve the accuracy of external, urban exposome assessments; this was previously limited to small validation and panel studies.
- The focus on new biological pathways and approaches, such as the microbiome, placental epigenetics, composite measures for aging and stress pathways and cross-omics risk prediction, in longitudinal datasets, will push the integration of omics data into environmental health studies beyond the state-of-the-art, which is currently largely limited to single exposures and single omics layers, often in small studies.
- New statistical and bioinformatics strategies will tackle the next set of analytical exposome challenges, including the evaluation of longitudinal exposome-health associations, the estimation of combined effects of exposures, causal structure models, and integration of cross-omics data.
- The documentation of systematic exposome-health relationships will move far beyond the traditional

"one-exposure-one-disease" approach, and include novel imaging-based outcome assessments during very early organ development, as well as the adolescent period for which knowledge is currently limited.
- The development of acceptable and feasible interventions to reduce personal exposures to the harmful effects of both the urban and the chemical exposome, and the health and economic impact assessment of a wide set of key chemical and urban exposures, will be important for translation of exposome results into practical recommendations.

Main challenges relate firstly to the many methodological issues inherent to the exposome concept, such as temporal exposure variability, differential measurement errors, mixture effects, false-positive findings, statistical power, and absence of causal structure in untargeted analyses, as discussed in detail above. Furthermore, the translation of exposome tools and findings to communities, stakeholders, and decision-makers is challenging due to the many complexities of the exposome; ATHLETE will make specific efforts to explain the unique features of exposome research and how they may contribute to a better understanding of the impact of environment on disease risk, and to the development of better prevention strategies. Lastly, ensuring the long-term sustainability of tools and data beyond the project's 5-year duration is an important challenge. Construction of the FAIR data infrastructure and online toolbox are aimed at ensuring such long-term sustainability, but future funding of these resources will need to be secured.

Ultimately, combining early life exposome data as gathered in this project, with data on the adult exposome gathered in other projects, would allow the study of how the exposome during different life stage affects disease trajectories spanning the entire life course. The European Human Exposome Network is in a unique position to initiate a platform for such future lifecourse exposome research.

Conclusions

The ATHLETE project has a strong focus on the vulnerable early stages of the life course. It will continue the implementation of the exposome in early life by developing a sustainable cohort data infrastructure, improving tools for exposure assessment and statistical analysis, generating longitudinal evidence linking the exposome to child and adolescent health, and translating exposome knowledge into policy. The assembly of data, tools and results in an openly accessible toolbox will lead to great opportunities for researchers, policymakers, and other stakeholders beyond the duration of the project. The results will help us to better understand health damage from environmental exposures and their mixtures from the earliest parts of the life course onward and will highlight opportunities for policy actions towards prevention and enhanced protection, including within the EU's Green Deal and Chemicals Strategy for Sustainability.

Collaboration

The authors encourage interested researchers to contact them to set up collaborations.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 874583—the Advancing Tools for Human Early Lifecourse Exposome Research and Translation (ATHLETE) project. This publication reflects only the authors' view and the European Commission is not responsible for any use that may be made of the information it contains.

Dr. Dadvand is funded by a Ramón y Cajal fellowship (RYC-2012-10995) awarded by the Spanish Ministry of Economy and Finance. Drs. Casas and Guxens are funded by the Instituto de Salud Carlos III (MS16/00128, CPII18/00018). Drs. Chatzi and Conti were supported by the National Institute of Environmental Health Sciences (R21ES029681, R01ES029944, R01ES030364, R01ES030691, and P30ES007048). Additional funding from National Institutes of Health supported Dr. Conti (P01CA196569, R01CA140561) and Dr. Stratakis (P30DK048522). Investigators Drs. McEachan and Wright receive funding from the National Institute for Health Research under its Applied Research Collaboration Yorkshire and Humber. Dr. Jaddoe received funding from a Consolidator Grant from the European Research Council (ERC-2014-CoG-648916). Dr. Duijts received funding from the European Union's Horizon 2020 co-funded programme European Research Area Net on Biomarkers for Nutrition and Health (European Research Area Healthy Diet for a Healthy Life) (Early life programming of childhood health project [number 696295; 2017], ZonMW, The Netherlands [number 529051014; 2017]). Dr. Guxens received funding from the Agence Nationale de Securite Sanitaire de l'Alimentation de l'Environnement et du Travail (EST-18 RF-25).
The views expressed are those of the author(s) and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care.

Conflicts of interest statement

The authors declare that they have no conflicts of interest with regard to the content of this report.

ACKNOWLEDGMENTS

ISGlobal acknowledges support from the Spanish Ministry of Science, Innovation and Universities through the "Centro de Excelencia Severo Ochoa 2019–2023" Program (CEX2018-000806-S) and support from the Generalitat de Catalunya through the Centres de Recerca de Catalunya Program. We acknowledge collaboration with European projects, specifically: Human Early Life Exposome (HELIX) (FP7 grant agreement number 308333), LifeCycle (H2020 grant agreement number 733206), and Connecting Europe and Canada in personalized health (EUCAN-Connect) (H2020 Grant Agreement number 824989). The Born in Bradford (BiB) cohort is only possible because of the enthusiasm and commitment of the Children and Parents in BiB. We are grateful to all the participants, health professionals, and researchers who have made BiB happen. The BiB cohort is only possible because of the enthusiasm and commitment of the children and parents in BiB. We are grateful to all the participants, health professionals, schools, and researchers who have made BiB happen. BiB has received funding from the Wellcome Trust (101597), a joint grant from the UK Medical Research Council and UK Economic and Social Science Research Council (MR/N024391/1), a British Heart Foundation Clinical Study grant (CS/16/4/32482). 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. We are grateful to all the participating families in Norway who take part in this ongoing cohort study. The Danish National Birth Cohort was established with a significant grant from the Danish National Research Foundation. Additional support was obtained from the Danish Regional Committees, the Pharmacy Foundation, the Egmont Foundation, the March of Dimes Birth Defects Foundation, the Health Foundation, and other minor grants. The Danish National Birth Cohort Biobank has been supported by the Novo Nordisk Foundation and the Lundbeck Foundation. Follow-up of mothers and children have been supported by the Danish Medical Research Council (SSVF 0646, 271-08-0839/06-066023, O602-01042B, 0602-02738B);

the Lundbeck Foundation (195/04, R100-A9193); The Innovation Fund Denmark 0603-00294B (09-067124); the Nordea Foundation (02-2013-2014); Aarhus Ideas (AU R9-A959-13-S804); University of Copenhagen Strategic Grant (IFSV 2012); and the Danish Council for Independent Research (DFF—4183-00594 and DFF—4183-00152). The general design of the Generation R Study is made possible by financial support from the Erasmus MC, University Medical Center, Rotterdam, Erasmus University Rotterdam, Netherlands Organization for Health Research and Development (ZonMw), Netherlands Organisation for Scientific Research (NWO), Ministry of Health, Welfare and Sport, and Ministry of Youth and Families. This project received funding from the European Union's Horizon 2020 research and innovation programme (LIFECYCLE, grant agreement number 733206, 2016; EUCAN-Connect grant agreement number 824989; ATHLETE, grant agreement number 874583). We gratefully acknowledge the contribution of participants, research collaborators, general practitioners, hospitals, midwives, and pharmacies in Rotterdam. The INfancia y Medio Ambiente cohort (INMA) Sababell study was funded by grants from Instituto de Salud Carlos III (Red INMA G03/176; CB06/02/0041; PI041436; PI081151 incl. Fondo Europeo de Desarollo Regional (FEDER) funds; PI12/01890 incl. FEDER funds; CP13/00054 incl. FEDER funds; PI15/00118 incl. FEDER funds; CP16/00128 incl. FEDER funds; PI16/00118 incl. FEDER funds; PI16/00261 incl. FEDER funds; PI17/01340 incl. FEDER funds, PI18/00547 incl. FEDER funds, PI20/01695 incl. FEDER funds), Centro de Investigación Biomédica en Red - Epidemiología y Salud Pública (CIBERESP), Generalitat de Catalunya-Comissió Interdepartamental de Recerca i Innovació Tecnològia 1999SGR 00241, Generalitat de Catalunya-Agencia de Gestión de Ayudas Universitarias y de Investigación (2009 SGR 501, 2014 SGR 822), Fundació La Marató de TV3 (090430), Spanish Ministry of Economy and Competitiveness (SAF2012-32991 incl. FEDER funds), Agence Nationale de Securite Sanitaire de l'Alimentation de l'Environnement et du Travail (1262C0010; EST-2016 RF-21, EST-19 RF-04), and the European Commission (261357, 308333, 603794 and 634453). The INMA Valencia study was supported by grants from Instituto de Salud Carlos III (FIS-FEDER: 13/2032, 13/1944, 14/00891, 16/1288, 17/00663, and 19/1338; Miguel Servet-FSE: CP15/0025 and MSII16/00051, 00051, and MSII20/0006), Alicia Koplowitz Foundation 2017, Generalitat Valenciana (AICO/2020/285), and Spanish Association Against Cancer (AECC) "Seed Ideas" 2019 (IDEAS19098LOPE). INMA Gipuzkoa was funded by grants from Instituto de Salud Carlos III (FIS-PI06/0867, FIS-PI09/00090, FIS-PI13/02187, and FIS-PI18/01142 incl. FEDER funds), CIBERESP, Department of Health of the Basque Government (2005111093, 2009111069, 2013111089, and 2015111065), and the Provincial Government of Gipuzkoa (DFG06/002, DFG08/001, and DFG15/221) and annual agreements with the municipalities of the study area (Zumarraga, Urretxu, Legazpi, Azkoitia y Azpeitia y Beasain). The Sepages cohort would like to thank the Grenoble University Hospital (CHU-GA) biobank (bb-0033-00069). We thank the Sepages Study group (E. Eyriey, P. Hoffmann, E. Hullo, J. Lepeule, C. Llerena, S. Lyon-Caen, X. Morin, A. Morlot, C. Philippat, I. Pin, J. Quentin, V. Siroux, R. Slama) and the participants of the Suivi de l'Exposition à la Pollution Atmosphérique durant la Grossesse et Effets sur la Santé (SEPAGES) study. We acknowledge support from Auvergne-Rhône-Alpes Région, Soutien aux coopérations universitaires et scientifiques internationales fund to support collaborations between Catalunya and Auvergne-Rhône-Alpes. SEPAGES cohort was supported by the European Research Council (consolidator grant N 311765-E-DOHaD, Principal Investigator [PI], R. Slama), by French National Research Agency (ANR), the ANR (Pregnancy, Air Pollution, Epigenetics and Respiratory health project ANR-12-PDOC-0029-01, PI, J. Lepeule; A Longitudinal Analysis of Effects of Early Life Exposure to Phenols on Health in Humans project, 14-CE21-0007-01, PI, R. Slama; Gut Microbiota in early childhood and Maternal Environmental exposures project, PI, R. Slama; Prenatal Exposure to Tobacco smoking and Air Pollution and Effects on offspring respiratory and neurodevelopmental outcomes ANR 18-CE36-005, PI, J. Lepeule). The Central European Longitudinal Study of Pregnacy and Childhood: The Next Generation (TNG) cohort is supported by the RECETOX research infrastructure (the Ministry of Education, Youth and Sports of the Czech Republic: LM2018121) and the CETOCOEN EXCELLENCE Teaming 2 project of European Union (EU) Horizon 2020 (857560) and the MEYS of the Czech Republic (02.1.01/0.0/0.0/18_ 046/0015975). The Barcelona Life Study Cohort has received funding from European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement number (785994) (Prenatal exposure to urban air pollution and pre and post-natal brain development project) and from the Health Effects Institute (HEI) under Agreement number 4959-RFPA15-1/18-1 (FRONTIER project). Genotyping in the HELIX study was supported by project PI17/01225 (Instituto de Salud Carlos III, co-funded by European Union [European Regional Development Fund], "A way to make Europe") and the Centro Nacional de Genotipado-CEGEN (PRB2-ISCIII). The Nascita e Infanzia: gli Effetti dell'Ambiente cohort was partially funded by the Compagnia SanPaolo Foundation and the Piedmont Region. The Piccolipiù project was funded by the Italian National Center for Disease Prevention and Control (National Centre for Disease Prevention and Control grants years 2010 and 2014) and by the Italian Ministry of Health (art 12 and 12 bis D.lgs 502/92). The Perturbateurs Endocriniens: Étude Longitudinale sur les Anomalies de la Grossesse, l'Infertilité et l'Enfance study is supported by National Institute of Health and Medical Research and has received multiple funds from the ANR, the French Agency for Food, Environmental and Occupational Health and Safety (ANSES), the Fondation de France, the National Institute for Public Health Surveillance (InVS), the French, Ministry of Labor, the French Ministry of Health, and the French Ministry of Ecology. DataSHIELD is funded under a group of projects that underpin a program of development and application of secure methods for co-analysis, data sharing, and visualization in the Population Health Sciences Institute at Newcastle University (United Kingdom). These include: the "Connected Health Cities project" (North East and North Cumbria) funded by the UK Department of Health (RES/0150/7943/202); the "EUCanConnect project" (European Commission H2020 Flagship Collaboration with Canada); the "58FORWARDS project" (Fostering new Opportunities for Researchers via Wider Access to Research Data and Samples) funded jointly by the Wellcome Trust and the Medical Research Council (108439/Z/15/Z); and the "METADAC project" (Managing Ethico-social, Technical and Administrative issues in Data ACcess) funded jointly by the Medical Research Council, the Wellcome Trust, and the Economic and Social Research Council (MR/N01104X/1 and MR/N01104X/2). We acknowledge the Molecular Genetics Information System team, including Fleur Kelpin, Tommy de Boer, Mariska Slofstra, Connor Stroomberg, Jelmer Veen, Jeroen van Veen, Fernanda de Andrade, Marije van der Geest, Dieuwke Roelofs-Prins, Dennis Hendriksen, Bart Charbon, Joeri van der Velde, Max Postema, Erik Schaberg, Christiaan Hilbrands, Alexander Kellmann, and Luuk Dijkhuis.

References

1. Kuh D; New Dynamics of Ageing (NDA) Preparatory Network. A life course approach to healthy aging, frailty, and capability. *J Gerontol A Biol Sci Med Sci*. 2007;62:717–721.

2. Blane D, Kelly-Irving M, D'Errico A, Bartley M, Montgomery S. Social-biological transitions: how does the social become biological? *Longitud Life Course Stud Int J*. 2013;3:136–146.
3. Godfrey KM, Gluckman PD, Hanson MA. Developmental origins of metabolic disease: life course and intergenerational perspectives. *Trends Endocrinol Metab*. 2010;21:199–205.
4. Durmuş B, Heppe DH, Taal HR, et al. Parental smoking during pregnancy and total and abdominal fat distribution in school-age children: the Generation R Study. *Int J Obes (Lond)*. 2014;38:966–972.
5. Taal HR, de Jonge LL, van Osch-Gevers L, et al. Parental smoking during pregnancy and cardiovascular structures and function in childhood: the Generation R Study. *Int J Epidemiol*. 2013;42:1371–1380.
6. Vogelezang S, Santos S, van der Beek EM, et al. Infant breastfeeding and childhood general, visceral, liver, and pericardial fat measures assessed by magnetic resonance imaging. *Am J Clin Nutr*. 2018;108:722–729.
7. Stratakis N, Roumeliotaki T, Oken E, et al. Fish intake in pregnancy and child growth: a pooled analysis of 15 European and US birth cohorts. *JAMA Pediatr*. 2016;170:381–390.
8. Morgen CS, Bjørk C, Andersen PK, Mortensen LH, Nybo Andersen AM. Socioeconomic position and the risk of preterm birth–a study within the Danish national birth cohort. *Int J Epidemiol*. 2008;37:1109–1120.
9. Sunyer J, Esnaola M, Alvarez-Pedrerol M, et al. Association between traffic-related air pollution in schools and cognitive development in primary school children: a prospective cohort study. *PLoS Med*. 2015;12:e1001792.
10. 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.
11. Zare Sakhvidi MJ, Zare Sakhvidi F, Mehrparvar AH, Foraster M, Dadvand P. Association between noise exposure and diabetes: a systematic review and meta-analysis. *Environ Res*. 2018;166:647–657.
12. McEachan RRC, Yang TC, Roberts H, et al. Availability, use of, and satisfaction with green space, and children's mental wellbeing at age 4 years in a multicultural, deprived, urban area: results from the Born in Bradford cohort study. *Lancet Planet Health*. 2018;2:e244–e254.
13. Dadvand P, Wright J, Martinez D, et al. Inequality, green spaces, and pregnant women: roles of ethnicity and individual and neighbourhood socioeconomic status. *Environ Int*. 2014;71:101–108.
14. Dadvand P, Nieuwenhuijsen MJ, Esnaola M, et al. Green spaces and cognitive development in primary schoolchildren. *Proc Natl Acad Sci U S A*. 2015;112:7937–7942.
15. de Keijzer C, Tonne C, Basagaña X, et al. Residential surrounding greenness and cognitive decline: a 10-year follow-up of the Whitehall II cohort. *Environ Health Perspect*. 2018;126:077003.
16. Agay-Shay K, Martinez D, Valvi D, et al. Exposure to endocrine-disrupting chemicals during pregnancy and weight at 7 years of age: a multi-pollutant approach. *Environ Health Perspect*. 2015;123:1030–1037.
17. Manzano-Salgado CB, Casas M, Lopez-Espinosa MJ, et al. Prenatal exposure to perfluoroalkyl substances and cardiometabolic risk in children from the Spanish INMA birth cohort study. *Environ Health Perspect*. 2017;125:097018.
18. Alderete TL, Jin R, Walker DI, et al. Perfluoroalkyl substances, metabolomic profiling, and alterations in glucose homeostasis among overweight and obese Hispanic children: a proof-of-concept analysis. *Environ Int*. 2019;126:445–453.
19. Vernet C, Pin I, Giorgis-Allemand L, et al; EDEN Mother–Child Cohort Study Group. In utero exposure to select phenols and phthalates and respiratory health in five-year-old boys: a prospective study. *Environ Health Perspect*. 2017;125:097006.
20. Gascon M, Casas M, Morales E, et al. Prenatal exposure to bisphenol A and phthalates and childhood respiratory tract infections and allergy. *J Allergy Clin Immunol*. 2015;135:370–378.
21. Agier L, Basagaña X, Maitre L, et al. Early-life exposome and lung function in children in Europe: an analysis of data from the longitudinal, population-based HELIX cohort. *Lancet Planet Health*. 2019;3:e81–e92.
22. Nieuwenhuijsen MJ, Agier L, Basagaña X, et al. Influence of the urban exposome on birth weight. *Environ Health Perspect*. 2019;127:47007.
23. Agier L, Basagaña X, Hernandez-Ferrer C, et al. Association between the pregnancy exposome and fetal growth. *Int J Epidemiol*. 2020;49:572–586.
24. Hu H, Zhao J, Savitz DA, Prosperi M, Zheng Y, Pearson TA. An external exposome-wide association study of hypertensive disorders of pregnancy. *Environ Int*. 2020;141:105797.
25. Steer CD, Bolton P, Golding J. Preconception and prenatal environmental factors associated with communication impairments in 9 year old children using an exposome-wide approach. *PLoS One*. 2015;10:e0118701.
26. Vrijheid M, Fossati S, Maitre L, et al. Environmental exposures and childhood obesity: an exposome approach. *EnvironHealth Perspect*. 2020;128:67009.
27. Tamayo-Uria I, Maitre L, Thomsen C, et al. The early-life exposome: description and patterns in six European countries. *Environ Int*. 2019;123:189–200;128:67009.
28. Robinson O, Basagaña X, Agier L, et al. The pregnancy exposome: multiple environmental exposures in the INMA-Sabadell birth cohort. *Environ Sci Technol*. 2015;49:10632–10641.
29. Chung MK, Kannan K, Louis GM, Patel CJ. Toward capturing the exposome: exposure biomarker variability and coexposure patterns in the shared environment. *Environ Sci Technol*. 2018;52:8801–8810.
30. Haug LS, Sakhi AK, Cequier E, et al. In-utero and childhood chemical exposome in six European mother-child cohorts. *Environ Int*. 2018;121(Pt 1):751–763.
31. Casas M, Basagaña X, Sakhi AK, et al. Variability of urinary concentrations of non-persistent chemicals in pregnant women and school-aged children. *Environ Int*. 2018;121(Pt 1):561–573.
32. Donaire-Gonzalez D, Curto A, Valentín A, et al. Personal assessment of the external exposome during pregnancy and childhood in Europe. *Environ Res*. 2019;174:95–104.
33. Robinson O, Tamayo I, de Castro M, et al. The urban exposome during pregnancy and its socioeconomic determinants. *Environ Health Perspect*. 2018;126:077005.
34. Montazeri P, Thomsen C, Casas M, et al. Socioeconomic position and exposure to multiple environmental chemical contaminants in six European mother-child cohorts. *Int J Hyg Environ Health*. 2019;222:864–872.
35. Papadopoulou E, Haug LS, Sakhi AK, et al. Diet as a source of exposure to environmental contaminants for pregnant women and children from six European countries. *Environ Health Perspect*. 2019;127:107005.
36. Agier L, Portengen L, Chadeau-Hyam M, et al. A systematic comparison of linear regression-based statistical methods to assess exposome-health associations. *Environ Health Perspect*. 2016;124:1848–1856.
37. Barrera-Gómez J, Agier L, Portengen L, et al. A systematic comparison of statistical methods to detect interactions in exposome-health associations. *Environ Health*. 2017;16:74.
38. Tanner E, Lee A, Colicino E. Environmental mixtures and children's health: identifying appropriate statistical approaches. *Curr Opin Pediatr*. 2020;32:315–320.
39. Agier L, Slama R, Basagaña X. Relying on repeated biospecimens to reduce the effects of classical-type exposure measurement error in studies linking the exposome to health. *Environ Res*. 2020;186:109492.
40. Lau CE, Siskos AP, Maitre L, et al. Determinants of the urinary and serum metabolome in children from six European populations. *BMC Med*. 2018;16:202.
41. Maitre L, Robinson O, Martinez D, et al. Urine metabolic signatures of multiple environmental pollutants in pregnant women: an exposome approach. *Environ Sci Technol*. 2018;52:13469–13480.
42. Plusquin M, Chadeau-Hyam M, Ghantous A, et al. DNA methylome marks of exposure to particulate matter at three time points in early life. *Environ Sci Technol*. 2018;52:5427–5437.
43. Robinson O, Keski-Rahkonen P, Chatzi L, et al. Cord blood metabolic signatures of birth weight: a population-based study. *J Proteome Res*. 2018;17:1235–1247.
44. Cadiou S, Bustamante M, Agier L, et al. Using methylome data to inform exposome-health association studies: an application to the identification of environmental drivers of child body mass index. *Environ Int*. 2020;138:105622.
45. Vives-Usano M, Hernandez-Ferrer C, Maitre L, et al. In utero and childhood exposure to tobacco smoke and multi-layer molecular signatures in children. *BMC Med*. 2020;18:243.
46. Stratakis N, Conti DV, Jin R, et al. Prenatal exposure to perfluoroalkyl substances associated with increased susceptibility to liver injury in children. *Hepatology*. 2020;72:1758–1770.
47. Wright J, Small N, Raynor P, et al; Born in Bradford Scientific Collaborators Group. Cohort profile: the Born in Bradford multi-ethnic family cohort study. *Int J Epidemiol*. 2013;42:978–991.
48. Heude B, Forhan A, Slama R, et al; EDEN mother-child cohort study group. Cohort profile: the EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *Int J Epidemiol*. 2016;45:353–363.
49. Grazuleviciene R, Danileviciute A, Nadisauskiene R, Vencloviene J. Maternal smoking, GSTM1 and GSTT1 polymorphism and susceptibility to adverse pregnancy outcomes. *Int J Environ Res Public Health*. 2009;6:1282–1297.
50. Guxens M, Ballester F, Espada M, et al; INMA Project. Cohort profile: the INMA–INfancia y Medio Ambiente–(environment and childhood) project. *Int J Epidemiol*. 2012;41:930–940.

51. Magnus P, Birke C, Vejrup K, et al. Cohort profile update: the Norwegian mother and child cohort study (MoBa). *Int J Epidemiol*. 2016;45:382–388.
52. Chatzi L, Leventakou V, Vafeiadi M, et al. Cohort profile: the mother-child cohort in crete, Greece (Rhea study). *Int J Epidemiol*. 2017;46:1392k–1393k.
53. 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.
54. Vrijheid M, Slama R, Robinson O, et al. The human early-life exposome (HELIX): project rationale and design. *Environ Health Perspect*. 2014;122:535–544.
55. Jaddoe VW, Mackenbach JP, Moll HA, et al. The generation R study: design and cohort profile. *Eur J Epidemiol*. 2006;21:475–484.
56. Olsen J, Melbye M, Olsen SF, et al. The Danish national birth cohort–its background, structure and aim. *Scand J Public Health*. 2001;29:300–307.
57. Richiardi L, Baussano I, Vizzini L, Douwes J, Pearce N, Merletti F; NINFEA cohort. Feasibility of recruiting a birth cohort through the Internet: the experience of the NINFEA cohort. *Eur J Epidemiol*. 2007;22:831–837.
58. Chevrier C, Limon G, Monfort C, et al. Urinary biomarkers of prenatal atrazine exposure and adverse birth outcomes in the PELAGIE birth cohort. *Environ Health Perspect*. 2011;119:1034–1041.
59. Lyon-Caen S, Siroux V, Lepeule J, et al. Deciphering the impact of early-life exposures to highly variable environmental factors on foetal and child health: design of SEPAGES couple-child cohort. *Int J Environ Res Public Health*. 2019;16:E3888.
60. Janssen BG, Madhloum N, Gyselaers W, et al. Corrigendum cohort profile: the ENVIRonmental influence ON early AGEing (ENVIRONAGE): a birth cohort study. *Int J Epidemiol*. 2017;46:1387m.
61. Farchi S, Forastiere F, Vecchi Brumatti L, et al. Piccolipiù, a multicenter birth cohort in Italy: protocol of the study. *BMC Pediatr*. 2014;14:36.
62. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018.
63. Merino-Martinez R, Norlin L, van Enckevort D, et al. Toward global biobank integration by implementation of the Minimum Information About BIobank Data Sharing (MIABIS 2.0 Core). *Biopreserv Biobank*. 2016;14:298–306.
64. van der Velde KJ, Imhann F, Charbon B, et al. MOLGENIS research: advanced bioinformatics data software for non-bioinformaticians. *Bioinformatics*. 2019;35:1076–1078.
65. Gaye A, Marcon Y, Isaeva J, et al. DataSHIELD: taking the analysis to the data, not the data to the analysis. *Int J Epidemiol*. 2014;43:1929–1944.
66. Wilson RC, Butters OW, Avraam D, et al. DataSHIELD – new directions and dimensions. *Data Sci J*. 2017;16:1–21.
67. Doiron D, Marcon Y, Fortier I, Burton P, Ferretti V. Software application profile: opal and mica: open-source software solutions for epidemiological data management, harmonization and dissemination. *Int J Epidemiol*. 2017;46:1372–1378.
68. Vernet C, Philippat C, Calafat AM, et al. Within-day, between-day, and between-week variability of urinary concentrations of phenol biomarkers in pregnant women. *Environ Health Perspect*. 2018;126:037005.
69. Perrier F, Giorgis-Allemand L, Slama R, Philippat C. Within-subject pooling of biological samples to reduce exposure misclassification in biomarker-based studies. *Epidemiology*. 2016;27:378–388.
70. Pourchet M, Debrauwer L, Klanova J, et al. Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: from promises to challenges and harmonisation issues. *Environ Int*. 2020;139:105545.
71. Dennis KK, Marder E, Balshaw DM, et al. Biomonitoring in the era of the exposome. *Environ Health Perspect*. 2017;125:502–510.
72. Ng CA, von Goetz N. The global food system as a transport pathway for hazardous chemicals: the missing link between emissions and exposure. *Environ Health Perspect*. 2017;125:1–7.
73. van der Voet H, Kruisselbrink JW, de Boer WJ, et al. The MCRA toolbox of models and data to support chemical mixture risk assessment. *Food Chem Toxicol*. 2020;138:111185.
74. Sprong C, Crépet A, Metruccio F, et al. Cumulative dietary risk assessment overarching different regulatory silos using a margin of exposure approach: a case study with three chemical silos. *Food Chem Toxicol*. 2020;142:111416.
75. van Klaveren JD, Kruisselbrink JW, de Boer WJ, et al. Cumulative dietary exposure assessment of pesticides that have acute effects on the nervous system using MCRA software. *EFSA Support Publ*. 2019;16:1708E.
76. Chen S, Bowman FD. A novel support vector classifier for longitudinal high-dimensional data and its application to neuroimaging data. *Stat Anal Data Min*. 2011;4:604–611.
77. Wang L, Zhou J, Qu A. Penalized generalized estimating equations for high-dimensional longitudinal data analysis. *Biometrics*. 2012;68:353–360.
78. Tsagris M, Lagani V, Tsamardinos I. Feature selection for high-dimensional temporal data. *BMC Bioinformatics*. 2018;19:17.
79. Tenenhaus M, Tenenhaus A, Groenen PJF. Regularized generalized canonical correlation analysis: a framework for sequential multiblock component methods. *Psychometrika*. 2017;82:737–777.
80. Tenenhaus A, Philippe C, Guillemot V, Le Cao KA, Grill J, Frouin V. Variable selection for generalized canonical correlation analysis. *Biostatistics*. 2014;15:569–583.
81. Tenenhaus A, Tenenhaus M. Regularized generalized canonical correlation analysis. *Psychometrika*. 2011;76:257–284.
82. Wu C, Zhou F, Ren J, Li X, Jiang Y, Ma S. A selective review of multilevel omics data integration using variable selection. *High Throughput*. 2019;8:E4.
83. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. *Int J Epidemiol*. 2013;42:1511–1519.
84. van der Laan MJ, Starmans RJCM. Entering the era of data science: targeted learning and the integration of statistics and computational data analysis. *Adv Stat*. 2014;2014:1–19.
85. Marcon Y, Bishop T, Avraam D, et al. Orchestrating privacy-protected big data analyses of data from different resources with R and DataSHIELD. *PLoS Comput Biol*. 2021;17:e1008880.
86. Hernandez-Ferrer C, Wellenius GA, Tamayo I, et al. Comprehensive study of the exposome and omic data using rexposome bioconductor packages. *Bioinformatics*. 2019;35:5344–5345.
87. Tenenhaus A, Guillemot V. *The RGCCA Package for Regularized/Sparse Generalized Canonical Correlation Analysis*. 2017:1–36. Available at: https://cran.r-project.org/package=RGCCA. Accessed 15 June 2020.
88. Vinken M. The adverse outcome pathway concept: a pragmatic tool in toxicology. *Toxicology*. 2013;312:158–165.
89. Leist M, Ghallab A, Graepel R, et al. Adverse outcome pathways: opportunities, limitations and open questions. *Arch Toxicol*. 2017;91:3477–3505.
90. Bui DS, Lodge CJ, Burgess JA, et al. Childhood predictors of lung function trajectories and future COPD risk: a prospective cohort study from the first to the sixth decade of life. *Lancet Respir Med*. 2018;6:535–544.
91. Franks PW, Hanson RL, Knowler WC, Sievers ML, Bennett PH, Looker HC. Childhood obesity, other cardiovascular risk factors, and premature death. *N Engl J Med*. 2010;362:485–493.
92. Hardy R, Lawlor DA, Kuh D. A life course approach to cardiovascular aging. *Future Cardiol*. 2015;11:101–113.
93. Juonala M, Magnussen CG, Berenson GS, et al. Childhood adiposity, adult adiposity, and cardiovascular risk factors. *N Engl J Med*. 2011;365:1876–1885.
94. Arango C, Díaz-Caneja CM, McGorry PD, et al. Preventive strategies for mental health. *Lancet Psychiatry*. 2018;5:591–604.
95. Khreis H, Kelly C, Tate J, Parslow R, Lucas K, Nieuwenhuijsen M. Exposure to traffic-related air pollution and risk of development of childhood asthma: a systematic review and meta-analysis. *Environ Int*. 2017;100:1–31.
96. de Bont J, Casas M, Barrera-Gómez J, et al. Ambient air pollution and overweight and obesity in school-aged children in Barcelona, Spain. *Environ Int*. 2019;125:58–64.
97. Forns J, Dadvand P, Esnaola M, et al. Longitudinal association between air pollution exposure at school and cognitive development in school children over a period of 3.5 years. *Environ Res*. 2017;159:416–421.
98. Stansfeld S, Clark C. Health effects of noise exposure in children. *Curr Environ Health Rep*. 2015;2:171–178.
99. Kmietowicz Z. Air pollution: ban schools and other children's facilities from hotspots, say campaigners. *BMJ*. 2018;363:k4489.
100. UK Government. *National Travel Survey Collection. Statistics and Data About the National Travel Survey, Based on a Household Survey to Monitor Trends in Personal Travel*. 2018. Available at: https://www.gov.uk/government/collections/national-travel-survey-statistics. Accessed 11 April 2019.
101. Philippat C, Mortamais M, Chevrier C, et al. Exposure to phthalates and phenols during pregnancy and offspring size at birth. *Environ Health Perspect*. 2012;120:464–470.
102. Philippat C, Nakiwala D, Calafat AM, et al; EDEN Mother–Child Study Group. Prenatal exposure to nonpersistent endocrine disruptors and behavior in boys at 3 and 5 years. *Environ Health Perspect*. 2017;125:097014.

103. Murray CJL, Aravkin AY, Zheng P, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396:1223–1249.
104. Rojas-Rueda D, Vrijheid M, Robinson O, et al. Environmental burden of childhood disease in Europe. *Int J Environ Res Public Health*. 2019;16:E1084.
105. Trasande L, Zoeller RT, Hass U, et al. Burden of disease and costs of exposure to endocrine disrupting chemicals in the European Union: an updated analysis. *Andrology*. 2016;4:565–572.
106. Maitre L, Villanueva CM, Lewis MR, et al. Maternal urinary metabolic signatures of fetal growth and associated clinical and environmental factors in the INMA study. *BMC Med*. 2016;14:177.