



HAL
open science

The LifeCycle Project-EU Child Cohort Network: a federated analysis infrastructure and harmonized data of more than 250,000 children and parents

Vincent W V Jaddoe, Janine F Felix, Anne-Marie Nybo Andersen, Marie-Aline Charles, Leda Chatzi, Eva Corpeleijn, Nina Donner, Ahmed Elhakeem, Johan G Eriksson, Rachel Foong, et al.

► To cite this version:

Vincent W V Jaddoe, Janine F Felix, Anne-Marie Nybo Andersen, Marie-Aline Charles, Leda Chatzi, et al.. The LifeCycle Project-EU Child Cohort Network: a federated analysis infrastructure and harmonized data of more than 250,000 children and parents. *European Journal of Epidemiology*, 2020, 35 (7), pp.709-724. 10.1007/s10654-020-00662-z . hal-02938634

HAL Id: hal-02938634

<https://hal.inrae.fr/hal-02938634v1>

Submitted on 4 Sep 2024

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

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



Distributed under a Creative Commons Attribution 4.0 International License



The LifeCycle Project-EU Child Cohort Network: a federated analysis infrastructure and harmonized data of more than 250,000 children and parents

Vincent W. V. Jaddoe^{1,2} · Janine F. Felix^{1,2} · Anne-Marie Nybo Andersen³ · Marie-Aline Charles^{4,5} · Leda Chatzi⁶ · Eva Corpeleijn⁷ · Nina Donner⁸ · Ahmed Elhakeem^{9,10} · Johan G. Eriksson^{11,12,13,14} · Rachel Foong^{15,16} · Veit Grote¹⁷ · Sido Haakma¹⁸ · Mark Hanson^{19,20} · Jennifer R. Harris^{21,22} · Barbara Heude⁴ · Rae-Chi Huang¹⁵ · Hazel Inskip^{20,23} · Marjo-Riitta Järvelin^{24,25,26,27} · Berthold Koletzko¹⁷ · Deborah A. Lawlor^{9,10,28} · Maarten Lindeboom²⁹ · Rosemary R. C. McEachan³⁰ · Tuija M. Mikkola¹² · Johanna L. T. Nader³¹ · Angela Pinot de Moira³ · Costanza Pizzi³² · Lorenzo Richiardi³² · Sylvain Sebert²⁴ · Ameli Schwalber⁸ · Jordi Sunyer^{33,34,35,36} · Morris A. Swertz^{18,37} · Marina Vafeiadi³⁸ · Martine Vrijheid^{33,34,35} · John Wright³⁰ · Liesbeth Duijts^{1,2} · LifeCycle Project Group

Received: 10 May 2020 / Accepted: 4 July 2020 / Published online: 23 July 2020
© The Author(s) 2020

Abstract

Early life is an important window of opportunity to improve health across the full lifecycle. An accumulating body of evidence suggests that exposure to adverse stressors during early life leads to developmental adaptations, which subsequently affect disease risk in later life. Also, geographical, socio-economic, and ethnic differences are related to health inequalities from early life onwards. To address these important public health challenges, many European pregnancy and childhood cohorts have been established over the last 30 years. The enormous wealth of data of these cohorts has led to important new biological insights and important impact for health from early life onwards. The impact of these cohorts and their data could be further increased by combining data from different cohorts. Combining data will lead to the possibility of identifying smaller effect estimates, and the opportunity to better identify risk groups and risk factors leading to disease across the lifecycle across countries. Also, it enables research on better causal understanding and modelling of life course health trajectories. The EU Child Cohort Network, established by the Horizon2020-funded LifeCycle Project, brings together nineteen pregnancy and childhood cohorts, together including more than 250,000 children and their parents. A large set of variables has been harmonised and standardized across these cohorts. The harmonized data are kept within each institution and can be accessed by external researchers through a shared federated data analysis platform using the R-based platform DataSHIELD, which takes relevant national and international data regulations into account. The EU Child Cohort Network has an open character. All protocols for data harmonization and setting up the data analysis platform are available online. The EU Child Cohort Network creates great opportunities for researchers to use data from different cohorts, during and beyond the LifeCycle Project duration. It also provides a novel model for collaborative research in large research infrastructures with individual-level data. The LifeCycle Project will translate results from research using the EU Child Cohort Network into recommendations for targeted prevention strategies to improve health trajectories for current and future generations by optimizing their earliest phases of life.

Keywords Consortium · Birth cohorts · Exposome · Life course · Non-communicable diseases

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10654-020-00662-z>) contains supplementary material, which is available to authorized users.

✉ Vincent W. V. Jaddoe
v.jaddoe@erasmusmc.nl

Extended author information available on the last page of the article

Rationale

Early life seems to be an important window of opportunity to improve health across the full lifecycle. An accumulating body of evidence suggests that exposure to adverse stressors during early life leads to developmental adaptations, which subsequently affect disease risk in later life [1]. Moreover,

geographical, socio-economic, and ethnic differences are related to health inequalities from early life onwards [1]. These research findings suggest that optimizing early-life conditions has the yet unfulfilled potential to improve life course health trajectories for individuals themselves and also for their offspring through transgenerational effects [2]. A better understanding of the causality, pathways and life course health trajectories explaining associations of early-life stressors with later life disease is urgently needed to translate results from observational studies into population-health prevention strategies.

Many European pregnancy and childhood cohorts have been established over the last years to assess the associations of early life with health across the lifecycle [3]. These cohorts are invaluable resources to obtain insight into societal, environmental, lifestyle and nutrition related determinants that may influence the onset and evolution of risk factors and diseases in later life. Cohort studies that started during pregnancy or early childhood provide the unique opportunity to study the potential for early-life interventions on factors that cannot be easily studied in experimental settings, such as socio-economic, migration, urban environment and lifestyle related determinants. Data from cohort studies can also be used for advanced analytical approaches such as sibling analyses and Mendelian randomization to assess causality of observed associations [4].

The impact of these cohorts and their data could be strongly increased by combining data from different cohorts. Combining data will lead to larger numbers and the opportunity to better identify risk groups and risk factors leading to disease across the lifecycle [3]. Also, it enables research for a better causal understanding and modelling of life course health trajectories. The enormous wealth of high-quality prospective cohort studies enables collaboration at individual participant data level. Meta-analyzing individual participant data has the advantage that it can identify smaller effect estimates, specific subgroups, and mediator effects and, maybe most importantly, capitalizes on existing published and unpublished data. Results from well-performed individual participant data meta-analyses suffer less from publication bias than meta-analyses based on published data. Multiple individual participant data meta-analyses on environmental exposures, lifestyle related and (epi)genetic associations have already been published as part of birth cohort collaborations [5–22].

The LifeCycle Project is a Horizon 2020-funded (2017–2022) international project. The general objective of the LifeCycle Project is to bring together pregnancy and childhood cohort studies into a new, open and sustainable EU Child Cohort Network, to use this network for identification of novel markers of early-life stressors affecting health trajectories throughout the life course, and to translate findings into policy recommendations for targeted prevention

strategies. The overall concepts, design and future perspectives are described in this paper. The logos of the LifeCycle Project are given in Fig. 1.

The EU Child Cohort Network

The EU Child Cohort Network, the main deliverable of the LifeCycle Project, brings together nineteen pregnancy and childhood cohorts. Together, they include more than 250,000 children and their parents (Fig. 2; Table 1). Recruitment to the cohorts of the EU Child Cohort Network began prior to and during pregnancy, as well as in childhood; together, the follow-up of these cohorts span the full life course and contain detailed phenotypic information and biological samples. The research potential of the EU Child Cohort Network is summarized in Table 2. The EU Child Cohort Network should be operational mid-2020. This network is open for other partners with population-based cohorts that started in early life and will be sustainable after the duration of the Horizon 2020 funded LifeCycle Project. The EU Child Cohort Network could contribute to future collaborations between different cohorts.

The LifeCycle Project and its EU Child Cohort Network do not stand on their own. By building on and collaborating with existing initiatives, we will create new synergies and form the basis of future initiatives. These synergies bring together principal investigators and their expertise of several international collaborations. These initiatives include:

- Cohort collaboration and data sharing platforms: BioSHaRe [23], CHICOS [24], DataSHIELD [25], DynaHEALTH [26], EarlyNutrition [27], ENRIECO [28], HELIX [29, 30], InterConnect [31] and NutriMenthe [32] (all EU-FP6, FP7 projects or Horizon2020).
- Genetic and epigenetic collaborations: Early Growth & Longitudinal Epidemiology (EAGLE), Early Growth Genetics (EGG) [33], Pregnancy And Childhood Epigenetics (PACE) [34] (no specific funds for the collaboration).

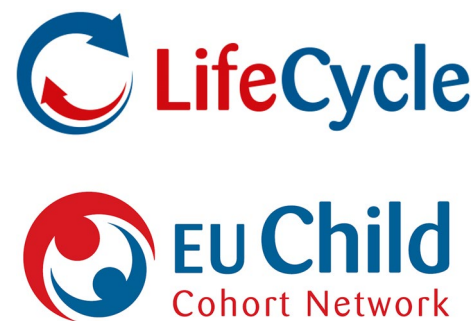


Fig. 1 Logos of the LifeCycle Project and EU Child Cohort Network

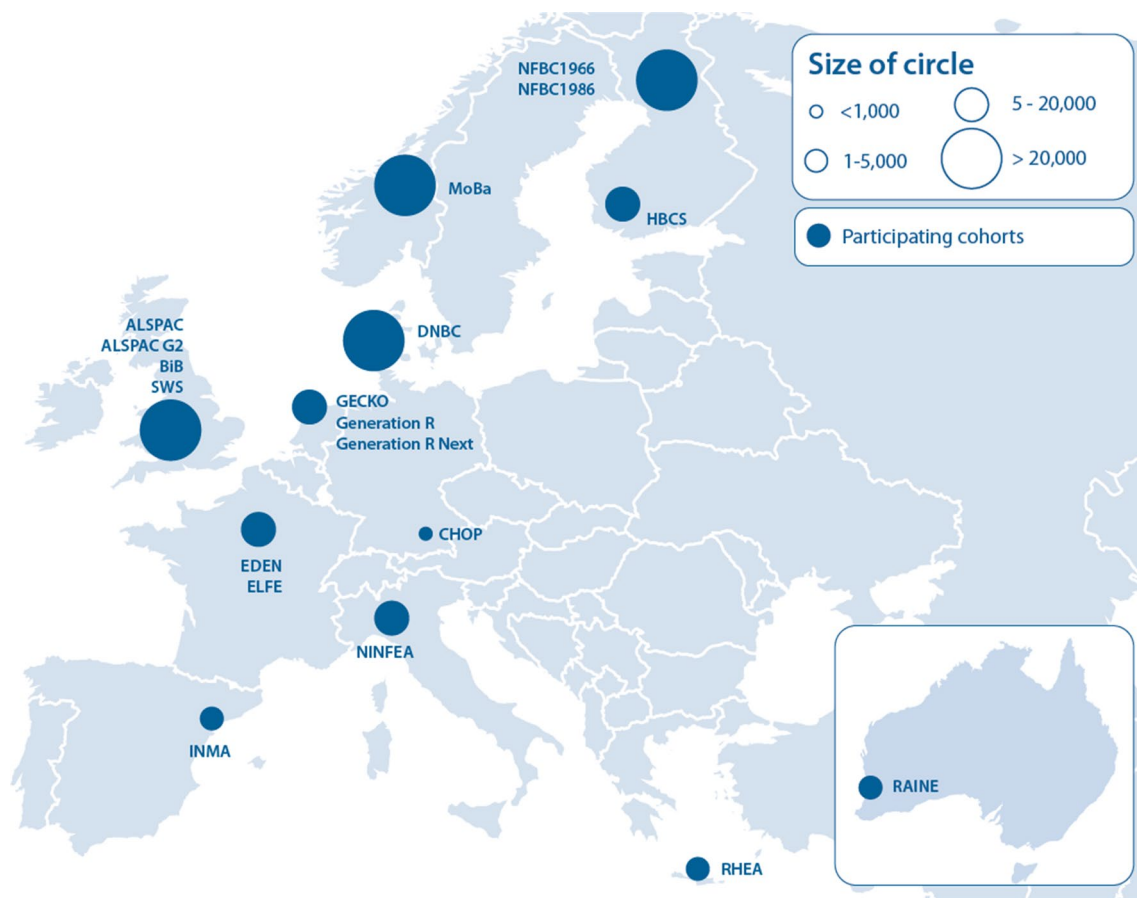


Fig. 2 LifeCycle Project core cohorts that established the EU Child Cohort Network

- E-Learning: Early Nutrition Academy [35] (EU-FP7 project).

Data harmonisation

The LifeCycle Project has developed a harmonized set of variables in each cohort necessary to perform multi-cohort analyses on different research questions. The harmonization work is performed by a data-harmonization group with representatives from each partner or cohort. Based on the primary research focus in the LifeCycle Project, a priority list of variables has been developed for harmonisation. The cohort studies participating in the EU Child Cohort Network will be further enriched with novel harmonized integrated data on early-life stressors related to socio-economic, migration, urban environment and lifestyle determinants, based on data availability within the cohorts and external data from registries [36]. Integrated data will also be used to construct a novel holistic ‘dynamic early-life exposome’ model, which will encompass many human environmental exposures during various stages

of early life [37–40]. The harmonized variables relate to the main research hypotheses (Fig. 3), and include:

- Main exposures: Socioeconomic, migration, urban environment, lifestyle and nutrition related factors, genome-wide association screen;
- Main mediators: Epigenetics, metabolomics, allergy, brain development;
- Main outcomes: Cardio-metabolic (body mass index (BMI), body composition, blood pressure, cardiac structure and function, lipids, insulin, glucose); respiratory (allergy, wheezing, infections, lung function, asthma), mental (behaviour, cognition, education, ASD, ADHD, anxiety, depression);

The availability of these data in different cohorts is given in Table 1.

Table 1 LifeCycle Project cohorts that together form the basis of the EU Child Cohort Network

Cohort, Country (N)	Design, birth years, Follow-up Main early-life stressors	Available mediators	Available outcomes
ALSPAC United Kingdom N = 14,500 [74, 75]	Prospective, 1991–1992 Pregnancy-25 yrs Socio-economic, migration, and life-style determinants, genome wide association screen	Epigenetics Metabolomics Allergy Brain development by MRI	Cardio-metabolic: BMI, blood pressure, cardiac structure and function, lipids, insulin, glucose Respiratory: wheezing, infections, asthma, lung function Mental: behaviour, cognition, education, ASD, ADHD, anxiety, depression
ALSPAC-G2 United Kingdom N = 2000 [76]	Prospective, from 2011 Preconception-2 yrs Socio-economic, migration and life-style determinants	Epigenetics Metabolomics Brain development by ultrasound	Cardio-metabolic: BMI, blood pressure Respiratory: wheezing, asthma Mental: behaviour, cognition
BIB United Kingdom N = 11,000 [77]	Prospective, 2007–2011 Pregnancy-9 yrs Socio-economic, migration, urban environment, and life-style determinants, genome wide association screen	Epigenetics Metabolomics Allergy Brain development by ultrasound	Cardio-metabolic: BMI, blood pressure, lipids, insulin, glucose Respiratory: wheezing, infections, asthma, lung function Mental: behaviour, cognition, education, ASD, ADHD, anxiety, depression
CHOP Germany N = 500 [78]	Prospective, 2002–2004 Pregnancy-11 yrs Socio-economic, life-style determinants, genome wide associa- tion screen	Epigenetics Metabolomics, Allergy	Cardio-metabolic: BMI, blood pressure, cardiac structure and function, lipids, insulin, glucose Respiratory: wheezing, asthma Mental: behaviour, cognition
DNBC Denmark N = 70,000 [79]	Prospective, 1996–2002 Pre-pregnancy-20 yrs Socio-economic, migration, urban environment, and life-style determinants, genome wide association screen	Allergy	Cardio-metabolic: BMI Respiratory: wheezing, infections, asthma, lung function Mental: behaviour, cognition, education, ASD, ADHD, anxiety, depression
EDEN France N = 2000 [80]	Prospective, 2003–2005 Pre-school-15 yrs Socio-economic, migration and life-style determinants	Allergy	Cardio-metabolic: BMI, blood pressure, lipids, insulin, glucose Respiratory: wheezing, lung function, asthma. Mental: behav- iour, cognition, education
ELFE France N = 18,000 [81]	Prospective, 2011 Pre-school-5 yrs Socio-economic, migration, urban environment, life-style determinants	Allergy	Cardio-metabolic: BMI Respiratory: wheezing, infections, asthma Mental: behaviour, cognition
GECKO the Netherlands N = 2500 [82]	Prospective, 2006–2007 Pregnancy-10 yrs Socio-economic, migration, life-style	Allergy	Cardio-metabolic: BMI, blood pressure Respiratory: wheezing, asthma Mental: behaviour, education
Generation R the Netherlands N = 7000 [83, 84]	Prospective, 2002–2006 Pregnancy-17 yrs Socio-economic, migration, urban environment, and life-style determinants, genome wide association screen	Epigenetics Metabolomics Allergy Brain development by ultrasound/MRI	Cardio-metabolic: BMI, blood pressure, cardiac structure and function, lipids, insulin, glucose Respiratory: wheezing, infections, lung function, asthma Mental: behaviour, cognition, education, ASD, ADHD, anxiety, depression
Generation R Next the Netherlands N = 2000	Prospective, 2016–2019 Pre-pregnancy-2 yrs Socio-economic, migration, urban environment, and life-style determinants	Epigenetics Metabolomics Allergy Brain development by ultrasound/MRI	Cardio-metabolic: body mass index, blood pressure, cardiac structure and function, lipids, insulin, glucose Respiratory: wheezing, infections, lung function, asthma Mental: behaviour, cognition

Table 1 (continued)

Cohort, Country (N)	Design, birth years, Follow-up Main early-life stressors	Available mediators	Available outcomes
HBSC Finland N = 13,000 [85]	Longitudinal, 1934–1944 Pregnancy-80 yrs Socio-economic, migration, and life-style determinants, genome wide association screen		Cardio-metabolic: BMI, blood pressure, lipids, insulin, glucose, hypertension, type 2 diabetes, dyslipidaemia Respiratory: asthma, COPD Mental: cognition, psychiatric illness
INMA Spain N = 3500 [86]	Prospective, 1997–2008 Pregnancy-20 yrs Socio-economic, migration, urban environment, and life-style determinants, genome wide association screen	Epigenetics Metabolomics Allergy	Cardio-metabolic: BMI, blood pressure, lipids, insulin, glucose Respiratory: wheezing, respiratory infections, lung function, asthma Mental: behaviour, cognition, ASD, ADHD, anxiety, depression
MoBa Norway N = 90,000 [87]	Prospective, 1999–2008 Pregnancy-14 yrs Socio-economic, urban environment and life-style determinants, genome wide association screen	Epigenetics Allergy Brain development by MRI	Cardio-metabolic: BMI, blood pressure Respiratory: wheezing, respiratory infections, asthma Mental: behaviour, cognition, ASD, ADHD, anxiety, depression
NFBC1966 Finland N = 12,000 [88]	Prospective, 1966 Pregnancy-50 yrs Socio-economic, migration, life-style determinants, genome wide association screen	Epigenetics Metabolomics Allergy Brain development by MRI	Cardio-metabolic: BMI, blood pressure, cardiac structure and function, lipids, insulin, glucose Respiratory: wheezing, respiratory infections, lung function, asthma, COPD Mental: behaviour, cognition, education, ASD, ADHD, anxiety, depression
NFBC1986 Finland N = 9500 [89]	Prospective, 1986 Pregnancy-30 yrs Socio-economic, migration, urban environment, and life-style determinants, genome wide association screen	Epigenetics Metabolomics Allergy Brain development by MRI	Cardio-metabolic: BMI, blood pressure, cardiac structure and function, lipids, insulin, glucose Respiratory: wheezing, respiratory infections, lung function, asthma, COPD Mental: behaviour, cognition, education, ASD, ADHD, anxiety, depression
NINFEA Italy N = 7500 [90]	Prospective, 2005–2016 Pregnancy-13 yrs Socio-economic, urban environment, and life-style determinants	Allergy	Cardio-metabolic: BMI Respiratory: wheezing, respiratory infections, asthma Mental: behaviour, education
RAINE Australia N = 2900 [91]	Prospective, 1989–1992 Pregnancy-25 yrs Socio-economic, migration, urban environment, and life-style determinants, genome wide association screen	Epigenetics Metabolomics Allergy Brain development	Cardio-metabolic: BMI, blood pressure, lipids, insulin, glucose Respiratory: wheezing, respiratory infections, lung function, asthma Mental: behaviour, cognition, education, ASD, ADHD, anxiety, depression
RHEA Greece N = 1300 [92]	Prospective, 2007–2008 Pregnancy-7 yrs Socio-economic, migration, urban environment, and life-style determinants	Epigenetics Metabolomics Allergy	Cardio-metabolic: BMI, blood pressure, lipids, insulin, glucose Respiratory: wheezing, respiratory infections, lung function, asthma Mental: behaviour, cognition, education, ASD, ADHD, anxiety, depression

Table 1 (continued)

Cohort, Country (N)	Design, birth years, Follow-up Main early-life stressors	Available mediators	Available outcomes
SWS United Kingdom N = 3000 [93]	Prospective, 1998–2007 Pregnancy-18 yrs Socio-economic and life-style determinants	Allergy	Cardio-metabolic: BMI, blood pressure, cardiac function and structure Respiratory: wheezing, respiratory infections, lung function, asthma Mental: behaviour, cognition, education, anxiety, depression

Federated data analysis approach

Analyses in the EU Child Cohort will be predominantly using DataSHIELD, developed as part of the EU-FP7 BioSHaRe Project [23, 25]. This is a safe and robust data analysis platform to perform joint multisite individual participant data meta-analyses, without physically transferring data (Fig. 4). DataSHIELD enables connections between local servers to analyze harmonized data located at different institutes. The major advantage of this approach is that the data from the different institutes, which together form the EU Child Cohort Network, are accessible for different researchers from various sites whilst they remain at the local sites.

Fair principles

The EU Child Cohort Network data management and access are based on the following key principles:

- Full compliance with best practice in data privacy and security;
- Use of coded data with appropriate institutional and participant consent;
- Use of privacy enhancing technologies such as filters;
- Use of policies that enable greater use of data in research;
- Approval of all procedures, policies and methods by the relevant local authorities.

Management of and access to all data is primarily the responsibility of each institution. The FAIR (findable, accessible, interoperable, reusable) principles are taken into account for the general data management approach.

Findable

The LifeCycle Project has revitalized the existing www.birthcohorts.net website. This website gives an overview of pregnancy and birth cohorts and the data available in these cohorts. Specific details of variables included in the EU child cohort network and their availability in the cohorts are presented in the open access EU Child Cohort Network Variable Catalogue. The catalogue was built using the MOLGENIS software platform for scientific data extending on BBMRI-ERIC directory of biobanks [41, 42]. It also documents how each cohort has harmonized these variables, including information about the source variables used by the cohorts. No actual data are given in the online catalogue. All relevant websites and their contents are presented in Table 3.

Table 2 Potential of the LifeCycle Project-EU Child Cohort Network

<p><i>Collaboration between prospective pregnancy/child cohort studies offers the opportunities to</i></p> <ul style="list-style-type: none"> Perform analyses in over 250,000 children and their parents Harmonize methods for data collection, biobanks, management, and analyses Perform analyses on published and unpublished data which limits publication bias Perform individual participant data meta-analyses with better statistical precision Stratify groups by geographical area or sex Compare determinants and outcomes between European populations Examine consequences of small variations in determinants from early life onwards Identify variations in geography and time periods for specific associations Infer causality from observed associations by advanced analytical approaches Enable analyses on life course trajectories on risk factors of non-communicable diseases Explore different life course models 	1
--	---

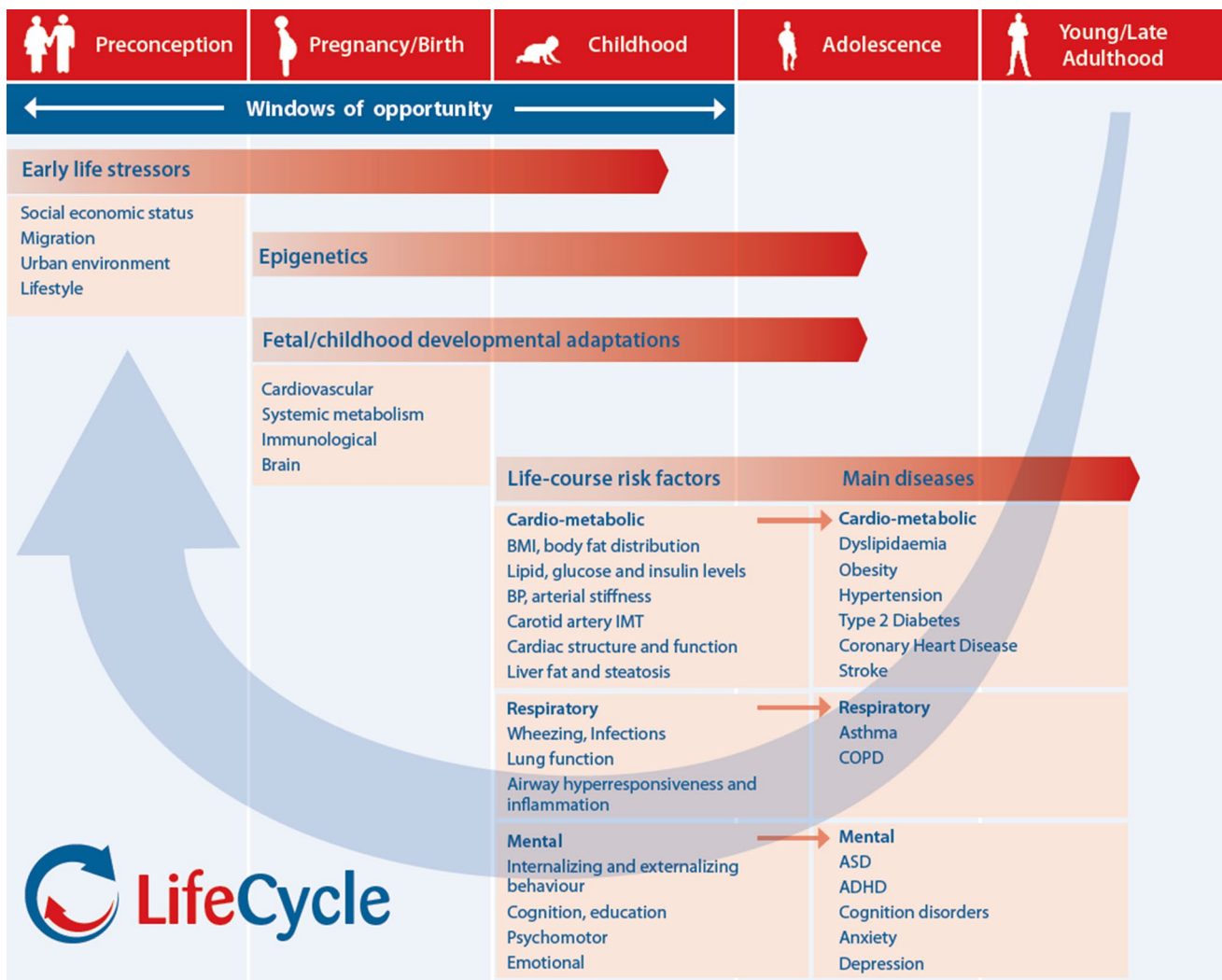


Fig. 3 Main concepts of the LifeCycle Project and related data in the EU Child Cohort Network

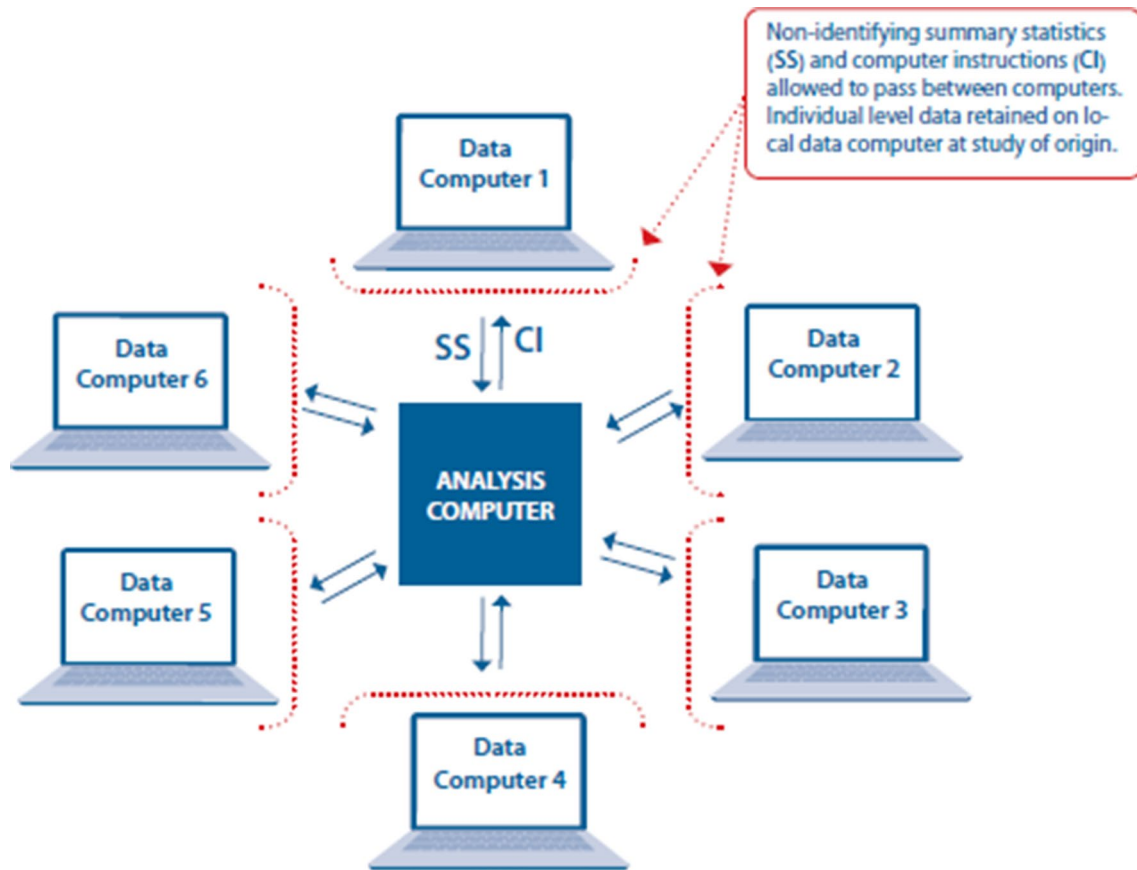


Fig. 4 Federated analysis approach using DataSHIELD approach

Accessible

A harmonized set of data for EU Child Cohort Network is available by a server controlled by or located at each specific institute. Harmonized data from each cohort are held on secure Opal servers (<http://opaldoc.obiba.org/en/latest/>) at their institution. Protocols for setting up this data infrastructure are available, together with YouTube instruction videos. Data are accessed via a central analysis server using the R-based platform DataSHIELD. Access to data is conditional on approval by the cohort. Partners and their cohorts can always decide to share research data without using DataSHIELD, conditional on relevant local ethical and legal approvals. This approach is used for analyses that are not yet possible in DataSHIELD [25]. The field of data sharing and cross study analyses is rapidly advancing. Although we start with using DataSHIELD, we recognise that over time this may change.

Interoperable

Existing data have been harmonized and integrated into exposure variables to make them interoperable. Protocols

for harmonization are available online. All harmonized data from different cohorts have been renamed into standardized variable names. A full list of the available variables per cohort is available in the EU Child Cohort Network Variable Catalogue.

Reusable

The EU Child Cohort Network reuses data that are already available within cohorts. The EU Child Cohort Network, with the harmonized set of variables and infrastructure, should be sustainable beyond the duration of the LifeCycle Project. During the last two years, four other European consortia have been funded, which are planning to build upon the harmonized data and federated analysis infrastructure in the EU Child Cohort Network. These consortia include the EUCAN-Connect, NutriPROGRAM, ATHLETE and LongITools Projects. Future collaborations may include not only European, but also global initiatives such as the NIH-Environmental influences on Child Health Outcomes (ECHO) Programme in the United States, which aims to build a virtual paediatric cohort based on new and existing birth cohorts, recognizing the enormous

Table 3 Websites of the LifeCycle Project–EU child cohort network

Data related to the LifeCycle Project is findable through different websites

LifeCyce Project

<https://lifecycle-project.eu> website

Overview of the LifeCycle Project

All protocols for harmonisation and setting up the data-servers

Open access

Links to other relevant websites

Birthcohorts.net

www.birthcohorts.net

Overview of all cohorts and their data

Open access, no restriction for access on cohort information

EU Child Cohort Network Variable Catalogue

<http://catalogue.lifecycle-project.eu>

Overview of harmonized data and variables in each cohort

Open access

Find function is included in website

EU Child Cohort harmonized data

Cohort websites via www.lifecycle-project.eu

Harmonized data from different cohorts

Data server is within institutional firewall

Access to data can only be given by data owner (LifeCycle Project partner)

opportunities in optimizing and networking existing resources [43, 44].

Data governance

The LifeCycle Project or EU Child Cohort Network do not own data, but bring data from other cohorts together via a federated data analysis platform. Ethical and legal responsibility for data management and security is maintained by the source studies or home institutions. The principal investigators or home institutions should always administer permission for external access to specific data on their server for addressing research questions. The EU Child Cohort Network cannot provide open access to researchers. The data sharing protocols and agreements will be updated regularly, according to new legal practices, such as the European General Data Protection Regulation 2016/679 (GDPR). All governance protocols will take not only the short-term, but also the long-term EU Child Cohort Network, beyond the LifeCycle Project duration, into account.

EU Child Cohort Network research proposals

Proposals for research using the EU Child Cohort Network can be put forward by both LifeCycle Project partners and other researchers. External researchers can send a request

for EU Child Cohort Network data use to the participating cohorts or lifecycle@erasmusmc.nl. Each LifeCycle Project proposal is discussed in the relevant coordinating work package (<https://lifecycle-project.eu/for-scientists/workpackages/>) and subsequently distributed among all cohorts participating in the LifeCycle Project and EU Child Cohort Network. Cohorts can opt-in or opt-out of each analysis, depending on the data availability, research interests or involvement in other projects. In the first phase, the focus of research projects is on those projects related to the LifeCycle Project research aims (see below). An efficient governance structure was organized and agreed upon by researchers and ethical and legal representatives. EU Child Cohort Network governance structure will be updated regularly where needed and will be made sustainable after the LifeCycle Project duration. Because there is no physical transfer of data needed, we are currently exploring the possibility of working with a short Data Access Agreement that replaces commonly used Data Transfer Agreements. When the EU Child Cohort Network is fully operational we aim to have regular EU Child Cohort Network meetings or telephone conferences to discuss:

- Research projects (novel proposals, progress of ongoing projects);
- Harmonization (novel proposals, progress of ongoing efforts);
- DataSHIELD analysis approaches (priorities for further development);
- Any relevant ethical or legal issues concerning federated analysis approaches;

Participants in these meetings or telephone conferences are not only LifeCycle Project Partners, but representatives of all institutes that have harmonized their data and set up the IT infrastructure needed for the federated analysis of data via DataShield.

LifeCycle Project primary research areas

The LifeCycle Project uses the integrated and harmonized set of variables from the EU Child Cohort Network for identification of early-life stressors influencing cardio-metabolic, respiratory and mental developmental adaptations and health trajectories during the full life course (Fig. 3).

Integrated early-life stressors approach and the exposome

Early-life stressors, including socio-economic, migration, urban environmental, and lifestyle related factors, have been associated with cardio-metabolic, respiratory, and mental

health and disease, which together contribute greatly to the global burden of non-communicable diseases [5–22]. An accumulating body of evidence suggests that exposure to these factors during fetal life and childhood affects later life health trajectories [38]. Thus far, studies focused on the effects of early-life environmental exposures on later life health outcomes have largely been using a ‘one-exposure at one-time point’ approach. Research from LifeCycle Project partners suggests that instead of exposure to single stressors that individually may have weak effects, exposure to a cluster or pattern of adverse early-life stressors in specific age windows is more likely to influence health during the lifecycle [39]. We will apply a holistic ‘early-life exposome’ model to encompass many human environmental exposures, which is dynamic from conception onwards and complements the genome. To develop this early-life exposome, we will specifically take into account measurements in the external environment (socio-economic, migration, urban environment, and lifestyle factors), and biological markers reflecting the internal environment (DNA methylation, RNA expression, and metabolomics), and the dynamic life course nature of the exposome. We will use available methods developed as part of the EU-FP7 HELIX Project for further development of the early-life exposome model [29].

Cardio-metabolic, respiratory and mental health outcomes

Embryonic life, fetal life and early childhood are characterized by high developmental rates and seem to be critical periods for developmental adaptations with long-term consequences. Research from LifeCycle Project partners have shown that specific maternal lifestyle factors and fetal growth variation in early pregnancy are related to non-communicable diseases and their risk factors [45–49]. We will use repeatedly measured exposure, mediator and outcome data from the EU Child Cohort Network to compare different potential life course models including those assuming specific critical periods and those assuming interactive and cumulative effects throughout the life course. We will relate early-life stressors measured in different early-life periods (preconception, fetal life, early childhood) with life course health trajectories. We specifically hypothesize that early-life stressors lead to developmental adaptations of:

- The cardiovascular system assessed in detail by advanced cardiac and great vessel ultrasound or Magnetic Resonance Imaging (MRI), and systemic metabolism, detected by measuring hundreds of metabolites using high-throughput approaches, which precede the development of cardio-metabolic diseases [50–60].
- Lung volumes, airway patency assessed by lung function measurements and clinical assessments, and immuno-

logical or allergy-related assessments, which precede the development of respiratory disease [61–63].

- Structural and functional brain development assessed by ultrasound in fetal life or early infancy, or brain MRI in later life, which precede the development of mental health outcomes [64–67].

Epigenetic pathways

An accumulating body of evidence suggests that epigenetic changes play a key role in the associations of early-life stressors with lifecycle health and disease trajectories [68]. DNA methylation, the most frequently studied epigenetic phenomenon in large populations, is a dynamic process, which may be influenced by environmental stressors such as urban environment, dietary factors and smoking [68]. DNA methylation changes are more common in early life. LifeCycle Project partners have identified DNA methylation markers related to specific early-life stressors including maternal BMI, smoking, dietary factors and birth weight [12, 17]. The EU Child Cohort Network brings together many pregnancy and childhood cohorts with information about epigenome-wide DNA methylation. Availability of repeatedly measured DNA methylation and of RNA expression data enables studies on persistence and functionality of DNA methylation markers potentially involved in early-life programming of non-communicable diseases.

Population impact

The concept that early life is critical for health and disease throughout the life course is well-acknowledged. However, there is still not much evidence for effective prevention or intervention strategies using early life as a window of opportunity to maximize the human developmental potential during the full life course. We will use different approaches to translate findings into population health recommendations. These include causal inference, aggregation of evidence for interventions based on reviews, dynamic microsimulation, and development of prediction models.

Causality cannot be directly concluded from observational studies. Advanced analytical approaches that can help to infer causality include sibling comparison studies, propensity score matching and Mendelian randomization studies, in which genetic variants are used as unconfounded proxies for adverse exposures [69]. The EU Child Cohort Network facilitates integration of different causal inference methods and comparison of their findings, which will strengthen causal inference needed for translation of findings from observational studies to public health recommendations.

We will review and summarize evidence based on findings both from observational studies in the EU Child Cohort Network and from published intervention studies to develop

recommendations for population and subgroup-specific interventions focused on the earliest phases of life. Dynamic microsimulation modelling using data from cohort studies enables policy evaluations and scenario analyses focused on early-life interventions when experimental studies are not possible [70, 71]. The EU Child Cohort Network provides a unique infrastructure for these analyses, because of the available data and variation in exposures and outcomes, life course trajectories of non-communicable diseases and various subpopulations with different baseline risks.

Data from observational studies can help to develop models to predict risk factors for non-communicable diseases. Previous studies suggested that pregnancy, birth and infancy characteristics have the potential to identify groups at risk for obesity [72, 73]. The EU Child Cohort Network is the ideal platform to develop models to predict from early-life stressor data the onset of risk factors for cardio-metabolic, respiratory and mental disease across the lifecycle. Models can include various background characteristics, which enable baseline risk estimation from socio-economic, migration, environment and lifestyle stressors, which may be difficult to modify in the short-term but help to predict the outcomes of interest.

Finally, we will develop E-learning modules and eHealth applications that will be made widely available to make the knowledge and research findings available for educational and health care purposes.

Conclusion

The LifeCycle Project and its EU Child Cohort Network lead to great opportunities for researchers to combine harmonized data from different cohorts by a federated analysis platform. It also provides a novel model for collaborative research in large research infrastructures with individual level data. The LifeCycle Project will translate results from research using the EU Child Cohort Network into recommendations for targeted prevention strategies to improve health trajectories for current and future generations by optimizing their earliest phases of life.

Acknowledgements The LifeCycle project received funding from the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 733206 LifeCycle). All study specific acknowledgements and funding are presented in the supplementary materials. This manuscript reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains

The LifeCycle Project Group Vincent W.V. Jaddoe^{1,2}, Janine F. Felix^{1,2}, Liesbeth Duijts^{1,2}, Hanan El Marroun^{1,3,4}, Romy Gaillard^{1,2}, Susana Santos^{1,2}, Madelon L. Geurtsen^{1,2}, Marjolein N. Kooijman^{1,2}, Sara M. Mensink-Bout^{1,2}, Florianne O.L. Vehmeijer^{1,2}, Ellis Voerman^{1,2}, Martine Vrijheid^{5,6,7}, Jordi Sunyer^{5,6,7,8}, Mark Nieuwenhuijsen^{5,6,7},

Xavier Basagaña^{5,6,7}, Mariona Bustamante^{5,6,7}, Maribel Casas^{5,6,7}, Montserrat de Castro^{5,6,7}, Lourdes Cirugeda^{5,6,7}, Sílvia Fernández-Barrés^{5,6,7}, Serena Fossati^{5,6,7}, Raquel Garcia^{5,6,7}, Jordi Júlvez^{5,6,9}, Aitana Lertxundi^{5,10,11}, Nerea Lertxundi^{10,11}, Sabrina Llop^{5,12}, Mònica López-Vicente^{2,3,6}, Maria-Jose Lopez-Espinosa^{5,12,13}, Lea Maitre⁶, Mario Murcia^{12,14}, Jose Urquiza^{5,6,7}, Charline Warembourg^{5,6,7}, Lorenzo Richiardi¹⁵, Costanza Pizzi¹⁵, Daniela Zugna¹⁵, Maja Popovic¹⁵, Elena Isaevska¹⁵, Milena Maule¹⁵, Chiara Moccia¹⁵, Giovenale Moirano¹⁵, Davide Rasella¹⁵, Mark A Hanson^{16,17}, Hazel M. Inskip^{17,18}, Chandni Maria Jacob^{16,17}, Theodosia Salika¹⁸, Deborah A. Lawlor^{19,20,21}, Ahmed Elhakeem^{19,21}, Tim Cadman^{19,21}, Anne-Marie Nybo Andersen²², Angela Pinot de Moira²², Katrine Strandberg-Larsen²², Marie Pedersen²², Johan L Vinther²², John Wright²³, Rosemary R.C. McEachan²³, Paul Wilson²⁴, Dan Mason²³, Tiffany C. Yang²³, Morris A. Swertz^{25,26}, Eva Corpeleijn²⁷, Sido Haakma²⁵, Marloes Cardol²⁷, Esther van Enckevoort^{25,26}, Eleanor Hyde^{25,26}, Salome Scholtens^{25,26}, Harold Snieder²⁷, Chris H.L. Thio²⁷, Marina Vafeiadi²⁸, Lida Chatzi²⁹, Katerina Margetaki²⁹, Theano Roumeliotaki²⁸, Jennifer R. Harris^{30,31}, Johanna L. Nader³², Gun Peggy Knudsen³³, Per Magnus³⁰, Marie-Aline Charles^{34,35}, Barbara Heude³⁴, Lidia Panico³⁶, Mathieu Ichou³⁶, Blandine de Lauzon-Guillain³⁴, Patricia Dargent-Molina³⁴, Maxime Cornet³⁴, Sandra M. Florian³⁶, Faryal Harrar³⁴, Johanna Lepeule³⁷, Sandrine Lioret³⁴, Maria Melchior³⁸, Sabine Plancoulaine³⁴, Marjo-Riitta Järvelin^{39,40,41,42}, Sylvain Sebert³⁹, Minna Männikkö⁴³, Priyanka Parmar³⁹, Nina Rautio³⁹, Justiina Ronkainen³⁹, Mimmi Tolvanen³⁹, Johan G Eriksson^{44,45,46,47}, Tuija M. Mikkola⁴⁵, Berthold Koltzko⁴⁸, Veit Grote⁴⁸, Nicole Aumüller⁴⁸, Ricardo Closa-Monasterolo⁴⁹, Joaquin Escribano⁴⁹, Natalia Ferré⁴⁹, Dariusz Gruszfeld⁵⁰, Kathrin Gürlich⁴⁸, Jean-Paul Langhendries⁵¹, Veronica Luque⁴⁹, Enrica Riva⁵², Philipp Schwarzfischer⁴⁸, Martina Totzauer⁴⁸, Elvira Verduci⁵², Annick Xhonneux⁵¹, Marta Zaragoza-Jordana⁴⁹, Maarten Lindeboom⁵³, Ameli Schwalber⁵⁴, Nina Donner⁵⁴, Rae-Chi Huang⁵⁵, Rachel E. Foong^{55,56}, Graham L. Hall^{55,56}, Ashleigh Lin⁵⁵, Jennie Carson⁵⁵, Phillip Melton^{57,58}, Sebastian Rauscher⁵⁵

¹Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. ²The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. ³Department of Child and Adolescent Psychiatry and Psychology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. ⁴Department of Psychology, Education and Child Studies, Erasmus School of Social and Behavioural Sciences, Rotterdam, the Netherlands. ⁵CIBER Epidemiología y Salud Pública (CIBERESP), Spain. ⁶ISGlobal, Barcelona, Spain. ⁷Universitat Pompeu Fabra (UPF), Barcelona, Spain. ⁸IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain. ⁹Institut d'Investigació Sanitària Pere Virgili (IISPV), Hospital Universitari Sant Joan de Reus, Reus, Spain. ¹⁰Biodonostia, Health research institute, San Sebastian, Spain. ¹¹University of Basque Country, Spain. ¹²Epidemiology and Environmental Health Joint Research Unit, FISABIO–Universitat Jaume I–Universitat de València, Valencia, Spain. ¹³Faculty of Nursing and Chiropody, Universitat de València, Valencia, Spain. ¹⁴Conselleria de Sanitat, Valencia, Spain. ¹⁵Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin, Turin, Italy. ¹⁶Institute of Developmental Sciences, Faculty of Medicine, University of Southampton, Southampton, United Kingdom. ¹⁷NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, United Kingdom. ¹⁸MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton General Hospital, Southampton, United Kingdom. ¹⁹MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, United Kingdom. ²⁰NIHR Bristol Biomedical Research Centre, Bristol, United Kingdom. ²¹Population Health Science, Bristol Medical School, University of Bristol, Bristol, United Kingdom. ²²Section of Epidemiology, Department of Public Health, University of Copenhagen, Copenhagen, Denmark. ²³Bradford Institute for Health Research,

Bradford Teaching Hospitals NHS Foundation Trust, Bradford, United Kingdom. ²⁴University of Manchester, Manchester, United Kingdom. ²⁵University of Groningen, University Medical Center Groningen, Genomics Coordination Center, Groningen, the Netherlands. ²⁶University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands. ²⁷Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands. ²⁸Department of Social Medicine, Faculty of Medicine, University of Crete, Heraklion, Crete, Greece. ²⁹Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. ³⁰Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway. ³¹Division of Health Data and Digitalization, Norwegian Institute of Public Health, Oslo, Norway. ³²Department of Genetics and Bioinformatics, Division of Health Data and Digitalisation, Norwegian Institute of Public Health, Oslo, Norway. ³³Norwegian Institute of Public Health, Oslo, Norway. ³⁴Université de Paris, Centre for Research in Epidemiology and Statistics (CRESS), INSERM, INRAE, Paris, France. ³⁵ELFE Joint Unit, French Institute for Demographic Studies (INED), French Institute for Medical Research and Health (INSERM), French Blood Agency, Aubervilliers, France. ³⁶Institut National d'Études Démographiques (INED), Aubervilliers, France. ³⁷Université Grenoble Alpes, Inserm, CNRS, Team of Environmental Epidemiology Applied to Reproduction and Respiratory Health, IAB, Grenoble, France. ³⁸Sorbonne Université, INSERM, Institut Pierre Louis d'Épidémiologie et de Santé Publique (IPLESP), Equipe de Recherche en Épidémiologie Sociale (ERES), Paris, France. ³⁹Center For Life-course Health research, Faculty of Medicine, University of Oulu, Oulu, Finland. ⁴⁰Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, United Kingdom. ⁴¹Department of Life Sciences, College of Health and Life Sciences, Brunel University London, London, United Kingdom. ⁴²Unit of Primary Health Care, Oulu University Hospital, OYS, Oulu, Finland. ⁴³Infrastructure for Population Studies, Faculty of Medicine, University of Oulu, Oulu, Finland. ⁴⁴Department of General Practice and Primary Health Care, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. ⁴⁵Folkhälsan Research Center, Helsinki, Finland. ⁴⁶Obstetrics & Gynecology, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore. ⁴⁷Singapore Institute for Clinical Sciences (SICS), Agency for Science and Technology (A*STAR), Singapore. ⁴⁸Department of Pediatrics, Dr. von Hauner Children's Hospital, University Hospital, LMU, Munich, Germany. ⁴⁹Universitat Rovira i Virgili, IISPV, Tarragona, Spain. ⁵⁰Neonatal Department, Children's Memorial Health Institute, Warsaw, Poland. ⁵¹CHC St Vincent, Liège-Rocourt, Belgium. ⁵²University of Milan, Milan, Italy. ⁵³Department of Economics, VU University Amsterdam, Amsterdam, the Netherlands. ⁵⁴Concentris Research Management GmbH, Fürstentfeldbruck, Germany. ⁵⁵Tel-ethon Kids Institute, Perth, Western Australia, Australia. ⁵⁶School of Physiotherapy and Exercise Science, Curtin University, Perth, Western Australia, Australia. ⁵⁷Curtin/UWA Centre for Genetic Origins of Health and Disease, School of Biomedical Sciences, The University of Western Australia, Australia. ⁵⁸School of Pharmacy and Biomedical Sciences, Curtin University, Perth, Western Australia, Australia

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions on adult health and disease. *N Engl J Med*. 2008;359(1):61–73.
2. Balbus JM, Barouki R, Birnbaum LS, et al. Early-life prevention of non-communicable diseases. *Lancet*. 2013;381(9860):3–4.
3. Larsen PS, Kamper-Jorgensen M, Adamson A, et al. Pregnancy and birth cohort resources in Europe: a large opportunity for aetiological child health research. *Paediatr Perinat Epidemiol*. 2013;27(4):393–414.
4. Santos S, Zugna D, Pizzi C, Richiardi L. Sources of confounding in life course epidemiology. *J Dev Orig Health Dis*. 2019;10(3):299–305. <https://doi.org/10.1017/s2040174418000582>.
5. Birks L, Casas M, Garcia AM, Alexander J, Barros H, Bergstrom A, et al. Occupational exposure to endocrine-disrupting chemicals and birth weight and length of gestation: a European meta-analysis. *Environ Health Perspect*. 2016;124(11):1785–93.
6. Casas M, den Dekker HT, Kruithof CJ, Reiss IK, Vrijheid M, Sunyer J, et al. The effect of early growth patterns and lung function on the development of childhood asthma: a population based study. *Thorax*. 2018;73(12):1137–45.
7. Reese SE, Xu CJ, den Dekker HT, Lee MK, Sikdar S, Ruiz-Arenas C, Merid SK, et al. Epigenome-wide meta-analysis of DNA methylation and childhood asthma. *J Allergy Clin Immunol*. 2019;143(6):2062–74.
8. Voerman E, Santos S, Inskip H, Amiano P, Barros H, et al. Association of gestational weight gain with adverse maternal and infant outcomes. *JAMA*. 2019;321(17):1702–15.
9. Gruzieva O, Xu CJ, Yousefi P, Relton C, Merid SK, Breton CV, et al. Prenatal particulate air pollution and DNA methylation in newborns: an epigenome-wide meta-analysis. *Environ Health Perspect*. 2019;127(5):57012.
10. Haworth S, Shapland CY, Hayward C, Prins BP, Felix JF, Medina-Gomez C, et al. Low-frequency variation in TP53 has large effects on head circumference and intracranial volume. *Nat Commun*. 2019;10(1):357.
11. Horikoshi M, Beaumont RN, Day FR, Warrington NM, Kooijman MN, Fernandez-Tajes J, et al. Genome-wide associations for birth weight and correlations with adult disease. *Nature*. 2016;538(7624):248–52.
12. Kupers LK, Monnereau C, Sharp GC, Yousefi P, Salas LA, Ghanthous A, et al. Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. *Nat Commun*. 2019;10(1):1893.
13. Leventakou V, Roumeliotaki T, Martinez D, Barros H, Brantsaeter AL, Casas M, et al. Fish intake during pregnancy, fetal growth, and gestational length in 19 European birth cohort studies. *Am J Clin Nutr*. 2014;99(3):506–16.
14. Patro Golab B, Santos S, Voerman E, Lawlor DA, Jaddoe VWV, Gaillard R, et al. Influence of maternal obesity on the association between common pregnancy complications and risk of childhood obesity: an individual participant data meta-analysis. *Lancet Child Adolesc Health*. 2018;2(11):812–21.
15. Santos S, Eekhout I, Voerman E, Gaillard R, Barros H, Charles MA, et al. Gestational weight gain charts for different body mass index groups for women in Europe, North America, and Oceania. *BMC Med*. 2018;16(1):201.


16. Santos S, Voerman E, Amiano P, Barros H, Beilin LJ, Bergstrom A, et al. Impact of maternal body mass index and gestational weight gain on pregnancy complications: an individual participant data meta-analysis of European, North American and Australian cohorts. *Bjog*. 2019;126(8):984–95.
17. Sharp GC, Salas LA, Monnereau C, Allard C, Yousefi P, Everson TM, et al. Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation: findings from the pregnancy and childhood epigenetics (PACE) consortium. *Hum Mol Genet*. 2017;26(20):4067–85.
18. van der Sonnenschein VAM, Arends LR, de Jongste JC, Annesi-Maesano I, Arshad SH, Barros H, et al. Preterm birth, infant weight gain, and childhood asthma risk: a meta-analysis of 147,000 European children. *J Allergy Clin Immunol*. 2014;133(5):1317–29.
19. Strandberg-Larsen K, Poulsen G, Bech BH, Chatzi L, Cordier S, Dale MTG, et al. Association of light-to-moderate alcohol drinking in pregnancy with preterm birth and birth weight: elucidating bias by pooling data from nine European cohorts. *Eur J Epidemiol*. 2017;32(9):751–64.
20. Stratakis N, Roumeliotaki T, Oken E, Barros H, Basterrechea M, Charles MA, et al. Fish intake in pregnancy and child growth: a pooled analysis of 15 European and US birth cohorts. *JAMA Pediatr*. 2016;170(4):381–90.
21. Voerman E, Santos S, Patro Golab B, Amiano P, Ballester F, Barros H, et al. Maternal body mass index, gestational weight gain, and the risk of overweight and obesity across childhood: an individual participant data meta-analysis. *PLoS Med*. 2019;16(2):e1002744.
22. Warrington NM, Beaumont RN, Horikoshi M, Day FR, Helgeland O, Laurin C, et al. Maternal and fetal genetic effects on birth weight and their relevance to cardio-metabolic risk factors. *Nat Genet*. 2019;51(5):804–14.
23. Doiron D, Burton P, Marcon Y, et al. Data harmonization and federated analysis of population-based studies: the BioSHaRE project. *Emerg Themes Epidemiol*. 2013;10(1):12.
24. Bousquet J, Anto J, Sunyer J, Nieuwenhuijsen M, Vrijheid M, Keil T. Pooling birth cohorts in allergy and asthma: European Union-Funded Initiatives—A MeDALL, CHICOS, ENRIECO, and GALEN joint paper. *Int Arch Allergy Immunol*. 2013;161(1):1–10. <https://doi.org/10.1159/000343018>.
25. 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(6):1929–44. <https://doi.org/10.1093/ije/dyu188>.
26. Sebert S, Lowry E, Aumüller N, et al. Cohort profile: the DYNHEALTH consortium—a European consortium for a life-course bio-psychosocial model of healthy ageing of glucose homeostasis. *Int J Epidemiol*. 2019. <https://doi.org/10.1093/ije/dyz056>.
27. Koletzko B, Brands B, Grote V, et al. Long-term health impact of early nutrition: the power of programming. *Ann Nutr Metab*. 2017;70(3):161–9. <https://doi.org/10.1159/000477781>.
28. Gehring U, Casas M, Brunekreef B, et al. Environmental exposure assessment in European birth cohorts: results from the ENRIECO project. *Environ Health*. 2013;12(1):8. <https://doi.org/10.1186/1476-069x-12-8>.
29. Vrijheid M, Slama R, Robinson O, et al. The human early-life exposome (HELIX): project rationale and design. *Environ Health Perspect*. 2014;122(6):535–44.
30. 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(9):e021311.
31. Pastorino S, Bishop T, Crozier SR, et al. Associations between maternal physical activity in early and late pregnancy and offspring birth size: remote federated individual level meta-analysis from eight cohort studies. *BJOG*. 2019;126(4):459–70. <https://doi.org/10.1111/1471-0528.15476>.
32. Anjos T, Altmäe S, Emmett P, et al. Nutrition and neurodevelopment in children: focus on NUTRIMENTHE project. *Eur J Nutr*. 2013;52(8):1825–42. <https://doi.org/10.1007/s00394-013-0560-4>.
33. Middeldorp CM, Felix JF, Mahajan A, et al. The early growth genetics (EGG) and early genetics and lifecourse epidemiology (EAGLE) consortia: design, results and future prospects. *Eur J Epidemiol*. 2019;34(3):279–300. <https://doi.org/10.1007/s10654-019-00502-9>.
34. Felix JF, Joubert BR, Baccarelli AA, et al. Cohort profile: pregnancy and childhood epigenetics (PACE) consortium. *Int J Epidemiol*. 2018;47(1):22–3.
35. Koletzko B, Benninga MA, Godfrey KM, et al. Public-private collaboration in clinical research during pregnancy, lactation, and childhood: joint position statement of the early nutrition academy and the European society for pediatric gastroenterology, hepatology, and nutrition. *J Pediatr Gastroenterol Nutr*. 2014;58(4):525–30. <https://doi.org/10.1097/mpg.0000000000000284>.
36. Pizzi C, Richiardi M, Charles MA, et al. Measuring child socioeconomic position in birth cohort research: the development of a novel standardized household income indicator. *Int J Environ Res Public Health*. 2020. <https://doi.org/10.3390/ijerph17051700>.
37. Wild CP. The exposome: from concept to utility. *Int J Epidemiol*. 2012;41(1):24–32.
38. Robinson O, Vrijheid M. The pregnancy exposome. *Curr Environ Health Rep*. 2015;2(2):204–13.
39. Vrijheid M. The exposome: a new paradigm to study the impact of environment on health. *Thorax*. 2014;69(9):876–8.
40. Santos S, Maitre L, Warembourg C, et al. Applying the exposome concept in birth cohort research: a review of statistical approaches. *Eur J Epidemiol*. 2020;35(3):193–204. <https://doi.org/10.1007/s10654-020-00625-4>.
41. Swertz MA, Dijkstra M, Adamusiak T, et al. The MOLGENIS toolkit: rapid prototyping of biosoftware at the push of a button. *BMC Bioinf*. 2010;11(Suppl 12):S12.
42. Holub P, Swertz MA, Reihls R, van Enkevort D, Müller H, Litton JE. BBMRI-ERIC directory: 515 biobanks with over 60 million biological samples. *Biopreserv Biobank*. 2016;14(6):559–62. <https://doi.org/10.1089/bio.2016.0088>.
43. NIH-Environmental influences on Child Health Outcomes (ECHO) Program. National Institutes of Health (NIH). (2015). <https://www.nih.gov/research-training/environmental-influences-child-health-outcomes-echo-program>. Accessed on 14 Aug 2019.
44. Trasande L, Ghassabian A, Kahn LG, Jacobson MH, Afanasieva Y, Liu M, Chen Y, Naidu M, Alcedo G, Gilbert J, Koshy TT. NYU children's health and environment study team. The NYU children's health and environment study. *Eur J Epidemiol*. 2020;35(3):305–20. <https://doi.org/10.1007/s10654-020-00623-6>.
45. Pike KC, Inskip HM, Robinson S, et al. Maternal late-pregnancy serum 25-hydroxyvitamin D in relation to childhood wheeze and atopic outcomes. *Thorax*. 2012;67(11):950–6. <https://doi.org/10.1136/thoraxjnl-2012-201888>.
46. Prady SL, Pickett KE, Croudace T, et al. Maternal psychological distress in primary care and association with child behavioural outcomes at age three. *Eur Child Adolesc Psychiatry*. 2016;25(6):601–13.
47. Roth C, Magnus P, Schjolberg S, et al. Folic acid supplements in pregnancy and severe language delay in children. *JAMA*. 2011;306(14):1566–73.
48. Suren P, Roth C, Bresnahan M, et al. Association between maternal use of folic acid supplements and risk of autism spectrum disorders in children. *JAMA*. 2013;309(6):570–7.
49. Valvi D, Casas M, Romaguera D, et al. Prenatal phthalate exposure and childhood growth and blood pressure: evidence from the Spanish INMA-sabadell birth cohort study. *Environ Health Perspect*. 2015;123(10):1022–9.

50. Hanson MA, Gluckman PD. Early developmental conditioning of later health and disease: physiology or pathophysiology? *Physiol Rev.* 2014;94(4):1027–76.
51. Kerkhof GF, Hokken-Koelega AC. Rate of neonatal weight gain and effects on adult metabolic health. *Nat Rev Endocrinol.* 2012;8(11):689–92.
52. Mook-Kanamori DO, Steegers EA, Eilers PH, Raat H, Hofman A, Jaddoe VW. Risk factors and outcomes associated with first-trimester fetal growth restriction. *JAMA.* 2010;303(6):527–34.
53. Gaillard R, Steegers EA, Franco OH, Hofman A, Jaddoe VW. Maternal weight gain in different periods of pregnancy and childhood cardio-metabolic outcomes. The generation R study. *Int J Obes (Lond).* 2015;39(4):677–85.
54. Jaddoe VWV, de Jonge LL, Hofman A, Franco OH, Steegers EAP, Gaillard R. First trimester fetal growth restriction and cardiovascular risk factors in school age children: population based cohort study. *BMJ Br Med J.* 2014;348:g14. <https://doi.org/10.1136/bmj.g14>.
55. de Jonge LL, Steegers EA, Ernst GD, et al. C-reactive protein levels, blood pressure and the risks of gestational hypertensive complications: the generation R study. *J Hypertens.* 2011;29(12):2413–21.
56. Lewandowski AJ, Augustine D, Lamata P, et al. Preterm heart in adult life: cardiovascular magnetic resonance reveals distinct differences in left ventricular mass, geometry, and function. *Circulation.* 2013;127(2):197–206.
57. Lewandowski AJ, Bradlow WM, Augustine D, et al. Right ventricular systolic dysfunction in young adults born preterm. *Circulation.* 2013;128(7):713–20.
58. Soininen P, Kangas AJ, Wurtz P, et al. High-throughput serum NMR metabolomics for cost-effective holistic studies on systemic metabolism. *Analyst.* 2009;134(9):1781–5.
59. Wurtz P, Havulinna AS, Soininen P, et al. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation.* 2015;131(9):774–85.
60. Voerman E, Jaddoe VWV, Uhl O, Shokry E, Horak J, Felix JF, Koletzko B, Gaillard R. A population-based resource for inter-generational metabolomics analyses in pregnant women and their children: the generation R study. *Metabolomics.* 2020;16(4):43. <https://doi.org/10.1007/s11306-020-01667-1>.
61. Duijts L, Reiss IK, Brusselle G, de Jongste JC. Early origins of chronic obstructive lung diseases across the life course. *Eur J Epidemiol.* 2014;29(12):871–85. <https://doi.org/10.1007/s10654-014-9981-5>.
62. Postma DS, Bush A, van den Berge M. Risk factors and early origins of chronic obstructive pulmonary disease. *Lancet.* 2015;385(9971):899–909.
63. Sears MR, Greene JM, Willan AR, et al. A longitudinal, population-based, cohort study of childhood asthma followed to adulthood. *N Engl J Med.* 2003;349(15):1414–22.
64. El Marroun H, Tiemeier H, Franken IH, et al. Prenatal cannabis and tobacco exposure in relation to brain morphology: a prospective neuroimaging study in young children. *Biol Psychiatry.* 2016;79(12):971–9.
65. Kok R, Thijssen S, Bakermans-Kranenburg MJ, et al. Normal variation in early parental sensitivity predicts child structural brain development. *J Am Acad Child Adolesc Psychiatry.* 2015;54(10):824–31.
66. Mous SE, Hammerschlag AR, Polderman TJ, et al. A population-based imaging genetics study of inattention/hyperactivity: basal ganglia and genetic pathways. *J Am Acad Child Adolesc Psychiatry.* 2015;54(9):745–52.
67. Muetzel RL, Mous SE, van der Ende J, et al. White matter integrity and cognitive performance in school-age children: a population-based neuroimaging study. *Neuroimage.* 2015;119:119–28.
68. Ng JW, Barrett LM, Wong A, Kuh D, Smith GD, Relton CL. The role of longitudinal cohort studies in epigenetic epidemiology: challenges and opportunities. *Genome Biol.* 2012;13(6):246.
69. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med.* 2008;27(8):1133–63. <https://doi.org/10.1002/sim.3034>.
70. De Stavola BL, Daniel RM, Ploubidis GB, Micali N. Mediation analysis with intermediate confounding: structural equation modeling viewed through the causal inference lens. *Am J Epidemiol.* 2015;181(1):64–80.
71. Li J, O'Donoghue C. A survey of dynamic microsimulation models: uses, model structure and methodology. *Int J Microsimul.* 2013;6(2):3–55.
72. Gillman MW, Ludwig DS. How early should obesity prevention start? *N Engl J Med.* 2013;369(23):2173–5. <https://doi.org/10.1056/NEJMp1310577>.
73. Robinson SM, Crozier SR, Harvey NC, et al. Modifiable early-life risk factors for childhood adiposity and overweight: an analysis of their combined impact and potential for prevention. *Am J Clin Nutr.* 2015;101(2):368–75.
74. Boyd A, Golding J, Macleod J, et al. Cohort profile: the 'children of the 90 s'—the index offspring of the avon longitudinal study of parents and children. *Int J Epidemiol.* 2013;42(1):111–27.
75. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, Henderson J, Macleod J, Molloy L, Ness A, Ring S, Nelson SM, Lawlor DA. Cohort profile: the avon longitudinal study of parents and children: ALSPAC mothers cohort. *Int J Epidemiol.* 2013;42(1):97–110. <https://doi.org/10.1093/ije/dys066> **Epub 2012 Apr 16**.
76. Lawlor DA, Lewcock M, Rena-Jones L, Rollings C, Yip V, Smith D, Pearson RM, Johnson L, Millard LAC, Patel N, Skinner A, Tilling K. ALSPAC Executive. The second generation of the avon longitudinal study of parents and children (ALSPAC-G2): a cohort profile. *Wellcome Open Res.* 2019;4:36. <https://doi.org/10.12688/wellcomeopenres.15087.2>.
77. Wright J, Small N, Raynor P, et al. Cohort Profile: the Born in Bradford multi-ethnic family cohort study. *Int J Epidemiol.* 2013;42(4):978–91.
78. Koletzko B, von Kries R, Closa R, et al. Lower protein in infant formula is associated with lower weight up to age 2 y: a randomized clinical trial. *Am J Clin Nutr.* 2009;89(6):1836–45.
79. Olsen J, Melbye M, Olsen SF, et al. The Danish national birth cohort—its background, structure and aim. *Scand J Public Health.* 2001;29(4):300–7.
80. Heude B, Forhan A, Slama R, Douhaud L, Bedel S, Saurel-Cubizolles MJ, Hankard R, Thiebaugeorges O, De Agostini M, Annesi-Maesano I, Kaminski M, Charles MA. 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(2):353–63.
81. Charles MA, Thierry X, Lanoe JL, Bois C, Dufourg MN, Popa R, Cheminat M, Zaros C, Geay B. Cohort profile: the French National cohort of children ELFE: birth to 5 years. *Int J Epidemiol.* 2020;49(2):368–9.
82. L'Abée C, Sauer PJ, Damen M, Rake JP, Cats H, Stolk RP. Cohort profile: the GECKO Drenthe study, overweight programming during early childhood. *Int J Epidemiol.* 2008;37(3):486–9.
83. Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van Ijzendoorn MH, de Jongste JC, Klaver CC, van der Lugt A, Mackenbach JP, Moll HA, Peeters RP, Raat H, Rings EH, Rivadeneira F, van der Schroeff MP, Steegers EA, Tiemeier H, Uitterlinden AG, Verhulst FC, Wolvius E, Felix JF, Jaddoe VW. The generation R study: design and cohort update 2017. *Eur J*

- Epidemiol. 2016;31(12):1243–64. <https://doi.org/10.1007/s10654-016-0224-9>.
84. White T, Muetzel RL, El Marroun H, Blanken LME, Jansen P, Bolhuis K, Kocovska D, Mous SE, Mulder R, Jaddoe VVW, van der Lugt A, Verhulst FC, Tiemeier H. Paediatric population neuroimaging and the generation R Study: the second wave. *Eur J Epidemiol.* 2018;33(1):99–125. <https://doi.org/10.1007/s10654-017-0319-y> **Epub 2017 Oct 24.**
85. Eriksson JG, Forsén T, Tuomilehto J, Osmond C, Barker DJP. Early growth and coronary heart disease in later life: longitudinal study. *BMJ.* 2001;322(7292):949. <https://doi.org/10.1136/bmj.322.7292.949>.
86. Guxens M, Ballester F, Espada M, Fernández MF, Grimalt JO, Ibarluzea J, Olea N, Rebagliato M, Tardón A, Torrent M, Vioque J, Vrijheid M, Sunyer J. INMA project. Cohort profile: the INMA-Infancia y Medio Ambiente-(Environment and Childhood) Project. *Int J Epidemiol.* 2012;41(4):930–40.
87. Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, Handal M, Haugen M, Høiseith G, Knudsen GP, Paltiel L, Schreuder P, Tambs K, Vold L, Stoltenberg C. Cohort profile update: the Norwegian mother and child cohort study (MoBa). *Int J Epidemiol.* 2016;45(2):382–8.
88. Järvelin MR, Hartikainen-Sorri AL, Rantakallio P. Labour induction policy in hospitals of different levels of specialisation. *Br J Obstet Gynaecol.* 1993;100(4):310–5.
89. Järvelin MR, Sovio U, King V, Lauren L, Xu B, McCarthy MI, Hartikainen AL, Laitinen J, Zitting P, Rantakallio P, Elliott P. Early life factors and blood pressure at age 31 years in the 1966 northern Finland birth cohort. *Hypertension.* 2004;44(6):838–46.
90. 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(12):831–7.
91. Newnham JP, Evans SF, Michael CA, Stanley FJ, Landau LI. Effects of frequent ultrasound during pregnancy: a randomised controlled trial. *Lancet.* 1993;342(8876):887–91.
92. Chatzi L, Plana E, Daraki V, et al. Metabolic syndrome in early pregnancy and risk of preterm birth. *Am J Epidemiol.* 2009;170(7):829–36.
93. Inskip HM, Godfrey KM, Robinson SM, Law CM, Barker DJ, Cooper C. SWS Study group. Cohort profile: the southampton women's survey. *Int J Epidemiol.* 2006;35(1):42–8.

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

Affiliations

Vincent W. V. Jaddoe^{1,2}  · Janine F. Felix^{1,2} · Anne-Marie Nybo Andersen³ · Marie-Aline Charles^{4,5} · Leda Chatzi⁶ · Eva Corpeleijn⁷ · Nina Donner⁸ · Ahmed Elhakeem^{9,10} · Johan G. Eriksson^{11,12,13,14} · Rachel Foong^{15,16} · Veit Grote¹⁷ · Sido Haakma¹⁸ · Mark Hanson^{19,20} · Jennifer R. Harris^{21,22} · Barbara Heude⁴ · Rae-Chi Huang¹⁵ · Hazel Inskip^{20,23} · Marjo-Riitta Järvelin^{24,25,26,27} · Berthold Koletzko¹⁷ · Deborah A. Lawlor^{9,10,28} · Maarten Lindeboom²⁹ · Rosemary R. C. McEachan³⁰ · Tuija M. Mikkola¹² · Johanna L. T. Nader³¹ · Angela Pinot de Moira³ · Costanza Pizzi³² · Lorenzo Richiardi³² · Sylvain Sebert²⁴ · Ameli Schwalber⁸ · Jordi Sunyer^{33,34,35,36} · Morris A. Swertz^{18,37} · Marina Vafeiadi³⁸ · Martine Vrijheid^{33,34,35} · John Wright³⁰ · Liesbeth Duijts^{1,2} · LifeCycle Project Group

¹ Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, The Generation R Study Group, (Na 29-18), PO Box 2040, 3000 CA Rotterdam, The Netherlands

² Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

³ Section of Epidemiology, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

⁴ Université de Paris, Centre for Research in Epidemiology and Statistics (CRESS), INSERM, INRAE, Paris, France

⁵ ELFE Joint Unit, French Institute for Demographic Studies (Ined), French Institute for Medical Research and Health (INSERM), French Blood Agency, Aubervilliers, France

⁶ Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

⁷ Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

⁸ Concentris Research Management GmbH, Fürstfeldbruck, Germany

⁹ MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

¹⁰ Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

¹¹ Department of General Practice and Primary Health Care, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

¹² Folkhälsan Research Center, Helsinki, Finland

¹³ Obstetrics and Gynecology, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore, Singapore

¹⁴ Singapore Institute for Clinical Sciences (SICS), Agency for Science and Technology (A*STAR), Singapore, Singapore

¹⁵ Telethon Kids Institute, Perth, WA, Australia

¹⁶ School of Physiotherapy and Exercise Science, Curtin University, Perth, WA, Australia

¹⁷ Department of Pediatrics, Dr. von Hauner Children's Hospital, University Hospital, LMU, Munich, Germany

¹⁸ University of Groningen, University Medical Center Groningen, Genomics Coordination Center, Groningen, The Netherlands

- 19 Institute of Developmental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK
- 20 NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK
- 21 Centre for Fertility and Health, The Norwegian Institute of Public Health, Oslo, Norway
- 22 Division of Health Data and Digitalization, Norwegian Institute of Public Health, Oslo, Norway
- 23 MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton General Hospital, Southampton, UK
- 24 Center for Life-Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland
- 25 Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK
- 26 Department of Life Sciences, College of Health and Life Sciences, Brunel University London, London, UK
- 27 Unit of Primary Health Care, Oulu University Hospital, OYS, Oulu, Finland
- 28 NIHR Bristol Biomedical Research Centre, Bristol, UK
- 29 Department of Economics, VU University Amsterdam, Amsterdam, The Netherlands
- 30 Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK
- 31 Department of Genetics and Bioinformatics, Division of Health Data and Digitalisation, Norwegian Institute of Public Health, Oslo, Norway
- 32 Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin, Turin, Italy
- 33 ISGlobal, Barcelona, Spain
- 34 Universitat Pompeu Fabra (UPF), Barcelona, Spain
- 35 CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
- 36 IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain
- 37 Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
- 38 Department of Social Medicine, Faculty of Medicine, University of Crete, Heraklion, Crete, Greece