University of Turin
Doctoral School in Life and Health Sciences

DOCTORAL PROGRAMME IN BIOMEDICAL SCIENCES AND ONCOLOGY
CURRICULUM IN HUMAN GENETICS

Gestational exposures and their association with newborn molecular markers and early childhood outcomes

Elena Isaevska

A thesis submitted for the degree of
Doctor of Philosophy (PhD)

February 2020

AUTHOR

Elena Isaevska, MD
Department of Medical Sciences,
University of Turin, Italy
elena.isaevska@unito.it

THESIS SUPERVISOR

Prof. Lorenzo Richiardi, MD, PhD
Department of Medical Sciences,
University of Turin, Italy
lorenzo.richiardi@unito.it

REVIEWERS

Prof. Michelle Plusquin, PhD
Center for Envionmental Sciences,
Hasselt University, Belgium
michelle.plusquin@uhasselt.be

Prof. Maribel Casas, PhD
Child Health research programme
Barcelona Institute for Global Health, ISGlobal, Spain
maribel.casas@isglobal.org.

Summary

Prenatal and early-life exposures are known to influence long term health trajectories. Several mechanisms are presumed to underlie these associations, among which the most probable and the most studied ones are: DNA methylation, the telomere maintenance system and the microbiome.

Since identifying early risk factors and understanding disease etiology are necessary for planning preventive health measures and creating targeted interventions, the aim of this thesis was to study (1) air pollution exposure and its relation to newborn molecular markers, in particular DNA methylation and telomere length (likely mediators in the association between early-exposure to air pollution and adverse birth and childhood outcomes) by conducting a systematic review to explore potential literature gaps and measuring these molecular markers in cord blood from children participants in the Italian birth cohort Piccolipiu and to (2) study antibiotic use and presence of vaginal infection in pregnancy, maternal factors known to influence offspring's microbiome and explore their association with preschool obesity in children from the Italian birth cohort NINFEA. In both instances, particular importance was given to specific gestational windows of susceptibility that might be of particular interest.

The systematic review provided evidence that air pollutants have the potential to cause alterations in DNA methylation, both on global scale and locus specific, and to unstabilise the delicate telomere maintenance system. These findings were confirmed in our own analysis in cord blood, in particular those relative to locus specific methylation, where we observed alterations in CpGs mapped to genes with roles in cell replication, differentiation and response to oxidative stress. This thesis underlined the beginning of the pregnancy as a particularly sensitive period for exposure to air pollution, when DNA methylation patterns are actively remodeling. Telomere length is also very sensitive to DNA damage and, similarly to previous studies, we observed indications that PM_{10} exposure might alter the telomere maintenance system, although the confidence intervals were wide.

With respect to childhood obesity, we observed an association between vaginal infections in third trimester and BMI at age of 4, while the association with antibiotics was less conclusive. Furthermore, we reported that maternal pre-pregnancy BMI modified the association between vaginal infections in the third trimester and obesity at the age of 4 years.

In conclusion, air pollution during a delicate period in early pregnancy can alter molecular markers at birth, that have the potential to leave consequences, years after the exposure. Improving air quality, therefore, has the potential to protect and promote health since conception. Our observations regarding vaginal infections and antibiotics, point to a potential role of the microbiome in childhood obesity, that needs to be further explored. Public health measures targeted at individual and societal level should also promote healthy BMI before pregnancy, as it seems to be an effect modifier for other common gestational exposures. Future studies should try to pinpoint some of the mechanisms between air pollution and/or other early risk factors with specific childhood outcomes.

List of papers based on this thesis

Study I

Isaevska E, Moccia E, Asta F, Cibella F, Gagliardi L, Ronfani L, Rusconi F, Stazi MA, Richiardi L. *Exposure to ambient air pollution in the first 1000 days of life and alterations in the DNA methylome and telomere length in children: a systematic review*. Environ Res. 2021 Feb; 193:110504. doi: 10.1016/j.envres.2020.110504

Study II

Isaevska E, Fiano V, Asta F, Stafoggia M, Morena Trevisan, Laura De Marco, Pizzi C, Popovic M, Polidoro S, Gagliardi L, Rusconi F, Brescianini S, Nistico L, Stazi MA, Ronfani L, Porta D, Richiardi L. *Air pollution exposure to PM10 and changes in DNA methylation and telomere length in cord blood* [manuscript in preparation]

Study III

Isaevska E, Popovic M, Pizzi C, Fiano V, Rusconi F, Merletti F, Richiardi L, Maule M. *Maternal antibiotic use and vaginal infections in the third trimester of pregnancy and the risk of obesity in preschool children*. Pediatr Obes 2020 Aug; 15(8): e12632. doi: 10.1111/ijpo.12632

Contents

Introduction

1. Developmental Origins of Health and Disease: a conceptual framework for understanding disease etiology and prevention

Today, it is a well-accepted notion among the scientific community that prenatal and early-life factors influence long-term health trajectories.[1] There are public health recommendations stemming from such findings that are even common knowledge among the general population. However, it was not so long ago when not much thought was given on possible long-terms effects of adverse prenatal exposures and the mechanisms underlying them, as it was believed that common non-communicable diseases were predominantly determined by genetic and lifestyle factors later in life.

The discovery in the 1960s that prenatal use of the drug thalidomide[2] caused a range or birth defects, followed by the report that drug diethylstilbestrol (DES)[3] increased the risk for vaginal cancer years after prenatal exposure to DES, caused a shift towards a greater monitoring of the safety of drugs used in pregnancy.[4] Around the same time, reports were published that pointed out the possible adverse effects of maternal smoking[5] and fetal alcohol syndrome[6] was recognized as a special clinical entity. This meant that for the first time, specific lifestyle factors were linked with birth and childhood outcomes. Nevertheless, the attention of the scientific community towards prenatal and early-life exposures did not change significantly until 1986 when a study by the epidemiologist David Barker was published[7], where he reported an association between prenatal nutrition and late-onset coronary heart disease, unraveling the now well-known relationship between low birth weight and later metabolic diseases. Since Barker's initial findings, the results have been replicated in diverse populations, such as children conceived during the 1918 flu pandemic[8], the Chinese Leap Forward[9], or the most studied example, children conceived during the Dutch Famine[10], that presented as well designed, although devastating, human experiments on the effects of adversity during pregnancy on adult health.

The Barker's theory on the specific mechanism behind these associations was that "*nutritional, hormonal and metabolic environment afforded by the mother may permanently programme the structure and physiology of her offspring*".[11] The driving principle behind this mechanism was the phenomenon of "*developmental plasticity*" and "*fetal programming*". Barker claimed that there is a critical period when a system is plastic and sensitive to the

environment (such as the period *in utero*), followed by loss of plasticity and a fixed functional capacity.[11]

This theory stimulated lot of interest in the fetal origins of adult disorders. The term "fetal programming" has been more recently replaced with the term "early life programming" as it was seen that adaptations to environmental insults may occur at different time points during prenatal and early postnatal life and they vary by the type, dosage, duration and timing. This early life period, from conception to the 2nd year of life, commonly referred to as "the first 1000 days of life", is the specific window of susceptibility studied in relation to the Barker's hypothesis, now known as the Developmental Origins of Health and Disease hypothesis, or DOHaD[1,12], now established as a conceptual framework and a research field in its own right. Although the initial studies were based mostly on epidemiological research, the DOHaD hypothesis expanded to include studies from various disciplines that integrated epidemiological, clinical, molecular and economic data[13–16], to name a few, with wade array of clinical implication for medical specialties such as gynecology, pediatrics and public health, as well as for developmental biology, anthropology, economics and social sciences.

In the past decades, studies nested within the DOHaD hypothesis revealed that many important environmental exposures during early development with a known effect on later disease outcomes, are in fact modifiable. For example, recommendations to avoid smoking in pregnancy, in addition to globe-wide public health measures to reduce smoking in general, came after smoking in pregnancy was consistently associated with wide-array of adverse pregnancy and birth outcomes.[17–19] The implementation of mandatory food fortification with folic acid in some countries and the recommendation that any woman planning to become pregnant should consume a folate dietary supplement before and during pregnancy, came after multiple studies confirmed that folate supplementation in the periconceptional period reduces the risk for neural tube defects.[20,21] More recently it was seen that maternal pre-pregnancy BMI increases the risk of pregnancy complications, adverse birth outcomes and childhood obesity, making it an additional critical point of pre-conception counseling.[22] Although, there is still a long way to go, these type of public health interventions were shown in the past to be remarkably successful in promoting and protecting the health of families and populations, through education and promotion of healthy behavior and by translating research findings into public health policies targeted towards changes on individual, as well as societal level. [23]

DOHaD research underlines that maternal and family circumstances, as well as the surrounding environment are of extreme importance during a vulnerable time of intense growth and development, such is the period during the first 1000 days of life, with rippling effects on health and disease trajectories throughout life course.[24] Having data on exposures during the first 1000 days of life and knowledge on the potential mechanisms involved is therefore essential to improve the health of individuals and plan prevention interventions.[24]

The primary candidate mechanisms that might help understand how the environment shapes trajectories of health and disease across the lifespan, include the involvement of the epigenome (mostly referring to DNA methylation)[1], the telomere maintenance system[25] and the microbiome.[26] They are all known to be sensitive to both genetic and early-life environmental influences, beginning *in utero*; are key players in many cellular and metabolic functions with particular involvement in growth and development, and their early-life change under environmental influence can leave long-lasting molecular fingerprint, possibly influencing risk or susceptibility to disease years after the exposure.

This thesis is focused on two common, but potentially modifiable, gestational exposures:

1) Air pollution, a ubiquitously present environmental factor, linked with a wide array of birth and childhood outcomes. The majority of this thesis will be dedicated to studying the association between prenatal and early-life exposure to air pollution and newborn molecular markers, in particular DNA methylation and telomere length, likely mediators in the association between early-exposure to air pollution and adverse birth and childhood outcomes.
2) Antibiotic use and presence of vaginal infection in pregnancy, maternal factors known to influence offspring's microbiome (a complex system known to influence wide array of metabolic functions) and study their association with preschool obesity, another important public heath challenge.

In both instances, particular importance was given to explore specific gestational windows of susceptibility that might be of particular interest, based on *a priori* hypothesis.

2. Exposures and outcomes studied within the context of this thesis

2.1 Studying air pollution exposure within the context of DOHaD

Air pollution exposure represents one of the main environmental challenges of the 21 century, with devastating consequences on climate and with alarming implications on public health. The latest report from the World Health Organization (WHO), called air pollution "*the major environmental health threat for children*" after estimating that globally 93% of all children are exposed to pollutant concentrations higher than the recommended WHO guidelines. The ubiquitous presence of air pollution, in both high and middle income countries, and the overwhelming evidence of its effect on child's health[27,28], makes acute and chronic exposure to elevated levels of air pollutants a global public health emergency.[29]

Pregnant women and young children are specifically susceptible to air pollutants. Women's respiratory rate increases during pregnancy to accommodate the new metabolic demands, increasing both her own exposure and that of her fetus.[30] Air pollution exposure during pregnancy may directly affect the fetus, as was seen that ultrafine particulate matter and heavy metals, can enter the mother's bloodstream, cross the placental barrier and reach the fetus, affecting growth and development.[31,32] As written earlier, the fetal period is characterized by accelerated growth and developmental plasticity, making the fetus especially vulnerable to environmental stimuli. Air pollution exposure is also known to increase the risk for hypertensive disorders in pregnancy, such as pre-eclampsia, indicating that prenatal air pollution exposure can have indirect effect of child's health by affecting maternal health during pregnancy.[33,34] Young children, on the other hand, are susceptible to the adverse effects of air pollution exposure[35] due to other physiological, environmental and behavioral factors. Children have higher respiratory rate than adults, their organs are still in development including their lungs[36] and brain[37], they spend more time outside and they move closer to the ground where some pollutants reach high concentrations.

There is substantial evidence that links prenatal exposure to air pollution to pregnancy complications,[34] preterm delivery and low birth weight.[33] Both prenatal and postnatal exposures were reported to increase the risk for asthma and respiratory diseases[38,39], metabolic disorders, such as obesity,[40] some types of childhood cancers[41,42] and some neurodevelopmental disorders.[43] It is hypothesized that air pollution exerts its main effects on human health via production of reactive oxygen species (ROS) and inflammation[44,45], that in turn can have direct or indirect damaging effect on DNA. Air pollutants are believed to

interfere with DNA integrity and stability as well as the response of DNA repair mechanisms.[46] Furthermore, air pollution was seen to interfere with DNA methylation and the telomerase maintenance machinery, responsible for key cellular functions.[47] In the past decades, many studies on air pollution and DNA methylation and/or telomere length in adults confirmed their association[48,49], however, until recently the number of studies on gestational air pollution exposure and its effect on important molecular markers was limited. Today, due to the availability of pollution data and advancements in cost-effective high-though output array technologies, this is an extremely active research area that is just beginning to provide evidence on the importance of monitoring and improving air pollution quality, due to the effect of pollutants on molecular markers measured as early as birth.[46,47,50–52]

Air pollution is complex mixture of pollutants that include solid and liquid particles suspended in the air (such as particulate matter) and various gases including nitrogen oxides, ozone, carbon monoxide and volatile organic compounds. Traditionally, the concentration of pollutants in ambient air is assessed using specific instruments at fixed monitoring stations, that allow daily measurement of air quality at high spatial resolution. Since it is a requirement for each country to monitor air quality indicators, many studies can now use data from fixed monitoring stations, or other sources, to study the effects on air pollution on human health.[53] The concentration and type of pollutant (and in the case of particulate matter their chemical composition), apart from their specific source, such as industrial or traffic-related pollution, are influenced by geographical, seasonal and meteorological factors, as well as background photochemical processes in the atmosphere.[53]

One of the most commonly measured air pollutant is particulate matter. Particulate matter represents air suspended microscopic particles, generated from combustion engines, mechanical processes such as construction, power plants, road dust resuspension and wind. The risk for many health and disease outcomes, including mortality, are known to increase in parallel with the increasing concentrations of particulate matter in atmospheric air. Furthermore, effects on health were seen both after short and long-term exposures and there does not seem to be any threshold bellow which particulate matter cannot influence human health. This lead the International Agency for Research of Cancer (IARC) and the WHO to classify particulate matter as Group 1 carcinogen.[54] Particulate matter is mostly measures as either as PM_{10}, that include particles with a diameter <10 micrometers (μm), or as $PM_{2.5}$ measuring fine particles with a diameter < 2.5μm. Since PM_{10} measures all particles that can enter the respiratory tract, by convention including both particles between 2.5 and 10μm and

particles measuring < 2.5μm, PM_{10} is often used as a proxy indicator of overall air pollution quality. [55]

Recently, several advanced statistical modelling approaches were developed, that allowed researchers to accurately estimate air pollution exposure based on residential/working address, by integrating data from fixed monitoring stations, satellite data, as well as other characteristics of the area such as meteorological conditions.[56–59] These methodological advancements allowed studying gestational exposures to air pollutants based solely on maternal residential address, rather than using direct exposure assessment methods, that are more expensive and time consuming, such as giving mothers personal air monitors or estimating pollutant metabolites in urine or other tissues. Furthermore, these models allow studying simultaneously both larger and smaller gestational windows of exposure. They also have some limitations, such as the potential exposure misclassification that may arise from determining exposure solely on residential address without taking into consideration individual time activity patterns.

2.2 Studying exposures linked with childhood obesity within the context of DOHaD

Childhood obesity is one of the most serious public health challenges of the 21st century due to its complex etiology, long-term consequences and high prevalence throughout the globe, with trends that plateau in high-income countries, with prevalence remaining relatively high, and ever increasing trends in low and middle income countries.[60,61] Childhood obesity is a public health issue difficult to tackle because it is a complex interplay of many genetic, nutritional, behavioral and socio-economic factors. Although there seems to be a strong genetic component for adiposity in childhood,[62] its increasing prevalence in the last 40 years[60] implies a strong environmental role in the onset of childhood obesity. The last decades were marked by many societal changes that included changes in food production and marketing, shift towards cheap energy-dense foods and sedentary behaviors, encompassed by increased urbanization and exposure to environmental toxins such as air pollution.[61,63]

Children are especially susceptible to such obesogenic environments and are heavily influenced by the burden of early-onset obesity that many of them carry well into adult life. [64] Adiposity in childhood is long known to be linked with high blood pressure and cholesterol, insulin resistance, asthma, sleep apnea, fatty liver disease, joint problems, musculoskeletal pain and precocious puberty, as well as increased risk for diabetes and cardiovascular

diseases later in life.[65–69] In addition, overweight children often suffer from low self-esteem, anxiety and depression, due to higher rates of bulling, social isolation and stigma.[70]

The BMI curve in childhood rises in two time points. The first rise starts after birth and peaks around the first year of life and then decreases. The second rise in the BMI curve, occurring between ages 5 and 7 years, is called "adiposity rebound".[71] Early adiposity rebound, before the age of 5, is associated with an increased risk of overweight, and is predictive of obesity in adolescence and adulthood,[72] as well as with a wide array of metabolic diseases.[71,73] Studying obesity in preschool children, therefore, is extremely important, since adiposity at that age suggests involvement of factors that operate very early in life. Several prenatal factors, such as high pre-pregnancy BMI, excess gestational weight gain and maternal tobacco use have already been clearly identified as early risk factors for obesity in children.[74]

Several studies analyzed prenatal antibiotic exposure in association with childhood obesity.[75–83] The rationale behind studying early-life antibiotic exposure in relation to childhood obesity is the ability of the antibiotics to interfere with metabolic pathways, probably though the human microbiome (see chapter "The microbiome"), leading to long lasting metabolic consequences.[84] Antibiotics are among the most commonly prescribed drugs in pregnancy[85] and their use in many circumstances is necessary for timely treatment of infections and reducing the risk of pregnancy complications. The antibiotics approved for use during pregnancy are considered generally save drugs, apart from limiting side-effects, and are commonly prescribed to treat mainly respiratory, ear, nose and throat, urinary, or genital infections during pregnancy.[86]

The link between antibiotic exposure and metabolic outcomes is not recent. The growth promoting effect of antibiotics was known 80 years ago, when farmers noticed that regular administration of small doses of antibiotics, will increase growth and weight gain in many species of farm animals, increasing their profits.[87] This was observed for different classes of antibiotics, indicating that the mechanisms might not be related to specific antibiotic type, but rather to their overall effect on microbial populations. Importantly, it seems that the effect on growth and weigh gain in strongly influenced by the timing of exposure, the earlier in life the exposure began, the greater the effect. The significance of timing of exposure indicates that antibiotics might interfere with metabolic programming during early development, with findings consistent with the DOHaD hypothesis.[84,88,89]

Apart from animal experimental models, the association between antibiotic use in infancy and obesity and other markers of metabolic syndrome, was also seen in epidemiological studies.[84] However, the association between antibiotic exposure during pregnancy and obesity in children is not well understood, possibly due to the complex interplay between many important factors that exert their effect in this period, with roles difficult to disentangle. Some of the studies on this topic reported positive associations between antibiotic exposure in some trimesters, or certain types of antibiotics and overweight or obesity, while others studies reported null associations.[75–83] Most of the studies on antibiotics either used BMI as a continuous variable or included children who are underweight in the reference category, potentially attenuating the association with overweight/obesity. Furthermore, women undergoing cesarean section are typically treated with antibiotics just before or during delivery to prevent infections, implying that, studying children born with cesarean section together with those born though vaginal delivery might further confound the results.

Since the first major microbial colonization of the newborn happens during the birthing process by bacteria colonizing the maternal vaginal and intestinal microbial communities,[90] in addition to antibiotics, we decided to study vaginal infections in pregnancy, that similarly to antibiotics could ultimately influence early-life metabolic programming.[26] Bacterial vaginal infections are defined as the presence of pathological microbial strain or as disbalance between commensal macrobiotic communities, often accompanied by symptoms such as abnormal vaginal discharge.[91]

Overgrown bacterial giving rise to vaginal infections during pregnancy, can also invade the *in utero* environment and stimulate an inflammatory cascade that can lead to maternal and fetal morbidity[26], in particular to increased risk for chorio-amnionitis and premature rupture of the membranes and preterm birth[92–94], which are all known risk factors for later metabolic diseases.[95] Vaginal infections and antibiotic use are often intertwined in pregnant women, as antibiotic use for other causes can cause vaginal disbalance, while in the same time antibiotics are drug of choice for many bacterial vaginal infections. To date, there are no studies that explore the possible association between vaginal infections during pregnancy and childhood BMI outcomes.

Identifying modifiable risk factors for obesity in children as early as in pregnancy might have long-term beneficial effect on the child's health. Antibiotic use and vaginal infections in

pregnancy are one of the most common exposures in pregnancy. Furthermore, both of them are believed to strongly influence neonatal gut microbiota.

3. Mechanism though which early exposures can affect the long-term health

In the following paragraphs, I will provide an introduction into the molecular mechanisms that are most likely to underlie the exposure-outcome associations seen within the DOHaD hypothesis, that were directly or indirectly studied within this thesis:

- The epigenome (with focus on DNA methylation) and the telomere maintenance system that will be measured in cord blood and studied with relation to air pollution exposure
- The microbiome that will be studied indirectly, by studying two exposures closely related to it: antibiotic use and vaginal infections

I will outline what makes them interesting within the context of the DOHaD hypothesis, as well as clarify some specific terminology that will be used throughout the Manuscript. Potential windows of susceptibility within the first 1000 days of life that raised some hypotheses tested in this thesis will also be underlined.

3.1 The epigenome

The epigenome is believed to be one of the key mediators within the framework of DOHaD. The term epigenetics (from the Greek prefix "epi" meaning "over", "outside of", "around of") was originally conceived in 1940s, by the British developmental biologist, Conrad Waddington to describe the existence of mechanisms of inheritance in addition to genetics. He claimed that individuals possess a developmental plasticity that is sensitive to environmental stimuli.[96] Although it was conceived around 40 years prior, the Waddington's definition of epigenetics interestingly overlaps with the Barker's definition of developmental plasticity as "*the phenomenon by which one genotype can give rise to a range of different physiological or morphological states in response to different environmental conditions during development*". In 1990s, after it was seen that changes in DNA methylation can influence gene expression, Robin Holiday updated the definition, by saying that epigenetics is the temporal and spacial control of gene activity during the development of an organism.[97] Further research into the epigenetic molecular patterns revealed that the same set of genes expressed in parent cell is also expressed in daughter cells, in order to maintain cellular

identity. This lead to the current definition of epigenetics by Arthur Riggs, that defines it as mitotically and/or meiotically heritable changes in gene function that cannot be explained by alterations in the sequence of DNA.[98]

The most commonly studied mechanism of epigenetic control is the DNA methylation and it is the primary candidate mechanism by which early-life conditions may leave stable molecular-level alternations with long-term health effects.[1] Several prenatal exposures were robustly linked with DNA methylation. Maternal smoking in pregnancy was consistently associated with newborn methylation of thousands of CpGs in genes relevant for child health and development.[99–101] A prediction score based on these CpGs was able to accurately assign prenatal exposure to smoke 30 years after the exposure.[102] Maternal body mass index (BMI),[103–105] folate levels[106,107], glucose concentrations and gestational diabetes[108,109], hypertensive disorders of the pregnancy[110,111] and maternal stress[112,113] during pregnancy were all reported to influence DNA methylation patterns at birth in more than one study.

DNA methylation is essential for cell viability, especially during the period studied by DOHaD, as we will see later. With the advancement in genome-wide methylation analyses, it is now known that the role of DNA methylation at cellular level goes far beyond that of a gene expression. DNA methylation is particularly important during early embryonic development, where the coordinated expression of gene ensures cell differentiation. The stability and mitotic hereditability of the methylation patterns prevents daughters cell to regress into an undifferentiated state. DNA methylation plays an important role in sex chromosome dosage compensation and the inactivation of one X chromosome in females, the suppression of repetitive elements that might lead to genome instability and the coordinated expression of imprinted genes necessary for intrauterine growth and development. In order to understand what functional consequences can arise from changes in DNA methylation, it is pertinent to explain what exactly is DNA methylation, where is it found across the genome and how its position can determine its function.[114,115]

DNA methylation is the addition of a methyl group to a cytosine base. It is predominantly found at CG dinucleotides i.e. when cytosine (C) is followed by guanine (G) along the DNA strand making the so-called CpG site, with the “p” representing the phosphate bond between them. The reason why methylation occurs exclusively at CpG dinucleotides is that dinucleotides are symmetrical when we look at double stranded DNA. During replication the double DNA helix “unzips” and both separated strands serve as templates for new DNA

strands. However, after replication, only the "parent" DNA strand will have the methylated cytosine, while the new "daughter" strand will not. This is when specific DNA methyltransferase enzymes, called DNMT1s, come in play. [116] They are responsible for the maintenance of the methylation patterns after replication. They have preference for hemi-methylated DNA and when they recognize the unmethylated "daughter" strand they lay down a methyl group to the unmethylated cytosine and restore the fully methylated CpG site. The symmetry of the CG dinucleotide allows methylation to be maintained through cell division and makes methylation a stable epigenetic mark. DNA methylation is found at specific places across the genome and has different functional consequences depending on both the methylation status and the genomic context where it is found.[117]

CpG dinucleotides are often found in clusters across the genome, to make the so called **CpG islands**. CpG islands are genomic regions where more CG dinucleotides are found than what might be expected by chance. They tend to be found around gene promotors where transcription factors bind to initiate gene transcription. The general rule is that CpG islands tend to be protected from methylation i.e. hypomethylated.[115,118] If the CpG island is methylated, this is almost universally associated with long-term gene silencing.[119] Methylation in CpG islands silences gene expression by attracting proteins that repress transcription, induce chromatin condensation and prevent transcription factors to bind. Apart from promoters of housekeeping genes that are always protected from methylation[120] (since they are needed for basic cellular functions, and are expressed in all cells under normal and pathological conditions), the gene silencing role of DNA methylation shows how a single genome can give rise to hundreds of different cell types during embryogenesis.[114]

In contrast to CpG islands, the function of **DNA methylation in the gene body** is still somewhat enigmatic, since CpGs in gene bodies are generally, but not exclusively, found to be methylated, but they do not seem to be associated with gene silencing. However, it is hypothesized that it has some silencing function on cryptic promoters hidden in gene bodies, as their activation might attract RNA polymerase to initiate transcription of aberrant peptides or interfere with the function of another RNA polymerase loaded on a canonical promoter.[119,121] Methylation in gene bodies might also be needed to silence cryptic splice sites. Splice sites usually surround gene exons (coding gene regions) to mark the place where splicing should occur, and are characterized by a specific DNA sequence recognized by the splicing machinery. If this sequence, by chance, is found somewhere else in the genome (called cryptic splice site) it also needs to be silenced by DNA methylation.[119,121]

The majority of DNA methylation in the genome, however, is not found neither in CpG islands nor gene bodies, but rather in **repetitive elements**. Repetitive elements make up for more than 50% of the entire genome. They are repetitive, non-coding sequences of DNA, previously called "junk" DNA, but now are known to have an evolutionary function. [122] In general, what repetitive elements do is that they make a copy of themselves, "jump" out and insert themselves somewhere else in the genome. They can move by using "copy and paste" mechanism or "cut and paste" mechanism, depending whether they leave a copy in their original position. Although these mobile genetic elements play important roles in shaping genomes during evolution, their actions are clearly mutagenic and can lead to genomic instability.[123] In order to maintain genomic stability, there is a need to silence these repetitive elements by DNA methylation.[124] To understand the functional consequences of altered DNA methylation in CpG islands, intergenic regions and repetitive elements, it is interesting to look at DNA methylation patterns in cancer cells. In fact, changes in DNA methylation are one of the hallmarks of cancer, where intergenic regions and repetitive elements tend to be unmethylated i.e. we see global hypomethylation, indicative of genomic instability. In cancer tissues there are also significant changes in CpG islands that induce activation of oncogenes and silencing of tumor-suppressor genes.[125]

The other two places in early embryo where methylation in necessary are imprinted genes and the extra X-chromosome in females. Females have two large X chromosomes and males have one X and one smaller Y chromosome. DNA methylation is responsible for the random inactivation of one X chromosome in females[126] and allows sex to be accurately predicted by measuring DNA methylation on the sex chromosomes.[127] Failure to inactivate the extra X chromosome in female results with early embryonic lethality.[128] DNA methylation also regulates the expression of imprinted genes, known to be expressed in a parent-of-origin-specific manner. DNA methylation silences the expression of the allele belonging to the other parent. The majority of imprinted genes are involved in embryonic growth and development, and the development of placenta, with paternally-expressed genes promoting intra-uterine growth and resource allocation, while maternally-expressed genes prevent unnecessary metabolic demands on her resources.[129,130]

3.1.1. Fetal programming of DNA methylation patterns

The establishment of DNA methylation patterns in early gestation and their stable maintenance throughout life perfectly encapsulates Barker's definition of "*developmental*

plasticity": there is a critical period of development when the system is plastic and sensitive to the environment (early gestational period), followed by loss of plasticity and a fixed functional capacity.[11] The zygote, the newly formed cell after fertilization, contains both maternal and paternal DNA and therefore all necessary genetic information to form a new individual. Unlike the genomic DNA sequence which is directly inherited from the parents, the methylation patterns are removed from both maternal and paternal DNA to ensure totipotency for the next generation in a process called "*epigenetic reprogramming*".[114,117] This process unfolds in several consecutive phases. The first wave of global epigenetic changes happens in the pre-implantation period when the embryo passed only though few cell replications. This pre-implantation period is characterized with a wave of demethylation. It was reported that demethylation of paternal methylation patterns happens very fast and most likely involves the active role of DNA methylation enxymes called "erasers". The erasure of the maternal DNA methylation patters happens more slowly and it is believed to be due to passive demethylation. DMNT1 enzyme known for the maintenance of methylation patterns with preference to hemi-methylated DNA, is translocated to the cytoplasm during this stage and therefore maternal methylation patterns are erased passively by being "diluted" slowly with each cell replication. It should be noted, the methylation of the parent-of-origin imprinted genes is protected from demethylation. It is also possible that other regions such as repetitive elements, may also be partially protected from this process.[114,115]

The second wave of global methylation changes happens around the time of implantation. At this point the entire genome goes through a dramatic wave of *de novo* methylation. This active process is guided by methylation enzymes specifically responsible for *de novo* methylation called DMNT3a-DMNT3c, present in large concentrations in the nuclei of the early embryo. CpG islands which regulate many housekeeping genes remain preferentially unmethylated. The re-methylation process also provides blanket inactivation of cryptic promoters and splicing cites. This wave of re-methylation creates the typical bimodal pattern: hypermethylation of CpG sites dispersed throughout the genome that need to be silenced, while CpG islands remain hypomethylated. The bimodal pattern created at this stage is then maintained and remains largely intact over the next cell divisions.[114,115,117]

The third stage of the "epigenetic reprogramming" of the early embryo happens during the post-implantation period. The changes are not on global scale, but they include discrete changes in the basal methylation patterns. This process corresponds to the initiation of embryonal cell differentiation and includes silencing of the pluripotency genes to prevent

regression into an undifferentiated state. Once the new methylation patterns are established, they create specific cellular identity and tissue-specific template that is extremely stable and generally maintained for the rest of the organism's life. This is made possible by the re-introduction of the DMNT1 enzymes responsible for methylation maintenance. Although these baseline patterns are established in early life, there are postnatal windows where epigenetic patterns might be subdued to further changes in order to respond to new requirements, such pubertal changes triggered by exposure to sex-specific hormones.[131,132] The process of "epigenetic reprogramming" is essential for embryonal survival.[114,115,117] If this ubiquitously important process goes wrong on a large scale, it can result with early embryonic lethality. Subtle changes compatible with life are known to exist, and they are believed to influence later susceptibility or trigger the emergence of different metabolic, immune and neurodevelopmental diseases.[1]

3.2 Telomere length maintenance system

Telomeres, as their names suggest (from the Greek words "telos" and "meros" translating as "end" and "part", respectively) are regions of non-coding repetitive nucleotide sequence 5′-TTAGGG-3′, that cap the ends of chromosomes and is bound by protective proteins.[133] They have a major function in maintaining chromosome stability, since the open end of chromosomes might be recognized as "broken" DNA, leading to fusion with other DNA ends, and/or initiating DNA recombination.[133] The second major role of telomeres takes place during cell replication. The DNA replication machinery cannot copy completely the DNA all the way to the end of the chromosome, so instead of genes being truncated at the end of the chromosome, the ends of telomeres are those being shorten with each cell division.[133] Telomeres are also very sensitive to environmental influences, such as oxidative stress. That is because telomeric DNA is enriched with G-clustered nucleotides that are more prone to oxidative damage than the rest of genome[134] and second, the resulting DNA damage cannot be resolved since DNA repair enzymes have limited access due to the telomeric protective proteins that cover them.[133,135] The interest in studying telomeres raised in the last decades, as evidence accumulated that telomere shortening is proportionate to age, and that the risk of common age-related diseases, as well as mortality risk, increases with shorter telomeres. Excessive telomere shortening induces cell senescence and genomic instability, that leads to mitochondrial malfunction, activation of pro-inflammatory processes and eventually apoptosis.[133,136]

The 2009 Nobel Prize in Physiology or Medicine went to three researchers that discovered the existence of the telomerase enzyme, the second key player in the telomerase maintenance system.[137] The telomerase adds telomere repeat sequence to the 3' end of telomeres elongating them. The telomerase is especially active is some cell types such as embryonic stem cells, since without the presence of telomerase, the few zygotic cells would not be able to pass through many replication cycles in order to become a fetus. In many adult somatic cells, with exception of those that need to divide regularly (such as active lymphocytes), the level of the telomerase is limited.[133,138]

There is strong genetic influence on the telomere maintenance system with hereditability estimates going from 30% to 80%.[133,139] Many genetic syndromes associated with premature aging, exhibit shorter telomeres. Telomeropathies, disorders with mutations in genes that lead to defects in the telomere maintenance machinery, also show characteristic signs of aging.[140] On the other hand, genetic determinants for longer telomeres raise the risk for specific cancer types and overall cancer mortality.[141] This goes in line with findings from several cancer cell types where the activity of the telomerase is up-regulated, providing cellular immortality. Taking all this into account, it seems that, on genetic level, there is some sort of trade-off between lowering the risk for age-related diseases, such as cardiovascular diseases, and increasing the risk for some cancers.[133,136] Although telomeres shorten with age, reflecting both limited cellular replicability and the cumulative effect of oxidative stress throughout life, many new evidence does not support the notion that telomere length is a "ticking mitotic/replicative clock", but rather a more complex phenotype.[138,142] The telomere length and the rate of telomere attrition is known to be variable among adults with the same age and, as many other complex phenotypes, telomere length is longitudinally influenced both from underlying genetics and the environment. In fact, many studies have confirmed that telomere length is strongly influenced by sex, body mass index, diet, physical activity, sleep, socio-economic adversity, psychologic stress and environmental exposures.[133,138,142]

The most used tissue for studying environmental influences on telomere length is peripheral blood (or cord blood in neonates). The telomere length in blood reflects the mean leucocyte telomere length, that in turn reflects systemic influences on telomere length in general, as well as immune cell senescence, that increase replications during inflammatory responses. The latter is important, since systemic inflammation is a hallmark of many non-communicable diseases, from cardiovascular to metabolic diseases and dementias. Furthermore, in some immune cells, such activated T-lymphocytes, the telomerase is still

active.[143] That is because is necessary to preserve the cell's ability to proliferate in response to external or internal stimuli. However, telomerase levels in T-lymphocytes are not static and they decrease in parallel with T-cell differentiation and age.[143] This further demonstrates that, leucocyte telomere length, at least in early life and adulthood, is actively modeled both by genetic and environmental influences.

3.2.1. Fetal programming of the telomere length

After it was seen that intrauterine exposures can influence the susceptibility to many non-communicable disease, it was hypothesized that early programming of the telomere maintenance, in addition to epigenetics, may be one of the molecular markers that might explain these findings.[144] Maternal psychological stress,[145–147] maternal vitamin D concentration,[148] maternal folate levels,[149] gestational diabetes[150] and pre-pregnancy BMI[151] are all linked with telomere length. Furthermore, *in utero* exposure to tobacco also seems linked with telomere length in a dose response pattern.[152]

Telomere length and telomerase activity are known to show early developmental plasticity.[25,133] This is supported by several findings. First, newborns exhibit variability in telomere length, and the degree of variability is comparable to that of adults from a same age, indicating that individual variation in both newborn and adult telomere length originates *in utero*. Second, studies that measured telomere length at several time points, showed that the majority of individuals maintained their telomere length ranking, indicating that that telomere length at birth may be the principal determinant of telomere length throughout life. Third, although the hereditability of the telomeres seems to be high[139], the known genetic variants known to influence telomere length account only for 2-3% of leucocyte telomere variability[141]. It is possible that heritability estimates may exaggerate the role of genetics, since they may also include intrauterine effects, suggesting a considerable role of prenatal exposures in the initial setting of telomere length.[138]

Telomere length at birth reflects the number of cell divisions (reflected in somatic growth and gestational age with preterm newborns having longer telomeres), exposure to oxidative stress and the activity of telomerase that hinders telomere shortening.[138,153,154] What we know about the telomeres and the telomerase maintenance system during the embryonal and fetal period comes from limited number of studies. It is believed that during early embryonic development all tissues have similar telomere length due to the universal activity of telomerase in all cells, that is needed to support embryonal growth.[155] An accelerated decline

in mean leucocyte telomere length is seen between 27 and 32 gestational weeks consistent with the high proliferation rate of hematopoietic cells before 32 gestational weeks.[153] Another period of significant telomere length shortening is during the first 3 years, that corresponds to the high proliferation rate of the immune cells during the process of developing acquired immunity.[153] After 3-4 years follows a period of relatively constant but moderate loss of telomeric repeats.[156–158]

3.3 The microbiome

The human body is host to a vast number of microbes, both commensal, and pathogenic, with population that counts around 10^{13}, with 500–1,000 species and 100 times more genes than our own human genome, that collectively constitute our microbiota, or most famously known the "microbime", a term used to refer to both microorganisms and their genes.[159] While bacteria colonize great number of surfaces on human body, such as skin, oral cavity and vagina, the majority of bacteria are found within the gastrointestinal tract, where they have important role in protection from pathogens, absorption of nutrients from food, synthesis of vitamins and enzymes, as well as the production of short-chain fatty acids, a primary products of the breakdown of non-digestible carbohydrates and a major source of energy for epithelial cells, pointing to the considerable involvement of the gut microbiome in metabolic processes. [159–161] The advancement in technologies for sequencing and studying the microbiome, including the International Human Microbiome Consortium, the European Commission's Metagenomics of the Human Intestinal Tract project, the US National Institutes of Health's Human Microbiome Project,[159] revealed many previously unknown interactions between the host and microbiome and vice versa, with effects on immunity, metabolism, and neuroendocrine responses that can modulate energy metabolism, fat accumulation, insulin resistance and systemic inflammation.[161,162] Studies suggest relationships between gut dysbiosis, defined as shift in microbial composition or diversity, with many multifactorial diseases such as obesity, cardiovascular diseases, colon cancer, inflammatory bowel disease, atopy and diseases such as autism and mood disorders.[160–163]

As many other mechanisms implicated in the development of multifactorial diseases, the human gut microbiome is influenced by both genetic and environmental factors. It seems that the physical architecture of the gut, that is genetically defined, defines the composition of the gut microbiome, since within the same species, there is a great similarity of the microbiome.

Apart from genetics, diet and lifestyle are implicated as the main reasons behind the human inter-individual variation of gut microbiome.

3.3.1. Programing of the neonatal microbiome

The development of the gut microbiota begins at birth, when newborns receive much of their initial colonizing microbiota. It is believed that there is are critical periods for the development of the neonatal and infant microbiome that might have long lasting metabolic consequences. These periods are pregnancy, birth, breastfeeding period and introduction of solid food.[26,164,165]

It is likely that programming of the neonatal microbiome starts *in utero*. Changes in the maternal gut and vaginal microbiome might have an effect on the initial neonatal colonizing bacteria, since the first major microbial colonization of the newborn during vaginal delivery happens during the birthing process.[26,84,164–166] Over the course of a normal, healthy pregnancy, the body undergoes though a variety of changes in order to allocate energy and increase metabolism to ensure and support the fetal growth. These changes are encompassed by changes in the gut microbiota that would allow higher energy extraction from food. In fact, when third trimester gut microbiota was transferred to germ-free mice it induced greater adiposity and insulin insensitivity compared to first trimester gut microbiota.[167,168] Pregnancy also induces changes in vaginal microbial communities, towards reducing the overall diversity and richness, with a dominance of Lactobacillus species[169,170] that protect the mother from infections and provides the newborn with bacteria that support growth and thriving. Imbalances in this vaginal microbiota during pregnancy, including patterns comparable with for vaginal dysbiosis, are linked with negative pregnancy outcomes, such as pregnancy loss and preterm birth.[171–173] Antibiotics use during pregnancy are known to alter the maternal vaginal flora.[174]

The second time point where individual's circumstances might influence neonatal microbiota is the mode of delivery.[77,175] It was reported that the microbiome of neonates born via vaginal delivery had closely resembles their mothers' vaginal microbiome, that include potentially beneficial microbiota such as Lactobacillus, Bifidobacterium, and Bacteroides, while the microbiome of neonates born via cesarean section has more similarities with skin flora.[26,164,165]

The third and fourth time points of intense changes and maturation when the gut microbiota undergoes important changes, are the period of breastfeeding and introduction of solid food.

Neonates that are breastfed and those fed with milk formula have different gut microbiome patterns. This is because neonates, apart from nutrients, though breastmilk, receive healthy bacteria, antimicrobial proteins, secretory IgA that has a protective effect against pathogens, as well as oligosaccharides and other breastmilk components that support the growth of healthy gut bacteria such as Bifidobacterium.[176,177] The infant microbiota then undergoes maturation and increases in diversity and stability, in correspondence with the correct introduction of solid food, resembling the microbiota of adults by the age of three years.[26,164,165]

In summary, the development of neonatal and infant microbiome is under strong influence of factors acting prenatally and during early infancy, that have the potential to influence growth trajectories and metabolic programming.

4. Birth cohort studies within the context of DOHaD

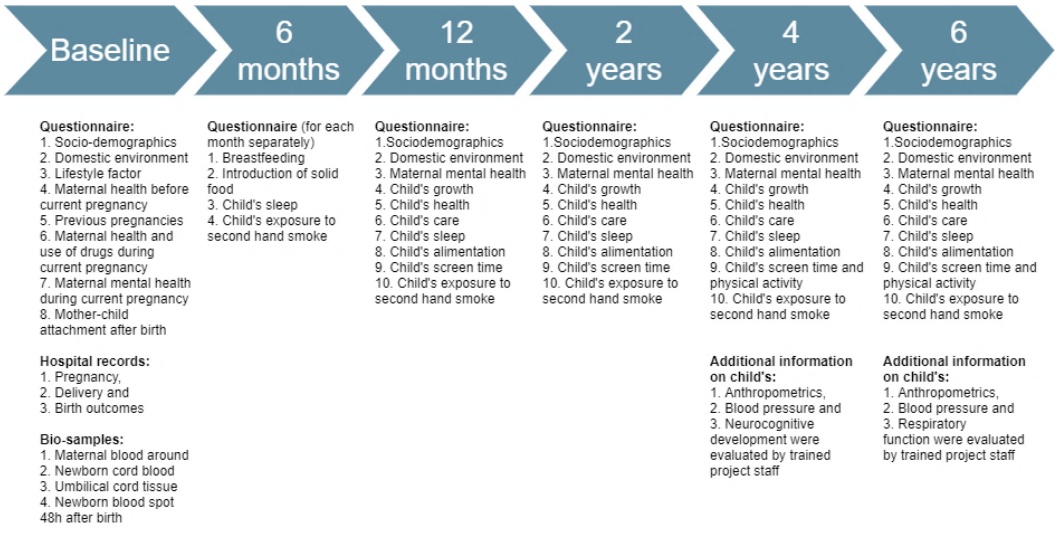
Historical birth cohorts, as described in the beginning of the Introduction, were the first that recognized the importance of the *in utero* environment on adult health and disease.[178] Since then, more than half of a century has passed, and there has been an incredible expansion of the number of newly established birth cohorts, that by using prospective study design and exhaustive data collection addressed previous limitations, such as recall bias and unmeasured confounding factors. [5]

Birth cohort studies are a powerful study design because they provide a unique opportunity to study numerous exposures that in humans could not be studied in experimental settings, such as exposure to early life stressors and adversity, as well as other obstetric, socioeconomic, environmental and lifestyle factors that are believed to have an adverse effect on health. The detailed follow-up, that in some cohorts passed to second generations participants, allowed collecting information on disease and health status though life course. The amount of scientific findings that birth cohorts provided in just few last decades is enormous, and their contribution to the knowledge on disease etiology is remarkable. Apart from traditional methods for data collection in environmental studies, such as questionnaires and administrative records linkages, most of the recent birth cohort collect biological samples. This, together with the huge advancements in high-throughput techniques, allowed birth cohorts to study genetic, epigenetic, metabolic and other "omic" markers, such as those studied within this thesis, and their influence on newborn growth and development, childhood health, and common complex diseases in adulthood.

Despite the advantages related to study design, the establishment and follow-up of birth cohort studies is expensive and many cohorts have now recognized the enormous potential that lies in collaborative projects that increase sample size and minimize publication bias.[16] The large sample size in consortia could be exploited to unravel novel risk factors that may have passed unnoticed due to small effect estimates, identify populations at risk and study mediating factors, capitalizing on already existing quality data. In the past years several successful birth cohort consortia were established,[13,15,179–181] with different study objectives and target research areas, that brought together scientists, health professionals and government policy makers with final aim to translate their research findings into public health policies acting that might help in health promotion as early as conception. The largest consortium is "The LifeCycle project"[182] (https://lifecycle-project.eu/) that aims to establish a European Union Child Cohort Network that would bring together 40 established birth cohorts, providing infrastructure for analysis of harmonized data of more than quarter of a million parents and children, opening endless possibilities for future research.

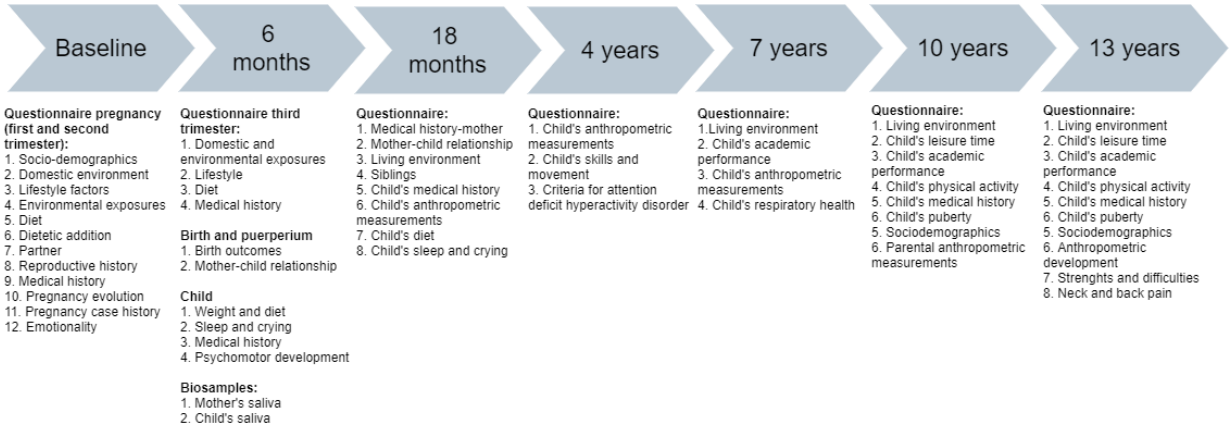
All data described and analyzed within this thesis are set within the framework of two birth cohort studies: Piccolipiù[183,184] and NINFEA (Nascita e INFanzia: gli Effetti dell'Ambiente).[185,186] Both cohorts are set up to investigate exposures acting during pre-natal and early post-natal life on infant and child health, describe complex interactions between genetic, epigenetic, lifestyle and environmental factors and promote infant and child health. The overview of data collection of PiccoliPiù is described in Figure 1 and includes baseline and follow-up questionnaires completed by the mother with help from a trained professional staff, and a biobank that stores maternal and cord blood samples.

Figure 1. Overview of the data collection in Piccolipiù



NINFEA is an internet-based cohort where the mother completes the baseline and follow-up questionnaires online, and maternal and child's saliva are collected during infancy. The overview of data collection in NINFEA is presented in Figure 2. Without their rich and curated data created by the long-term collaborative effort of many professionals from different disciplines and thousands of families, this thesis would not be possible.

Figure 2. Overview of the data collection in NINFEA



Bibliography:

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. doi:10.1056/NEJMra0708473

2. Vargesson N. Thalidomide-induced teratogenesis: History and mechanisms. *Birth Defects Res Part C - Embryo Today Rev*. 2015;105(2):140-156. doi:10.1002/bdrc.21096

3. Goodman A, Schorge J, Greene MF. The Long-Term Effects of In Utero Exposures — The DES Story. *N Engl J Med*. 2011;364(22):2083-2084. doi:10.1056/nejmp1104409

4. Saunders EJ, Saunders JA. Rug therapy in pregnancy: The lessons of diethylstilbestrol, thalidomide, and Bendectin. *Health Care Women Int*. 1990;11(4):423-432. doi:10.1080/07399339009515912

5. Underwood P, Hester LL, Laffitte T, Gregg K V. The relationship of smoking to the outcome of pregnancy. *Am J Obstet Gynecol*. 1965;91(2):270-276. doi:10.1016/0002-9378(65)90211-5

6. Jones KL, Smith DW. RECOGNITION OF THE FETAL ALCOHOL SYNDROME IN EARLY INFANCY. *Lancet*. 1973. doi:10.1016/S0140-6736(73)91092-1

7. Barker DJP, Osmond C. INFANT MORTALITY, CHILDHOOD NUTRITION, AND ISCHAEMIC HEART DISEASE IN ENGLAND AND WALES. *Lancet*. 1986;327(8489):1077-1081. doi:10.1016/S0140-6736(86)91340-1

8. Acquah JK, Dahal R, Sloan FA. 1918 influenza pandemic: In utero exposure in the United States and long-term impact on hospitalizations. *Am J Public Health*. 2017;107(9):1477-1483. doi:10.2105/AJPH.2017.303887

9. Kim S, Fleisher B, Sun JY. The Long-term Health Effects of Fetal Malnutrition: Evidence

from the 1959-1961 China Great Leap Forward Famine. *Health Econ*. 2017;26(10):1264-1277. doi:10.1002/hec.3397

10. Schulz LC. The Dutch hunger winter and the developmental origins of health and disease. *Proc Natl Acad Sci U S A*. 2010;107(39):16757-16758. doi:10.1073/pnas.1012911107

11. Barker DJP. The origins of the developmental origins theory. In: *Journal of Internal Medicine*. Vol 261. J Intern Med; 2007:412-417. doi:10.1111/j.1365-2796.2007.01809.x

12. Wadhwa PD, Buss C, Entringer S, Swanson JM. Developmental Origins of Health and Disease: Brief History of the Approach and Current Focus on Epigenetic Mechanisms. *Semin Reprod Med*. 2009;27(5):358-368. doi:10.1055/s-0029-1237424

13. Felix JF, Joubert BR, Baccarelli AA, et al. Cohort profile: Pregnancy and childhood epigenetics (PACE) consortium. *Int J Epidemiol*. 2018;47(1):22-23u. doi:10.1093/ije/dyx190

14. 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):21311. doi:10.1136/bmjopen-2017-021311

15. Vineis P, Chadeau-Hyam M, Gmuender H, et al. The exposome in practice: Design of the EXPOsOMICS project. *Int J Hyg Environ Health*. 2017;220(2):142-151. doi:10.1016/j.ijheh.2016.08.001

16. Jaddoe VWV, Felix JF, Andersen AMN, et al. The LifeCycle Project-EU Child Cohort Network: a federated analysis infrastructure and harmonized data of more than 250,000 children and parents. *Eur J Epidemiol*. 2020;35(7):709-724. doi:10.1007/s10654-020-00662-z

17. Castles A, Adams EK, Melvin CL, Kelsch C, Boulton ML. Effects of smoking during pregnancy: Five meta-analyses. *Am J Prev Med*. 1999;16(3):208-215. doi:10.1016/S0749-3797(98)00089-0

18. Mohsin M, Bauman AE. Socio-demographic factors associated with smoking and smoking cessation among 426,344 pregnant women in New South Wales, Australia. *BMC Public Health*. 2005;5(1):1-9. doi:10.1186/1471-2458-5-138

19. Pollack HA. Sudden infant death syndrome, maternal smoking during pregnancy, and the cost-effectiveness of smoking cessation intervention. *Am J Public Health*. 2001;91(3):432-436. doi:10.2105/AJPH.91.3.432

20. Grosse SD, Waitzman NJ, Romano PS, Mulinare J. Reevaluating the benefits of folic acid fortification in the United States: Economic analysis, regulation, and public health. *Am J Public Health*. 2005;95(11):1917-1922. doi:10.2105/AJPH.2004.058859

21. Eichholzer M, Tönz O, Zimmermann R. Folic acid: a public-health challenge. *Lancet*. 2006;367(9519):1352-1361. doi:10.1016/S0140-6736(06)68582-6

22. Santos S, Voerman E, Amiano P, 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 An Int J Obstet Gynaecol*. 2019. doi:10.1111/1471-0528.15661

23. Swanson JM, Entringer S, Buss C, Wadhwa PD. Developmental origins of health and disease: Environmental exposures. *Semin Reprod Med*. 2009;27(5):391-402. doi:10.1055/s-0029-1237427

24. Heindel JJ, Vandenberg LN. Developmental origins of health and disease: A paradigm for understanding disease cause and prevention. *Curr Opin Pediatr*. 2015;27(2):248-253. doi:10.1097/MOP.0000000000000191

25. Entringer S, de Punder K, Buss C, Wadhwa PD. The fetal programming of telomere biology

hypothesis: An update. *Philos Trans R Soc B Biol Sci*. 2018;373(1741). doi:10.1098/rstb.2017.0151

26. Robertson RC, Manges AR, Finlay BB, Prendergast AJ. The Human Microbiome and Child Growth – First 1000 Days and Beyond. *Trends Microbiol*. 2019;27(2):131-147. doi:10.1016/j.tim.2018.09.008

27. Vrijheid M, Casas M, Gascon M, Valvi D, Nieuwenhuijsen M. Environmental pollutants and child health-A review of recent concerns. *Int J Hyg Environ Health*. 2016;219(4-5):331-342. doi:10.1016/j.ijheh.2016.05.001

28. Kim JJ. Ambient air pollution: health hazards to children. *Pediatrics*. 2004.

29. WHO/CED/PHE/18.01 WR number: Air pollution and child health: prescribing clean air. 2018.

30. Hegewald MJ, Crapo RO. Respiratory Physiology in Pregnancy. *Clin Chest Med*. 2011;32(1):1-13. doi:10.1016/j.ccm.2010.11.001

31. Bové H, Bongaerts E, Slenders E, et al. Ambient black carbon particles reach the fetal side of human placenta. *Nat Commun*. 2019. doi:10.1038/s41467-019-11654-3

32. Muoth C, Aengenheister L, Kucki M, Wick P, Buerki-Thurnherr T. Nanoparticle transport across the placental barrier: Pushing the field forward. *Nanomedicine*. 2016. doi:10.2217/nnm-2015-0012

33. Klepac P, Locatelli I, Korošec S, Künzli N, Kukec A. Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges. *Environ Res*. 2018;167:144-159. doi:10.1016/j.envres.2018.07.008

34. Assibey-Mensah V, Glantz JC, Hopke PK, et al. Wintertime Wood Smoke, Traffic Particle Pollution, and Preeclampsia. *Hypertension*. 2020;75(3):851-858. doi:10.1161/HYPERTENSIONAHA.119.13139

35. Van Leeuwen DM, Pedersen M, Hendriksen PJM, et al. Genomic analysis suggests higher susceptibility of children to air pollution. *Carcinogenesis*. 2008;29(5):977-983. doi:10.1093/carcin/bgn065

36. Latzin P, Röösli M, Huss A, Kuehni CE, Frey U. Air pollution during pregnancy and lung function in newborns: A birth cohort study. *Eur Respir J*. 2009;33(3):594-603. doi:10.1183/09031936.00084008

37. Ha S, Yeung E, Bell E, et al. Prenatal and early life exposures to ambient air pollution and development. *Environ Res*. 2019;174:170-175. doi:10.1016/j.envres.2019.03.064

38. Boezen HM, Van Der Zee SC, Postma DS, et al. Effects of ambient air pollution on upper and lower respiratory symptoms and peak expiratory flow in children. *Lancet*. 1999;353(9156):874-878. doi:10.1016/S0140-6736(98)06311-9

39. Goldizen FC, Sly PD, Knibbs LD. Respiratory effects of air pollution on children. *Pediatr Pulmonol*. 2016;51(1):94-108. doi:10.1002/ppul.23262

40. Jerrett M, McConnell R, Wolch J, et al. Traffic-related air pollution and obesity formation in children: A longitudinal, multilevel analysis. *Environ Heal A Glob Access Sci Source*. 2014;13(1):1-9. doi:10.1186/1476-069X-13-49

41. Heck JE, Wu J, Lombardi C, et al. Childhood Cancer and Traffic-Related Air Pollution Exposure in Pregnancy and Early Life. *Environ Health Perspect*. 2013;121(11-12):1385-1391. doi:10.1289/ehp.1306761

42. Amigou A, Sermage-Faure C, Orsi L, et al. Road traffic and childhood leukemia: The

ESCALE study (SFCE). *Environ Health Perspect*. 2011;119(4):566-572. doi:10.1289/ehp.1002429

43. Oudin A, Frondelius K, Haglund N, et al. Prenatal exposure to air pollution as a potential risk factor for autism and ADHD. *Environ Int*. 2019;133:105149. doi:10.1016/j.envint.2019.105149

44. Lodovici M, Bigagli E. Oxidative stress and air pollution exposure. *J Toxicol*. 2011;2011. doi:10.1155/2011/487074

45. Leni Z, Künzi L, Geiser M. Air pollution causing oxidative stress. *Curr Opin Toxicol*. 2020;20-21:1-8. doi:10.1016/j.cotox.2020.02.006

46. Saenen ND, Martens DS, Neven KY, et al. Air pollution-induced placental alterations: An interplay of oxidative stress, epigenetics, and the aging phenotype? *Clin Epigenetics*. 2019;11(1):124. doi:10.1186/s13148-019-0688-z

47. Martens DS, Nawrot TS. Air Pollution Stress and the Aging Phenotype: The Telomere Connection. *Curr Environ Heal reports*. 2016;3(3):258-269. doi:10.1007/s40572-016-0098-8

48. Miri M, Nazarzadeh M, Alahabadi A, et al. Air pollution and telomere length in adults: A systematic review and meta-analysis of observational studies. *Environ Pollut*. 2019;244:636-647. doi:10.1016/j.envpol.2018.09.130

49. Ferrari L, Carugno M, Bollati V. Particulate matter exposure shapes DNA methylation through the lifespan. *Clin Epigenetics*. 2019. doi:10.1186/s13148-019-0726-x

50. Desai G, Chu L, Guo Y, Myneni AA, Mu L. Biomarkers used in studying air pollution exposure during pregnancy and perinatal outcomes: a review. *Biomarkers*. 2017;22(6):489-501. doi:10.1080/1354750X.2017.1339294

51. Alvarado-Cruz I, Alegría-Torres JA, Montes-Castro N, Jiménez-Garza O, Quintanilla-Vega B. Environmental epigenetic changes, as risk factors for the development of diseases in children: A systematic review. *Ann Glob Heal*. 2018;84(2):212-224. doi:10.29024/aogh.909

52. Luyten LJ, Saenen ND, Janssen BG, et al. Air pollution and the fetal origin of disease: A systematic review of the molecular signatures of air pollution exposure in human placenta. *Environ Res*. 2018;166:310-323. doi:10.1016/j.envres.2018.03.025

53. World Health Organization. *Air Quality Guidelines. Global Update 2005*.; 2006. doi:10.1007/BF02986808

54. IARC. *Air Pollution and Cancer, IARC Scientific Publications NO. 161*.; 2013.

55. World Health Organization. *Ambient Air Pollution: A Global Assessment of Exposure and Burden of Disease*.; 2016.

56. Ryan PH, Lemasters GK. A review of land-use regression models for characterizing intraurban air pollution exposure. In: *Inhalation Toxicology*. Vol 19. NIH Public Access; 2007:127-133. doi:10.1080/08958370701495998

57. Janssen S, Dumont G, Fierens F, Mensink C. Spatial interpolation of air pollution measurements using CORINE land cover data. *Atmos Environ*. 2008. doi:10.1016/j.atmosenv.2008.02.043

58. Lipfert FW. A critical review of the ESCAPE project for estimating long-term health effects of air pollution. *Environ Int*. 2017. doi:10.1016/j.envint.2016.11.028

59. Stafoggia M, Bellander T, Bucci S, et al. Estimation of daily PM10 and PM2.5 concentrations in Italy, 2013–2015, using a spatiotemporal land-use random-forest model. *Environ Int*. 2019;124:170-179. doi:10.1016/j.envint.2019.01.016

60. Abarca-Gómez L, Abdeen ZA, Hamid ZA, et al. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *Lancet*. 2017;390(10113):2627-2642. doi:10.1016/S0140-6736(17)32129-3

61. WHO. *Report of the Commision on Ending Childhood Obesity*.; 2016. http://www.who.int/end-childhood-obesity/news/launch-final-report/en/.

62. Llewellyn CH, Trzaskowski M, Plomin R, Wardle J. Finding the missing heritability in pediatric obesity: The contribution of genome-wide complex trait analysis. *Int J Obes*. 2013. doi:10.1038/ijo.2013.30

63. Jones RE, Jewell J, Saksena R, Ramos Salas X, Breda J. Overweight and Obesity in Children under 5 Years: Surveillance Opportunities and Challenges for the WHO European Region. *Front Public Heal*. 2017;5(April):1-12. doi:10.3389/fpubh.2017.00058

64. Dev DA, McBride BA, Fiese BH, Jones BL, Cho, on behalf of the STRONG Kids R H. Risk Factors for Overweight/Obesity in Preschool Children: An Ecological Approach. *Child Obes*. 2013;9(5):399-408. doi:10.1089/chi.2012.0150

65. Must A, Strauss RS. Risks and consequences of childhood and adolescent obesity. *Int J Obes*. 1999;23:S2-S11. doi:10.1038/sj/ijo/0800852

66. Dietz WH. Health consequences of obesity in youth: Childhood predictors of adult disease. *Pediatrics*. 1998.

67. Nathan BM, Moran A. Metabolic complications of obesity in childhood and adolescence: More than just diabetes. *Curr Opin Endocrinol Diabetes Obes*. 2008. doi:10.1097/MED.0b013e3282f43d19

68. Salgin B, Norris SA, Prentice P, et al. Even transient rapid infancy weight gain is associated with higher BMI in young adults and earlier menarche. *Int J Obes*. 2015. doi:10.1038/ijo.2015.25

69. Kelsey MM, Zaepfel A, Bjornstad P, Nadeau KJ. Age-related consequences of childhood obesity. *Gerontology*. 2014. doi:10.1159/000356023

70. Rankin J, Matthews L, Cobley S, et al. Psychological consequences of childhood obesity: psychiatric comorbidity and prevention. *Adolesc Health Med Ther*. 2016;Volume 7:125-146. doi:10.2147/AHMT.S101631

71. Rolland-Cachera MF, Deheeger M, Maillot M, Bellisle F. Early adiposity rebound: Causes and consequences for obesity in children and adults. *Int J Obes*. 2006. doi:10.1038/sj.ijo.0803514

72. Hughes AR, Sherriff A, Ness AR, Reilly JJ. Timing of adiposity rebound and adiposity in adolescence. *Pediatrics*. 2014. doi:10.1542/peds.2014-1908

73. Koyama S, Ichikawa G, Kojima M, Shimura N, Sairenchi T, Arisaka O. Adiposity rebound and the development of metabolic syndrome. *Pediatrics*. 2014. doi:10.1542/peds.2013-0966

74. Woo Baidal JA, Locks LM, Cheng ER, Blake-Lamb TL, Perkins ME, Taveras EM. Risk Factors for Childhood Obesity in the First 1,000 Days: A Systematic Review. *Am J Prev Med*. 2016;50(6):761-779. doi:10.1016/j.amepre.2015.11.012

75. Zhang M, Differding MK, Benjamin-Neelon SE, Østbye T, Hoyo C, Mueller NT. Association of prenatal antibiotics with measures of infant adiposity and the gut microbiome. *Ann Clin Microbiol Antimicrob*. 2019;18(1):18. doi:10.1186/s12941-019-0318-9

76. Mueller NT, Rifas-Shiman SL, Blaser MJ, Gillman MW, Hivert MF. Association of prenatal antibiotics with foetal size and cord blood leptin and adiponectin. *Pediatr Obes*. 2017;12(2):129-136. doi:10.1111/ijpo.12119

77. Mueller NT, Whyatt R, Hoepner L, et al. Prenatal exposure to antibiotics, cesarean section and risk of childhood obesity. *Int J Obes*. 2015;39(4):665-670. doi:10.1038/ijo.2014.180.Prenatal

78. Mor A, Antonsen S, Kahlert J, et al. Prenatal exposure to systemic antibacterials and overweight and obesity in Danish schoolchildren: A prevalence study. *Int J Obes*. 2015;39(10):1450-1455. doi:10.1038/ijo.2015.129

79. Poulsen MN, Pollak J, Bailey-Davis L, Hirsch AG, Glass TA, Schwartz BS. Associations of prenatal and childhood antibiotic use with child body mass index at age 3 years. *Obesity*. 2017;25(2):438-444. doi:10.1002/oby.21719

80. Cassidy-Bushrow AE, Burmeister C, Havstad S, et al. Prenatal antimicrobial use and early-childhood body mass index. *Int J Obes*. 2018;42(1):1-7. doi:10.1038/ijo.2017.205

81. Wang B, Liu J, Zhang Y, et al. Prenatal Exposure to Antibiotics and Risk of Childhood Obesity in a Multicenter Cohort Study. *Am J Epidemiol*. 2018;187(10):2159-2167. doi:10.1093/aje/kwy122

82. Heerman WJ, Daley MF, Boone-Heinonen J, et al. Maternal antibiotic use during pregnancy and childhood obesity at age 5 years. *Int J Obes*. January 2019:1. doi:10.1038/s41366-018-0316-6

83. Jess T, Morgen CS, Harpsøe MC, et al. Antibiotic use during pregnancy and childhood overweight: A population-based nationwide cohort study. *Sci Rep*. 2019;9(1). doi:10.1038/s41598-019-48065-9

84. Cox LM, Blaser MJ. Antibiotics in early life and obesity. *Nat Rev Endocrinol*. 2015;11(3):182-190. doi:10.1038/nrendo.2014.210

85. Broe A, Pottegård A, Lamont RF, Jørgensen JS, Damkier P. Increasing use of antibiotics in pregnancy during the period 2000-2010: Prevalence, timing, category, and demographics. *BJOG An Int J Obstet Gynaecol*. 2014;121(8):988-996. doi:10.1111/1471-0528.12806

86. de Tejada BM. Antibiotic use and misuse during pregnancy and delivery: Benefits and risks. *Int J Environ Res Public Health*. 2014. doi:10.3390/ijerph110807993

87. Taylor JH, Gordon WS. Growth-promoting activity for pigs of inactivated penicillin [12]. *Nature*. 1955. doi:10.1038/176312a0

88. Cho I, Yamanishi S, Cox L, et al. Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature*. 2012. doi:10.1038/nature11400

89. Blaser MJ. Antibiotic use and its consequences for the normal microbiome. *Science (80-).* 2016. doi:10.1126/science.aad9358

90. Dinan TG, Cryan JF, Development N, et al. The Infant Microbiome: Implications for Infant Health and Neurocognitive Development. *Pediatrics*. 2013;46(1):1-16. doi:10.1097/NNR.0000000000000133.The

91. Govender L, Hoosen AA, Moodley J, Moodley P, Sturm AW. Bacterial vaginosis and associated infections in pregnancy. *Int J Gynecol Obstet*. 1996. doi:10.1016/0020-7292(96)02744-0

92. Brown RG, Al-Memar M, Marchesi JR, et al. Establishment of vaginal microbiota composition in early pregnancy and its association with subsequent preterm prelabor rupture of the fetal membranes. *Transl Res*. 2019;207:30-43. doi:10.1016/J.TRSL.2018.12.005

93. Fettweis JM, Serrano MG, Brooks JP, et al. The vaginal microbiome and preterm birth. *Nat Med*. 2019;25(6):1012-1021. doi:10.1038/s41591-019-0450-2

94. Hočevar K, Maver A, Vidmar Šimic M, et al. Vaginal Microbiome Signature Is Associated

With Spontaneous Preterm Delivery. *Front Med*. 2019;6. doi:10.3389/fmed.2019.00201

95. Markopoulou P, Papanikolaou E, Analytis A, Zoumakis E, Siahanidou T. Preterm Birth as a Risk Factor for Metabolic Syndrome and Cardiovascular Disease in Adult Life: A Systematic Review and Meta-Analysis. *J Pediatr*. 2019;210:69-80.e5. doi:10.1016/j.jpeds.2019.02.041

96. Noble D. Conrad Waddington and the origin of epigenetics. *J Exp Biol*. 2015;218(6):816-818. doi:10.1242/jeb.120071

97. Holliday R. Epigenetics: A Historical Overview. *Epigenetics*. 1(2):76-80. doi:10.4161/epi.1.2.2762

98. Villota-Salazar NA, Mendoza-Mendoza A, González-Prieto JM. Epigenetics: from the past to the present. *Front Life Sci*. 2016;9(4):347-370. doi:10.1080/21553769.2016.1249033

99. Monasso GS, Jaddoe VWV, de Jongste JC, Duijts L, Felix JF. Timing- and Dose-Specific Associations of Prenatal Smoke Exposure With Newborn DNA Methylation. *Nicotine Tob Res*. 2020. doi:10.1093/ntr/ntaa069

100. Joubert BR, Felix JF, Yousefi P, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am J Hum Genet*. 2016;98(4):680-696. doi:10.1016/j.ajhg.2016.02.019

101. Küpers LK, Xu X, Jankipersadsing SA, et al. DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. *Int J Epidemiol*. 2015. doi:10.1093/ije/dyv048

102. Richmond RC, Suderman M, Langdon R, Relton CL, Smith GD. DNA methylation as a marker for prenatal smoke exposure in adults. *Int J Epidemiol*. 2018;47(4):1120-1130. doi:10.1093/ije/dyy091

103. Sharp GC, Lawlor DA, Richmond RC, et al. Maternal pre-pregnancy BMI and gestational weight gain, offspring DNA methylation and later offspring adiposity: Findings from the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2015. doi:10.1093/ije/dyv042

104. Sharp GC, Salas LA, Monnereau C, 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. doi:10.1093/hmg/ddx290

105. Liu X, Chen Q, Tsai HJ, et al. Maternal preconception body mass index and offspring cord blood DNA methylation: Exploration of early life origins of disease. *Environ Mol Mutagen*. 2014. doi:10.1002/em.21827

106. Joubert BR, Den Dekker HT, Felix JF, et al. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. *Nat Commun*. 2016. doi:10.1038/ncomms10577

107. Pauwels S, Ghosh M, Duca RC, et al. Maternal intake of methyl-group donors affects DNA methylation of metabolic genes in infants. *Clin Epigenetics*. 2017. doi:10.1186/s13148-017-0321-y

108. Geurtsen ML, Jaddoe VWV, Gaillard R, Felix JF. Associations of maternal early-pregnancy blood glucose and insulin concentrations with DNA methylation in newborns. *Clin Epigenetics*. 2020. doi:10.1186/s13148-020-00924-3

109. Howe CG, Cox B, Fore R, et al. Maternal gestational diabetes mellitus and newborn DNA Methylation: Findings from the pregnancy and childhood epigenetics consortium. *Diabetes Care*. 2020. doi:10.2337/dc19-0524

110. Kazmi N, Sharp GC, Reese SE, et al. Hypertensive Disorders of Pregnancy and DNA Methylation in Newborns: Findings from the Pregnancy and Childhood Epigenetics

Consortium. *Hypertension*. 2019;74(2):375-383. doi:10.1161/HYPERTENSIONAHA.119.12634

111. Herzog EM, Eggink AJ, Willemsen SP, et al. Early- and late-onset preeclampsia and the tissue-specific epigenome of the placenta and newborn. *Placenta*. 2017. doi:10.1016/j.placenta.2017.08.070

112. Rijlaarsdam J, van IJzendoorn MH, Verhulst FC, et al. Prenatal stress exposure, oxytocin receptor gene (OXTR) methylation, and child autistic traits: The moderating role of OXTR rs53576 genotype. *Autism Res*. 2017. doi:10.1002/aur.1681

113. Vidal AC, Neelon SEB, Liu Y, et al. Maternal stress, Preterm birth, And dna methylation at imprint regulatory sequences in humans. *Genet Epigenetics*. 2014. doi:10.4137/GEG.S18067

114. Cedar H, Bergman Y. Programming of DNA Methylation Patterns. *Annu Rev Biochem*. 2012. doi:10.1146/annurev-biochem-052610-091920

115. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. *Lancet*. 2018. doi:10.1016/S0140-6736(18)31268-6

116. Tollefsbol TO. *Handbook of Epigenetics*.; 2011. doi:10.1016/C2009-0-30488-2

117. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*. 2007. doi:10.1038/nature05918

118. Long HK, King HW, Patient RK, Odom DT, Klose RJ. Protection of CpG islands from DNA methylation is DNA-encoded and evolutionarily conserved. *Nucleic Acids Res*. 2016. doi:10.1093/nar/gkw258

119. Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13(7):484-492. doi:10.1038/nrg3230

120. Gardiner-Garden M, Frommer M. CpG Islands in vertebrate genomes. *J Mol Biol*. 1987. doi:10.1016/0022-2836(87)90689-9

121. Jeziorska DM, Murray RJS, De Gobbi M, et al. DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. *Proc Natl Acad Sci U S A*. 2017. doi:10.1073/pnas.1703087114

122. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over Two-Thirds of the human genome. *PLoS Genet*. 2011. doi:10.1371/journal.pgen.1002384

123. Beck CR, Garcia-Perez JL, Badge RM, Moran J V. LINE-1 elements in structural variation and disease. *Annu Rev Genomics Hum Genet*. 2011. doi:10.1146/annurev-genom-082509-141802

124. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. *Science (80-)*. 2001. doi:10.1126/science.1063852

125. Das PM, Singal R. DNA methylation and cancer. *J Clin Oncol*. 2004. doi:10.1200/JCO.2004.07.151

126. Singer-Sam J, Riggs AD. X chromosome inactivation and DNA methylation. *EXS*. 1993;64:358-384. doi:10.1007/978-3-0348-9118-9_16

127. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363-1369. doi:10.1093/bioinformatics/btu049

128. Sado T, Okano M, Li E, Sasaki H. De novo DNA methylation is dispensable for the initiation

and propagation of X chromosome inactivation. *Development*. 2004. doi:10.1242/dev.00995

129. Tycko B, Morison IM. Physiological functions of imprinted genes. *J Cell Physiol*. 2002;192(3):245-258. doi:10.1002/jcp.10129

130. Morison IM, Reeve AE. A catalogue of imprinted genes and parent-of-origin effects in humans and animals. *Hum Mol Genet*. 1998;7(10):1599-1609. doi:10.1093/hmg/7.10.1599

131. Reizel Y, Spiro A, Sabag O, et al. Gender-specific postnatal demethylation and establishment of epigenetic memory. *Genes Dev*. 2015. doi:10.1101/gad.259309.115

132. Reizel Y, Sabag O, Skversky Y, et al. Postnatal DNA demethylation and its role in tissue maturation. *Nat Commun*. 2018. doi:10.1038/s41467-018-04456-6

133. Blackburn EH, Epel ES, Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science (80-)*. 2015;350(6265):1193-1198. doi:10.1126/science.aab3389

134. Singh A, Kukreti R, Saso L, Kukreti S. Oxidative stress: Role and response of short guanine tracts at genomic locations. *Int J Mol Sci*. 2019. doi:10.3390/ijms20174258

135. Coluzzi E, Colamartino M, Cozzi R, et al. Oxidative stress induces persistent telomeric DNA damage responsible for nuclear morphology change in mammalian cells. *PLoS One*. 2014. doi:10.1371/journal.pone.0110963

136. Aviv A, Shay JW. Reflections on telomere dynamics and ageing-related diseases in humans. *Philos Trans R Soc B Biol Sci*. 2018. doi:10.1098/rstb.2016.0436

137. Varela E, Blasco MA. 2009 Nobel Prize in Physiology or Medicine: Telomeres and telomerase. *Oncogene*. 2010. doi:10.1038/onc.2010.15

138. Entringer S, de Punder K, Buss C, Wadhwa PD. The fetal programming of telomere biology hypothesis: An update. *Philos Trans R Soc B Biol Sci*. 2018. doi:10.1098/rstb.2017.0151

139. Dugdale HL, Richardson DS. Heritability of telomere variation: It is all about the environment! *Philos Trans R Soc B Biol Sci*. 2018. doi:10.1098/rstb.2016.0450

140. Holohan B, Wright WE, Shay JW. Telomeropathies: An emerging spectrum disorder. *J Cell Biol*. 2014. doi:10.1083/jcb.201401012

141. Haycock PC, Burgess S, Nounu A, et al. Association between telomere length and risk of cancer and non-neoplastic diseases a mendelian randomization study. *JAMA Oncol*. 2017. doi:10.1001/jamaoncol.2016.5945

142. Monaghan P, Eisenberg DTA, Harrington L, Nussey D. Understanding diversity in telomere dynamics. *Philos Trans R Soc B Biol Sci*. 2018. doi:10.1098/rstb.2016.0435

143. Patrick MS, Cheng NL, Kim J, et al. Human T cell differentiation negatively regulates telomerase expression resulting in reduced activation-induced proliferation and survival. *Front Immunol*. 2019. doi:10.3389/fimmu.2019.01993

144. Entringer S, de Punder K, Buss C, Wadhwa PD. The fetal programming of telomere biology hypothesis: an update. *Philos Trans R Soc B Biol Sci*. 2018;373(1741):20170151. doi:10.1098/rstb.2017.0151

145. Entringer S, Epel ES, Lin J, et al. Maternal psychosocial stress during pregnancy is associated with newborn leukocyte telomere length. *Am J Obstet Gynecol*. 2013;208(2). doi:10.1016/j.ajog.2012.11.033

146. Salihu HM, King LM, Nwoga C, et al. Association Between Maternal-Perceived Psychological Stress and Fetal Telomere Length. *South Med J*. 2016;109(12):767-772.

doi:10.14423/SMJ.0000000000000567

147. Marchetto NM, Glynn RA, Ferry ML, et al. Prenatal stress and newborn telomere length. In: *American Journal of Obstetrics and Gynecology*. Vol 215. ; 2016:94.e1-94.e8. doi:10.1016/j.ajog.2016.01.177

148. Kim JH, Kim GJ, Lee D, et al. Higher maternal vitamin D concentrations are associated with longer leukocyte telomeres in newborns. *Matern Child Nutr*. 2018;14(1). doi:10.1111/mcn.12475

149. Entringer S, Epel ES, Lin J, et al. Maternal Folate Concentration in Early Pregnancy and Newborn Telomere Length. *Ann Nutr Metab*. 2015;66(4):202-208. doi:10.1159/000381925

150. Cross JA, Temple RC, Hughes JC, et al. Cord blood telomere length, telomerase activity and inflammatory markers in pregnancies in women with diabetes or gestational diabetes. *Diabet Med*. 2010;27(11):1264-1270. doi:10.1111/j.1464-5491.2010.03099.x

151. Martens DS, Plusquin M, Gyselaers W, De Vivo I, Nawrot TS. Maternal pre-pregnancy body mass index and newborn telomere length. *BMC Med*. 2016;14(1). doi:10.1186/s12916-016-0689-0

152. Salihu HM, Pradhan A, King L, et al. Impact of intrauterine tobacco exposure on fetal telomere length. *Am J Obstet Gynecol*. 2015;212(2):205.e1-205.e8. doi:10.1016/j.ajog.2014.08.026

153. Friedrich U, Schwab M, Griese EU, Fritz P, Klotz U. Telomeres in neonates: New insights in fetal hematopoiesis. *Pediatr Res*. 2001;49(2):252-256. doi:10.1203/00006450-200102000-00020

154. Okuda K, Bardeguez A, Gardner JP, et al. Telomere length in the newborn. *Pediatr Res*. 2002;52(3):377-381. doi:10.1203/01.PDR.0000022341.72856.72

155. Gorenjak V, Petrelis AM, Stathopoulou MG, Visvikis-Siest S. Telomere length determinants in childhood. *Clin Chem Lab Med*. 2019. doi:10.1515/cclm-2019-0235

156. Wojcicki JM, Shiboski S, Heyman MB, et al. Telomere length change plateaus at 4 years of age in Latino children: associations with baseline length and maternal change. *Mol Genet Genomics*. 2016. doi:10.1007/s00438-016-1191-2

157. Zeichner SL, Palumbo P, Feng YR, et al. Rapid telomere shortening in children. *Blood*. 1999. doi:10.1182/blood.v93.9.2824

158. Frenck RW, Blackburn EH, Shannon KM. The rate of telomere sequence loss in human leukocytes varies with age. *Proc Natl Acad Sci U S A*. 1998. doi:10.1073/pnas.95.10.5607

159. Bäckhed F, Fraser CM, Ringel Y, et al. Defining a healthy human gut microbiome: Current concepts, future directions, and clinical applications. *Cell Host Microbe*. 2012. doi:10.1016/j.chom.2012.10.012

160. Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. *Curr Opin Gastroenterol*. 2015. doi:10.1097/MOG.0000000000000139

161. Khan MJ, Gerasimidis K, Edwards CA, Shaikh MG. Role of Gut Microbiota in the Aetiology of Obesity: Proposed Mechanisms and Review of the Literature. *J Obes*. 2016. doi:10.1155/2016/7353642

162. Kinross JM, Darzi AW, Nicholson JK. Gut microbiome-host interactions in health and disease. *Genome Med*. 2011. doi:10.1186/gm228

163. Tilg H, Kaser A. Gut microbiome, obesity, and metabolic dysfunction. *J Clin Invest*. 2011. doi:10.1172/JCI58109

164. Gritz EC, Bhandari V. The Human Neonatal Gut Microbiome: A Brief Review. *Front Pediatr*. 2015;3. doi:10.3389/fped.2015.00017

165. Tamburini S, Shen N, Wu HC, Clemente JC. The microbiome in early life: Implications for health outcomes. *Nat Med*. 2016. doi:10.1038/nm.4142

166. Cox LM, Yamanishi S, Sohn J, et al. Window Has Lasting Metabolic Consequences. 2015;158(4):705-721. doi:10.1016/j.cell.2014.05.052.Altering

167. Koren O, Goodrich JK, Cullender TC, et al. Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell*. 2012;150(3):470-480. doi:10.1016/j.cell.2012.07.008

168. Nuriel-Ohayon M, Neuman H, Koren O. Microbial changes during pregnancy, birth, and infancy. *Front Microbiol*. 2016;7(JUL):1-13. doi:10.3389/fmicb.2016.01031

169. Farr A, Kiss H, Hagmann M, et al. Role of Lactobacillus Species in the Intermediate Vaginal Flora in Early Pregnancy : A Retrospective Cohort Study. 2015:1-12. doi:10.1371/journal.pone.0144181

170. Freitas AC, Chaban B, Bocking A, et al. The vaginal microbiome of pregnant women is less rich and diverse, with lower prevalence of Mollicutes, compared to non-pregnant women. *Sci Rep*. 2017;7(1). doi:10.1038/s41598-017-07790-9

171. Wilks M, Wiggins R, Whiley A, et al. Identification and H2O2 Production of Vaginal Lactobacilli from Pregnant Women at High Risk of Preterm Birth and Relation with Outcome. *J Clin Microbiol*. 2004. doi:10.1128/JCM.42.2.713-717.2004

172. Stout MJ, Zhou Y, Wylie KM, Tarr PI, Macones GA, Tuuli MG. Early pregnancy vaginal microbiome trends and preterm birth. *Am J Obstet Gynecol*. 2017. doi:10.1016/j.ajog.2017.05.030

173. Aagaard K, Riehle K, Ma J, et al. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS One*. 2012. doi:10.1371/journal.pone.0036466

174. Stokholm J, Schjørring S, Eskildsen CE, et al. Antibiotic use during pregnancy alters the commensal vaginal microbiota. *Clin Microbiol Infect*. 2014. doi:10.1111/1469-0691.12411

175. Dominguez-Bello MG, Costello EK, Contreras M, et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci*. 2010;107(26):11971-11975. doi:10.1073/pnas.1002601107

176. Pannaraj PS, Li F, Cerini C, et al. Association between breast milk bacterial communities and establishment and development of the infant gut microbiome. *JAMA Pediatr*. 2017. doi:10.1001/jamapediatrics.2017.0378

177. Lemas DJ, Yee S, Cacho N, et al. Exploring the contribution of maternal antibiotics and breastfeeding to development of the infant microbiome and pediatric obesity. *Semin Fetal Neonatal Med*. 2016. doi:10.1016/j.siny.2016.04.013

178. Lawlor DA, Andersen AMN, Batty GD. Birth cohort studies: Past, present and future. *Int J Epidemiol*. 2009. doi:10.1093/ije/dyp240

179. Van Cauwenberge P, Watelet JB, Van Zele T, Bousquet J, Burney P, Zuberbier T. Spreading excellence in allergy and asthma: The GA2LEN (Global Allergy and Asthma European Network) project. *Allergy Eur J Allergy Clin Immunol*. 2005. doi:10.1111/j.1398-9995.2005.00822.x

180. Bousquet J, Anto J, Auffray C, et al. MeDALL (Mechanisms of the Development of ALLergy): An integrated approach from phenotypes to systems medicine. *Allergy Eur J Allergy Clin Immunol*. 2011. doi:10.1111/j.1398-9995.2010.02534.x

181. Vrijheid M, Casas M, Bergström A, et al. European birth cohorts for environmental health research. *Environ Health Perspect*. 2012. doi:10.1289/ehp.1103823

182. Jaddoe VWV, Felix JF, Andersen AMN, et al. The LifeCycle Project-EU Child Cohort Network: a federated analysis infrastructure and harmonized data of more than 250,000 children and parents. *Eur J Epidemiol*. 2020. doi:10.1007/s10654-020-00662-z

183. Toccaceli V, Serino L, Stazi MA. Informed consent, and an ethico- legal framework for paediatric observational research and biobanking: the experience of an Italian birth cohort study. *Cell Tissue Bank*. 2014. doi:10.1007/s10561-014-9431-3

184. Farchi S, Forastiere F, Vecchi Brumatti L, et al. Piccolipiù, a multicenter birth cohort in Italy: Protocol of the study. *BMC Pediatr*. 2014;14(1). doi:10.1186/1471-2431-14-36

185. Firestone R, Cheng S, Pearce N, et al. Internet-Based Birth-Cohort Studies: Is This the Future for Epidemiology? *JMIR Res Protoc*. 2015;4(2):e71. doi:10.2196/resprot.3873

186. Blumenberg C, Zugna D, Popovic M, Pizzi C, Barros AJD, Richiardi L. Questionnaire Breakoff and Item Nonresponse in Web-Based Questionnaires: Multilevel Analysis of Person-Level and Item Design Factors in a Birth Cohort. *J Med Internet Res*. 2018;20(12):e11046. doi:10.2196/11046

Objectives and study hypotheses

The aim of this thesis was to study common gestational exposures and how they influence newborn molecular markers and later childhood outcomes.

Three important molecular mechanisms that are believed to underlie the exposure-outcome associations seen within the DOHaD hypothesis were directly or indirectly studied within this thesis: the epigenome (DNA methylation in particular), telomere maintenance system and the neonatal microbiome.

Particular attention was given to assess and identify gestational windows of exposure that might be of particular relevance for future health outcomes.

Study I

The first study is a systematic review on all published studies on air pollution exposure during the first 1000 days of life and changes in somatic cell DNA, particularly DNA methylation and telomere length, since (1) they are influenced by the environment, (2) can survive cell replications and (3) are key players in biological processes i.e. have the potential to influence health across lifetime.

The aim of the systematic review was to accumulate knowledge on what is already known and established in this field, what needs further assessment, and what are potential gaps in literature, so we can address, at least some of them in our original investigation.

Study II

The second study is an independent study nested within the Italian birth cohort study Piccolipiù. The aim of the study was to assess whether PM_{10} exposure during the gestational period can alter overall or locus specific DNA methylation patterns and/or cause alteration in the telomere maintenance system in cord blood. Potential windows of vulnerability during pregnancy were explored, with particular attention to the beginning of the pregnancy, that we believed to be of particular importance for air pollution exposure because it is a period when the most active remodeling of the epigenome occurs.

Study III

The third study is nested within the Italian birth cohort NINFEA and focuses on the relationship between two common gestational exposures linked with the maternal and initial neonatal microbiome: antibiotics and vaginal infections, and their association with childhood obesity at the age of 4. We focused on exposures during the third trimester of pregnancy because we hypothesized that exposures that alter the maternal microbiome closer to the time of delivery might be the most relevant ones for obesity risk.

Study I Exposure to ambient air pollution in the first 1000 days of life and alterations in the DNA methylation and telomere length in children: a systematic review

[Environ Res. 2021 Feb; 193:110504. doi: 10.1016/j.envres.2020.110504]

Elena Isaevska[1], Chiara Moccia[1], Federica Asta[2], Fabio Cibella[3], Luigi Gagliardi[4], Luca Ronfani[5], Franca Rusconi[6], Maria Antonietta Stazi[7], Lorenzo Richiardi[1]

1. Department of Medical Sciences, University of Turin, CPO Piemonte, Turin, Italy
2. Department of Epidemiology, Lazio Regional Health Service, ASL Roma 1, Rome, Italy
3. Institute for Biomedical Research and Innovation (IRIB), National Research Council of Italy, Palermo, Italy
4. Division of Neonatology and Pediatrics, Ospedale Versilia, Viareggio, AUSL Toscana Nord Ovest, Pisa, Italy
5. Clinical Epidemiology and Public Health Research Unit, Institute for Maternal and Child Health - IRCCS "Burlo Garofolo", Trieste, Italy.
6. Unit of Epidemiology, Meyer Children's University Hospital, Florence, Italy
7. Center "Behavioral Sciences and Mental Health", Istituto Superiore di Sanità, Rome, Italy.

Funding: The project was realized with the financial support of the Italian Ministry of Health – CCM. / Progetto realizzato con il supporto finanziario del Ministero della Salute – CCM.

Abstract

Background. Exposure to air pollution during the first 1000 days of life (from conception to the 2nd year of life) might be of particular relevance for long-term child health. Changes in molecular markers such as DNA methylation and telomere length could underlie the association between air pollution exposure and pollution-related diseases as well as serve as biomarkers for past exposure. The objective of this systematic review was to assess the association between air pollution exposure during pregnancy and the first two years of life and changes in DNA methylation or telomere length in children.

Methods. PubMed was searched in October 2020 by using terms relative to ambient air pollution exposure, DNA methylation, telomere length and the population of interest: mother/child dyads and children. Screening and selection of the articles was completed independently by two reviewers. Thirty-two articles matched our criteria. The majority of the articles focused on gestational air pollution exposure and measured DNA methylation/telomere length in newborn cord blood or placental tissue, to study global, candidate-gene or epigenome-wide methylation patterns and/or telomere length. The number of studies in children was limited.

Results. Ambient air pollution exposure during pregnancy was associated with global loss of methylation in newborn cord blood and placenta, indicating the beginning of the pregnancy as a potential period of susceptibility. Candidate gene and epigenome-wide association studies provided evidence that gestational exposure to air pollutants can lead to locus-specific changes in methylation, in newborn cord blood and placenta, particularly in genes involved in cellular responses to oxidative stress, mitochondrial function, inflammation, growth and early life development. Telomere length shortening in newborns and children was seen in relation to gestational pollutant exposure.

Conclusions. Ambient air pollution during pregnancy is associated with changes in both global and locus-specific DNA methylation and with telomere length shortening. Future studies need to test the robustness of the association across different populations, to explore potential windows of vulnerability and assess the role of the methylation and telomere length as mediators in the association between early exposure to ambient air pollutants and specific childhood health outcomes.

Introduction

Air pollution is one of the today's main environmental and public health challenges in both high, and low income countries, with well documented health effects even at low exposure levels.[1] The period during pregnancy and the first years of life is an important window of susceptibility characterized by accelerated growth and developmental plasticity. Epidemiological evidence on the effects of exposures during early life lead to the formation Developmental Origins of Health and Disease (DOHaD) hypothesis, according to which the adaptive responses of the fetus/child to adverse early-life exposures could permanently shape the molecular programming and contribute to later disease predisposition.[2]

Growing evidence links exposure to air pollutants during early life to adverse pregnancy outcomes, including low birth weight and preterm birth,[3] reduced lung function, impaired neurodevelopment and susceptibility to later metabolic diseases.[4] The biological mechanisms that underlie these associations are still not well understood, although studies suggest that one mechanism may include changes in somatic cell DNA that (1) are influenced by the environment (2) can survive cell replications and (3) have the potential to influence biological processes. The most commonly studied biological markers that satisfy these criteria are DNA methylation and telomere length.

DNA methylation is the most well-known epigenetic mechanism that involves adding a methyl group to the cytosine (C) base of the DNA when next to the guanine (G) base, forming a so called CpG site. The methylation pattern represents a layer of molecular information atop of the DNA sequence that has an important role in wide array of functions, especially in the early embryonic and fetal development, including cell differentiation, regulation of gene expression, imprinting, X-chromosome inactivation and maintenance of genome stability.[5] The majority of the DNA methylation patterns are established around the period of implantation, making the early gestational period a possible window of susceptibility.[6,7] Telomeres, on the other hand, are nucleoprotein complexes located at the end of each chromosome to ensure complete chromosomal replication and prevent genomic instability.[8] Telomere length normally decreases with each cellular replication and variations in telomere length among adults seem to be largely attributed to genetic and environmental determinants that start their effect in utero.[9–11] Their vulnerability to reactive oxygen species makes them a plausible biomarker, not just for age and cellular replicability, but to for overall

exposure to oxidative stress and inflammation. Additionally, they may play an important role mediating the chronic health effects of early-life air pollution exposure.[12–15]

Several systematic reviews[16–19] were published on air pollution exposure during the course of life and molecular markers, including early life exposure, but they did not include the majority of the studies on this topic that have been published only recently.

Therefore, our aim was to evaluate the association between exposure to air pollutants during the 1000 days of life, from conception to 2 years, and changes in the DNA methylation patterns and telomere length in children.

Methods

To conduct the systematic review, a search strategy was first prepared. It included Medical Subject Headings (MESH) terms and keywords, based on our population, exposure and outcome of interest, Supplementary Table S1. We focused on most common air pollutants measured in atmospheric air such as $PM_{2.5}$, PM_{10}, polycyclic aromatic hydrocarbons (PAH), CO, SO_2, NO, NO_2, O_3, volatile organic compounds, black carbon, elemental or organic carbon. The population of interest included mother-child dyads during the gestational period and children. The outcomes were DNA methylation and telomere length.

We limited our search to articles written in English without limitations on the publication date. The search was not restricted to specific exposure assessment methods, tissue sample, and laboratory methods used to measure the outcomes, in order to assess the methodological variability between the selected studies and identify potential gaps in literature. The list of eligibility criteria is listed in Table 1. The literature search was conducted in the electronic database PubMed, and lastly updated on October, 2020.

Manual search of the references of the articles selected for full reading and systematic reviews previously published on the topic was also performed to identify additional articles that could match our selection criteria and one was found. Two investigators (EI and CM) conducted the literature search, read all papers and extracted relevant information independently. The discrepancies were resolved by consensus.

From each study that met the eligibility criteria we extracted the following information: study design, country of origin and population size, studied pollutants, method for exposure assessment, concentration levels of the pollutants, studied molecular marker, laboratory

technique used to assess the marker, effect estimates for the major findings, covariates considered in the analyses, and relevant results from any additional analyses.

Table 1 Criteria used to assess the eligibility of the articles

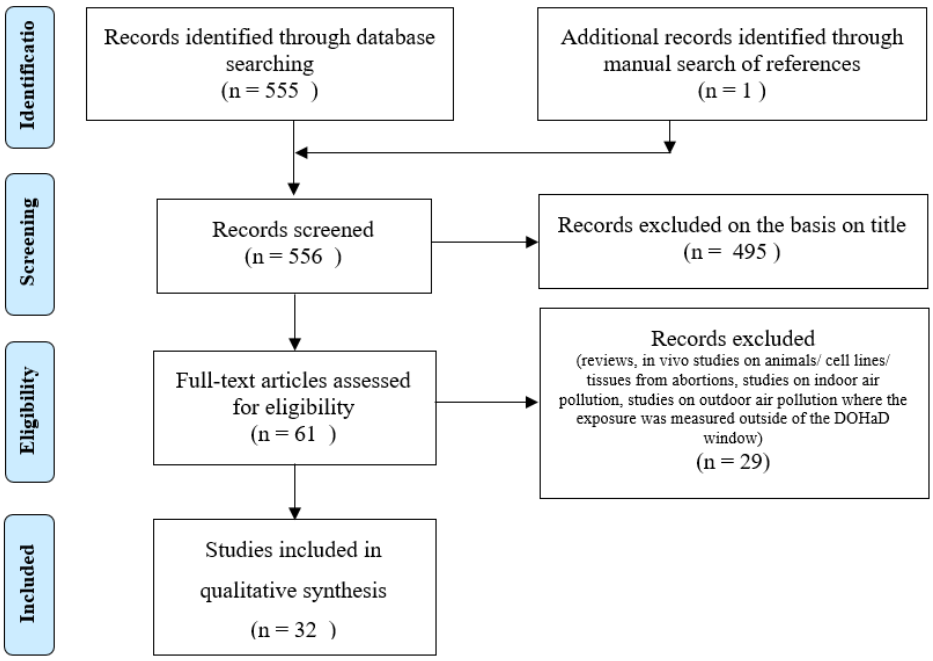
Study exposure	Particulate Matter ($PM_{2.5}$, PM_{10}), Nitrogen oxides (NO_2, NO), Ozone (O_3), Carbon Monoxide (CO), Sulfur Dioxide (SO_2), Volatile Organic Compounds (VOC), Black Carbon, Elemental carbon, Organic Carbon, Polycyclic Aromatic Hydrocarbons (PAH)
Time of exposure	Pregnancy, the first 2 years of life
Outcome(s)	DNA Methylation, telomere length
Population	Mother/child dyads, children
Study design	Observational studies on singletons.
Time frame	No time frame
Other criteria	Articles in English. No geographical restrictions.

Results

Study characteristics

Our search identified 556 articles; 495 were excluded on the basis of the title or the abstract, and the remaining 61 articles were selected for full reading. Thirty-two studies[20–50] met our selection criteria, Figure 3. All of them were ordered according to publication date (ranging from 2009 to 2020) and summarized in details in Supplementary Table S2.

Figure 1. Flow chart



Thirty articles measured gestational exposures to air pollutants and DNA methylation/ telomere length in cord blood/newborn blood or placenta. Limited number of studies (n=5) included DNA methylation/ telomere length analysis in children in relation to pollutant exposure during the DOHaD window.[26,27,35,43,48]

The most commonly studied pollutants were particulate matter (PM) ($PM_{2.5}$ or PM_{10}, 21 studies), NO_2 (10 studies) and PAH (7 studies). Some studies performed trimester specific analyses, other analyzed smaller predefined gestational windows or used distributed-lag model to study weekly exposures during pregnancy. Most of the studies used indirect methods for exposure assessment based on the residential address. Fewer studies measured the exposure by using personal air monitors [41,42,51] or by measuring PAH-DNA adducts[28,36] (cord blood) or PAH metabolites (maternal urine).

The articles were based on mother-child dyads from different continents, mostly Europe, North America and Asia. A number of articles included data from the same birth cohort including: eight studies from the ENVIRonmental influence ON early AGEing (ENVIRONAGE)[25,26,29,33,34,39,40,48], four from the Children's Health Study (CHS)[35,37,48,52], four from the Etude de cohorte généraliste, menée en France sur les Déterminants pré et post natals précoces du développement psychomoteur et de la santé de l'ENfant (EDEN) cohort[23,35,43,48], three from Columbia Center for Child's Environmental Health (CCCEH) study[41,42,51], three from Chinese cohort from Zhengzhou[21,22,45] and two articles from birth cohorts enrolled before and after closing of a coal plant in China[28,36]. Four studies[26,35,43,48] meta-analysed data from multiple European and American birth cohorts.

We classified each of the thirty-two studies into at least one of the following categories: (1) studies of global methylation patterns, (2) studies on candidate-gene methylation that focus on targeted genes of interest, usually with a hypothesized role in the association between ambient air pollution exposure and human diseases (3) epigenome-wide association studies (EWAS) that used untargeted methylation analysis of thousands of CpGs across the genome to discover unknown associations between air pollutants and CpG methylation and (4) telomere length studies.

Twenty-five studies focused on DNA methylation. Of them, ten measured global DNA methylation (Table 2) using different methods, including quantifying total genomic methylation and measuring the methylation in repetitive elements (RE), such as LINE1 and/or Alu. In this group we additionally included two studies did not measure global

methylation based on traditional methods, but summarized methylation data across all loci targeted on the Infinium HumanMethylation450 BeadChip (Illumina450K platform). This platform is mainly used in EWAS studies (Table 4) where it provides a cost-efficient measurement of DNA methylation of more than 450 thousand CpGs across the entire genome. The CpGs included in the platform account for 2% of the total genomic CpG content, but are enriched with potentially relevant CpGs clustered near transcription start sites (called CpG islands) and in the body of the majority of human genes. EWAS studies used CpG-based, region-based approach (differentially methylated region, DMR analysis) and/ or enriched pathway analysis to discover associations between air pollutants and untargeted CpGs or regions across the genome. Candidate gene methylation (n=12 studies) was estimated either by pyrosequencing or by using CpG data from EWAS studies, Table 3. Seven studies analyzed telomere length (Table 5) by using the quantitative polymerase chain reaction (qPCR) protocol developed by Cawthon and expressed the telomere length as relative T/S ratio.[53]

Findings in newborn blood, cord blood and placenta

1. Global DNA methylation studies

Table 2 summarizes the main findings of the studies[20–26,28–42,44–50,52] that assessed the link between air pollution during pregnancy and global methylation patterns. Most studies reported global loss of methylation following increased gestational exposure to $PM_{2.5}$ and PM_{10}, mostly due to exposures in the first trimester [31,4037] Exposure to PAH was also associated with decreased global loss of methylation in cord blood. The exposure to PAH was measured in maternal urine only at one-time point (mainly during the third trimester), and therefore data on other potential important windows of exposure are lacking.[30,36,42] The findings regarding O_3 and NO_2 were less conclusive.[37,49]

Table 2. Studies on global DNA methylation

Author	Method	Sample	Pollutant	Main findings
Liu, 2019	LINE1	Cord blood, n=258	PM_{10}, $PM_{2.5}$ and PM_1 (ambient air)	PM_{10}, $PM_{2.5}$ and PM_1 exposure was associated with ↓LINE1-DNAm. The analyses were conducted only for the period between 12 to 20 gestational week, previously identified as window of exposure associated with preterm birth.
Ladd-Acosta, 2019	Illumina450K	Placenta, n=124; cord blood n=163	NO_2, O_3 (ambient air)	O_3 associated with ↓DNAm at open sea regions (in cord blood) and shelf regions (in both cord blood and placenta). O_3 associated with ↑DNAm at CpG islands (in placenta) and shore regions (both cord blood and placenta). NO_2 associated with ↓DNAm in placenta, mostly at CpG islands.

Abraham, 2018	LINE1, Alu, Illumina450K	Placenta, n=668	PM_{10}, NO_2 (ambient air)	No association, except PM_{10} exposure the day before birth and ↑Alu-DNAm.
Maghbooli, 2018	HPLC	Placenta, n=92	$PM_{2.5}$, PM_{10} (ambient air)	Exposure to $PM_{2.5}$ and PM_{10} in T1 positively correlated with global DNAm
Yang, 2017	LINE1, Alu	Cord blood, n=106	PAH (maternal urine)	PAH (measured only in T3) associated with ↓Alu and ↓LINE1-DNAm
Cai, 2017	LINE1	Placenta, n=181	PM_{10} (ambient air)	PM_{10} exposure in T1 associated with ↓LINE1-DNAm, mostly in newborns with FGR
Lee, 2016	LINE1	Cord blood, n=217	PAH (cord blood DNA adducts)	PAH-DNA adducts associated with ↓LINE1-DNAm
Breton, 2016	LINE1, AluYb8	Newborn blood; n=392 in LINE1, n=181 in AluYb8 analyses	$PM_{2.5}$, O_3, PM_{10}, NO_2 (ambient air)	PM_{10} and O_3 exposure in T1 associated with ↓LINE1-DNAm. O_3 exposure in T3 associated with ↑LINE1-DNAm
Janssen, 2013	LC/MS-MS	Placenta, N=240	$PM_{2.5}$ (ambient air)	$PM_{2.5}$ associated with ↓global DNAm, mostly driven by exposures in T1 (during implantation)
Herbstman, 2012	ELISA-based	Cord blood, N=168	PAH (ambient air, maternal urine)	Ambient PAH (measured only in T3) associated with ↓global DNAm.

Abbreviations: DNAm: DNA methylation; T1, T2 and T3: first, second and third trimester, respectively; Illumina450K: Illumina's Infinium HumanMethylationBeadChip; HPLC: High Performance Liquid Chromatography; LC/MS-MS: Liquid chromatography coupled with tandem mass spectrometry. ELISA: enzyme-linked immunosorbent assay; PM: Particulate Matter; PAH: Polycystic Aromatic Hydrocarbons.

2. Candidate genes studies

Exposure to air pollution was associated with altered methylation (mostly, but not exclusively, with gene-promoter hypermethylation) of a number of targeted genes with names reported in Table 3. Unfortunately, no gene was analyzed in more than one study. Briefly, exposure to air pollutants (the studied pollutant and analyzed tissue are shown in brackets) was associated with altered methylation of several genes involved in pre-eclampsia (NO_2; placenta)[23], circadian rhythm regulation ($PM_{2.5}$; placenta)[25], growth (PM_{10}, NO_2, SO_2; cord blood)[22], obesity (PM_{10}, NO_2, SO_2; cord blood)[21], DNA repair and tumor suppression ($PM_{2.5}$, BC; placenta)[29], glucocorticoid metabolism (PM_{10}; placenta),[31] energy regulation ($PM_{2.5}$; placenta),[34] pro-allergic immune responses (PAH; cord blood),[41] synthesis of antioxidant enzymes ($PM_{2.5}$, NO_2; cord blood),[35] and mitochondrial functions ($PM_{2.5}$, NO2, cord blood, placenta)[39,45]

Table 3. Studies on candidate gene methylation

Author	Candidate gene	Method	Sample	Pollutant	Main findings
Fang, 2020	GPR61 gene	QMS-PCR	Cord blood, n=568	PM_{10}, NO_2, SO_2 (ambient air)	PM_{10} and SO_2 exposure in pregnancy was associated with ↓GPR61-DNAm, while NO_2 exposure with ↑GPR61-DNAm.
Zhou, 2019	SOD2 gene	QMS-PCR	Cord blood, n=568	PM_{10}, NO_2, SO_2 (ambient air)	PM_{10} in T2 associated with ↑DNAm, NO_2 in T3 with ↓DNAm.
He, 2018	H19 gene	QMS-PCR	Cord blood, n=527	PM_{10}, NO_2, SO_2 (ambient air)	PM_{10} and SO_2 exposure in pregnancy was associated with ↓DNAm in H19 promoter, while $NO_2$with ↑DNAm.
Abraham, 2018	Genes with specific expression	Illumina 450K	Placenta, n=668	PM_{10}, NO_2 (ambient	NO_2 associated with ↓DNAm in ADORA2B, CAPN10 and in PXT1/ KCTD20. PM_{10} associated with

	patterns in the placenta (18972 CpGs in total)			air)	↑DNAm of SLC44A5, ADCK5 and TMG6 genes and ↓DNAm in KYNU.
Nawrot, 2018	Circadian pathway genes: CLOCK, NPAS2, BMAL1, CRY1, CRY2, PER1, PER2, PER3	pyro	Placenta, n=407	$PM_{2.5}$ (ambient air)	$PM_{2.5}$ associated with ↑BMAL1-DNAm. T1 exposure with ↓CLOCK-DNAm. T3 exposure with ↑NPAS2 and CRY1 and ↓PER2 and PER3 DNAm.
Lee, 2018	GSTP1 gene	pyro	Nasal epithelia at 7 years, n=131	$PM_{2.5}$ (ambient air)	$PM_{2.5}$ exposure >37 gestational weeks associated with ↑DNAm in GSTP1 in NE at 7 years.
Neven, 2018	DNA repair and tumor suppressor genes:APEX1, OGG1, PARP1, ERCC1, ERCC4, p53, DAPK1	pyro	Placenta, n=463	$PM_{2.5}$, BC, NO_2 (ambient air)	$PM_{2.5}$ associated with ↑DNAm in APEX1, OGG1, ERCC4 and p53 and with ↓DNAm of DAPK1. BC was associated with ↑DNAm in APEX1 and ERCC4. NO_2: no associations.
Gruzieva, 2017	Antioxidant and inflammatory genes (38 genes in total, 739 CpGs)	Illumina 450K	Cord blood, n=1508	NO_2 (ambient air)	NO_2 exposure in pregnancy was associated with ↑DNAm in CAT gene and ↓DNAm in TPO gene
Cai, 2017	Fetal growth related genes: HSD11B2 and NR3C1	pyro	Placenta, n=181	PM_{10} (ambient air)	Exposure in T1 and T2 associated with ↑DNAm in HSD11B2 gene
Saenen, 2017	LEP gene	pyro	Placenta, n=361	$PM_{2.5}$ (ambient air)	$PM_{2.5}$ exposure in T2 associated with ↓LEP-DNAm
Janssen, 2015	Mitochondrial DNA regions: D-loop and MT-RNR1 region	pyro	Placenta, n=381	$PM_{2.5}$ (ambient air)	$PM_{2.5}$ exposure, mostly in T1, was associated with ↑mtDNAm in both D-loop and MT-RNR1 region
Tang, 2012	Asthma-related genes: IFNγ and IL4	BGS	Cord blood, n=53	PAH (ambient air)	PAH measured in T3 associated with ↑ IFNγ promoter DNAm

Abbreviation: pyro: pyrosequencing; BGS: Bisulfite Genomic Sequencing; Illumina450K: Illumina's Infinium HumanMethylation BeadChip; QMS-PCR: Quantitative Methylation-Specific-Polymerase Chain Reaction; DNAm: DNA methylation; PM: Particulate Matter; PAH: Polycystic Aromatic Hydrocarbons

3. Epigenome-wide association studies

Table 4 describes the main findings of the studies that conducted an epigenome wide analysis.[23,26,35,38,48,49,51,52] The largest study on NO_2 exposure meta-analyzed data from 1508 mother-child dyads from nine separate cohorts from Europe and United States by measuring epigenome-wide methylation in cord blood.[35] The top three CpGs associated with gestational NO_2 exposure were mapped to genes important for mitochondrial functions. One of the CpGs (cg08973675, in the SLC25A28 gene) showed similar direction of association in an another cohort of newborns.[49] Gestational NO_2 exposure was also associated with placental methylation, mainly mapped to genes linked to preeclampsia[23] and inflammatory processes.[49]

Particulate matter exposure in pregnancy was assessed in four epigenome-wide studies.[23,26,48,52] The largest and most recent study conducted by Gruzieva and colleagues[48] studied cord blood methylation and included information on 1,949 and 1,551 mother-child

dyads in the corresponding PM_{10} and $PM_{2.5}$ analyses.[48] The study reported associations between gestational PM_{10} or $PM_{2.5}$ exposures and the methylation of 20 CpGs in cord blood. The robustness of the associations of the 6 PM_{10}-related CpGs was tested in an independent cohort of newborns and only the PM_{10}-related CpG, cg18640183 in the P4HA2 gene showed consistent direction of association. (Supplementary Table S3). The region-based analysis identified large number of DMRs related to air pollutants, and two PM_{10}-related DMRs in the genes H19 and MARCH11 were replicated in newborns. Other two studies on particulate matter and DNA methylation in newborns[26,52] had smaller sample size and/or were based on cohorts already included in the meta-analysis by Gruzieva and colleagues.

Only one epigenome-wide study estimated PAH exposure.[51] The study was published in 2009 and used a slightly older method to perform unbiased methylation profiling. The top finding was the change in methylation of the ACSL3 gene in relation to PAH exposure.

Table 4. Epigenome-wide association studies

Author	Sample	Pollutant	Main findings
Ladd-Acosta, 2019	Cord blood, n=163; placenta, n=124	NO_2, O_3 (ambient air)	Several DMR associated with NO_3 and O_3, some of which were sex specific. The DMRs in the placenta seemed to be tissue-specific, while those reported in cord blood, showed similar direction in the placenta.
Gruzieva, 2019[a]	Discovery analyses in cord/ newborn blood: n=1,949 (PM_{10}) and n=1,551 ($PM_{2.5}$). Replication analyses in cord blood (n=688), peripheral blood of 7-9yrs (n1=692, n2=525 n3=901) and 15-16yrs (n1=198, n2=903)	$PM_{2.5}$, PM_{10} (ambient air)	Gestational exposure to ether $PM_{2.4}$ or PM_{10} was associated with 20 CpGs at birth and hundreds of DMRs. Enriched pathways: NOTCH signaling pathway, Rho GTPase cycle, neuro-transmitter release cycle, GABA synthesis, release, reuptake and degradation. The CpG cg18640183, and two DMRs (H19 and MARCH1) showed consistent direction in an independent cohort of newborns. Three CpGs cg00905156 (FAM13A), cg06849931 (NOTCH4) and cg06849931 (P4HA2) showed consistent association in at least one of the independent cohorts of older children aged 7-9.
Abraham, 2018	Placenta, n=668	PM_{10}, NO_2 (ambient air)	Out of the 4 identified PM_{10} or NO_2-related CpGs, 2 were in the ADORA2B gene linked with hypoxia and pre-eclampsia. Strongest association was seen after exposures in T2. More than 20 DMRs were also identified.
Gruzieva, 2017[a]	Discovery analyses in cord/newborn blood, n=1508. Replication analyses in peripheral blood of 4 yrs (n=733) and 8 yrs (n=786)	NO_2 (ambient air)	Gestational NO_2 exposure was associated with 3 CpG sites in mitochondria-related genes: cg12283362 (LONP1), cg24172570 (HIBADH), and cg08973675 (SLC25A28). Enriched pathways: negative regulation of cellular process, negative regulation of biological process and integrin-linked kinase signaling pathway. The cg08973675 replicated in an independent sample of older children.
Goodrich, 2016	Cord blood, n=22	NOx (ambient air)	No CpG passed the FDR threshold. Enriched pathways were found related to xenobiotic metabolism, oxygen and gas transport, and sensory perception of chemical stimuli
Perera, 2009	Cord blood, n=22	PAH (ambient air)	Top finding was the ACSL3 gene whose association with PAH was further confirmed in a slightly larger sample (N=53)

Two studies by Plusquin et al. 2018 and Breton et al. 2016 were excluded from the main summary of the findings since they included cohorts included in a meta-analysis by Gruzieva et al. 2019. [a]The study included children in their replication analysis.
All studies, except for Perera et al, 2008 (that used Methylation Sensitive Restriction Fingerprinting), used Illumina's Infinium HumanMethylationBeadChip to assess epigenome wide methylation patterns. Abbreviations: T1, T2 and T3: first, second and third trimester, respectively; PAH: Polycystic Aromatic Hydrocarbons

4. Telomere length studies

The association between air pollution exposure in pregnancy and telomere length at birth was assessed in seven studies.[28,33,43,46,47,50] The main findings are presented in Table 5. The studies were generally consistent in reporting an inverse association between gestational exposure to air pollutants (mainly PM or PAH) and telomere length in newborn blood and placenta, [28,33,46,47,50] although there were studies that also reporter longer telomeres in later pregnancy.[33,50]

Table 5. Telomere length studies

Author	Sample	Pollutant	Conclusion
Lee, 2020	Cord blood, n=152	$PM_{2.5}$ (ambient air)	Exposures during pregnancy was associated with ↓TL, mostly due to exposures in mid-gestation between 12-20 gestational week.
Clemente, 2019[a]	Peripheral blood, 8 yrs; n=1396	NO_2, $PM_{2.5}$ (ambient air)	Gestational NO_2 exposure was associated with shorter TL across all trimesters. 1 year-childhood exposure to NO_2 and $PM_{2.5}$ was associated with shorter TL.
Song, 2019	Cord blood, n=743	$PM_{2.5}$, PM_{10}, SO_2, CO, NO (ambient air)	Exposures to $PM_{2.5}$, PM_{10}, CO, and SO_2 during T3 were related to shorter TL. Associations were stronger in males.
Nie, 2019	Cord blood, n=247	PAH (maternal urine)	Association with shorter TL.
Rosa, 2019	Cord blood, n=423	$PM_{2.5}$ (ambient air)	Exposure during gestational weeks 4-9 associated with shorter TL. Exposure during weeks 14-19 and 34-36 associated with longer TL. Associations were stronger in girls.
Perera, 2018	Cord blood, n=225	PAH (ambient air)	Association with shorter TL.
Martens, 2017	Cord blood, n=698; placenta, n=660	$PM_{2.5}$ (ambient air)	Exposure during mid-gestation (weeks 12-25 for cord blood and weeks 15-27 for placenta) associated with shorter TL. Exposure in late pregnancy (weeks 32-34) associated with longer telomeres in cord blood. No effect modification by sex.

All studies used the same method for estimating telomere length (quantitative polymerase chain reaction)
[a]The study was the only one conducted in children. It was also the only study where exposure after pregnancy was assessed, in particular, during the first year of life.

Findings in children

The number of studies that conducted analysis in children is limited. Briefly, an association was found between late gestation $PM_{2.5}$ exposure and nasal epithelia methylation of a candidate gene (GSTP1 gene, involved in xenobiotic metabolism) at age 7, with strongest effects seen in boys.[27] A large meta-analysis on early life air pollution and telomeres included more than 1300 8-year old children from six European birth cohort studies reported that prenatal exposure and exposure during the first year of life to $PM_{2.5}$ and NO_2 was associated with shorted telomeres at age 8.[43] There were no studies exploring the association between early-life exposure to air pollution and global methylation.

The two large meta-analyses on PM[48] and NO_2[35] exposure, described previously, tried to replicate in several cohorts of older children the association seen in newborns, in order to see whether they are stable throughout childhood. The results were inconclusive. Only one NO_2-related CpG cg08973675 in the SLC25A28 gene showed robust association in two cohorts of older children aged 4 and 8 years. Out of total 20 CpGs associated with PM_{10} or $PM_{2.5}$, none showed clear association across different cohorts of older children (6 PM_{10}-related CpGs were tested in three cohorts of children aged 7-9 years, and two cohorts of children aged 15-16 years, while the 14 $PM_{2.5}$-related CpGs were tested in two cohorts of children aged 7-9 years and in one cohort of teenagers aged 15-16 years). It should be noted that three PM_{10} associated CpGs (cg00905156, cg06849931 and cg06849931 mapped to three genes important for respiratory health: FAM13A, NOTCH4 and P4HA2 gene, respectively) showed consistent direction in at least one of the three independent cohorts 7-9 year-olds.

Discussion

The studies included in this review provided evidence that prenatal exposure to air pollutants is linked with global and locus-specific alterations in DNA methylation as well as telomere length shortening in newborn cord blood and placenta. Further studies are needed to elucidate whether these changes can influence childhood outcomes years after the exposure. The number of studies that studied air pollution exposure during the first 1000 days of life by measuring DNA methylation or telomere length in older children was limited.

Global loss of methylation is linked with genomic instability and can predispose to the development of human diseases.[54] Gestational exposure to air pollutants ($PM_{2.5/10}$ and PAH) was generally associated with global loss of methylation in different cohorts and different tissues (placenta and cord/ newborn blood), independently of the exposure assessment method and the laboratory method used to measure the global methylation patterns. Some studies on PM exposure that conducted trimester-specific analyses, identified the beginning of the pregnancy as a potential period of susceptibility.[31,37,40] It is plausible that exposures in early pregnancy might be strongly associated with global loss of methylation since the period around the implantation is the period when the epigenetic reprogramming occurs *de novo* methylation takes place.[6] Exposure to air pollutants in such a vulnerable period might interfere with the DNA methylation machinery and lead to generalized loss of methylation.[55] Whether these changes persist into childhood is unknown.

Air pollution is believed to influence human health through the generation of reactive oxygen species, as increased oxidative stress is known to trigger number of redox-sensitive cellular signaling pathways.[56] Although, the heterogeneity of the published candidate gene studies (all studies analyzed different sets of genes with different sets of pollutants in different tissues) did not provide enough evidence to draw strong conclusions regarding a specific gene, the overall findings suggest that gestational exposure to pollutants can lead to methylation changes in cord blood and placenta, in genes involved in key cellular responses to oxidative stress,[29,35,39,45] and genes with known role in growth, early life development and hypertensive disorders of the pregnancy.[22,23,31,34]

Epigenome-wide association studies independently tested the association between gestational air pollution exposure and more than 400 000 CpGs throughout the genome. This agnostic approach allows to identify novel genomic regions associated with the exposure. The strongest and most robust associations were seen for CpGs or DMRs mapped to genes with roles in mitochondrial[35,49], respiratory functions[48] and fetal growth[48]. The rest of the CpGs were mapped to genes with known roles in auto-immunity[49], inflammation[49], inter/intracellular signaling[48,49], cell cycle regulation[48,49], embryonal development and adverse pregnancy outcomes[23,48,52]. Gestational exposure to PM_{10} and altered methylation in the NOTCH signaling pathway with an important role in embryonal development, while[48] gestational NO_2 exposure was associated with altered methylation in pathways that downregulate cellular functions and are involved in cell migration, proliferation and survival.[35]

These findings compliment those from global methylation and candidate gene studies, and provide further evidence that gestational air pollution exposure can have an impact on early-life global and locus-specific methylation patterns. The gestational period, especially early pregnancy, is the period when the DNA methylation pattern undergo most dramatic changes: active and passive de-methylation of nearly all maternal and paternal patterns, process of epigenome-wide re-methylation and, finally, gene-specific changes that initiate embryonal cell differentiation.[6] Since the majority of these patterns are believed to be largely maintained in the next cell replications it is possible that air pollution exposure during this dynamic period can leave epigenetic fingerprints that might influence later health and disease outcomes.

It should be noted however, that the identified CpGs/DMRs/enriched pathways were quite heterogeneous between studies and it seems difficult to replicate the EWAS findings across different populations. This could be partially explained by pollutant-specific effects that trigger different biological cascades, as suggested by the lack of overlap between the top NO_2-related[35] and PM_{10}-related CpGs[48] and the different enriched pathways found in the NO_2 and PM_{10} analyses. Particulate matter- specific effects might be even more difficult to replicate due to the possible differences in the source and chemical composition of the particulate matter particles in different populations, although this probably is not the major cause. In the context of air pollution, environmental mixtures and different confounding pattern across study populations may be contributing factors to baseline differences in laboratory conditions, unmeasured batch effects and different pre-processing pipelines.[57,58] For example, different studies use different methods of estimating pollutant concentration and misclassification of exposure is possible when studying exposure based on residential address. Two PM_{10}-related DMRs[48] (including the imprinted growth-related gene H19, that showed associations with prenatal PM_{10} exposure in a previous candidate gene study[22]) showed promising results by replication in an independent cohort of newborns. This could mean that future studies should consider expanding the search from single CpG level to genomic regions that contain multiple CpGs, or even to epi-signatures based on methylation levels of hundreds of CpGs spread across the genome to find patterns predictive of the exposure. However, due to the relatively small effect sizes and the variable chemical composition of PM, advanced statistical methods would be needed to appropriately model the exposure (or the concurrent exposure to multiple pollutants that would better reflect real-life exposure), as well as large sample size, in order to detect robust associations on population level and/or create scores that could accurately predict early-life exposure to pollution, as was previously done with prenatal exposure to smoke.[59] Future studies should also assess whether the changes seen at birth are stable throughout childhood. Moreover, it is known that methylation patterns are tissue-specific. For example, cord blood and placenta are expected to have different methylation patterns due to their different biological function and cell composition. According to one study[49], DMRs identified in cord blood showed consistent direction of effect in the placental tissue, while the DMRs identified in placenta seem tissue-specific. Further studies are needed to confirm these findings.

Telomere length at birth is a reflection of the complex interplay between genetics, number of cell divisions (dependent of both somatic growth and gestational age), exposure to oxidative

stress and the counter-regulatory effect of the telomerase.[60] Findings from studies included in this review indicate that prenatal exposure to air pollution can lead to telomere attrition, as seen at birth and in childhood. It is known that the variability in telomere length in adults most likely originates *in utero* and that short telomeres in adults are associated with higher risk for chronic-non communicable diseases. Therefore, the possible effect of prenatal exposure to air pollution on early telomere maintenance system might not be negligible when talking about the lifetime risk of chronic non-communicable diseases.[60] Results regarding possible windows of exposure during pregnancy and the effect modification by sex are unclear. The authors of two studies that reported longer telomeres in late gestation hypothesized that prolonged exposure to air pollution might increase activity of the telomerase.[33,50]

It is known that DNA methylation and the telomere maintenance system are interrelated on cellular lever.[13,14] This is especially true during the early gestational period. Short telomeres in embryonic cells might led to downregulation of the *de novo* DNA methyl transferases, that in turn might induce genomic instability and impair embryonic stem cell differentiation. DNA methylation is can also influence telomere length via the regulation of the telomerase activity.[61,62]

Considering the both DNA methylation and the telomere system are key players in many cellular functions, future studies need to assess their potential to leave long-term consequences in the context of the fetal origins of health and disease hypothesis. Unfortunately, only few studies included in this review analyzed data in relation to some specific birth or childhood outcomes. Some of them provided preliminary evidence, that global methylation, methylation at specific CpGs and/or genes and telomere length, might mediate the association between prenatal exposure to air pollution and birth outcomes[31,32], childhood respiratory outcomes[27,51,52] and neurodevelopmental scores[28,47], respectively.

Conclusion

Prenatal exposure to air pollution was associated with global loss of methylation, telomere shortening and epigenetic alterations mapped to key genes involved in oxidative stress response, mitochondrial function, inflammation, fetal growth and development. Additional studies are needed to test the robustness of the associations across different populations and explore potential windows of vulnerability during pregnancy and early-life, as well as to

confirm the role of DNA methylation and telomere length as mediators in the association between prenatal and early-life exposure to air pollution and later childhood outcomes.

Bibliography:

1. WHO Global update 2005 - Summary of risk assessment. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. 2005.
2. 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. doi:10.1056/NEJMra0708473
3. Klepac P, Locatelli I, Korošec S, Künzli N, Kukec A. Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges. *Environ Res*. 2018. doi:10.1016/j.envres.2018.07.008
4. Capello F, Gaddi AV. *Clinical Handbook of Air Pollution-Related Diseases*.; 2018. doi:10.1007/978-3-319-62731-1
5. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. *Lancet*. 2018. doi:10.1016/S0140-6736(18)31268-6
6. Cedar H, Bergman Y. Programming of DNA Methylation Patterns. *Annu Rev Biochem*. 2012. doi:10.1146/annurev-biochem-052610-091920
7. Slieker RC, Roost MS, van Iperen L, et al. DNA Methylation Landscapes of Human Fetal Development. *PLoS Genet*. 2015. doi:10.1371/journal.pgen.1005583
8. Blackburn EH, Epel ES, Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science (80-)*. 2015;350(6265):1193-1198. doi:10.1126/science.aab3389
9. Entringer S, de Punder K, Buss C, Wadhwa PD. The fetal programming of telomere biology hypothesis: An update. *Philos Trans R Soc B Biol Sci*. 2018;373(1741). doi:10.1098/rstb.2017.0151
10. Okuda K, Bardeguez A, Gardner JP, et al. Telomere length in the newborn. *Pediatr Res*. 2002;52(3):377-381. doi:10.1203/01.PDR.0000022341.72856.72
11. Benetos A, Dalgård C, Labat C, et al. Sex difference in leukocyte telomere length is ablated in opposite-sex co-twins. *Int J Epidemiol*. 2014;43(6):1799-1805. doi:10.1093/ije/dyu146
12. Martens DS, Nawrot TS. Ageing at the level of telomeres in association to residential landscape and air pollution at home and work: a review of the current evidence. *Toxicol Lett*. 2018;298:42-52. doi:10.1016/j.toxlet.2018.06.1213
13. Saenen ND, Martens DS, Neven KY, et al. Air pollution-induced placental alterations: An interplay of oxidative stress, epigenetics, and the aging phenotype? *Clin Epigenetics*. 2019;11(1):124. doi:10.1186/s13148-019-0688-z
14. Martens DS, Nawrot TS. Air Pollution Stress and the Aging Phenotype: The Telomere Connection. *Curr Environ Heal reports*. 2016;3(3):258-269. doi:10.1007/s40572-016-0098-8
15. Miri M, Nazarzadeh M, Alahabadi A, et al. Air pollution and telomere length in adults: A systematic review and meta-analysis of observational studies. *Environ Pollut*. 2019;244:636-647. doi:10.1016/j.envpol.2018.09.130
16. Desai G, Chu L, Guo Y, Myneni AA, Mu L. Biomarkers used in studying air pollution exposure during pregnancy and perinatal outcomes: a review. *Biomarkers*. 2017;22(6):489-

501. doi:10.1080/1354750X.2017.1339294

17. Rider CF, Carlsten C. Air pollution and DNA methylation: Effects of exposure in humans. *Clin Epigenetics*. 2019;11(1). doi:10.1186/s13148-019-0713-2

18. Luyten LJ, Saenen ND, Janssen BG, et al. Air pollution and the fetal origin of disease: A systematic review of the molecular signatures of air pollution exposure in human placenta. *Environ Res*. 2018;166:310-323. doi:10.1016/j.envres.2018.03.025

19. Ferrari L, Carugno M, Bollati V. Particulate matter exposure shapes DNA methylation through the lifespan. *Clin Epigenetics*. 2019;11(1). doi:10.1186/s13148-019-0726-x

20. Lee AG, Cowell W, Kannan S, et al. Prenatal particulate air pollution and newborn telomere length: Effect modification by maternal antioxidant intakes and infant sex. *Environ Res*. 2020;187:109707. doi:10.1016/j.envres.2020.109707

21. Feng F, Huang L, Zhou G, et al. GPR61 methylation in cord blood: a potential target of prenatal exposure to air pollutants. *Int J Environ Health Res*. 2020. doi:10.1080/09603123.2020.1773414

22. He T, Zhu J, Wang J, et al. Ambient air pollution, H19/DMR methylation in cord blood and newborn size: A pilot study in Zhengzhou City, China. *Chemosphere*. 2018;212:863-871. doi:10.1016/j.chemosphere.2018.08.140

23. Abraham E, Rousseaux S, Agier L, et al. Pregnancy exposure to atmospheric pollution and meteorological conditions and placental DNA methylation. *Environ Int*. 2018. doi:10.1016/j.envint.2018.05.007

24. Maghbooli Z, Hossein-nezhad A, Adabi E, et al. Air pollution during pregnancy and placental adaptation in the levels of global DNA methylation. *PLoS One*. 2018. doi:10.1371/journal.pone.0199772

25. Nawrot TS, Saenen ND, Schenk J, et al. Placental circadian pathway methylation and in utero exposure to fine particle air pollution. *Environ Int*. 2018. doi:10.1016/j.envint.2018.02.034

26. 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. doi:10.1021/acs.est.7b06447

27. Lee AG, Le Grand B, Hsu HHL, et al. Prenatal fine particulate exposure associated with reduced childhood lung function and nasal epithelia GSTP1 hypermethylation: Sex-specific effects. *Respir Res*. 2018;19(1). doi:10.1186/s12931-018-0774-3

28. Perera F, Lin C jung, Qu L, Tang D. Shorter telomere length in cord blood associated with prenatal air pollution exposure: Benefits of intervention. *Environ Int*. 2018. doi:10.1016/j.envint.2018.01.005

29. Neven KY, Saenen ND, Tarantini L, et al. Placental promoter methylation of DNA repair genes and prenatal exposure to particulate air pollution: an ENVIRONAGE cohort study. *Lancet Planet Heal*. 2018. doi:10.1016/S2542-5196(18)30049-4

30. Yang P, Gong YJ, Cao WC, et al. Prenatal urinary polycyclic aromatic hydrocarbon metabolites, global DNA methylation in cord blood, and birth outcomes: A cohort study in China. *Environ Pollut*. 2018. doi:10.1016/j.envpol.2017.11.082

31. Cai J, Zhao Y, Liu P, et al. Exposure to particulate air pollution during early pregnancy is associated with placental DNA methylation. *Sci Total Environ*. 2017. doi:10.1016/j.scitotenv.2017.07.029

32. Liu X, Ye Y, Chen Y, et al. Effects of prenatal exposure to air particulate matter on the risk of preterm birth and roles of maternal and cord blood LINE-1 methylation: A birth cohort study

in Guangzhou, China. *Environ Int*. 2019;133. doi:10.1016/j.envint.2019.105177

33. Martens DS, Cox B, Janssen BG, et al. Prenatal air pollution and newborns' predisposition to accelerated biological aging. *JAMA Pediatr*. 2017. doi:10.1001/jamapediatrics.2017.3024

34. Saenen ND, Vrijens K, Janssen BG, et al. Lower placental leptin promoter methylation in association with fine particulate matter air pollution during pregnancy and placental nitrosative stress at birth in the ENVIRONAGE cohort. *Environ Health Perspect*. 2017. doi:10.1289/EHP38

35. Gruzieva O, Xu CJ, Breton C V., et al. Epigenome-wide meta-analysis of methylation in children related to prenatal NO2 air pollution exposure. *Environ Health Perspect*. 2017. doi:10.1289/EHP36

36. Lee J, Kalia V, Perera F, et al. Prenatal airborne polycyclic aromatic hydrocarbon exposure, LINE1 methylation and child development in a Chinese cohort. *Environ Int*. 2017. doi:10.1016/j.envint.2016.12.009

37. Breton C V., Yao J, Millstein J, et al. Prenatal air pollution exposures, DNA methyl transferase genotypes, and associations with newborn line1 and Alu methylation and childhood blood pressure and carotid intima-media thickness in the children's health study. *Environ Health Perspect*. 2016. doi:10.1289/EHP181

38. Goodrich JM, Reddy P, Naidoo RN, Asharam K, Batterman S, Dolinoy DC. Prenatal exposures and DNA methylation in newborns: A pilot study in Durban, South Africa. *Environ Sci Process Impacts*. 2016;18(7):908-917. doi:10.1039/c6em00074f

39. Janssen BG, Byun HM, Gyselaers W, Lefebvre W, Baccarelli AA, Nawrot TS. Placental mitochondrial methylation and exposure to airborne particulate matter in the early life environment: An ENVIRONAGE birth cohort study. *Epigenetics*. 2015. doi:10.1080/15592294.2015.1048412

40. Janssen BG, Godderis L, Pieters N, et al. Placental DNA hypomethylation in association with particulate air pollution in early life. *Part Fibre Toxicol*. 2013. doi:10.1186/1743-8977-10-22

41. Tang WY, Levin L, Talaska G, et al. Maternal exposure to polycyclic aromatic hydrocarbons and 5'-CpG methylation of interferon-γ in cord white blood cells. *Environ Health Perspect*. 2012. doi:10.1289/ehp.1103744

42. Herbstman JB, Tang D, Zhu D, et al. Prenatal exposure to polycyclic aromatic hydrocarbons, benzo[a]pyrene-DNA adducts, and genomic DNA methylation in cord blood. *Environ Health Perspect*. 2012. doi:10.1289/ehp.1104056

43. Clemente DBP, Vrijheid M, Martens DS, et al. Prenatal and childhood traffic-related air pollution exposure and telomere length in european children: The HELIX project. *Environ Health Perspect*. 2019;127(8). doi:10.1289/EHP4148

44. Perera F, Tang WY, Herbstman J, et al. Relation of DNA methylation of 5′-CpG island of ACSL3 to transplacental exposure to airborne polycyclic aromatic hydrocarbons and childhood asthma. *PLoS One*. 2009. doi:10.1371/journal.pone.0004488

45. Zhou G, He T, Huang H, et al. Prenatal ambient air pollution exposure and SOD2 promoter methylation in maternal and cord blood. *Ecotoxicol Environ Saf*. 2019;181:428-434. doi:10.1016/j.ecoenv.2019.06.039

46. Song L, Zhang B, Liu B, et al. Effects of maternal exposure to ambient air pollution on newborn telomere length. *Environ Int*. 2019. doi:10.1016/j.envint.2019.04.064

47. Nie J, Li J, Cheng L, et al. Prenatal polycyclic aromatic hydrocarbons metabolites, cord blood telomere length, and neonatal neurobehavioral development. *Environ Res*. 2019;174:105-113.

doi:10.1016/j.envres.2019.04.024

48. Gruzieva O, Xu C-J, Yousefi P, et al. Prenatal Particulate Air Pollution and DNA Methylation in Newborns: An Epigenome-Wide Meta-Analysis. *Environ Health Perspect*. 2019;127(5):057012. doi:10.1289/EHP4522

49. Ladd-Acosta C, Feinberg JI, Brown SC, et al. Epigenetic marks of prenatal air pollution exposure found in multiple tissues relevant for child health. *Environ Int*. 2019. doi:10.1016/j.envint.2019.02.028

50. Rosa MJ, Hsu HHL, Just AC, et al. Association between prenatal particulate air pollution exposure and telomere length in cord blood: Effect modification by fetal sex. *Environ Res*. 2019. doi:10.1016/j.envres.2019.03.003

51. Perera F, Tang WY, Herbstman J, et al. Relation of DNA methylation of 5′-CpG island of ACSL3 to transplacental exposure to airborne polycyclic aromatic hydrocarbons and childhood asthma. *PLoS One*. 2009;4(2). doi:10.1371/journal.pone.0004488

52. Breton C V., Gao L, Yao J, Siegmund KD, Lurmann F, Gilliland F. Particulate matter, the newborn methylome, and cardio-respiratory health outcomes in childhood. *Environ Epigenetics*. 2016. doi:10.1093/eep/dvw005

53. Cawthon RM. Telomere measurement by quantitative PCR. *Nucleic Acids Res*. 2002. doi:10.1093/nar/30.10.e47

54. Pogribny IP, Beland FA. DNA hypomethylation in the origin and pathogenesis of human diseases. *Cell Mol Life Sci*. 2009;66(14):2249-2261. doi:10.1007/s00018-009-0015-5

55. Teneng I, Montoya-Durango DE, Quertermous JL, Lacy ME, Ramos KS. Reactivation of L1 retrotransposon by benzo(a)pyrene involves complex genetic and epigenetic regulation. *Epigenetics*. 2011;6(3):355-367. doi:10.4161/epi.6.3.14282

56. Kelly FJ. Oxidative stress: Its role in air pollution and adverse health effects. *Occup Environ Med*. 2003;60(8):612-616. doi:10.1136/oem.60.8.612

57. Breton C V., Marsit CJ, Faustman E, et al. Small-magnitude effect sizes in epigenetic end points are important in children's environmental health studies: The children's environmental health and disease prevention research center's epigenetics working group. *Environ Health Perspect*. 2017;125(4):511-526. doi:10.1289/EHP595

58. Pekkanen J, Pearce N. Environmental epidemiology: Challenges and opportunities. *Environ Health Perspect*. 2001. doi:10.2307/3434913

59. Richmond RC, Suderman M, Langdon R, Relton CL, Smith GD. DNA methylation as a marker for prenatal smoke exposure in adults. *Int J Epidemiol*. 2018;47(4):1120-1130. doi:10.1093/ije/dyy091

60. Entringer S, de Punder K, Buss C, Wadhwa PD. The fetal programming of telomere biology hypothesis: An update. *Philos Trans R Soc B Biol Sci*. 2018. doi:10.1098/rstb.2017.0151

61. Harrington L, Pucci F. In medio stat virtus: Unanticipated consequences of telomere dysequilibrium. *Philos Trans R Soc B Biol Sci*. 2018;373(1741). doi:10.1098/rstb.2016.0444

62. Joyce BT, Zheng Y, Nannini D, et al. DNA methylation of telomere-related genes and cancer risk. *Cancer Prev Res*. 2018;11(8):511-522. doi:10.1158/1940-6207.CAPR-17-0413

Supplementary Material: Study I

Table S1. PubMed search strategy with keywords used to define the exposure, outcome and the population of interest.

Search date: 2020 October 23rd		**# records**
1. Exposure	"Air pollution"[mesh] OR "air pollution"[tiab] OR "air pollutant"[tiab] OR "air pollutants"[tiab] OR "Carbon Monoxide"[mesh] OR "sulfur dioxide"[mesh] OR "sulphur dioxide"[tiab] OR "SO2"[tiab] OR "Nitrogen Dioxide"[Mesh] OR "nitrogen oxide*" OR "NOx"[tiab] OR "NO2"[tiab] OR ozone[mesh] OR ozone[tiab] OR "O3"[tiab] OR "Particulate Matter"[mesh] OR "particulate matter"[tiab] OR "PM"[tiab] OR "PM*"[tiab] OR particle*[tiab] OR "Polycyclic Aromatic Hydrocarbons"[Mesh] OR "PAH"[tiab] OR "Volatile Organic Compounds"[tiab] OR "VOC"[tiab] OR "black carbon"[tiab] OR "elemental carbon"[tiab] OR "organic carbon"[tiab]	1023800
2. Outcome:	"DNA Methylation"[mesh] OR "Telomere Shortening"[mesh] OR "epigenetic process"[tiab] OR "DNA methylation"[tiab] OR "Epigenome-Wide"[tiab] OR telomere[tiab] OR "CpG"[tiab] OR "Genome-Wide Association Study"[Mesh] OR "gene expression"[tiab]	520273
3 Population	Infant[tiab] OR newborn[tiab] OR toddler[tiab] OR child*[tiab] OR "Child"[Mesh] OR "Infant"[Mesh] OR preschool*[tiab] OR young[tiab] OR prenatal[tiab] OR pregnancy[tiab] OR conception[tiab] OR birth[tiab] OR perinatal[tiab] OR gestation[tiab] OR fetal[tiab]	3891838
4	1 AND 2 AND 3	809
5	"Animals"[Mesh] NOT "Humans"[Mesh]	4748697
6	4 NOT 5	555

Table S2. List of all studies included in the systematic review, ordered chronologically, with all extracted data.

Author, Country, Study population, Sample Size	Pollutant, Exposure assessment method; Windows of exposure, Exposure levels	Type of analyses conducted, Outcome assessment method, Sample type	Main results and/or effect estimates	Covariates	Additional relevant analyses
Lee, 2020 USA PRISM cohort N=152	**Pollutant:** $PM_{2.5}$ **Method:** hybrid LUR and satellite-based model. **Windows:** whole pregnancy, by gestational week. **Mean levels:** 8.8μg/m^3	**Analysis:** TL **Methods:** qPCR**. Sample:** CB**. Effect presented as:** change in DNAm for 1μg/m^3 increase in $PM_{2.5}$ (in brackets 95%CI)	$PM_{2.5}$ exposures during pregnancy was associated with ↓TL: −0.29 (95% CI -0.49 to −0.10). Sensitive period of exposure was the period between 12-20 weeks.	Maternal age, self-reported race/ethnicity, marital status, education level, lifetime stressors, antioxidant intake during pregnancy and child sex.	Children born to mothers reporting low antioxidant intake were most vulnerable to $PM_{2.5}$ exposure.
Feng, 2020 China, Zhengzhou cohort N=568	**Pollutant:** PM_{10}, NO_2, SO_2**. Method:** fixed site monitoring stations. **Windows**: whole pregnancy, T1, T2, T3. **Mean levels (in μg/m3):** 104.6 for PM_{10}, 45.3 for NO_2, 51.8 for SO_2	**Analysis:** GPR61 gene **Methods:** QMS-PCR. **Sample:** CB**. Effect presented as:** change in DNAm per ??? units increase in prenatal exposure (in brackets p-value)	PM10 and SO2 exposure during pregnancy was associated with ↓GPR61-DNAm: -0.15 (<0.001) and -0.42 (<0.001), while NO_2 exposure with ↑GPR61-DNAm: 0.40 (<0.001)	Passive smoking, maternal height and weight.	Adjustment for maternal GPR61-DNAm did not change the results.
Liu, 2019 China, PEOH cohort N=258	**Pollutants:** PM_{10}, $PM_{2.5}$, $PM_{1.}$ **Method**: LUR. **Windows**: 12th to 20th gestational week (only this period was assessed; see last column). **Mean levels** (in μg/m^3): 49.3 for PM_{10}, 32.9 for $PM_{2.5}$, 28.9 for PM_1	**Analysis**: LINE1-DNAm. **Method**: pyro **Sample**: CB. **Effect estimates presented as**: % LINE1-DNAm change for a 10 μg/m^3 increase in exposure (in brackets 95%CI)	PM_{10}, $PM_{2.5}$ and PM_1 exposure was associated with ↓LINE1-DNAm: -0.51% (-0.91 to -0.11), -0.66% (-1.25 to -0.06) and -0.67% (-1.28 to -0.06), respectively.	Maternal age, parity, pre-pregnancy BMI, adverse pregnancy history, GD, gestational hypertension, paternal smoking newborn's sex, GA	$PM_{2.5}$ exposure between 12th to 20th gestational week was associated with preterm birth. Maternal and fetal LINE1-DNAm might underlie the association with preterm birth
Clemente, 2019 EU countries HELIX project (birth cohorts: BiB, EDEN, INMA, MoBa, Rhea, KANC) N=1237 for NO_2, N=1307 for $PM_{2.5}$	**Pollutants:** $PM_{2.5}$, $NO_{2.}$ **Method: LUR. Windows:** whole pregnancy, T1, T2, T3, 1st year of life. **Mean levels** (in μg/m^3): 15.1 for $PM_{2.5}$, 25.0 for NO_2 for the whole pregnancy	**Analysis:** TL**. Method** qPCR. **Sample:** PB (8yrs). **Effect estimates presented as:** % difference in TL for SD increment in exposure, SD=2.6 for $PM_{2.5}$, SD=13.9 for NO_2. (in brackets 95% CI)	NO_2 exposure during pregnancy was associated with ↓TL: -1.5 (-2.8 to -0.2). NO_2 exposure during the first year of life was associated with ↓TL: −1.6 (−2:9, −0:4). Suggestive evidence for exposure to PM2.5 in the first year of life and ↓TL: −1.4 (−2.9 to 0.1). Similar findings across trimesters were seen.	Maternal age, education, smoking in pregnancy, child's age, sex, ethnicity, BMI, parental smoking at 8y and qPCR batch	$PM_{2.5}$ and NO_2 exposure in the year before the sampling (8y) was also associated with ↓TL
Zhou, 2019 China Zhengzhou cohort N=568	**Pollutants:** PM_{10}, SO_2, $NO_{2.}$ **Method:** fixed site monitoring stations. **Windows**: whole pregnancy, T1, T2, T3. **Mean levels** (in μg/m^3): 106.7 for PM_{10}, 44.2 for NO_2.and 50.1 for SO_2, for whole pregnancy	**Analysis:** SOD2 gene. **Method:** QMS-PCR. **Sample:** CB. **Effect estimates presented as:** % change in SOD2-DNAm for 10μg/m^3 increase in exposure (in brackets 95%CI)	PM_{10} exposure during pregnancy was associated with ↑SOD2-DNAm: 2.69% (1.99 to 4.17), mostly due to exposures in T2. T	Maternal age, education, pre-pregnancy BMI, gestational weight gain, family income, smoking, folic acid intake, newborn's sex, GA, birth season.	The found association was partly mediated by maternal SOD2-DNAm.
Song, 2019 China Wuhan Children's Hospital cohort N=743	**Pollutants:** $PM_{2.5}$, PM_{10}, SO_2, CO, NO. **Method:** LUR. **Window:** T1, T2, T3. **Mean levels:** (in μg/m3): around 76 for $PM_{2.5}$, 140 for PM_{10}, 16.0 for SO_2, 49 NO_2, and 998 for CO	**Analysis:** TL. **Method:** qPCR. **Sample:** CB. **Effect estimates presented as:** change in TL for 10 μg/m3 increase in exposure to $PM_{2.5}$, PM_{10}, SO, and a 100 μg/m3 increase in CO	Exposure to each of the pollutants was associated with ↓TL. The effect estimates were: -3.71% (-6.06 to -1.30), -3.24% (-5.29 to -1.14), -11.07% (-18.86 to -2.53) and -3.67% (-6.27 to, -1.00) for	Maternal age, education, pre-pregnancy BMI, parity, passive smoking[a], GA, GD, hypertensive disorders of pregnancy, newborn's sex,	Associations were stronger in males.

			$PM_{2.5}$, PM_{10}, SO_2, NO_2, and CO, respectively.	birth weight, birth season	
Nie, 2019 China Taiyuan Mother and Child Cohort N=247	**Pollutants**: PAH. **Method**: PAH metabolites in T3 urine. **Window**: T3. **Mean levels** (geometric mean, in ng/mL): 2-OH Nap=2.412, 1-OH Nap=0.647, 2-OH Phe=0.159	**Analysis**: TL. **Method**: qPCR. **Sample**: CB. **Effect estimates presented as**: % change in TL for urinary PAH levels in the 3rd tertile relative to the 1st tertile (in brackets 95%CI)	Maternal PAH urine metabolites associated with ↓TL. 2-OH Nap, 2-OH Flu, 9-OH Phe, and 2-OH Phe concentrations in the 3rd tertile had 69.72% (-115.55 to33.51), 31.06% (-46.31 to -11.49), 23.66% (-40.31 to -2.47) and 30.02% (-45.12 to -10.68) reduction in TL relative to the 1st tertile, respectively.	Maternal age, pre-pregnancy BMI, paternal age, average monthly income, delivery type, newborns' sex, GA, birth length and head circumference	Mediation analysis showed TL could explain 21.74% of the effect of 2-OH Phe on neonatal behavioral neurological assessment score
Gruzieva, 2019 EU countries, USA. PACE consortium (Discovery cohorts: INMA, GenR, CHS, EARLI, PRISM, ENVIRONAGE; Replication cohorts: ALSPAC, MeDALL=BAMSE+PIAMA, HELIX=MoBa, EDEN, KAUNAS, BiB) Discovery N = 1,949 Replication: (1)CB N=688, (2) 7-9yrs N_1=692, N_2=525, N_3=901; (3) 15-16yrs N_1=198, N_2=903	**Pollutants**: $PM_{2.5}$, PM_{10}. **Method**: LUR, SPT, hybrid models. **Window**: entire pregnancy. **Median levels:** Varied across cohorts. $PM_{2.5}$ from 8.1 to 30.6 μg/m^3; PM_{10} from 17.2 to 48.5 μg/m^3	***Meta-analysis:*** EWAS, candidate CpGs previously related with in utero tobacco exposure (total of 6073 CpGs). **Method**: Illumina 450K. **Sample**: CB, NB, PB (for analyses in 6-7yrs, 15-16yrs**). Effect estimates presented as**: increase in average IQR of PM10 (5.6) and PM2.5 (2.0). All presented data passed the FDR threshold of 0.05	**EWAS:** 6 CpGs associated with PM10 and 14 CpGs associated with PM2.5. The absolute value of the estimates ranged from 0.001 to 0.004. The cg06849931 was the only that showed consistent direction in newborns. 3 CpGs cg00905156 (FAM13A), cg06849931 (NOTCH4) and cg06849931 (P4HA2) mapped genes important for respiratory health, were replicated in at least one of the independent cohorts of older children aged 7-9 (although the results were not consistent across all cohorts) Two PM_{10}-related DMRs, H19 and MARCH11, replicated in newborns. **Candidate genes**: no findings	Child´s gender, maternal smoking during pregnancy, cohort-specific batch indicator(s), and ancestry (in CHS).	Enriched pathways: NOTCH signaling pathway, Rho GTPase cycle, Neuro- transmitter ReleaseCycle, GABA synthesis, release, reuptake and degradation
Ladd-Acosta, 2019 USA. EARLI cohort N=163 (CB), N=124 (PT)	**Pollutants:** NO_2, O_3. **Method**: spatial interpolation using IDSW. **Window**: whole pregnancy. **Mean levels** (whole pregnancy, in ppb): 12.3 for NO_2, 37.7 for O_3	**Analyses:** EWAS, global DNAm (mean beta value), candidate CpGs from Gruzieva et al, 2017. **Method:** Illumina 450K. **Sample:** PT and CB. **Effect estimates presented as:** average difference in DMR-DNAm between samples in the highest and lowest quartiles of exposure, reported as a percent (all presented DMR passed the FWER threshold of <0.05)	**EWAS:** 9 DMRs in associated NO_2 and/or O_3. (a) Top findings in CB: RNF39 (2 DMRs. 3.8% and -5.6%), CYP2E1 (-9.3%), PM20D1 (13.9%) (b) Top findings in PT: ZNF442 (-9.1%); PTPRH (-0.09%); SLC25A44 (-1.7%); F11R (-12.6%); STK38 (-0.7%). **Global DNAm:** O_3 associated with ↓DNAm at open sea and/or shelf regions and ↑DNAm in CGI and/or shore regions. NO_2 associated only with ↓DNAm in CGI regions in PT. **Candiate CpGs for replication (Gruzieva, 2017):** cg08973675 (pval 0.06) showed consistent direction, but smaller magnitude of effect.	SVA used to remove unwanted technical and biological sources of variation	5/ 9 DMRs were child-sex specific. The DMRs identified in cord blood also showed consistent direction of effect in the placenta, while DMRs identified in the placenta appear to be tissue-specific.
Rosa, 2019 Mexico PROGRESS cohort N=423	**Pollutant**: $PM_{2.5}$. **Method**: LUR **Window**: T1, T2, T3. **Median levels**: 22.8 μg/m3	**Analyses**: TL. **Method**: qPCR. **Sample**: CB. **Effect estimates presented as:** % change in TL for 10 μg/m3 of $PM_{2.5}$.	Exposure in early pregnancy (4-9 weeks) associated with ↓TL (approximately -0.025% shorter for 10 μg/m3 of PM2.5. Exposure in weeks 14–19 and 34–36	Maternal age, passive smoking, pre-pregnancy BMI, newborn's sex, GA, birth season and batch.	In analyses stratified by sex the association between PM2.5 and shorter

			associated with ↑TL, but with a smaller magnitude.		TL was stronger in girls than in boys.
He, 2018 **China,** **Zhengzhou cohort** **N=527**	**Pollutant:** PM10, NO2, SO2. **Method:** fixed site monitoring stations. **Windows:** whole pregnancy, T1, T2, T3. **Mean levels (in µg/m3):** 104.6 for PM10, 45.3 for NO2, 51.8 for SO2	**Analysis:** H19 gene **Methods:** QMS-PCR. **Sample:** CB**. Effect presented as: change in DNAm per ??? units increase in prenatal exposure**	PM_{10} and SO_2 exposure in pregnancy was associated with ↓DNAm in H19 promoter (-0.16, p-val 0.001) and H19 DMR (-0.18, p-val 0.014), respectively; while NO_2 exposure was associated with ↑DNAm in H19 promoter.	Methylation level of H19/DMR in maternal blood and season of conception.	H19 promoter and DMR DNAm does not mediate the relationship between pollutant exposure and birth size.
Abraham, 2018 France EDEN cohort N=668	**Pollutants:** PM_{10}, NO_2. **Method:** quasi-Gaussian ADMS-Urban. **Window:** whole pregnancy, T1, T2, T3, one month before delivery; one week before delivery; 1, 2, 3 days before delivery. **Mean levels**, whole pregnancy (in µg/m^3): 19 for NO_2 and 20 for PM_{10}	**Analyses:** EWAS, candidate genes specifically expressed or repressed in the placenta, global DNAm. **Method:** Illumina 450K (EWAS, candidate gene, global DNAm-GAMP analysis), pyro (LINE1, Alu). **Sample:** PT. **Effect estimates presented as:** change in DNAm per 10 units increase in prenatal exposure (all presented results passed the FDR threshold of 0.05)	**Top finding from the candidate genes approach and EWAS:** Increase in NO_2 in T2 was associated with -0.004 and -0.005 ↓DNAm of CpGs in the ADORA2B gene, linked with hypoxia and pre-eclampsia **DMR analysis**: 27 DMRs associated with air pollution exposure to PM_{10} or NO_2. **Global DNAm.** PM_{10} 1d before birth ↑Alu DNAm, suggestive results for exposure in T1 and ↓LINE1 and Alu-DNAm,	Maternal age at delivery and end of education, parity, pre-pregnancy BMI, smoking, season of conception, study centre, GA, newborn's sex, technical factors (batch, plate and chip) and estimated cell-type proportions	Temperature exposure during T1 was also associated global DNAm. Humidity was associated with DNAm of 2 CpGs:cg16917193 and cg16075020
Maghbooli, 2018 Iran Tehran University cohort N=92	**Pollutants:** $PM_{2.5}$, PM_{10}. **Method:** fixed site monitoring stations. **Window:** whole pregnancy, T1, T2, T3. **Mean levels** (whole pregnancy, in µg/m^3): 37.1 for $PM_{2.5}$, 91.5 for PM_{10}.	**Analyses:** Global DNAm **Method:** HPLC. **Sample:** PT. **Effect estimates presented as:** Spearman correlation coefficients ($PM_{2.5}$/PM_{10} and global DNAm)	Exposure in T1 positively correlated with global DNAm: $PM_{2.5}$ 0.26 ($p<0.01$) and PM_{10} 0.38 ($p<.0001$)	Only results from univariate analysis were shown	Mild negative correlation between global DNAm and gene expression of SAMe and DNMT-1α
Nawrot, 2018 Belgium **ENVIRONAGE** cohort N=407	**Pollutant**: $PM_{2.5}$. **Method**: STI method in combination with dispersion model. **Window**: whole pregnancy, T1, T2, T3 last month of pregnancy. **Median levels**: 16.4µg/m^3 for whole pregnancy	**Analysis**: 8 circadian pathway genes: CLOCK, NPAS2, BMAL1, CRY1, CRY2, PER1, PER2, PER3). **Method**: PYR. **Sample**: PT. **Effect estimates presented as:** log-fold change in methylation for an IQR increment in $PM_{2.5}$ (in brackets 95%CI)	Exposure in T1 and T2 was not associated with DNAm of the candidate genes, except for CLOCK gene in T1 (−0.59; −0.93 to −0.25). Exposure in T3 was associated with ↑DNAm in the NPAS2 gene (0.16; 0.06 to 0.27), CRY1 gene (0.59; 0.22 to 0.95), PER2 gene (0.36; 0.16 to 0.57), PER3 gene (0.42; 0.18 to 0.67) and ↓DNAm in PER1 gene (−0.51, −0.90 to −0.13). Similar results for the last month of pregnancy.	Maternal age, education, parity, ethnicity, smoking status, prepregnancy BMI, newborn's sex, GA, season at conception, cord blood vitamin D, hour or delivery, time window-specific apparent temperature	Multi-gene model to combine the methylation status of all genes in the Circadian pathway: $PM_{2.5}$ in T1 was associated with ↓DNAm, and exposure in T3 with ↑DNAm.
Plusquin, 2018 EU countries, USA. ALSPAC, ENVIRONAGE, INMA, Piccoli-più, Rhea cohorts N=1235	**Pollutant**: PM_{10}. **Method**: Dispersion modeling; LUR, STI method. **Window**: entire pregnancy. **Mean levels** (in µg/m^3): ranged from 17.7 to 59.2	**Meta-analysis:** EWAS, candidate CpGs (6246 linked with smoking in pregnancy and 25 CpGs linked with NO_2 exposure in the study by Gruzieva, 2017). **Method:** Illumina 450K. **Sample**: CB.	EWAS: no CpG passed the FDR threshold. Candidate CpGs: no CpG passed the FDR threshold (3 CpGs previously associated with smoking, cg07571337, cg26500033 and cg10704395, had the same direction as in the original study (p_{adj}=0.24)	Technical variables, newborn's sex, maternal smoking during pregnancy, and estimated blood cell composition	Longitudinal analysis at 3 time points (pregnancy, 7 and 15−17 years): PM_{10} associated with one CpG (cg21785536), enriched pathways GABA-ergic synapse, p53 and NOTCH1 pathway
Lee, 2018	**Pollutant**: $PM_{2.5}$. **Method**: hybrid	**Analysis**: GSTP1 gene. **Method**: pyro.	$PM_{2.5}$ in late pregnancy (>37 gestational	Maternal age, education,	↑GSTP1-DNAm

USA ACCESS cohort N=131	spatial temporal method **Window**: whole pregnancy, by gestational weeks. **Median levels**: 11.0μg/m^3	**Sample**: NE (at age 7).	weeks) associated with ↑GSTP1-DNAm (data shown only in figure)	child's age at spirometry test, sex, and race/ethnicity	associated with FEV1 at age 7. The results were more evident in boys.
Perera, 2018 China Two Tongliang cohorts: before and after closure of a power plant N=225	**Pollutant**: PAH. **Method**: PAH-DNA adducts in CB. **Window**: Pregnancy. **Mean levels** (in 10^{-8} nucleotides): 0.33 and 0.20 for the first and second cohort, respectively	**Analysis**: TL. **Method**: qPCR. **Sample**: CB. **Effect estimates presented as**: change in TL per SD increase in PAH-DNA cord adduct: SD for the first and second cohort was 0.07 and 0.13, respectively (in brackets 95% CI)	Increase in PAH-DNA adduct was associated with ↓TL: −0.019 (-0.032 to -0.006)	Maternal age, education, newborn's sex, GA and cord blood mercury	TL also associates with (ln) BDNF (95% CI): 0.167 (0.092 to 0.242)
Neven, 2018 Belgium ENVIRONAGE cohort N=463	**Pollutants**: $PM_{2.5}$, BC, NO_2. **Method**: STI in combination with dispersion model. **Window**: whole pregnancy, T1, T2, T3. **Median levels** (in μg/m^3): 13.6 for $PM_{2.5}$, 0.9 for BC and 18.6 for NO_2, for whole pregnancy.	**Analysis**: candidate genes (7 DNA repair and tumor suppressor genes: APEX1, OGG1, PARP1, ERCC1, ERCC4, p53, DAPK1). **Method**: PYR. **Sample**: PT. **Effect estimates presented as**: relative % change in DNAm for an IQR increment in exposure (in bracket 95% CI)	$PM_{2.5}$ associated with changes in methylation of several genes. ↑APEX1: 9.0% (0.5 to 17.5), ↑OGG1: 13.9% (22.6 to 5.1), ↑ERCC4: 16.3% (27.2 to 5.4), ↑p53: 10.6% (16.7 to 4.5) and ↓DAPK1: -12.9% (-3.5 to -22.4). BC associated with ↑APEX1: 9.2% (4.1 to 14.2), ↑ERCC4 27.6% (17.6 to 37.6). NO_2: no associations,	Maternal age, education, parity, ethnicity, smoking status, prepregnancy BMI, newborn's sex, GA, season at conception and batch effect	Association between $PM_{2.5}$ with an ↑Alu mutation rate
Yang, 2018 China Tongji Medical College Cohort N=106	**Pollutant:** PAH. **Method:** PAH metabolites in T3 urine. **Window**: T3. **Median levels** (mg/g creatinine): 6.7 for 2-OHNa, 10.1 for 1-OHPh	**Analyses:** LINE1, Alu DNAm. **Method:** pyro **Sample:** CB. **Effect estimates presented as:** % change in DNAm for urinary PAH levels in the 3rd tertile relative to the 1st tertile (in brackets 95%CI)	2-OHNa and 1-OHPh associated with ↓Alu-DNAm: -1.88% (-3.73 to -0.10%) and -2.57% (-4.30 to -0.80%), respectively. Similar results for LINE1 (data not shown).	Maternal age, pre-pregnancy BMI, gestational weight gain, parity, education, household income, passive smoking[a], newborn's sex[a]	PAH exposure also associated with birth length. DNAm was not a mediator.
Cai, 2017 China Wenzhou Medical College cohort N=181	**Pollutant:** PM_{10}. **Method**: fixed site monitoring stations. **Windows**: whole pregnancy, T1, T2, T3. **Median levels** (in μg/m^3): 62.7, 68.1, 65.3, 63.5 for T1, T2, T3 and whole pregnancy, respectively.	**Analysis**: LINE1-DNAm, candidate genes DNAm (HSD11B2, NR3C1). **Method**: PYR. **Sample**: PT. **Effect estimates presented as**: relative difference in DNAm for 10 μg/m^3 increase in PM_{10} (in brackets 95% CI)	PM_{10} exposure in T1 associated with ↓LINE1-DNAm in FGR newborns: −1.78 (−3.35 to −0. 22%). PM_{10} in T2 associated with ↑HSD11B2-DNAm in both FGR cases and controls: 1.42 (0.24 to 2.57). NR3C1: no associations	Maternal age, education, pre-pregnancy BMI, passive smoking[a], delivery mode, newborn's sex, GA, day and month of birth, NO_2, SO_2. Models by trimester also adjusted for season and PM_{10} in other trimesters.	CpGs within LINE1 and within HSD11B2, were highly correlated.
Martens, 2017 Belgium ENVIRONAGE cohort N=641	**Pollutant**: $PM_{2.5}$. **Method**: SPT method in combination with dispersion model. **Window**: whole pregnancy, by gestational week. **Mean weekly levels:** 13.4 μg/m3	**Analysis**: TL. **Method**: qPCR **Sample**: CB and PT. **Effect estimates presented as**: % change in TL per 5-μg/m3 increment in $PM_{2.5}$ (in brackets 95% CI)	Exposure during pregnancy was associated with ↓TL: -8.8% (-14.1 to −3.1%) and -13.2% (−19.3 to −6.7%) for CB and PT, respectively. Strongest association in the mid-gestation period (weeks 12-25 for cord blood and weeks 15-27 for placenta). Exposure in late pregnancy (from weeks 32-34) associated with longer telomeres in cord blood	Maternal age, BMI, smoking status, education, parity, pregnancy complications, paternal age, newborn's sex, GA, ethnicity, date of delivery, season of delivery and ambient temperature.	No modification by newborn sex for cord and placental TL
Saenen, 2017 Belgium ENVIRONAGE cohort	**Pollutant**: $PM_{2.5}$. **Method**: STI in combination with dispersion model. **Window**: T1, T2, T3. **Median level** (in μg/m^3): 13.9, 14.6 and 16.9 for	**Analysis**: LEP gene. **Method**: pyro. **Sample**: PT. **Effect estimates presented as**: % change in DNAm for a IQR increment (7.5 μg/m^3) in $PM_{2.5}$ (in	$PM_{2.5}$ exposure in T2 was associated with ↓LEP-DNAm: -1.4% (–2.7 to –0.19%)	Maternal age, education, parity, ethnicity, smoking status, prepregnancy BMI, newborn's sex, GA, season	LEP DNAm also associated with 3-NTp content, marker of oxidative

N=361	T1, T2 and T3, respectively.	brackets 95% CI)		at conception	stress.
Gruzieva, 2017 EU countries, USA. PACE consortium (Discovery cohort: MeDALL pooled= EDEN+INMA, Gen. R, CHS, MoBA; Replication cohorts: MeDALL pooled: EDEN+INMA+ BAMSE+PIAMA) Discovery N=1508 Replication (1) 4yrs N=733 (2)8yrs N_1=444, N_2=342	**Pollutant**: NO_2 **Method**: LUR, spatial interpolation using IDSW **Window**: entire pregnancy. **Median levels**: Varied across cohorts. Ranged from 10.3 to 38.7 μg/m³	**Analysis**: EWAS, candidate genes (38 antioxidant and inflammatory genes). **Method**: Illumina 450K (EWAS and candidate gene). **Sample**: CB, PB (4-5y, 8y). **Effect estimates in brackets presented as:** change in DNAm per 10 $\mu g/m^3$ of increase in NO_2. All presented CpGs passed the FDR threshold > 0.05.	**EWAS:** In CB, NO^2 was associated with 3 CpGs: cg12283362 (-0.007), cg24172570 (-0.004) and cg08973675 (0.005) in the LONP1, HIBADH and SLC25A28 genes, respectively, all involved in mitochondrial functions. The CpG cg08973675 was also associated with NO2 in 4 and 8year olds. **Candidate-genes:** NO2 exposure was associated with ↑DNAm of the CAT gene (0.003 and 0.002 for cg03728580 and cg17034036, respectively) and ↓DNAm of the TPO gene (-0.003 for cg01385533).	Newborn's sex, maternal smoking during pregnancy, cohort-specific batch indicator(s), cohort indicator (in the pooled MeDALL sample set), and ancestry (in CHS).	Enriched pathways: pathways related to negative regulation of cellular process, and biological process, integrin-linked kinase signaling. NO2 exposure at the time of biosampling in childhood had a impact on CAT and TPO expression.
Lee, 2016 China Two Tongliang cohorts: before and after closure of a power plant N=217	**Pollutant**: PAH. **Method**: PAH-adducts in CB. **Window**: pregnancy. **Mean levels** (in adducts/10^8 nucleotides): 0.33 and 20 for the first and second cohort, respectively	**Analysis:** LINE1-DNAm. **Method:** pyro. **Sample**: CB. **Effect estimates presented as:** change in LINE1-DNAm for unit increase in lnPAH-DNA adducts (in brackets 95%CI)	PAH cord adducts associated with ↓LINE1-DNAm : -0.009 (-0.019 to -0.000)	Maternal age, education, household income, newborn's sex [a]	LINE1-DNAm associated with IQ WISCscores. but was not a mediator between PAH-adducts and IQ.
Breton, 2016 USA CHS cohort N=392 (LINE1), N=181 (AluYb8)	**Pollutants**: $PM_{2.5}$, PM_{10}, NO_2, O_3 **Method**: Spatial interpolation using IDSW. **Windows**: T1, T2 and T3. **Median levels:** around 25 $\mu g/m^3$ for $PM_{2.5}$, 40 $\mu g/m^3$ for PM_{10}, 30 ppb for NO_2 and 40 ppb for O_3	**Analyses**: LINE1, AluYb8-DNAm. **Method**: Pyrosequencing. **Sample**: newborn blood spot. **Effect estimates presented as:** change in DNAm for 2 SD change in exposure (in brackets 95% CI)	PM_{10} in T1 associated with ↓LINE1-DNAm: -0.66% (–1.22 to –0.09). O_3 in T1 associated with ↓LINE1-DNAm: -0.86% (-1.42 to -0.30). O_3 in T3 associated with ↑LINE1-DNAm: 0.60% (0.01 to 1.19). AluYb8: no associations	Newborn's sex, admixture, plate, smoking and maternal education level	Genotype of methylatransferase genes modified the association between T1 air pollutants and LINE1-DNAm
Goodrich, 2016 South Africa MACE cohort N=22	**Pollutant**: NOx **Method**: LUR. **Window**: whole pregnancy**. Mean levels.** 26.5 ppb	**Analysis**: EWAS. **Method**: Illumina 450K. **Sample**: CB	No single CpGs passed the FDR threshold.	Newborn's sex and GA	Enriched pathways: pathways related to xenobiotic metabolism, oxygen and gas transport, and sensory perception of chemical stimuli
Breton, 2016 USA CHS cohort N=240 for PM10, N=185 for PM2.5 Replication cohort: N=280 for PM10, N=149 for PM2.5	**Pollutants**: $PM_{2.5}$, PM_{10}. **Method**: spatially interpolation using IDSW. **Window**: T1, T2 and T3. **Median levels:** around 25 $\mu g/m^3$ for $PM_{2.5}$, 40 $\mu g/m^3$ for PM_{10}	**Analyses**: EWAS restricted to promoter regions. **Method**: Illumina 450K + pyro for replication. **Sample**: NB. Effect estimates presented as: change in DNAm for a 2SD increment in exposure (2SD for PM_{10} exposure were 32.4 and 31.6, respectively)	31 CpGs associated with PM_{10} or $PM_{2.5}$ for exposures mostly in T2 and T3. PM_{10} was associated with ↑DNAm of cg03579365 (COLEC11) accross trimesters, with strongest association seen in T3 (β=0.14) and replicated in an independent sample from the same cohort, showing consistent direction only in T1 (β=0.08).	Adjusted for gender, plate, and cell types [a]	Some of the 31 loci associated with PM2.5 and PM10 were found to be associated with cardio-respiratory outcomes in childhood
Janssen, 2015	**Pollutant**: $PM_{2.5}$. **Method**: STI	**Analysis**: candidate regions in mtDNAm:	$PM_{2.5}$ during the whole pregnancy was	Maternal age, education,	MtDNAm mediated

Belgium ENVIRONAGE cohort N=381	method in combination with dispersion model. **Window**: whole pregnancy, T1, T2, T3. **Mean levels**: 16.7µg/m3	D-loop and MT-RNR1. **Method**: pyro. **Sample:** PT. **Effect estimates presented as**: absolute change in % mtDNAm for an IQR increment in $PM_{2.5}$ exposure (in brackets 95%CI). IGR for T1 was 7.8µg/m^3.	associated with ↑mtDNAm in both the MT-RNR1 region and D-loop. The associations were strongest for T1: 1.27% (0.23 to 2.32%) and 0.44% (0.12 to 0.75%) for MT-RNR1 and D-loop, respectively.	parity, ethnicity, smoking status, newborn's sex, GA, season at conception	54% (MT-RNR1) and 27% (D-loop) of the inverse association between $PM_{2.5}$ and mtDNA content.
Janssen, 2013 Belgium ENVIRONAGE cohort N=240	**Pollutant**: $PM_{2.5}$. **Method**: STI method. **Window**: whole pregnancy, T1, T2, T3, pre-implantation, implantation, implantation range and post-implantation. **Mean levels:** 17.4 µg/m³ for the whole pregnancy	**Analysis**: Global DNAm. **Method**: LC/MS-MS. **Sample**: PT. **Effect estimates presented as**: absolute change in percentage of global DNAm for 5µg/m3 increase in $PM_{2.5}$	$PM_{2.5}$ associated with ↓global DNAm: -2.19% (-3.65 to -0.73), mostly due to exposures in T1: −2.41% (-3.62 to -1.20) and in the implantation range: −1.08% (-1.80 to -0.36)	Maternal age, education, smoking status and parity, newborn's sex, GA, season at conception, trimester-specific apparent temperature	No effect modification by infant's sex or birth weight.
Tang, 2012 USA CCCEH cohort N=53	**Pollutant:** PAH. **Method:** personal air monitors in T3. **Window**: T3. **Median levels**: 2.3 ng/m^3	**Analysis**: candidate genes (IFNγ and IL4). **Method**: bisulfite sequencing **Sample**: CB. **Effect estimates presented as:** regression coefficient between lnPAH and lnPercent DNAm	IFNγ region 1: PAH was associated with ↑DNAm, 1.07 (0.46 to 1.68). IFNγ region 2: DNAm ↑ until PAH reached 3.5 ng/m^3 and then decreased: 1.15 (0.54 to 1.76) and -0.57 (-1.12 to -0.02) for the linear and nonlinear spline model, respectively.	Newborn's sex, ethnicity, environmental tobacco smoke [a], maternal age, receipt of public assistance	Cell line experiments*: In vitro* PAH exposure ↑DNAm of IFNγ promoter and ↓expression
Herbstman, 2012 USA CCCEH cohort N=164 (N=87 for urinary PAH metabolites)	**Pollutant:** PAH. **Method:** personal air monitors in T3, PAH metabolites in T3 urine. **Window:** T3. **Median levels:** PAH in air (in ng/m^3) 2.5 for total PAH, 2.6 for pyrene and 0.2 for BaP. PAH metabolites in urine (in ng/L urine): 1870.2 for 1-OHPh and 154.1 for 2-OHNa	**Analysis**: Global DNAm. **Method**: ELISA-based **Sample**: CB. **Effect estimates presented as:** regression coefficient between lnPAH and lnPercent DNAm (in brackets 95%CI)	Total PAH, pyrene and BaP in air associated with ↓global DNAm: -0.11 (0.21 to 0.00); -0.18 (-0.32 to -0.04); -0.09 (-0.18 to 0.00). PAH metabolites showed slightly positive association with global DNAm (p values > 0.1)	Maternal age, ethnicity, marital status, education, household income, parity, newborn's sex	Global DNAm were positively associated with presence of detectable PAH DNA adducts in CB
Perera, 2009 USA CCCEH cohort Discovery N=20 Validation: N=56 (includes the previous 20 subjects)	**Pollutant**: PAH. **Method**: personal air monitors in T3. **Window**: T3. **Median level:** 2.3 ng/m^3.	**Analysis**: EWAS. **Method**: MSRF + bisulfite GS and QMS-PCR for the validation. **Sample**: CB, PT	Identified 31 candidates with different methylation according to dichotomized PAH exposure. 6 candidate genes (ACSL3, RAD 21, DUSP22, SCD5, SFMBT2 and WWOX) further validated with bisulfite GS. Top finding: association between ACSL3-DNAm and PAH confirmed by MS-PCR: OR= 13.8 (3.8, 50.2) (2.41 ng/m3 cut point value identified with ROC analysis)		Association between ACSL3 and asthma: OR = 3.9 (1.1, 14.3) odds of asthma given ACSL3 was methylated vs unmethylated

Abbreviations: T1, T2 and T3: first, second and third trimester, respectively; CB: cord blood; PT: placental tissue; EWAS: Epigenome-Wide Association Study, TL: telomere length; SE: Standard Error; CI: Confidence Intervals; GA: Gestational Age; FGR: Fetal Growth Restriction; MSRF: Methylation Sensitive Restriction Fingerprinting; DMR: Differentially Methylated Regions; FWER: Family-Wise Error Rate; FDR: False Discovery Rate, PAH: Polycystic Aromatic Hydrocarbons; GAMP: Global Analysis of Methylation Profiles; ppb: part per billion (equals to 1 µg/L); CGI: CpG Island; IQR: Inter Quintile Range; SVA: Surrogate Variable Analysis; IDSW: Inverse Distance Squared Weighting; ADMS: Atmospheric Dispersion Modeling System; STI: Spatial Temporal Interpolation; mtDNAm: mitochondrial DNA methylation. LUR: Land Use Regression; MSP: Quantitative Methylation-Specific-Polymerase Chain Reaction; pyro: pyrosequencing; Illumia450K: Illumina's Infinium 450K Bead Chip; ELISA: enzyme linked immunosorbent assay; qPCR: Quantitative Polymerase Chain Reaction; HPLC: High-Pressure Liquid Chromatography

[a] Only non-smoking women were included in the study

Study II Prenatal exposure to PM_{10} and changes in DNA methylation and telomere length in cord blood

[Manuscript in preparation]

Isaevska E[1], Fiano V[1], Asta F[2], Stafoggia M[2,3], Popovic M[1], Pizzi C[1], Morena Trevisan[1], Laura De Marco[1], Polidoro S[4,5], Gagliardi L[6], Rusconi F[7], Brescianini S[8], Nistico L[8], Stazi MA[8], Ronfani L[9], Porta D[2], Richiardi L[1]

1. Department of Medical Sciences, University of Turin, CPO Piemonte, Turin, Italy
2. Department of Epidemiology, Lazio Regional Health Service, ASL Roma 1, Rome, Italy
3. Karolinska Institutet, Institute of Environmental Medicine, Stockholm, Sweden
4. Italian Institute for Genomic Medicine (IIGM), Candiolo, Italy.
5. MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College, London, UK.
6. Division of Neonatology and Pediatrics, Ospedale Versilia, Viareggio, AUSL Toscana Nord Ovest, Pisa, Italy
7. Unit of Epidemiology, Meyer Children's University Hospital, Florence, Italy
8. Center "Behavioral Sciences and Mental Health", Istituto Superiore di Sanità, Rome, Italy.
9. Clinical Epidemiology and Public Health Research Unit, Institute for Maternal and Child Health - IRCCS "Burlo Garofolo", Trieste, Italy.

Funding: The project was realized with the financial support of the Italian Ministry of Health – CCM. / Progetto realizzato con il supporto finanziario del Ministero della Salute- CCM. The Department of Medical Sciences, University of Turin, is funded by the Italian Ministry for Education, University and Research (Ministero dell'Istru-zione, dell'Università e della Ricerca- MIUR) under the programme "Dipartimenti di Eccellenza 2018-2022"

Abstract

Background: Gestational air pollution exposure can cause molecular level alterations, including changes in DNA methylation and telomere length shortening, that can survive cell divisions and influence later disease susceptibility. We aimed to investigate changes in two molecular markers in cord blood in relation to gestational PM_{10} exposure and explore potential gestational windows of susceptibility.

Methods: We measured cord blood epigenome-wide DNA methylation (N=384) and telomere length (N=500) in children of the Italian birth cohort Piccoli Più using the Illumina's Infinium Methylation EPIC BeadChip and qPCR, respectively. PM_{10} exposure levels were estimated for different gestational periods based on maternal residential address, using advanced methods based on satellite data with particular interest in exposures in the beginning of the pregnancy.

Results: The level of PM_{10} exposure during the implantation and post-implantation period was associated with variation in DNA methylation of more than 100 unique CpGs, mapped to genes with relevant functions in cell replication, differentiation and response to environmental stressors. We were able to test half of the CpGs in an independent subsample from the same cohort and Six CpGs showed robust associations. There was also suggestive evidence that PM_{10} exposure in early pregnancy is related to shorter offspring cord blood telomeres. Second trimester PM_{10} exposure was associated with CpG island hypomethylation.

Conclusions: The beginning of the pregnancy seems to be a particularly important PM_{10} exposure window that should be accounted for in future analyses. DNA methylation and telomere length are fundamental regulators of cellular processes during early life, and their alterations could have major implications for later disease susceptibility.

Introduction

Environmental exposures during pregnancy can cause changes at molecular level that can survive cell divisions and potentially influence the cellular regulation of many biological and genetic processes.[1–4] Several studies have reported an association between exposure to ambient air pollutants during pregnancy and global[5–11] or locus-specific changes in DNA methylation[12–16] as well as telomere length shortening in newborns.[17–22] These findings are of relevance considering today's ubiquitous presence of ambient air pollutants and the potential implications for health across the lifespan after adverse exposures during intrauterine life.[23]

DNA methylation is an epigenetic control mechanism essential for normal cell function. During early gestation DNA methylation coordinates cell differentiation, sex chromosome dosage compensation, expression of imprinted genes and repression of retrotransposons that might influence genome stability.[24,25] Telomeres, on the other hand, are repetitive nucleotide sequences at the end of chromosomes that ensure complete chromosomal replication and progressively shorten with each cell division.[24] Short telomeres are considered a risk factor for age-related diseases and, since they are sensitive to oxidative stress, a potential biomarker for cumulative exposure.[26]

The early pregnancy period is of particular interest because during this time, methylation patterns undergo dramatic changes to establish the baseline methylation patterns that will largely maintain intact throughout life.[24] However, most studies on global[5,6,8,9] or epigenome-wide methylation patterns[12–16,27] either analyzed trimester specific exposure or averaged the pollutant concentrations across the entire pregnancy. Only one study analyzed smaller windows of exposure and reported that air pollution exposure during the implantation period was strongly associated with global methylation loss.[11]

Telomere length at birth is shaped by genetics, environmental exposures, number of cell divisions (which in turn are related to growth and gestational age), and the counter-regulatory effect of telomerase. Similar to DNA methylation, the initial *in utero* setting of the telomere system is particularly important for its long-term effects.[2] Previous studies on telomere length that analyzed trimester-specific[18,19] or weekly[17,21] exposures during gestation, showed overall inconclusive findings regarding potential susceptibility windows.

Our objective was to study overall and locus-specific changes in umbilical cord blood DNA methylation and telomere length in relation to maternal exposure to PM_{10} (particulate matter < 10μm) in specific time widows during pregnancy, with particular focus on the beginning of

the pregnancy. By studying two molecular markers sensitive to environmental exposures, we aimed to capture different cellular responses to such stressors that might act as indicators for past exposure and/or predictors for future disease risk trajectories.

Methods

1. Study population

Piccolipiù is a multicentric Italian birth cohort set up to investigate the effects of exposures acting during pre-natal and early post-natal life on infant and child health, describe complex interactions between genetic, epigenetic, lifestyle and environmental factors and promote infant and child health.[28] Between 2011 and 2015, 3,358 mother-child pairs were recruited in five Italian centers (Turin, Trieste, Florence, Viareggio and Rome). Pregnant women with singleton pregnancy were eligible for inclusion if they were at least 18 years old, were scheduled to give birth in one of the selected units participating in the study, had residence in the catchment area of the maternity center, ability to fill out the informed consent and the questionnaire in Italian and had a telephone number to be reached at. Ethical approvals were obtained from the Ethics committees of the Local Health Unit Roma E (management center), of the Istituto Superiore di Sanità (National Institute of Public Health) and of each local center. All parents provided written informed consent. Information on parents and children was collected using questionnaires and medical examinations at birth and at the follow up visits. At recruitment, women completed a baseline questionnaire including information on several sociodemographic, lifestyle, environmental and medical factors. Follow-up information through questionnaires is also collected at 6, 12 months and when children turned 2, 4 and 6 years after delivery. For each newborn, whole blood was withdrawn from cord vessels and collected in a tube with EDTA, fractionated in buffy coat, plasma and erythrocytes and stored in a bio-bank at -80 °C.

2. Air pollution exposure assessment

Daily PM_{10} concentrations were estimated at 1-km^2 grid using the Random Forest (RF) method, as described in details elsewhere.[29] Briefly, the RF method combined data from PM monitoring sites and satellites with spatial and spatio-temporal predictors.

Daily data on 24-hour mean PM_{10} concentrations over the period 2006–2015 from all the available monitoring sites was provided by the Italian Institute for Environmental Protection and Research (ISPRA). Daily worldwide Aerosol Optical Depth (AOD) data, were provided by the Moderate-Resolution Imaging Spectro-radiometer, positioned on NASA satellites.

Missing satellite AOD data were imputed using ensemble atmospheric models to obtain PM_{10} predictions. The models showed good predictive performance (R^2~0.95) with small prediction errors (root mean squared prediction error ~0.02) and negligible bias (intercepts = 0 and slopes ~ 1). Predictions were equally good in capturing annual and daily PM_{10} variability.

We used daily data to obtain reliable exposure estimates within different gestational windows of exposure. We obtained the mean PM_{10} exposure levels for eight gestational window of exposure: a) pre-implantation (days 1-5 since conception), b) implantation (6-12 days after conception), c) post-implantation (22-28 days after conception), d) first trimester 1 (1-13 gestational week) e) second trimester (14-26 gestational week), f) third trimester (27 gestational week- delivery), g) last month of pregnancy (30 days before delivery) g) whole pregnancy.

3. Molecular analyses

3.1. Selection of samples for molecular analyses

The DNA methylation analyses included a set of 384 children of the Piccolipiù cohort, that were selected from each center based on the following criteria: pre-extracted DNA with sufficient quality and quantity (DNA was extracted using the kit QIAsymphony DSP DNA, Qiagen, Hilden, Germany), mother with georeferenced address at enrollment, European origin (defined as having 4 grandparents born in Europe) and a complete follow up of at least 24 months. Out of samples that satisfied the criteria, 96 samples were from the Florence center, 96 from the Viareggio center, 96 from the Trieste center, 75 samples from the Turin center and 21 samples from the Rome center. The smaller number of samples from the Rome center was due to smaller number of samples with extracted DNA at the time when methylation analyses were conducted.

A total of 500 children were included in the telomere length analysis. All samples included in the DNA methylation analyses were also included in the telomere length analysis. Samples from each center, based on the previously mentioned criteria were selected to top up the number of children included in the analyses to hundred from each center. Therefore, four additional samples were selected for Florence, Viareggio and Trieste centers, 25 from the Turin center and 79 samples from the Rome center. The 25 samples from Turin had DNA extracted within the framework of the EXPOsOMICS project[30] by using the QIAamp 96 DNA Blood Kit (Qiagen, Hilden, Germany).

3.2. DNA methylation analyses

DNA methylation analyses started by using the Zymo EZ-96 DNA Methylation™ Kit (Zymo, Irvine, CA) for bisulfite conversion. Then we used Illumina's Infinium MethylationEPIC BeadChip Kit (Illumina Inc., San Diego, CA, USA), (IlluminaEPIC) to assess DNA methylation levels.[31] The array targets over 850,000 methylation sites across the genome at single-nucleotide resolution, accounting for roughly 4% of all CpG content in the genome, covering gene promoters of the majority of known genes, transcription start sites, untranslated regions and regulatory elements, such as FANTOM5 enhancers and ENCODE regulatory elements.

The InfiniumMethylationEPIC BeadChips used in this study accommodate 8 samples which are organized in one column. Our 384 samples fitted into 4 plates, each with a grid containing 12 columns and 8 rows. Therefore, in order to ensemble one plate, 12 Illumina's EPIC BeadChips were ordered one next to another, to be analyzed together in one experimental batch. To control for possible batch effect[32], we randomized the samples by study center and newborn's sex. In order to balance the number of subjects from each center and sex within a chip, we first stratified the samples by sex and study center then randomly distributed them into the 48 chips. Then the 48 chip were randomly distributed in 4 plates. And finally the position of the sample within each chip was randomized.

The Illumina's software BeadArray Controls Reporter was used to run the initial assessment of the raw signal quality. Ten samples were illegible and therefore excluded from further analysis. The *minfi* package[33] was used to import the data from IDAT files, the proprietary format outputted by the scanner, into the R software for further quality control and processing. Detection p-value was calculated for each of the 865869 probes for each sample by comparing the total signal for each probe to the background signal level, estimated from the negative control probes. Very small p-values are indicative of a reliable probe signal while large p-values indicate a poor signal quality. Only one sample had mean detection value > 0.01, was flagged as low-quality and was excluded from further analysis.

To minimize the unwanted variation within and between samples, we normalized the data using the quintile normalization. Then we filtered out probes that failed in one or more samples (n=10014, 1.16%), probes on the sex chromosomes (n=18739), probes with SNPs at CpG site (n=27690) and cross reactive probes i.e probes that have been demonstrated to map to multiple places in the genome (n=39288),[34,35] leaving a total of 770128 CpGs for the

analysis. Before filtering probes on sex chromosomes, we estimated the sex based on the median values of measurements on the X and Y chromosomes. Three samples were found to be discordant by the reported sex (males) and predicted sex (females) and were excluded, leaving total of 370 samples for further analysis. Methylation levels were expressed as beta values, ranging from 0 to 1, that represent the ratio of the intensity of the methylated-probe signal to the total locus signal intensity, methylated and unmethylated.

Many studies have demonstrated the importance of adjusting for cell type heterogeneity in epigenome-wide studies. In this study, for cell type deconvulsion we used the filtered reference dataset "FlowSorted.CordBloodCombined.450k", that combined and optimized four previously used reference datasets for cord blood cell type deconvulsion and selected a set of 517 CpGs present on both the IlluminaEPIC array and the Illumina's Infinium Methylation450K BeadChip (Illumina 450K) that emerged as the ideal set of probes for cell type estimation in cord blood.[36] We estimated cord blood proportions of 7 cell types: CD8 T cells, CD4 T cells, Natural Killer cells, B cells, monocytes, granulocytes and nucleated red blood cells, by using the function *estimateCellCounts2* from the *FlowSorted.Blood.EPIC* R-package[37] according to the authors' instructions.[36]

3.3. Telomere length

We analyzed cord blood telomere length in 500 children. The average relative telomere length was measured using the monochrome multiplex quantitative PCR (qPCR) method developed by Cawthon.[38] Each sample was run in triplicate on Rotor-Gene Q (Qiagen, Hilden, Germany) with 72 well rotor. The mastermix used was iQ Syber Green supermix (Bio-Rad). All samples were randomized by sex and study center into 27 batches, each containing the triplicates of up to 19 samples. There was no need to randomize by row and columns since the Rotor-Gene plate is circular and rotates during the analyses, thus eliminating the plate- position effect.

On each run, a five-point serial dilution of pooled DNA was run to assess the PCR efficiency and same pooled DNA was used as inter run calibrator. We assessed telomere length by each child by calculating the mean relative T/S ratio using the method $2^{-\Delta\Delta Ct}$, that represents the ratio of telomere repeat copy number (T) to single copy gene (S) copy number, relative to average T/S ratio of the entire sample set.

Before starting the analysis, a protocol, was developed to establish the reproducibility of the technique. Ten samples (male and female from each study center) were randomly chosen. In

order to account for the plate-to-plate and day-to-day variation of the assay, the samples were run in triplicates on two separate occasions several days apart. The 6 data points per sample were used to assess the intra and inter-assay variation and calculate the intra and inter assay coefficient of variation (CV). The relative T/S ratio for each sample was expressed as relative to the mean ΔCt of the entire dataset from the first and second run, accordingly. The geometric mean of the intra-assay CV of the telomere Ct's, albumin Ct's and rT/S from the first run was 0.71%, 0.64% and 6.98%, respectively. The geometric mean of the inter-assay CV was 1.48%. The scatter plot and the Bland-Altman plot plotting the agreement between two runs can be seen in Supplementary Figure S1.

Telomere length is known to be highly influenced by genetics. Genotype data were available for all children included in the telomere length analysis. Genotype analysis was performed using an Illumina assay technology and a custom made panel by Genomix4Life™ containing 730059 SNPs. We selected 16 SNPs that were known to be strongly associated with telomere length[39] to account for them in our analyses. The SNPs and their corresponding genes are outlined in Supplementary Table S1. Collectively they are believed to account for 2% to 3% of the variance in leukocyte telomere length.[39]

4. Statistical analyses

4.1 DNA methylation analyses

First, we examined the associations between exposure to PM_{10} in each window of exposure and epigenome-wide methylation levels using linear regression models (lm in the R package *stats* version 3.6.1; R Core Team). The results from the linear regression are presented as change in DNA methylation for $10\mu g/m^3$ increase in PM_{10} concentration. False discovery rate (FDR) was used to adjust for multiple testing. All analyses were adjusted for an a priori selected set of covariates: technical variables (plate and DNA buffer type – ATE vs H2O), study center, maternal age (continuous), education (high- tertiary education, medium- upper secondary education, low-lower secondary or less), parity (nulliparous vs multiparous), pre-pregnancy BMI (continuous), smoking during pregnancy (yes/no) and child's sex. Additional analysis was performed by adjusting also for cell type proportionss. We decided not to adjust for gestational duration and pregnancy complications such as eclampsia or intrauterine growth restriction since they may lie on the causal pathway between the exposure and the outcome. We also did not adjust for season of conception, variable that strongly affect the exposure of interest, with weak or null expected effects on DNA

methylation. We searched all CpGs that passed the FDR threshold in the EWAS Atlas[40] to examine whether they were associated with any specific traits in previous studies. We were able to test whether our main EWAS findings showed similar strength and direction of association in a dataset of 99 Piccolipiù children, analysed with the Illumina 450K platform within the EXPOsOMICS project, which are not included in the current analysis on DNA methylation. To increase the comparability, we used the same preprocessing and quality control pipeline on the raw Illumina450K data and estimated the exposure to single pollutant, PM_{10}, by using the methods described in the present study. The quality control, preprocessing and filtering are described in Supplementary Text S1.

Second, to explore which pathways are targeted by the top most significant CpG sites (arbitrary p-value<0.0001, used in a previous epigenome wide study on PM_{10} exposure and methylation signatures in cord blood[15]) we used the function *gometh* from the *missMetyl*[41] package that performs gene set testing using the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases, taking into account the different number of targeted CpGs per gene.

Third, we conducted an analysis to identify differentially methylated regions (DMRs) in relation to the exposure. We used the R package *DMRcate* that identifies DMRs across the genome based on kernel smoothing of the differential methylation signal. We used the following specifications for calling DMRs: a DMR should contain ≥2 CpGs and lambda was set to 1000, meaning that regions are agglomerated when the distance between is <1000 base pairs. The FDR threshold was set to 0.05.

Fourth, we selected 6 candidate CpGs found to be differentially methylated in relation to PM_{10} exposure in the epigenome-wide meta-analysis conducted by Gruzieva and colleagues.[15] The meta-analysis included the 99 Piccolipiù children from the Turin center.

Lastly, for each sample we derived a measure for overall methylation by averaging the mean beta value across all measured probes (n=770128). We will be using the term "overall methylation" to differentiate from the term "global methylation" that is usually reserved for global methylation profiles measured by high performance liquid chromatography or variants, or by measuring LINE1 or Alu repetitive elements. In addition to the mean beta values across all probes, we averaged the mean beta values across probes in the following regions: (a) CpG island, (b) shore (0–2 kb from CpG islands), (c) shelf (2–4 kb from CpG islands) and (d) open sea regions. To compliment these analyses, we used the R package

GAMP[42] (Global Analysis of Methylation Profiles) that uses a functional regression approach to estimate these two measures for each individual using B-spline based functions. The package is targeted towards capturing more comprehensive, modest changes in methylation globally and is robust to very strong differential methylation in a few CpGs of interest, which may affect summary measures such as mean methylation. Using the *GAMP* package we tested weather PM_{10} exposure influences the cumulative distribution function (CDF) and the density of the methylation values across all probes and probes by genomic regions.

4.2 Telomere length analyses

We used linear regression to assess the association between exposure to PM_{10} in each window of exposure and telomere length in cord blood. The relative T/S ratio measure showed skewed distribution and therefore it was log10 transformed prior to analysis. In the first model we used the same set of covariates used in the DNA methylation analyses. Telomere length measured in cord blood represents the average TL of cells with different replicative histories. To account for differences in cell type proportions, we used the subset of children with both telomere and methylation data. Sensitivity analyses were conducted by excluding 55 samples (11%) with high CV (CV>20). Out of the 16 SNPs associated with telomere length, four SNPs (rs11125529, rs3027234, rs412658 and rs7253490) had large number of missing values and were therefore excluded from further analysis. Information on the remaining 12 SNPs was used to derive a polygenic score summing up the number of effect alleles (alleles associated with longer SNPs) weighted by the beta coefficients estimated in the original genome-wide analysis.[39]

We conducted complete case analysis, meaning that after quality control, apart from the three samples with mismatched sex, we excluded all samples with missing data on any of the covariates. The final number of children included in the EWAS and overall methylation analysis was 365. The number of children included in the telomere length analysis was 453 in the main and 399 in the sensitivity analyses, noting that information on cell types was available for 365 and 324 children in the main and the sensitivity analyses, respectfully.

Results

The characteristics of the study population are described in Table 1. The mean (standard deviation, SD) PM_{10} levels through the course of the pregnancy were 32.1μg/m^3 (10.1) and 32.2μg/m^3 (9.7) for the methylation and telomere length analyses respectively.

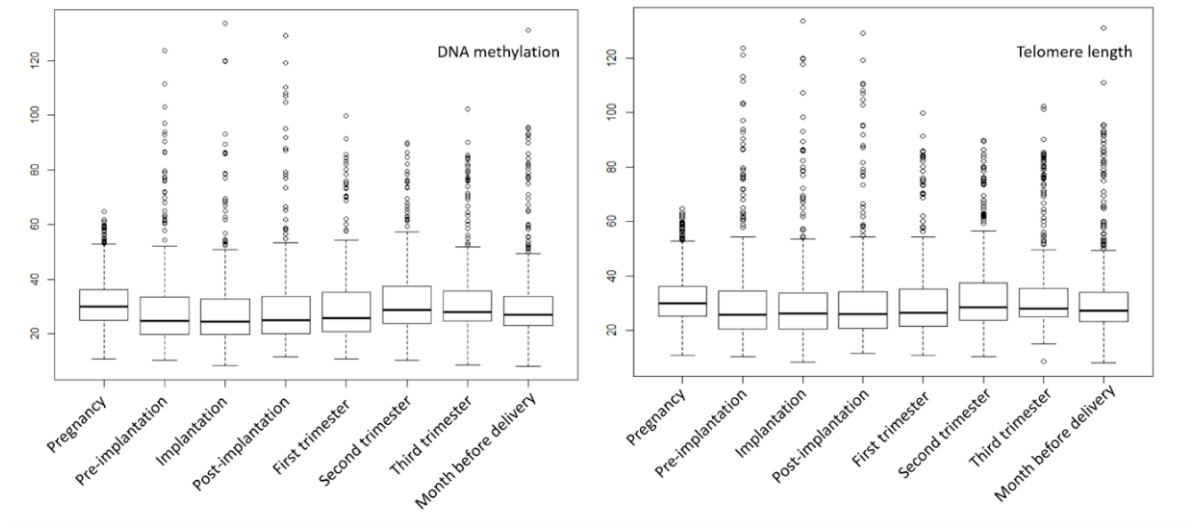
In both analyses, for around 97% of children, the average maternal exposure during the whole pregnancy exceeded 20μg/m^3 (World Health Organization (WHO) annual mean recommended reference level for PM_{10}) and for 17% of the children mean exposure levels during the pregnancy were above 40μg/m^3 (WHO's PM_{10} annual mean limit value for human health). The shorter was the exposure window the more variable were the PM_{10} levels (Table 1, Figure 1), due to the number of days from which the average concentration was taken.

Concentrations in the beginning of the pregnancy, namely pre-implantation, implantation and post implantation seem highly correlated with each other and with first trimester exposure, while their correlation to mid and late-pregnancy PM_{10} concentration is low, Supplementary Figure S2.

Table 1. Characteristics of the participants included in the study

	DNA methylation (n=384)		**Telomere length (n=500)**	
	N	**Mean (SD) or %**	**N**	**Mean (SD) or %**
Study center				
Florence	96	25.0%	100	20.0%
Rome	21	5.5%	100	20.0%
Trieste	96	25.0%	100	20.0%
Turin	75	19.5%	100	20.0%
Viareggio	96	25.0%	100	20.0%
Maternal age	383	34.7 (4.7)	499	34.8 (4.7)
Missing	1		1	
Maternal education				
High	185	48.2%	253	50.6%
Medium	157	40.9%	196	39.2%
Low	42	10.9%	51	10.2%
Maternal pre-pregnacy BMI (in kg/m^2)	381	22.4 (3.5)	496	22.4 (3.5)
Missing	3		4	
Parity				
Nulliparous	225	58.6%	306	61.2%
Multiparous	159	41.4%	194	38.8%
Smoking in pregnancy				
No	306	79.7%	384	77.2%
Yes	78	20.3%	114	22.8%
Child's sex				
Female	204	46.9%	239	47.8%
Male	180	53.1%	261	52.2%
PM10 estimates (in μg/m^3)				
Whole pregnancy	383	32.1 (10.1)	499	32.2 (9.7)
Pre-implantation	383	29.9 (16.8)	499	30.7 (16.9)
Implantation	383	29.1 (15.8)	499	30.1 (16.2)
Post-implantation	383	30.0 (17.3)	499	30.6 (17.1)
First trimester	383	30.2 (14.3)	499	30.5 (13.9)
Second trimester	383	32.8 (14.8)	499	32.8 (14.5)
Third trimester	383	33.3 (14.7)	499	33.4 (14.9)
Last month of pregnancy	383	31.7 (15.9)	499	32.1 (16.3)
Missing	1		1	

Figure 1. Box plot of PM_{10} levels across different windows of exposure in a) the study population with methylation data and b) the study population with telomere length data



1. Agnostic epigenome-wide association analysis on single CpG level

After correction for multiple testing, no CpG was associated with the average PM_{10} exposure during the entire course of the pregnancy. PM_{10} exposure level during the pre-implantation, implantation, the post-implantation period, the first trimester, and the last month of pregnancy was associated (FDR<0.05) with DNA methylation in 4, 64, 46, 1 and 2 CpGs, respectively, amounting to 111 unique CpGs (Table 2).

Table 2. Results from the epigenome-wide association analysis at single CpG level. Across all exposure windows 111 CpGs passed the FDR threshold.

N	Exposure Window	ProbeID	Coef	p-value	Chr	Position	Relation to CpG island	Gene	Location in gene	Promoter
1	PreI, I	cg01688211	-0.0057	1.2E-08	chr20	49337863	OpenSea			
2	Prel, I	cg14643648	-0.0057	5.2E-08	chr16	23426236	OpenSea	COG7	Body	
3	Prel, I	cg09270435	-0.0051	9.7E-08	chr21	36507812	OpenSea			
4	Prel, I	cg18318655	-0.0052	1.2E-07	chr13	31713793	OpenSea	HSPH1	Body	
5	I	cg11317661	-0.0084	2.8E-11	chr3	9343708	OpenSea			
6	I	cg18243764	-0.0065	1.2E-10	chr20	35924185	OpenSea	MANBAL	5'UTR	
7	I	cg13289118	-0.0068	2.9E-10	chr15	75255681	S_Shelf			
8	I	cg24894563	-0.0064	5.4E-10	chr1	227150327	OpenSea	ADCK3	Body	
9	I	cg18980939	-0.0060	1.3E-09	chr8	140999901	OpenSea	TRAPPC9	Body	
10	I	cg13696148	-0.0058	1.4E-09	chr3	139199845	OpenSea			
11	I	cg01200186	-0.0059	2.6E-09	chr10	52765418	OpenSea	PRKG1	Body	
12	I	cg10472067	0.0029	3.3E-09	chr20	35402298	Island	DSN1	TSS200	Yes
13	I	cg18241780	0.0045	5.4E-09	chr12	118620216	OpenSea	TAOK3	Body	
14	I	cg15828158	-0.0059	6.3E-09	chr5	122152681	OpenSea	SNX2	Body	
15	I	cg10026495	-0.0104	1.0E-08	chr4	77631215	OpenSea	SHROOM3	Body	
16	I	cg02283353	-0.0076	1.4E-08	chr19	622406	Island	POLRMT	Body	
17	I	cg18091960	-0.0054	1.5E-08	chr2	43197057	OpenSea			
18	I	cg19853848	-0.0054	1.7E-08	chr5	138857955	OpenSea	TMEM173	Body	
19	I	cg05958000	-0.0057	1.9E-08	chr9	138280729	OpenSea			
20	I	cg18447143	-0.0038	2.9E-08	chr11	74461705	S_Shore	RNF169	Body	
21	I	cg04775569	-0.0064	2.9E-08	chr7	195656	Island	FAM20C	Body	
22	I	cg08100576	-0.0063	3.6E-08	chr8	9627834	OpenSea	TNKS	Body	
23	I	cg01791236	0.0043	5.3E-08	chr19	4815877	OpenSea			
24	I	cg06469440	-0.0048	5.8E-08	chr7	1000336	OpenSea			

25	I	cg09371530	-0.0055	5.9E-08	chr10	131640006	S_Shore	EBF3	Body	
26	I	cg09522226	-0.0067	8.3E-08	chr22	28501343	OpenSea	TTC28	Body	
27	I	cg12066409	-0.0051	1.1E-07	chr2	101595462	OpenSea	NPAS2	Body	
28	I	cg18917544	0.0037	1.5E-07	chr6	99281070	Island			
29	I	cg25133533	0.0075	1.8E-07	chr7	100611097	S_Shore			
30	I	cg03808324	-0.0055	1.9E-07	chr18	64215948	OpenSea	CDH19	Body	
31	I	cg08396875	-0.0051	2.3E-07	chr2	225847045	OpenSea	DOCK10	Body	
32	I	cg10409786	0.0042	2.5E-07	chr20	35402300	Island	DSN1	TSS200	Yes
33	I	cg13583523	-0.0048	2.5E-07	chr5	24764190	OpenSea			
34	I	cg08563994	0.0030	2.6E-07	chr3	49823664	Island	IP6K1	5'UTR	Yes
35	I	cg10262931	0.0033	2.9E-07	chr10	79807251	OpenSea	RPS24	Body	
36	I	cg06527535	-0.0056	2.9E-07	chr7	157633771	S_Shelf	PTPRN2	Body	
37	I	cg08454563	-0.0082	3.4E-07	chr22	31040187	OpenSea	SLC35E4	Body	
38	I	cg04933990	0.0024	3.7E-07	chr16	10133501	OpenSea	GRIN2A	Body	
39	I	cg00257497	-0.0054	5.4E-07	chr5	176396997	OpenSea	UIMC1	Body	
40	I	cg14468007	-0.0051	5.4E-07	chr14	96176384	N_Shelf	TCL1A	3'UTR	
41	I	cg09195657	-0.0049	5.7E-07	chr16	2321187	S_Shelf	MIR940	TSS1500	
42	I	cg25198579	0.0010	8.3E-07	chr15	67358136	Island	SMAD3	TSS200	Yes
43	I	cg18798495	-0.0050	8.9E-07	chr17	77957004	S_Shore	TBC1D16	Body	
44	I	cg00531796	-0.0050	9.0E-07	chr5	140042779	N_Shore			
45	I	cg03175206	-0.0034	1.1E-06	chr3	194978283	N_Shelf	C3orf21	Body	
46	I	cg07485754	0.0044	1.2E-06	chr7	116511339	OpenSea	CAPZA2	Body	
47	I	cg14334911	0.0012	1.3E-06	chr17	44271018	Island			Yes
48	I	cg00493609	-0.0056	1.4E-06	chr16	13704393	OpenSea			
49	I	cg17806989	-0.0058	1.4E-06	chr13	25338287	OpenSea	RNF17	TSS200	
50	I	cg13851412	-0.0050	1.4E-06	chr14	95922175	OpenSea	SYNE3	Body	
51	I	cg21683659	-0.0028	1.5E-06	chr2	242844943	Island			
52	I	cg20523029	-0.0067	1.5E-06	chr6	143928427	OpenSea	PHACTR2	TSS1500	
53	I	cg27343976	-0.0053	1.6E-06	chr6	168766332	S_Shore			
54	I, T1	cg00599628	-0.0116	1.7E-06	chr20	17763350	OpenSea			
55	I	cg09628503	-0.0049	2.0E-06	chr1	1842629	S_Shelf			
56	I	cg24583704	-0.0051	2.5E-06	chr13	21047671	OpenSea	CRYL1	Body	
57	I, PostI	cg09634226	0.0007	2.9E-06	chr8	52811768	Island	PCMTD1	TSS200	Yes
58	I	cg09153950	-0.0022	3.0E-06	chr2	67623673	N_Shore	ETAA1	TSS1500	
59	I	cg19222405	-0.0027	3.2E-06	chr19	1826867	Island	REXO1	Body	
60	I	cg18342425	-0.0054	3.8E-06	chr13	25337983	OpenSea	RNF17	TSS1500	
61	I	cg05408607	-0.0053	3.8E-06	chr12	116653902	OpenSea	MED13L	Body	
62	I	cg13474011	0.0009	3.9E-06	chr10	49515374	Island			
63	I	cg24491780	0.0036	3.9E-06	chr15	41107391	OpenSea			
64	I	cg25423145	-0.0056	4.1E-06	chr17	43513233	S_Shelf			
65	PostI	cg01509792	0.0021	3.3E-09	chr21	43608929	OpenSea			
66	PostI	cg00121374	-0.0055	5.3E-09	chr1	3392638	N_Shelf	ARHGEF16	Body	
67	PostI	cg00325528	-0.0042	2.8E-08	chr11	2819073	S_Shore	KCNQ1	Body	
68	PostI	cg26173959	-0.0050	3.2E-08	chr9	124575779	OpenSea			
69	PostI	cg07390478	-0.0053	3.2E-08	chr7	150646082	N_Shore	KCNH2	Body	
70	PostI	cg23830969	-0.0041	3.3E-08	chr3	15474768	OpenSea	EAF1	Body	
71	PostI	cg05324638	-0.0048	3.6E-08	chr9	90428397	OpenSea			
72	PostI	cg02603756	-0.0052	4.1E-08	chr11	75909015	S_Shelf	WNT11	Body	
73	PostI	cg13797723	-0.0048	5.8E-08	chr11	57078109	OpenSea	TNKS1BP1	Body	
74	PostI	cg21797500	-0.0049	1.1E-07	chr13	84406289	OpenSea			
75	PostI	cg10399764	0.0020	2.0E-07	chr9	115517223	S_Shelf	SNX30	Body	
76	PostI	cg01971227	0.0095	4.2E-07	chr16	3988694	N_Shore			
77	PostI	cg09721840	-0.0048	4.4E-07	chr13	99114023	OpenSea	STK24	ExonBnd; Body	
78	PostI	cg19140429	0.0039	4.5E-07	chr17	48858746	Island			
79	PostI	cg00903584	0.0048	5.0E-07	chr1	202128682	OpenSea	PTPN7	5'UTR; 1stExon	Yes
80	PostI	cg12194267	-0.0047	5.4E-07	chr9	98473994	OpenSea			
81	PostI	cg22068476	-0.0045	5.6E-07	chr12	104271473	OpenSea	TTC41P	Body	
82	PostI	cg14565439	-0.0049	5.7E-07	chr5	107702185	OpenSea	FBXL17	Body	
83	PostI	cg18669139	0.0017	5.7E-07	chr7	50346768	S_Shelf	IKZF1	5'UTR	
84	PostI	cg26305606	-0.0044	5.9E-07	chr10	76997231	S_Shore	COMTD1	TSS1500	
85	PostI	cg01132839	-0.0039	5.9E-07	chr10	73083954	OpenSea	SLC29A3	Body	
86	PostI	cg15997497	0.0018	6.5E-07	chr11	95948847	OpenSea	MAML2	Body	
87	PostI	cg02961575	0.0008	6.8E-07	chr3	186285036	Island	TBCCD1	5'UTR	Yes
88	PostI	cg24646748	0.0058	7.5E-07	chr5	82281494	OpenSea			
89	PostI	cg17002039	-0.0045	7.9E-07	chr9	91955646	OpenSea	SECISBP2	Body	
90	PostI	cg18857759	-0.0045	7.9E-07	chr1	248110661	OpenSea	OR2L8;OR2L13	TSS1500	
91	PostI	cg09968723	0.0020	8.5E-07	chr8	143545789	Island	BAI1	1stExon	
92	PostI	cg16624646	0.0014	9.3E-07	chr19	2474993	Island	GADD45B	TSS1500	Yes
93	PostI	cg21487550	0.0051	9.4E-07	chr3	127056955	Island			
94	PostI	cg22810059	-0.0044	9.4E-07	chr6	167031282	N_Shore	RPS6KA2	Body	
95	PostI	cg07656744	-0.0042	9.6E-07	chr1	174945070	OpenSea	RABGAP1L	Body	
96	PostI	cg21222426	0.0058	1.0E-06	chr9	20339790	OpenSea			
97	PostI	cg00677562	-0.0048	1.2E-06	chr9	111206648	OpenSea			

98	PostI	cg25292359	-0.0046	1.4E-06	chr20	50820973	OpenSea			
99	PostI	cg20806143	-0.0046	1.4E-06	chr1	89567209	OpenSea			
100	PostI	cg22240773	-0.0027	1.6E-06	chr11	15074269	OpenSea			
101	PostI	cg05040872	-0.0042	1.7E-06	chr5	107156463	OpenSea			
102	PostI	cg23471777	0.0026	1.8E-06	chr13	95254279	Island	GPR180	1stExon	Yes
103	PostI	cg23480855	-0.0037	2.0E-06	chr7	101554900	OpenSea	CUX1	Body	
104	PostI	cg11691189	-0.0062	2.0E-06	chr2	47743741	OpenSea			
105	PostI	cg22223441	-0.0041	2.4E-06	chr6	11920499	OpenSea			
106	PostI	cg23707521	-0.0107	2.6E-06	chr3	159585493	OpenSea	SCHIP1	Body	
107	PostI	cg18526161	0.0027	2.8E-06	chr14	64854366	Island	MTHFD1	TSS1500	Yes
108	PostI	cg12876838	-0.0055	2.9E-06	chr4	5897223	S_Shelf			
109	PostI	cg17387122	0.0060	3.0E-06	chr22	18507711	S_Shore	MICAL3	TSS1500	
110	M1	cg26030063	0.0033	6.1E-08	chr14	42078342	S_Shelf	LRFN5	5'UTR	
111	M1	cg06230418	-0.0056	1.1E-07	chr12	133129990	N_Shore	FBRSL1	Body	

Adjusted for technical variables, study center, maternal age, education, parity, pre-pregnancy BMI, smoking in pregnancy, child's sex. The coefficient represents DNA methylation change per 10μg/m^3 increase in PM_{10}.
Abbreviations: Prel, I, PostI, T1 and M1 correspond to "Pre-implantation", "Implantation", "Post-implantation", "First trimester" and "Last month of pregnancy", respectively.

The 4 CpGs identified in the analysis on the pre-implantation period were also present among those identified for the implantation analyses. The CpGs associated with exposure during implantation, post-implantation and the last month of pregnancy were independent, with the exception of one CpG (cg09634226) emerging in both implantation and post-implantation analyses. All CpGs that passed the FDR threshold in more than one window of exposure, showed a consistent direction of association across exposure windows.

Eleven CpGs were mapped to gene promotors and all were hypermethylated with relation to PM_{10} exposure. The rest of the CpGs were found in gene bodies or in non-coding regions far from CpG islands, and were generally hypomethylated. Many of our top CpGs were novel findings, meaning they were either not found in the EWAS atlas or they were not associated with a known trait. Twenty-eight out of 111 CpGs were previously reported in other studies, including studies on prenatal exposures. Table S2 in the Supplementary material lists the CpGs found in the EWAS Atlas and their corresponding traits.

2. Enriched pathway analysis

No pathways were clearly identified after correction for multiple testing in the pathway analysis on 230 and 269 top significant CpGs that had p-value<0.0001 in the implantation and post-implantation analysis respectively. The top 25 GO and KEGG pathways are listed in the Supplementary Tables S3 and S4.

3. Testing the robustness of our findings in another subset of children from the Piccolipiu cohort.

Out of the initial 99 samples from the Turin center analyzed with the Illumina450K platform within the EXPOsOMICs project, only 92 were included in this analysis. The pre-processing

pipeline and the reasons for exclusion are available in Supplementary Text S1. The subsample (in comparison to the sample population with methylation data used in the main analysis) had higher mean levels of PM_{10}, higher percentage of multiparous women and higher percentage of males, Supplementary Table S5. Methylation values were available for 58 of our top 111 CpGs. Around half of the CpGs (n=32) showed consistent direction within the same exposure window, Supplementary Table S6. Six CpGs (10.3%) showed strong and consistent direction of association across all windows of exposure in beginning of the pregnancy (pre-implantation, implantation, post-implantation and first trimester), Table 3. A search on the four genes (SHROOM3, PTPN7, BAI1 and RPs6KA2) in the EWAS atlas revealed previous associations between NO_2 air pollution exposure and DNA methylation in all four genes, as reported in a previous study.

Table 3. Top six CpGs identified in the main epigenome-wide analysis (n=365) that showed the same direction of association in the PiccoliPiù EXPOsOMICS dataset (n=92)

ProbeID	Gene	Main analysis		Replication analysis													
				PreI		I		PostI		T1		T2		T3		M1	
		W	Dir	Dir	p-val	Dir	p-val	Dir	p-val	Dir	p-val	Dir	p-val	Dir	p-val	Dir	p-val
cg10026495	SHROOM3	I	-	-	0.01	-	0.01	-	0.00	-	0.00	+	0.16	+	0.00	+	0.03
cg27343976	unmapped	I	-	-	0.01	-	0.00	-	0.00	-	0.00	+	0.08	+	0.02	+	0.24
cg00903584	PTPN7	PostI	+	+	0.24	+	0.01	+	0.00	+	0.00	+	0.79	-	0.04	-	0.07
cg09968723	BAI1	PostI	+	+	0.21	+	0.04	+	0.07	+	0.24	-	0.16	-	0.05	-	0.36
cg19140429	unmapped	PostI	+	+	0.01	+	0.02	+	0.04	+	0.02	-	0.13	-	0.05	-	0.43
cg22810059	RPS6KA2	PostI	-	-	0.11	-	0.14	-	0.08	-	0.05	+	0.11	+	0.30	+	0.62

Adjusted for technical variables, study center, maternal age, education, parity, pre-pregnancy BMI, smoking in pregnancy, child's sex and cell heterogeneity. Abbreviations: W: Window of exposure; Dir: Direction of association; Prel, I, Post-I, T1, T2, T3 and M1 correspond to "Pre-implantation", "Implantation", "Post-implantation", "First trimester", "Second trimester", "Third trimester" and "Last month of pregnancy", respectively.

4. Results from the DMR analysis

The DMR analysis identified 3, 4 and 1 DMR when analyzing implantation, post-implantation and first trimester PM_{10} exposures, Table 4.

Table 4. Results from the differentially methylated region analysis

DMR	Region coordinates	N of CpGs	$Coef_{mean}$	FDR_{min}
Implantation				
RNF17	chr13:25337649-25338289	3	-0.0045	1.20E-13
SOD3	chr4:24797007-24797176	3	-0.0055	1.76E-08
unmapped	chr6:30094960-30095728	24	0.0036	4.07E-10
Post-implantation				
KIAA0101	chr15:64673790-64673872	6	0.0016	1.33E-12
unmapped	chr6:32064497-32064588	5	0.0035	1.33E-12
First trimester				
HIVEP3	chr1:42384474-42384647	4	0.0063	8.94E-12

Adjusted for technical variables, study center, maternal age, education, parity, pre-pregnancy BMI, smoking in pregnancy, child's sex. The coefficient represents mean DNA methylation change per 10μg/m3 increase in PM10. The FDR value represents the FDR value for the CpG most strongly associated with the outcome.

The DMRs were mapped to genes involved in spermatogenesis and regulation of transcriptional activity of MYC (RNF17 gene), response to oxidative stress (SOD3 gene), regulation of DNA repair during DNA replication (KIAA0101 gene) and with involvement in cell progression and differentiation (HIVEP3).

5. Candidate CpG analysis

We conducted the candidate CpG analysis on 4 out of 6 PM_{10}-related CpGs from the meta-analysis by Gruzieva and colleagues[43] that were available in our main dataset. All 4 CpGs showed variable effect estimates across different windows of exposure. None of them showed the same direction of association as the original study when we analyzed PM_{10} exposure during the whole pregnancy, Table 5.

Table 5. Testing for association the 4 PM_{10}-related CpGs (Gruzieva et al. 2019) in our complete case dataset of 365 children.

Gruzieva et al. 2019 N=1949		Main DNA amalysis (n=365 in the complete case analysis)															
		Preg		PreI		I		PostI		T1		T2		T3		M1	
CpG	Direction	Dir	pval	Dir	pval	Dir	pval	Dir	pval	Dir	pval	Dir	pval	Dir	pval	Dir	pval
cg15082635	-+-++++	-	0.44	-	0.76	-	0.57	-	0.57	-	0.26	-	0.30	+	0.26	+	0.29
cg24127244	-++++++	-	0.78	+	0.72	+	0.87	+	0.80	+	0.69	-	0.34	+	0.71	+	0.54
cg06849931	---++--	+	0.15	+	0.71	-	0.46	-	0.76	+	0.85	+	0.15	+	0.89	+	0.81
cg18640183	+++++++	-	0.15	-	0.71	-	0.12	+	0.80	+	0.99	-	0.58	-	0.20	+	0.95

DirMeta: Direction of methylation change in relation to PM_{10} for each cohort included in the original meta-analysis: INMA, ProjectViva, CHS, PRISM, ENVIRONAGE, Rhea, Piccolipiù and EARLI. The estimates from the current study are adjusted for technical variables, study center, maternal age, education, parity, pre-pregnancy BMI, smoking in pregnancy, child's sex and cell heterogeneity. Preg, Prel, I, Post-I, T1, T2, T3 and M1 correspond to "Pregnancy", "Pre-implantation", "Implantation", "Post-implantation", "First trimester", "Second trimester", "Third trimester" and "Last month of pregnancy", respectively.

6. Overall DNA methylation

We averaged and plotted the methylation density across all probes and across probes in different genomic regions (Supplementary Figures S3-S4) and, as expected, there was a notable difference between regions, while the variation within regions was small, Supplementary Figure S5. Estimates were relatively small ranging from $1x10^{-6}$ to $1x10^{-7}$ change in methylation for $10\mu g/m^3$ increase in PM_{10}. Exposure to PM_{10} during pregnancy was associated with hypomethylation in CpGs islands ($-8.4x10^{-6}$, 95% CI: $-1.7x10^{-5}$ to $-6.4x10^{-7}$), that seems to be mostly due to exposures in the second trimester. This association was also confirmed with the GAMP-CDF analysis, Supplementary Table S7.

7. Telomere length

Although the confidence intervals for most analyses were wide, children exposed to high PM_{10} concentrations during the beginning of the pregnancy showed tendency to have shorter telomeres in cord blood, with strongest effect sizes seen for the post-implantation period (-0,0044, 95% CI -0,0095 to 0,0007), Table 10.

The estimates further attenuated after removing samples with high CV (pre-implantation period: -0,0009, 95% CI -0,0064 to 0,0046), but showed a consistent direction of association, Table 6. Adjusting for cell type proportions did not affect the estimates, Supplementary Table S8. The polygenic score was associated with longer telomeres which is in line with the literature.[39] Adding it to the final model did not alter the PM_{10} estimates.

Table 6. Association between gestational PM_{10} exposure and telomere length

	N=453				N=399 (samples with CV<20)			
	Coef	95% CI		p value	Coef	95% CI		p value
		Lower	Upper			Lower	Upper	
All pregnancy	-0.0005	-0.0144	0.0134	0.943	0.0078	-0.0064	0.0219	0.281
Pre-implantation	-0.0036	-0.0088	0.0016	0.175	-0.0021	-0.0073	0.0032	0.444
Implantation	-0.0033	-0.0087	0.0021	0.232	-0.0011	-0.0066	0.0045	0.704
Post-implantation	-0.0045	-0.0096	0.0006	0.081	-0.0013	-0.0065	0.0040	0.640
First trimester	-0.0047	-0.0115	0.0021	0.173	-0.0003	-0.0073	0.0066	0.926
Second trimester	0.0007	-0.0059	0.0073	0.834	0.0040	-0.0027	0.0106	0.240
Third trimester	0.0046	-0.0021	0.0112	0.175	0.0028	-0.0040	0.0095	0.424
Last month of pregnancy	0.0023	-0.0034	0.0080	0.429	-0.0001	-0.0061	0.0059	0.975

Adjusted for technical variables, study center, maternal age, education, parity, pre-pregnancy BMI, smoking in pregnancy, child's sex and polygenic score

Discussion

We observed that gestational PM_{10} exposure was associated with DNA methylation and possibly telomere length in the cord blood. The reported associations were mostly due to exposures during the beginning of the pregnancy, namely around the period of implantation.

More than 100 unique CpGs and 4 DMRs were associated with PM_{10} exposure during the beginning of the pregnancy with roles mostly in cell replication, differentiation and response to oxidative stress.[44] Hypermethylation of CpG islands (located within gene promotors) is almost always synonymous with long-term gene silencing.[25,45] In our study, the majority of CpGs in CpG islands and all CpGs associated with gene promoters (n=11) were hypermethylated with respect to PM_{10}. Some of them were in genes that, according to the Human Gene Database,[44] have roles in cell replication (*DSN1* gene), tumor suppression (SMAD3 gene), differentiation of T and B-lymphocytes (*PTPN7* gene), response to environmental stress and DNA damage (*GADD45B* gene) and involvement in one-carbon metabolism providing methyl groups needed for DNA methylation (*MTHFD1* gene). The rest of the CpGs were hypomethylated in response to PM_{10} exposure. Most of these genes were in gene bodies and non-coding genomic regions. Contrary to CpG islands, gene bodies generally need to be hypermethylated to provide blanket silencing of cryptic promoters or cryptic splice

sites. Exposure to PM_{10} might cause activation of these genomic elements that might lead to transcription of aberrant peptides.[46,47]

Prenatal exposures such as maternal smoking[48] and toxic environmental substances[49–51] were previously linked with methylation levels of some of the CpGs identified in this study, further confirming their involvement in responses linked to oxidative stress. We conducted a more in-depth search for the 4 CpGs mapped to a known gene that showed robust associations in the EXPOsOMICS dataset, by looking for traits not only associated with the CpG locus, but also with gene they are mapped to. We did the same thing for the 4 identified DMRs. Interestingly, changes in methylation of all four genes and two DMRs (SOD3 and HIVEP3) were associated with NO_2 exposure in previous study.[52] All these findings indicate that locus and region-specific changes in DNA methylation at birth could have a functional relevance and act as a mediator through which air pollution exposure could leave a molecular fingerprint in early life and influence health outcomes years after the exposure.

Although there seems to be abundant number of studies that report locus-specific changes in DNA methylation in relation to prenatal exposure to air pollution[12,15,43,53], epigenome-wide studies often suffer from lack of robust associations at single CpG level across cohorts, possibly due to environmental mixtures, as well as technical and true variation across study populations.[54,55] Even in the present study, we did not replicate findings for 4 PM_{10}-related CpGs emerging from a previous meta-analysis including 1949 children.[15] There may be several reasons for this lack of replication, such as different methods of estimating PM_{10} concentrations (annual average concentrations using land-use regression models/other methods in the meta-analysis vs RF estimations based on daily satellite data in the current study) and misclassification of exposure (air pollution exposure based on residential address cannot fully capture individual behaviors and daily pattern), differences in laboratory conditions, assays (Illumina450K vs IlluminaEPIC) and other batch variability uncaptured by the variables chosen for adjustment, different pre-processing pipeline, different methods for estimating cell types, as well as different confounding pattern not accounted for in the analyses. We ran sensitivity analyses by removing maternal age, education, parity and pre-pregnancy BMI from the model, since these covariates were not considered in the meta-analysis, and the results did not change significantly.

We were able to detect robust signals in six CpGs when were tested our top CpGs in the same subsample of Piccolipiù included in the aforementioned meta-analysis. This could be due to

fact that we addressed some of the issues underlined previously, such as using the same method to re-estimate PM_{10} concentrations, running the pre-processing and filtering pipeline on raw data and estimating cell counts with the same method. This however does not indicate that the signal will be stable across different populations, and future studies should replicate these findings in other cohorts paying attention to the period of exposure, since the associations of these CpGs seems to be specific for exposure around pregnancy.

We did not observe an association between prenatal PM_{10} exposure and overall methylation patterns when we averaged all probes on the IlluminaEPIC assay. Previous studies measuring global methylation using traditional methods generally reported an association between prenatal exposure to particulate matter and global hypomethylation.[6,7,9,11] However, a study similar to ours, that used array-based data to derive a measure of overall methylation, also did not report differences in methylation in relation to PM_{10} exposure.[53] When we averaged the methylation levels across different genomic regions, we observed a lower mean methylation in CpG islands in relation to gestational PM_{10} exposure, mostly due to exposures during the second trimester. This association was confirmed by the GAMP-CDF analysis, meaning that the association likely reflects true mean shifts in methylation patterns. This is at odds with the hypermethylation at CpG islands we observed in our CpG-based EWAS analysis. There is no other study that analyzed methylation patterns by genomic regions in relation to prenatal PM_{10} exposure. Previous study[12] on gestational NO_2 and O_3 exposure, that averaged probes across different regions, did not observe an association with methylation in CpG islands in cord blood.

Our study provided some indication that air pollution exposure during the beginning of the pregnancy could influence telomere length, although the confidence intervals were wide and included the null. The estimates further attenuated after removing 11% of the samples that had high CV. We believe that the observed estimates do not reflect a lack of association, but lack of power to detect one, considering the consistent direction of association. Our findings indicate a possible association between exposures in the beginning of the pregnancy and telomere length shortening, which is in line with a study that reported short telomeres in relation to exposures during the 4^{th} to 9^{th} gestational week, and are in odds with studies that reported telomere length shortening in relation to exposure in mid[17,56] or late pregnancy.[19] Similarly to two other studies[17,21], we also observed an indication that exposures in late pregnancy might lead to longer telomeres. Disturbance in the telomere length maintenance system leading to both or long telomeres, conveys some sort of risk. Shorter telomeres

increase the risk for many age-related diseases, while longer telomeres increase the risk for some types of cancer.[39,57]

One of the main findings of our study is that the beginning of the pregnancy is particularly susceptible period during which PM_{10} exposure can influence fetal DNA methylation programming and, possibly, telomere length shortening. This is in line with what is already know about the establishment of DNA methylation patterns during early cell divisions.[24] The embryo undergoes three consecutive waves of epigenome-wide pattern change during the pre-implantation, implantation and post-implantation period. In the preimplantation embryo, almost all parental methylation patterns are erased, with the exception of the imprinting loci that remain methylated on one allele. At the time of implantation, the entire genome undergoes a dramatic wave of *de novo* methylation with most loci becoming highly methylated, while CpG islands remain unmethylated, creating the typical bimodal pattern that will largely maintained intact in all subsequent cell divisions. In the post-implantation period, changes in the bimodal pattern do not occur on a global scale but there are tissue or gene-specific changes that initiate embryonal cell differentiation and including silencing of the pluripotency genes to prevent regression into an undifferentiated state.[24,25] It should be noted that, although the baseline methylation patterns are established before birth, the methylome remains a dynamic system throughout life and is influenced both by the underlying genetic make-up and environment in which it is immersed.[58]

Our study has several strengths. We measured two molecular markers, that capture different biological responses to cellular stress and we reported that PM_{10} exposure during pregnancy has the potential to leave molecular marks at birth that could potentially influence later disease susceptibility. We used daily PM_{10} exposure data based on machine learning method to obtain exposure estimates within smaller gestational windows of exposure and we reported the beginning of the pregnancy as potential window of vulnerability. In addition, we were able to test whether our top findings show robust associations in another subsample of the same cohort and we reported six CpGs that showed robust associations. There are also some weaknesses. While we were able to reliably estimate air pollution exposure based on residential address, these estimates do not necessarily reflect true individual exposure and time-activity patterns might introduce exposure misclassification. Secondly, we studied only exposure to PM_{10}, since daily exposure predictions for other pollutants, such as $PM_{2.5}$, were not available for our study period. Although our study included relatively large number of subjects, probably more power was needed to detect stronger association with telomere

length and while we carefully chose the set of confounders included in the models, there is a possibility for residual confounding. Finally, we were not able to assess the stability of the found epigenetic changes in later childhood.

Conclusion

Exposure to PM_{10} during the beginning of the pregnancy was associated with altered methylation in more than 100 loci and 4 DMR with relevant function in cell replication, differentiation and response to environmental stressors. Six of them showed strong and robust associations in an independent subsample from the same cohort.

The beginning of pregnancy seems to be a window period of particulate importance for DNA methylation and PM_{10} exposure, and possibly for telomere length. Further studies should confirm these findings and integrate them in their analyses when studying gestational exposures and DNA methylation patterns.

In the beginning of the pregnancy, methylation and telomere length are fundamental regulators of cellular processes and their alteration could have major implications for health outcomes and disease susceptibility for complex diseases.

Bibliography:

1. Alvarado-Cruz I, Alegría-Torres JA, Montes-Castro N, Jiménez-Garza O, Quintanilla-Vega B. Environmental epigenetic changes, as risk factors for the development of diseases in children: A systematic review. *Ann Glob Heal*. 2018;84(2):212-224. doi:10.29024/aogh.909
2. Entringer S, de Punder K, Buss C, Wadhwa PD. The fetal programming of telomere biology hypothesis: An update. *Philos Trans R Soc B Biol Sci*. 2018. doi:10.1098/rstb.2017.0151
3. Ferrari L, Carugno M, Bollati V. Particulate matter exposure shapes DNA methylation through the lifespan. *Clin Epigenetics*. 2019;11(1). doi:10.1186/s13148-019-0726-x
4. Martin EM, Fry RC. Environmental Influences on the Epigenome: Exposure- Associated DNA Methylation in Human Populations. *Annu Rev Public Health*. 2018;39(1):309-333. doi:10.1146/annurev-publhealth-040617-014629
5. Yang P, Gong YJ, Cao WC, et al. Prenatal urinary polycyclic aromatic hydrocarbon metabolites, global DNA methylation in cord blood, and birth outcomes: A cohort study in China. *Environ Pollut*. 2018. doi:10.1016/j.envpol.2017.11.082
6. Cai J, Zhao Y, Liu P, et al. Exposure to particulate air pollution during early pregnancy is associated with placental DNA methylation. *Sci Total Environ*. 2017. doi:10.1016/j.scitotenv.2017.07.029
7. Liu X, Ye Y, Chen Y, et al. Effects of prenatal exposure to air particulate matter on the risk of preterm birth and roles of maternal and cord blood LINE-1 methylation: A birth cohort study in Guangzhou, China. *Environ Int*. 2019;133. doi:10.1016/j.envint.2019.105177
8. Lee J, Kalia V, Perera F, et al. Prenatal airborne polycyclic aromatic hydrocarbon exposure, LINE1 methylation and child development in a Chinese cohort. *Environ Int*. 2017.

doi:10.1016/j.envint.2016.12.009

9. Breton C V., Yao J, Millstein J, et al. Prenatal air pollution exposures, DNA methyl transferase genotypes, and associations with newborn line1 and Alu methylation and childhood blood pressure and carotid intima-media thickness in the children's health study. *Environ Health Perspect*. 2016. doi:10.1289/EHP181

10. Janssen BG, Byun HM, Gyselaers W, Lefebvre W, Baccarelli AA, Nawrot TS. Placental mitochondrial methylation and exposure to airborne particulate matter in the early life environment: An ENVIRONAGE birth cohort study. *Epigenetics*. 2015. doi:10.1080/15592294.2015.1048412

11. Janssen BG, Godderis L, Pieters N, et al. Placental DNA hypomethylation in association with particulate air pollution in early life. *Part Fibre Toxicol*. 2013. doi:10.1186/1743-8977-10-22

12. Ladd-Acosta C, Feinberg JI, Brown SC, et al. Epigenetic marks of prenatal air pollution exposure found in multiple tissues relevant for child health. *Environ Int*. 2019. doi:10.1016/j.envint.2019.02.028

13. Abraham E, Rousseaux S, Agier L, et al. Pregnancy exposure to atmospheric pollution and meteorological conditions and placental DNA methylation. *Environ Int*. 2018;118:334-347. doi:10.1016/j.envint.2018.05.007

14. Gruzieva O, Xu CJ, Breton C V., et al. Epigenome-wide meta-analysis of methylation in children related to prenatal NO2 air pollution exposure. *Environ Health Perspect*. 2017. doi:10.1289/EHP36

15. Gruzieva O, Xu C-J, Yousefi P, et al. Prenatal Particulate Air Pollution and DNA Methylation in Newborns: An Epigenome-Wide Meta-Analysis. *Environ Health Perspect*. 2019;127(5):057012. doi:10.1289/EHP4522

16. Breton C V., Gao L, Yao J, Siegmund KD, Lurmann F, Gilliland F. Particulate matter, the newborn methylome, and cardio-respiratory health outcomes in childhood. *Environ Epigenetics*. 2016. doi:10.1093/eep/dvw005

17. Martens DS, Cox B, Janssen BG, et al. Prenatal air pollution and newborns' predisposition to accelerated biological aging. *JAMA Pediatr*. 2017. doi:10.1001/jamapediatrics.2017.3024

18. Clemente DBP, Vrijheid M, Martens DS, et al. Prenatal and childhood traffic-related air pollution exposure and telomere length in european children: The HELIX project. *Environ Health Perspect*. 2019;127(8). doi:10.1289/EHP4148

19. Song L, Zhang B, Liu B, et al. Effects of maternal exposure to ambient air pollution on newborn telomere length. *Environ Int*. 2019. doi:10.1016/j.envint.2019.04.064

20. Nie J, Li J, Cheng L, et al. Prenatal polycyclic aromatic hydrocarbons metabolites, cord blood telomere length, and neonatal neurobehavioral development. *Environ Res*. 2019;174:105-113. doi:10.1016/j.envres.2019.04.024

21. Rosa MJ, Hsu HHL, Just AC, et al. Association between prenatal particulate air pollution exposure and telomere length in cord blood: Effect modification by fetal sex. *Environ Res*. 2019. doi:10.1016/j.envres.2019.03.003

22. Perera F, Lin C jung, Qu L, Tang D. Shorter telomere length in cord blood associated with prenatal air pollution exposure: Benefits of intervention. *Environment International*. 2018.

23. 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. doi:10.1056/NEJMra0708473

24. Cedar H, Bergman Y. Programming of DNA Methylation Patterns. *Annu Rev Biochem*. 2012.

doi:10.1146/annurev-biochem-052610-091920

25. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. *Lancet*. 2018. doi:10.1016/S0140-6736(18)31268-6

26. Aviv A, Shay JW. Reflections on telomere dynamics and ageing-related diseases in humans. *Philos Trans R Soc B Biol Sci*. 2018. doi:10.1098/rstb.2016.0436

27. 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. doi:10.1021/acs.est.7b06447

28. Farchi S, Forastiere F, Vecchi Brumatti L, et al. Piccolipiù, a multicenter birth cohort in Italy: Protocol of the study. *BMC Pediatr*. 2014;14(1). doi:10.1186/1471-2431-14-36

29. Stafoggia M, Bellander T, Bucci S, et al. Estimation of daily PM10 and PM2.5 concentrations in Italy, 2013–2015, using a spatiotemporal land-use random-forest model. *Environ Int*. 2019;124:170-179. doi:10.1016/j.envint.2019.01.016

30. Vineis P, Chadeau-Hyam M, Gmuender H, et al. The exposome in practice: Design of the EXPOsOMICS project. *Int J Hyg Environ Health*. 2017;220(2):142-151. doi:10.1016/j.ijheh.2016.08.001

31. Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*. 2016. doi:10.1186/s13059-016-1066-1

32. Price EM, Robinson WP. Adjusting for batch effects in DNA methylation microarray data, a lesson learned. *Front Genet*. 2018;9(MAR). doi:10.3389/fgene.2018.00083

33. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363-1369. doi:10.1093/bioinformatics/btu049

34. Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*. 2016;17(1). doi:10.1186/s13059-016-1066-1

35. McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. *Genomics Data*. 2016;9:22-24. doi:10.1016/j.gdata.2016.05.012

36. Gervin K, Salas LA, Bakulski KM, et al. Systematic evaluation and validation of reference and library selection methods for deconvolution of cord blood DNA methylation data. *Clin Epigenetics*. 2019;11(1):125. doi:10.1186/s13148-019-0717-y

37. Salas LA, Koestler DC, Butler RA, et al. An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. *Genome Biol*. 2018;19(1):64. doi:10.1186/s13059-018-1448-7

38. Cawthon RM. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res*. 2009;37(3). doi:10.1093/nar/gkn1027

39. Haycock PC, Burgess S, Nounu A, et al. Association between telomere length and risk of cancer and non-neoplastic diseases a mendelian randomization study. *JAMA Oncol*. 2017. doi:10.1001/jamaoncol.2016.5945

40. Li M, Zou D, Li Z, et al. EWAS Atlas: A curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Res*. 2019;47(D1):D983-D988. doi:10.1093/nar/gky1027

41. Phipson B, Maksimovic J, Oshlack A. MissMethyl: An R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics*. 2016. doi:10.1093/bioinformatics/btv560

42. Zhao N, Bell DA, Maity A, et al. Global Analysis of Methylation Profiles From High Resolution CpG Data. *Genet Epidemiol*. 2015;39(2):53-64. doi:10.1002/gepi.21874

43. Gruzieva O, Xu CJ, Breton C V, et al. Epigenome-Wide Meta-Analysis of Methylation in Children Related to Prenatal NO2 Air Pollution Exposure. *Env Heal Perspect*. 2017;125(1):104-110. doi:10.1289/ehp36

44. Database GeneCards. GeneCards - Human Genes | Gene Database | Gene Search. *Genecards.org*. 2017.

45. Long HK, King HW, Patient RK, Odom DT, Klose RJ. Protection of CpG islands from DNA methylation is DNA-encoded and evolutionarily conserved. *Nucleic Acids Res*. 2016. doi:10.1093/nar/gkw258

46. Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13(7):484-492. doi:10.1038/nrg3230

47. Jeziorska DM, Murray RJS, De Gobbi M, et al. DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. *Proc Natl Acad Sci U S A*. 2017. doi:10.1073/pnas.1703087114

48. Sikdar S, Joehanes R, Joubert BR, et al. Comparison of smoking-related DNA methylation between newborns from prenatal exposure and adults from personal smoking. *Epigenomics*. 2019;11(13):1487-1500. doi:10.2217/epi-2019-0066

49. Rojas D, Rager JE, Smeester L, et al. Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. *Toxicol Sci*. 2015. doi:10.1093/toxsci/kfu210

50. Gliga AR, Engström K, Kippler M, et al. Prenatal arsenic exposure is associated with increased plasma IGFBP3 concentrations in 9-year-old children partly via changes in DNA methylation. *Arch Toxicol*. 2018;92(8):2487-2500. doi:10.1007/s00204-018-2239-3

51. Rager JE, Bailey KA, Smeester L, et al. Prenatal arsenic exposure and the epigenome: Altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood. *Environ Mol Mutagen*. 2014. doi:10.1002/em.21842

52. Lichtenfels AJDFC, Van Der Plaat DA, De Jong K, et al. Long-term air pollution exposure, genome-wide DNA methylation and lung function in the lifelines cohort study. *Environ Health Perspect*. 2018;126(2). doi:10.1289/EHP2045

53. Abraham E, Rousseaux S, Agier L, et al. Pregnancy exposure to atmospheric pollution and meteorological conditions and placental DNA methylation. *Environ Int*. 2018. doi:10.1016/j.envint.2018.05.007

54. Pekkanen J, Pearce N. Environmental epidemiology: challenges and opportunities. *Environ Health Perspect*. 2001;109(1):1-5. doi:10.1289/ehp.011091

55. Breton C V., Marsit CJ, Faustman E, et al. Small-magnitude effect sizes in epigenetic end points are important in children's environmental health studies: The children's environmental health and disease prevention research center's epigenetics working group. *Environ Health Perspect*. 2017;125(4):511-526. doi:10.1289/EHP595

56. Lee AG, Cowell W, Kannan S, et al. Prenatal particulate air pollution and newborn telomere length: Effect modification by maternal antioxidant intakes and infant sex. *Environ Res*.

2020;187:109707. doi:10.1016/j.envres.2020.109707

57. Blackburn EH, Epel ES, Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science (80-).* 2015;350(6265):1193-1198. doi:10.1126/science.aab3389

58. Reizel Y, Spiro A, Sabag O, et al. Gender-specific postnatal demethylation and establishment of epigenetic memory. *Genes Dev*. 2015. doi:10.1101/gad.259309.115

Supplementary Material: Study II

Figure S1. *Reproducibility of the telomere length measures. Ten samples were in triplicates run on two separate occasions, several days apart. A) Scatter plot, B) Bland Altman plot*

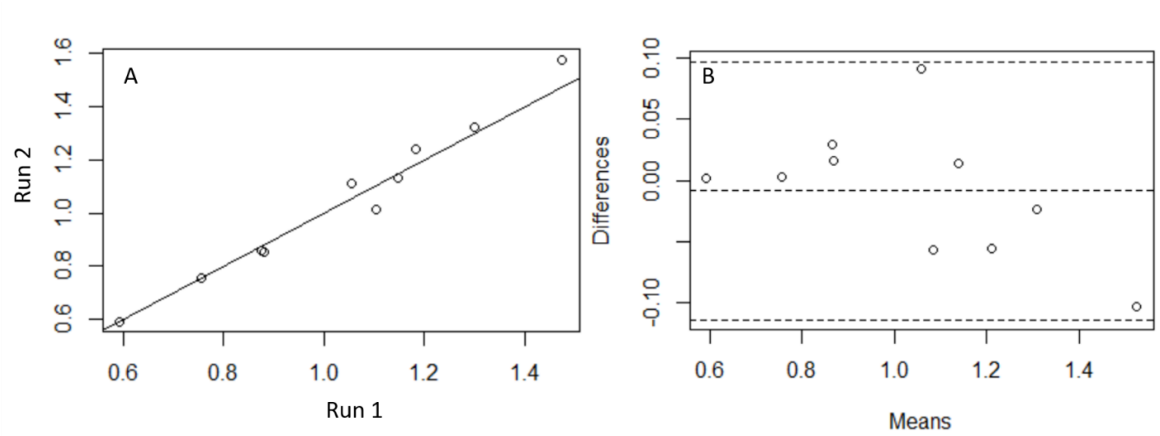


Table S1. The list of SNPs associated with telomere length

SNP	Chromosome	Position	Gene
rs11125529	2	54248729	ACYP2
rs6772228	3	58390292	PXK
rs12696304	3	169763483	TERC
rs10936599	3	169774313	TERC
rs1317082	3	169779797	TERC
rs10936601	3	169810661	TERC
rs7675998	4	163086668	NAF1
rs2736100	5	1286401	TERT
rs9419958	10	103916188	OBFC1
rs9420907	10	103916707	OBFC1
rs4387287	10	103918139	OBFC1
rs3027234	17	8232774	CTC1
rs8105767*	19	22032639	ZNF208
rs412658	19	22176638	ZNF676
rs6028466	20	39500359	DHX35
rs755017	20	63790269	ZBTB46

*rs8105767 was unavailable and replaced with rs7253490. The two SNPs were in linkage disequilibrium r2=0.69.

Text S1. Pre-processing and quality control of the 99 samples analyzed within EXPOsOMICS project

Within the EXPOsOMICS collaborative European project (Vineis et al., 2017), a combination of three population based birth cohorts ENVIRonmental influence ON AGEing in early life (ENVIRONAGE), Rhea and Piccolipiu was established to conduct DNA methylation analyses. Aliquots of Piccoli Più cord blood samples (collected and frozen at birth at -80°C) were shipped on dry ice to the Epigenetics Group at the International Agency for Research on Cancer (IARC), Lyon, France, where DNA was extracted (QIAamp 96 DNA Blood Kit, Qiagen 51161), quantified (Quant-iT PicoGreen dsDNA Assay Kit, Molecular Probes P7589) and bisulfite converted (600 ng of DNA using EZ-96 DNA Methylation kit, Zymo Research D5004). DNA methylation was measured at 485512 GpGs using Illumina Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, USA). The phenotypic variables were harmonized across the three cohorts, and their biospecimens were semi-randomized on the DNA methylation arrays such that the latter would incorporate proportional representations of the three cohorts and that batch effects do not completely confound with biological covariates of interest.

In the current study, we used the minfi package to import the raw data from IDAT files, into the R software for quality control and processing. Detection p-value was calculated for each of the 485512 probes for each sample by comparing the total signal for each probe to the background signal level, estimated from the negative control probes. All 99 samples had mean detection p value < 0.01. The dataset was normalized using quantile normalization (preprocessQuantile) from the minfi package. We filtered out probes that failed in one or more samples (n=2964, 0.6%), probes on the sex chromosomes (n=11234), probes with SNPs at CpG site (n=16469) and cross reactive probes i.e probes that have been demonstrated to map to multiple places in the genome (n=35044) (Chen et al., 2013, Benton et al., 2015), leaving a total of 419801 CpGs for the analysis. Before filtering probes on sex chromosomes we estimated the sex based on the median values of measurements on the X and Y chromosomes. One sample was found to be discordant by the reported sex (male) and predicted sex (female) and was excluded, leaving total of 98 samples for further analysis. Methylation levels were expressed as beta values, ranging from 0 to 1, that represent the ratio of the intensity of the methylated-probe signal to the total locus signal intensity, methylated and unmethylated.

We used the same method as the discovery dataset to estimate 7 cell types: CD8 T cells, CD4 T cells, Natural Killer cells, B cells, monocytes, granulocytes and nucleated red blood cells, by using the filtered reference dataset "FlowSorted.CordBloodCombined.450k" (Gervin et al., 2019) for deconvulsion. We used the function estimateCellCounts2 from the FlowSorted.Blood.EPIC R package (Salas et al., 2018) according to the authors' instructions.

The 99 PiccoliPiù samples were distributed in 6 plates together with samples from other cohorts. The first four plates contained more than 20 PiccoliPiù samples per plate, while the last two plates contained 2 and 4 samples, respectively. The 6 samples from the last two plates were excluded from the analyses in order to be able to adequately adjust for plate in the regression models, leaving 92 samples for analysis. While in the original EXPOsOMICS study the PM_{10} concentrations were estimated using land-use regression modeling described elsewhere (Gruzieva et al., 2019), to increase the comparability between the two datasets we re-estimated the PM_{10} concentrations using the same method as the main analysis. (Stafoggia et al., 2016) Using linear regression, we assessed the association between PM_{10} exposure and the methylation of the top CpGs from the discovery analysis available in the Piccoli Più EXPOsOMICS dataset. Fifty-eight out of 111 CpGs were available on the Illumina450K array. All analyses were adjusted for an a priori selected set of covariates: technical variables (plate and DNA buffer type – ATE vs H2O), study center, maternal age (continuous), education (high- tertiary education, medium- upper secondary education, low-lower secondary or

less), parity (nulliparous vs multiparous), pre-pregnancy BMI (continuous), smoking during pregnancy (yes/no) and child's sex.

Figure S2. Correlation matrix between PM_{10} concentrations in different exposure windows. "Preg", "PreImp", "Imp", "PostImp", "T1", "T2", "T3" and "M1" refer to PM_{10} exposures during the entire pregnancy, pre-implantation, implantation, post-implantation, first trimester, second trimester, third trimester and last month of pregnancy, respectively.

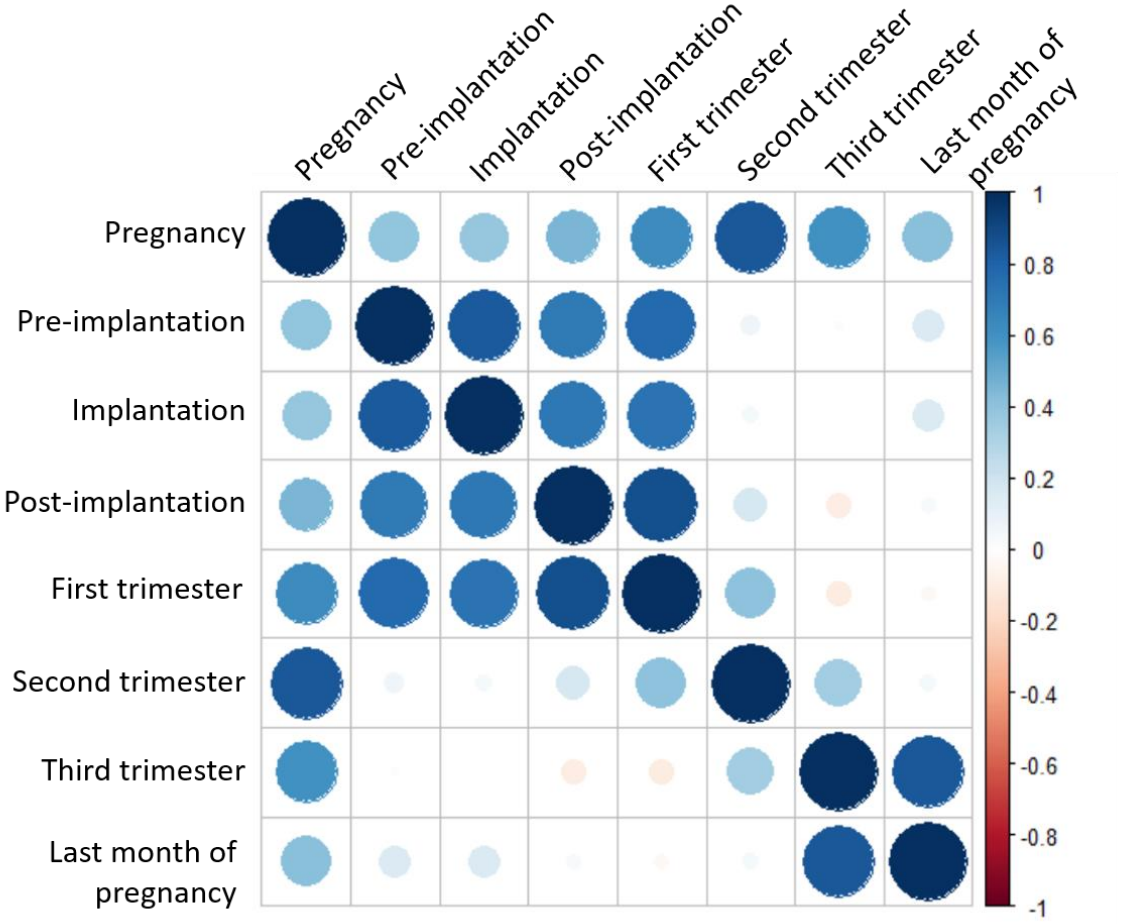


Table S2. Traits associated with 28/111 CpGs found in EWAS Atlas and the studies where the original associations were reported.

Probe ID	Trait	PMID of the corresponding study
cg00121374	obesity	26977391
cg00903584	psoriasis	30092825
cg02283353	smoking, human herpesvirus 6B infection	24559495, 28298607
cg02603756	prenatal arsenic exposure, Crohn's disease	25304211, 27279921
cg04933990	B Acute Lymphoblastic Leukemia with t(1;19)(q23;p13.3); E2A-PBX1 (TCF3-PBX1), oral squamous cell carcinoma	26237075, 28890207
cg06230418	obesity	26977391
cg07485754	air pollution (NO2), smoking	29410382, 27651444, 31536415
cg08454563	smoking	27651444, 31536415
cg09968723	B Acute Lymphoblastic Leukemia with t(1;19)(q23;p13.3); E2A-PBX1 (TCF3-PBX1), Helicobacter pylori infection, oral squamous cell carcinoma, colorectal cancer, colorectal laterally spreading tumor, insufficient sleep, Tetralogy of Fallot	29574700, 28890207, 26237075, 30183087, 30718923, 30977211
cg10026495	right insular surface area	30830696
cg11691189	breast cancer, type 2 diabetes	26418287, 28899398
cg12876838	obesity, perinatally acquired HIV	29692867, 31324826
cg13583523	asthma, adenocortical carcenoma, follicular thyroid carcinoma, facial aging, adenoma, short-term diesel exhaust inhalation, chemotherapy for breast cancer	27942592, 26963385, 28938489, 30409196, 30183087, 25487561, 30867049
cg14334911	child abuse	23332324
cg16624646	asthma	27942592
cg17806989	absolute fat free mass, childhood adversity	29084944, 30905381
cg18241780	smoking, aging	31536415, 27651444, 30626398
cg18318655	alcohol consuption	31910897
cg18342425	Gulf war Illness	30920300
cg18798495	preterm birth, B Acute Lymphoblastic Leukemia with t(12;21)(p13.2;q22.1) ETV6-RUNX1	28428831, 26237075
cg18857759	hepatocellular carcinoma, colorectal cancer, adenoma	29988590, 30183087
cg21222426	aging, B Acute Lymphoblastic Leukemia with t(1;19)(q23;p13.3) E2A-PBX1 (TCF3-PBX1), fetal alcohol spectrum disorder, prenatal arsenic exposure, preterm birth, ancestry, perinatally- acquired HIV, gestational age	29947889, 29064478, 28428831, 27358653, 26237075, 30563547, 31324826, 30966880
cg21487550	gestational diabetes	29596946
cg22223441	breast cancer	29776342
cg23480855	polychlorinated biphenyl exposure	32660331
cg24583704	obesity	26977391
cg24646748	Down syndrome	31843015
cg26030063	aging	30626398

Table S3. Top 25 GeneOntology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways among the top CpGs associated with PM_{10} exposure during the period of implantation. N, number of genes in the pathway; DM, number of differentially methylated genes in the pathway, P, p-value and FDR, false discovery rate.

GO TERM	N	DM	P	FDR
transcription initiation from mitochondrial promoter	3	2	0,0001	1
integral component of presynaptic active zone membrane	14	3	0,0005	1
T cell homeostasis	28	3	0,0010	1
growth cone membrane	6	2	0,0017	1
protein kinase binding	401	12	0,0018	1
cell morphogenesis	52	4	0,0022	1
mitochondrial transcription	9	2	0,0028	1
calcium-transporting ATPase activity	5	2	0,0029	1
leukocyte chemotaxis	11	2	0,0035	1
protein dimerization activity	158	6	0,0037	1
superoxide anion generation	14	2	0,0038	1
negative regulation of cardiac muscle hypertrophy in response to stress	8	2	0,0041	1
coenzyme binding	10	2	0,0043	1
positive regulation of potassium ion transport	9	2	0,0050	1
cell recognition	11	2	0,0050	1
kinase activity	67	4	0,0054	1
negative regulation of potassium ion transmembrane transporter activity	8	2	0,0055	1
detection of tumor cell	1	1	0,0055	1
calcium ion transmembrane import into cytosol	7	2	0,0056	1
cyclic-di-GMP binding	1	1	0,0059	1
cyclic-GMP-AMP binding	1	1	0,0059	1
integral component of ERGIC membrane	1	1	0,0059	1
hormone biosynthetic process	12	2	0,0063	1
cholestenol delta-isomerase activity	2	1	0,0067	1
negative regulation of potassium ion transmembrane transport	10	2	0,0068	1

KEGG term	N	DM	P	FDR
Synaptic vesicle cycle	78	4	0,0120	1
Phagosome	152	4	0,0464	1
Thyroid hormone signaling pathway	121	4	0,0650	1
Adherens junction	71	3	0,0730	1
Viral protein interaction with cytokine and cytokine receptor	100	2	0,0898	1
Other types of O-glycan biosynthesis	47	2	0,0972	1
PPAR signaling pathway	77	2	0,1369	1
Endocrine and other factor-regulated calcium reabsorption	53	2	0,1371	1
Endocytosis	248	5	0,1404	1
Epithelial cell signaling in Helicobacter pylori infection	70	2	0,1518	1
Human papillomavirus infection	331	6	0,1620	1
AMPK signaling pathway	120	3	0,1667	1
Pentose and glucuronate interconversions	34	1	0,1724	1
Folate biosynthesis	26	1	0,1771	1
Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate	20	1	0,2066	1
Thyroid hormone synthesis	75	2	0,2069	1
Collecting duct acid secretion	27	1	0,2127	1
Salivary secretion	93	2	0,2276	1
mTOR signaling pathway	155	3	0,2369	1
Herpes simplex virus 1 infection	491	5	0,2401	1
Cytokine-cytokine receptor interaction	294	3	0,2470	1
Biosynthesis of unsaturated fatty acids	27	1	0,2476	1
TGF-beta signaling pathway	94	2	0,2484	1
Pancreatic secretion	102	2	0,2526	1
Apoptosis - multiple species	32	1	0,2582	1

Table S4. Top 25 GeneOntology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways among the top CpGs associated with PM_{10} exposure during the period of post-implantation. N, number of genes in the pathway; DM, number of differentially methylated genes in the pathway, P, p-value and FDR, false discovery rate.

GO TERM	N	DM	P	FDR
ventricular cardiac muscle cell action potential	15	3	0,0011	1
chemosensory behavior	5	2	0,0016	1
methionine metabolic process	6	2	0,0019	1
membrane repolarization during action potential	5	2	0,0022	1
negative regulation of mesenchymal cell proliferation	5	2	0,0028	1
astrocyte differentiation	7	2	0,0036	1
ketohexokinase activity	1	1	0,0043	1
voltage-gated potassium channel activity involved in cardiac muscle cell action potential repolarization	7	2	0,0049	1
thyroid hormone generation	11	2	0,0053	1
voltage-gated potassium channel activity involved in ventricular cardiac muscle cell action potential repolarization	8	2	0,0062	1
adenine transmembrane transporter activity	2	1	0,0062	1
potassium ion export across plasma membrane	9	2	0,0063	1
mitochondrial genome maintenance	11	2	0,0066	1
ATP-dependent 5'-3' DNA/RNA helicase activity	1	1	0,0075	1
male germ cell nucleus	15	2	0,0077	1
regulation of mesoderm development	1	1	0,0079	1
regulation of membrane repolarization	10	2	0,0081	1
hormone biosynthetic process	12	2	0,0081	1
membrane repolarization during ventricular cardiac muscle cell action potential	12	2	0,0082	1
histidine decarboxylase activity	1	1	0,0083	1
histidine metabolic process	1	1	0,0083	1
positive regulation of potassium ion transmembrane transport	12	2	0,0087	1
transverse filament	1	1	0,0089	1
lateral element assembly	1	1	0,0089	1
membrane repolarization during cardiac muscle cell action potential	9	2	0,0090	1

KEGG term	N	DE	P.DE	FDR
Non-small cell lung cancer	68	5	0,0035	1
Glycerophospholipid metabolism	98	4	0,0259	1
Mannose type O-glycan biosynthesis	23	2	0,0289	1
Fc gamma R-mediated phagocytosis	93	4	0,0372	1
Sulfur relay system	8	1	0,0485	1
Thyroid hormone synthesis	75	3	0,0716	1
Metabolic pathways	1492	20	0,0807	1
Valine, leucine and isoleucine degradation	48	2,5	0,0884	1
Thyroid cancer	37	2	0,0890	1
PD-L1 expression and PD-1 checkpoint pathway in cancer	89	3	0,0928	1
N-Glycan biosynthesis	50	2	0,1022	1
Gastric cancer	149	4	0,1286	1
Phosphatidylinositol signaling system	97	3	0,1391	1
Necroptosis	159	3	0,1401	1
Notch signaling pathway	53	2	0,1470	1
RNA transport	185	3	0,1492	1
Choline metabolism in cancer	98	3	0,1535	1
Inflammatory mediator regulation of TRP channels	100	3	0,1639	1
Histidine metabolism	22	1	0,1670	1
VEGF signaling pathway	59	2	0,1864	1
One carbon pool by folate	20	1	0,1868	1
Basal cell carcinoma	63	2	0,1870	1
Linoleic acid metabolism	29	1	0,1872	1
Pentose and glucuronate interconversions	34	1	0,1908	1
Mismatch repair	23	1	0,1940	1

Table S5. Study characterististics of the Piccoli Più subsample from the Turin (N=99) center analyzed within the EXPOsOMICS project.

	N	Mean (SD) or %
Maternal age	99	34.1 (4.4)
Maternal education		
High	50	50.5%
Medium	41	41.4%
Low	8	8.1%
Maternal pre-pregnacy BMI	99	22.6 (3.9)
Parity		
Nulliparous	45	45.5%
Multiparous	54	54.6%
Smoking in pregnancy		
No	78	78.8%
Yes	21	21.2%
Child's sex		
Female	54	54.6%
Male	45	45.5%
PM10 estimates		
Whole pregnancy	383	45.9 (9.6)
Pre-implantation	383	43.6 (25.8)
Implantation	383	43.6 (15.8)
Post-implantation	383	38.9 (24.6)
First trimester	383	39.0 (20.5)
Second trimester	383	43.6 (21.5)
Third trimester	383	55.0 (22.1)
Last month of pregnancy	383	57.1 (26.8)

Table S6. Testing our top findings in a subsample from the Turin center of the Piccoli Più cohort analysed within the EXPOsOMICS project to assess whether they show consistent direction of association in the corresponding window of exposure. Fifty-eight CpGs were available for testing. Asterisk (*) indicates p-value < 0.1 and two asterisks (**) p-value<0.05.

ProbeID	Discovery dataset		Pre-impl			Impl			Post-impl			T1			T2			T3			M1			Preg		
	Window of exposure	Dir	Dir	p-val		Dir	p-val		Dir	p-val		Dir	p-val		Dir	p-val		Dir	p-val		Dir	p-val		Dir	p-val	
cg09634226	Implantation and post-implantation	+	+	0,32		+	0,91		+	0,99		+	0,72		+	0,82		+	0,76		+	0,50		+	0,86	
cg00493609	Implantation	-	-	0,88		+	0,65		-	0,65		-	0,95		+	0,28		+	0,94		-	0,54		+	0,27	
cg01200186	Implantation	-	-	0,09	*	-	0,14		-	0,78		-	0,60		+	0,51		-	0,83		-	0,44		+	0,95	
cg02283353	Implantation	-	+	0,11		+	0,24		+	0,35		+	0,62		-	0,01	**	-	0,34		+	0,86		-	0,03	**
cg03175206	Implantation	-	+	0,53		+	0,69		+	0,48		+	0,08	*	+	0,49		-	0,09	*	-	0,21		-	0,96	
cg04933990	Implantation	+	-	0,10		-	0,32		-	0,39		-	0,69		+	0,71		+	0,49		+	0,67		+	0,66	
cg06469440	Implantation	-	+	0,63		+	0,93		-	0,86		-	0,74		-	0,26		-	0,50		-	0,79		-	0,17	
cg06527535	Implantation	-	+	0,23		+	0,04	**	+	0,14		+	0,10	*	-	0,42		-	0,25		-	0,54		-	0,29	
cg07485754	Implantation	+	-	0,44		-	0,26		-	0,12		-	0,01	**	-	0,41		+	0,20		+	0,09	*	-	0,19	
cg08454563	Implantation	-	+	0,89		+	0,44		+	0,37		+	0,68		-	0,51		-	0,31		-	0,37		-	0,23	
cg08563994	Implantation	+	+	0,17		+	0,77		+	0,47		+	0,76		-	0,47		+	0,72		+	0,37		-	0,75	
cg09153950	Implantation	-	+	0,18		+	0,01	**	+	0,01	**	+	0,08	*	-	0,53		-	0,21		-	0,85		-	0,88	
cg09195657	Implantation	-	+	0,40		+	0,52		+	0,82		-	0,95		-	0,25		+	0,63		+	0,16		-	0,90	
cg09371530	Implantation	-	-	0,38		-	0,17		-	0,07	*	-	0,14		+	0,40		+	0,32		+	0,74		+	0,59	
cg10026495	Implantation	-	-	0,01	**	-	0,01	**	-	0,00	**	-	0,00	**	+	0,16		+	0,00	**	+	0,03	**	+	0,31	
cg10409786	Implantation	+	-	0,93		+	0,73		+	0,78		+	0,36		+	0,91		-	0,38		-	0,05	**	-	0,64	
cg10472067	Implantation	+	+	0,84		+	0,13		+	0,40		+	0,19		-	0,82		-	0,31		-	0,11		-	0,48	
cg13289118	Implantation	-	+	0,08	*	+	0,00	**	+	0,05	*	+	0,12		-	0,55		-	0,05	**	-	0,56		-	0,91	
cg13474011	Implantation	+	-	0,69		+	0,21		+	0,23		+	0,33		+	0,83		-	0,30		-	0,48		+	0,75	
cg13583523	Implantation	-	-	0,68		+	0,87		+	0,71		+	0,28		+	0,24		-	0,80		-	0,32		+	0,20	
cg13696148	Implantation	-	-	0,46		-	0,26		-	0,10	*	-	0,33		+	0,64		+	0,28		+	0,18		+	0,47	
cg14334911	Implantation	+	+	0,11		-	0,73		+	0,88		-	0,56		-	0,06	*	+	0,98		+	0,44		-	0,08	*
cg14468007	Implantation	-	+	0,18		+	0,00	**	+	0,04	**	+	0,11		-	0,11		-	0,03	**	-	0,10	*	-	0,07	*
cg17806989	Implantation	-	-	0,22		-	0,11		-	0,09	*	-	0,06	*	+	0,21		+	0,16		+	0,87		+	0,57	
cg18241780	Implantation	+	-	0,12		-	0,05	**	-	0,12		-	0,82		+	0,01	**	+	0,64		-	0,35		+	0,02	**
cg18447143	Implantation	-	+	0,99		-	0,29		-	0,98		-	0,67		+	0,99		+	0,84		-	0,67		+	0,97	
cg18798495	Implantation	-	-	0,84		-	0,61		-	0,49		-	0,87		+	0,79		+	0,74		-	0,50		+	0,58	

cg18917544	Implantation	+	+	0,09	*	+	0,76		-	0,86		-	0,41		-	0,13		+	0,40		+	0,06	*	-	0,38	
cg18980939	Implantation	-	+	0,24		+	0,14		+	0,06	*	+	0,41		-	0,06	*	-	0,35		-	0,77		-	0,07	*
cg19222405	Implantation	-	-	0,53		-	0,75		-	0,53		-	0,26		-	0,54		+	0,40		+	0,18		-	0,26	
cg21683659	Implantation	-	-	0,81		-	0,72		-	0,97		-	0,25		-	0,28		+	0,80		+	0,18		-	0,29	
cg24583704	Implantation	-	+	0,55		-	0,97		-	0,97		-	0,29		-	0,11		+	0,71		+	0,13		-	0,11	
cg25198579	Implantation	+	+	0,62		+	0,54		+	0,48		+	0,62		+	0,91		-	0,63		-	0,86		+	0,98	
cg27343976	Implantation	-	-	0,01	**	-	0,00	**	-	0,00	**	-	0,00	**	+	0,08	*	+	0,02	**	+	0,24		+	0,12	
cg00121374	Post-implantation	-	+	0,83		+	0,21		+	0,84		+	0,61		+	0,08	*	-	0,92		-	0,84		+	0,16	
cg00325528	Post-implantation	-	+	0,43		+	0,71		+	0,93		-	0,87		-	0,94		-	0,33		-	0,40		-	0,53	
cg00903584	Post-implantation	+	+	0,24		+	0,01	**	+	0,00	**	+	0,00	**	+	0,79		-	0,04	**	-	0,07	*	+	0,41	
cg01971227	Post-implantation	+	-	0,27		-	0,62		-	0,49		-	0,93		+	0,15		+	0,41		-	0,46		+	0,33	
cg02603756	Post-implantation	-	-	0,93		-	0,92		-	0,79		-	0,34		+	0,32		+	0,21		+	0,16		+	0,17	
cg02961575	Post-implantation	+	+	0,11		+	0,10		+	0,22		+	0,07	*	-	0,69		-	0,11		-	0,56		-	0,88	
cg07390478	Post-implantation	-	+	0,76		-	0,28		-	0,63		-	0,63		-	0,37		+	0,38		+	0,08	*	-	0,70	
cg09968723	Post-implantation	+	+	0,21		+	0,04	**	+	0,07	*	+	0,24		-	0,16		-	0,05	**	-	0,36		-	0,09	*
cg11691189	Post-implantation	-	-	0,51		-	0,63		-	0,42		-	0,18		-	0,90		+	0,59		+	0,91		-	0,40	
cg12876838	Post-implantation	-	+	0,48		-	0,88		+	0,78		+	0,59		+	0,62		+	0,74		+	0,80		+	0,21	
cg14565439	Post-implantation	-	+	0,43		+	0,60		+	0,78		+	0,45		+	0,88		-	0,78		+	0,90		+	0,90	
cg16624646	Post-implantation	+	+	0,35		+	0,39		+	0,39		+	0,26		-	0,16		-	0,27		-	0,85		-	0,35	
cg18526161	Post-implantation	+	+	0,22		+	0,54		+	0,74		-	0,18		-	0,01	**	+	0,56		+	0,03	**	-	0,04	**
cg18857759	Post-implantation	-	-	0,97		+	0,27		+	0,21		+	0,89		-	0,32		-	0,51		-	0,56		-	0,15	
cg19140429	Post-implantation	+	+	0,01	**	+	0,02	**	+	0,04	**	+	0,02	**	-	0,13		-	0,05	**	-	0,43		-	0,22	
cg20806143	Post-implantation	-	+	0,90		+	0,11		+	0,35		+	0,16		+	0,50		-	0,13		-	0,02	**	-	0,82	
cg21222426	Post-implantation	+	-	0,33		-	0,46		+	0,68		+	0,87		+	0,60		+	0,61		-	0,85		+	0,19	
cg21487550	Post-implantation	+	+	0,55		+	0,78		+	0,19		+	0,23		-	0,95		-	0,24		-	0,36		-	0,94	
cg21797500	Post-implantation	-	+	0,29		+	0,23		+	0,22		+	0,16		+	0,96		-	0,37		-	0,48		-	0,92	
cg22223441	Post-implantation	-	+	0,20		+	0,22		+	0,27		+	0,30		-	0,45		-	0,17		-	0,43		-	0,44	
cg22810059	Post-implantation	-	-	0,11		-	0,14		-	0,08	*	-	0,05	**	+	0,11		+	0,30		+	0,62		+	0,36	
cg23471777	Post-implantation	+	-	0,88		+	0,76		+	0,69		-	0,62		+	0,58		-	0,53		-	0,98		+	0,45	
cg06230418	Month before birth	-	+	0,29		+	0,04	**	+	0,19		+	0,35		-	0,29		-	0,13		-	0,60		-	0,14	
cg26030063	Month before birth	+	+	0,16		+	0,05	*	+	0,08	*	+	0,15		-	0,18		-	0,35		-	0,76		-	0,21	

The estimates are adjusted for technical variables, maternal age, education, parity, pre-pregnancy BMI, smoking in pregnancy, child's sex and cell heterogeneity. Pre-impl, imp, post-impl, T1, T2, T3, M1 and preg: pre-implantation, implantation, post-implantation, first, second and third trimester, last month of pregnancy and entire pregnancy respectively.

Figure S3. The density plot showing the distribution of the normalized beta values for each sample across all probes included after quality control and filtering (N=770128).

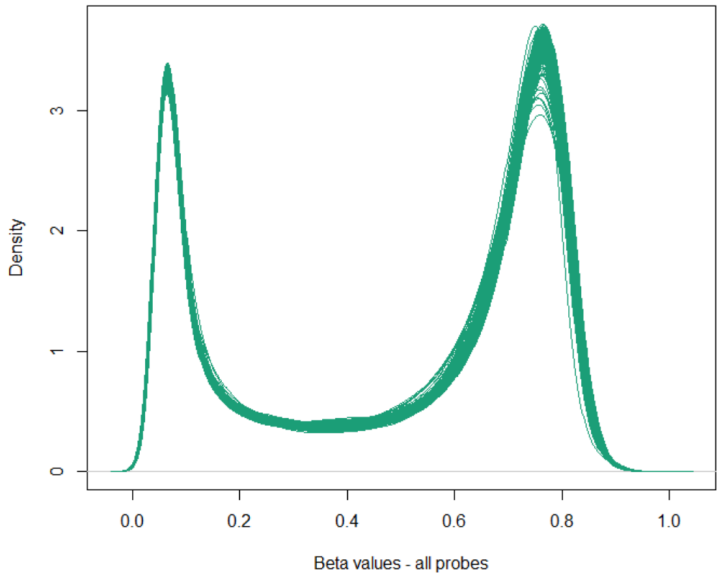


Figure S4. The density plot showing the distribution of the normalized beta values for each sample across different genomic regions: a) CpG islands (number of included probes after quality control and filtering, n=146684, 19.0%), b) shore regions (n=139339, 18.1%), c) shelf regions (n=53062, 6.9%) and d) open sea region (n=431043, 60.0%)

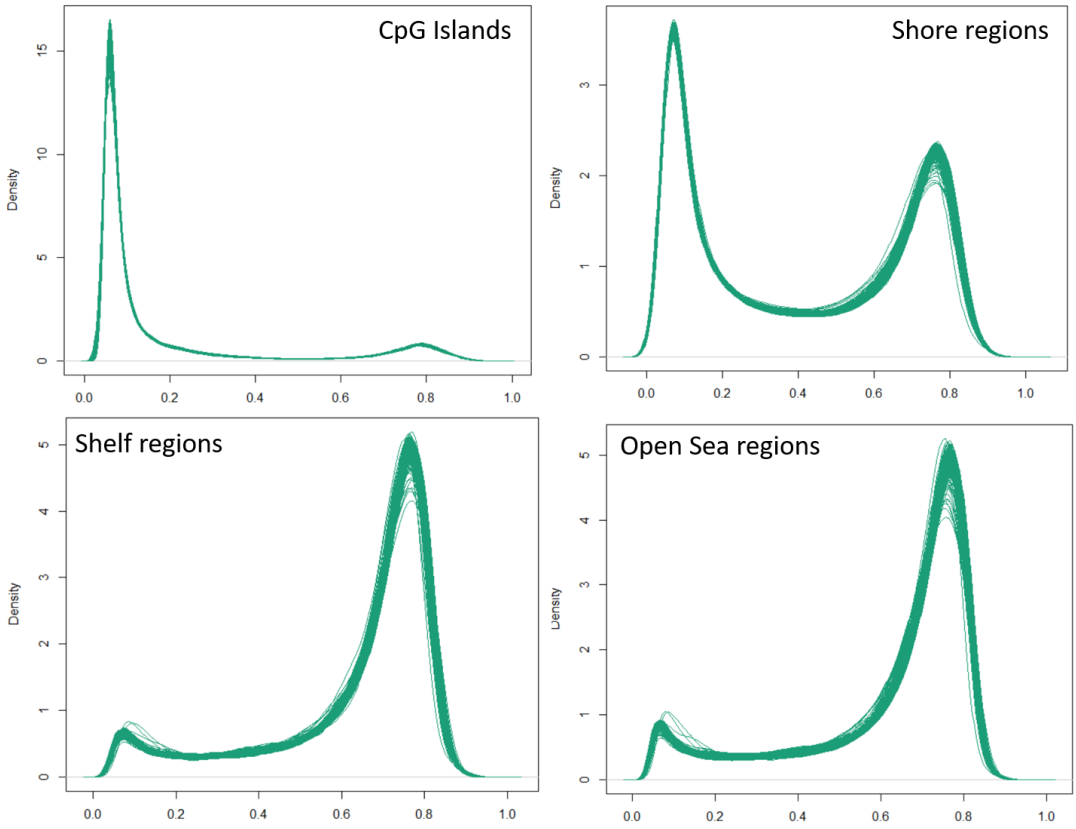


Figure S5. Box plot showing the median methylation across different genomic regions

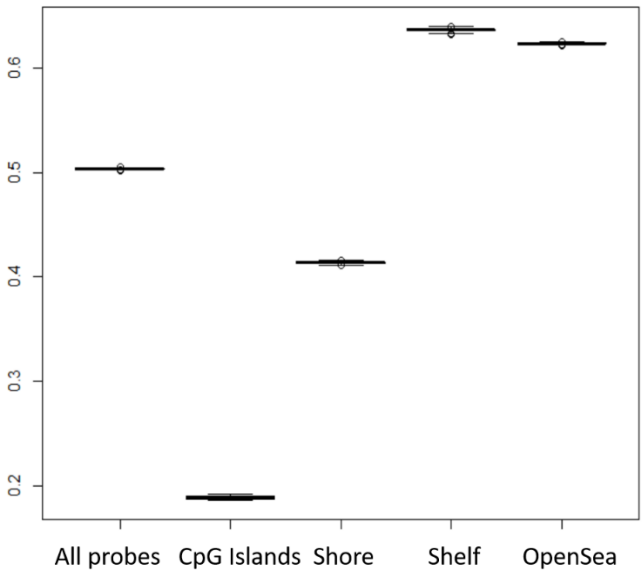


Table S7. Association between ambient air pollution exposure to PM_{10} and measures of overall methylation across all probes and across probes in different genomic regions.

	Linear regression (mean DNAm)				GAMP	
	Coef	95% CI		p-val	CDF	Density
Whole pregnancy		Lower	Upper		p-val	p-val
All probes	5.0E-05	-1.6E-05	1,2E-04	0,136	0,429	0,554
CpG islands	-8.4E-05	-1.7E-04	6,4E-06	0,068	0,071	0,122
Shore	5.3E-05	-6.8E-05	1,7E-04	0,387	0,414	0,153
Shelf	5.5E-05	-1.3E-04	2,4E-04	0,565	0,381	0,283
Open Sea	9.4E-05	2.8E-05	1,6E-04	0,005	0,421	0,326
Pre-implantation						
All probes	-1.2E-06	-2.5E-05	2.3E-05	0.919	0.567	0.547
CpG islands	9.4E-06	-2.3E-05	4.2E-05	0.572	0.602	0.581
Shore	2.7E-06	-4.1E-05	4.6E-05	0.904	0.825	0.526
Shelf	-3.9E-05	-1.1E-04	2.9E-05	0.264	0.676	0.657
Open Sea	-1.5E-06	-2.6E-05	2.3E-05	0.901	0.457	0.366
Implantation						
All probes	1.0E-05	-1.5E-05	3.6E-05	0.432	0.736	0.442
CpG islands	2.4E-05	-1.1E-05	5.9E-05	0.179	0.368	0.398
Shore	2.3E-05	-2.4E-05	6.9E-05	0.338	0.764	0.355
Shelf	-4.2E-05	-1.1E-04	3.1E-05	0.258	0.795	1.000
Open Sea	7.8E-06	-1.8E-05	3.4E-05	0.550	0.593	0.552
Post-implantation						
All probes	-5.6E-06	-2.9E-05	1.8E-05	0.638	0.221	0.461
CpG islands	1.7E-05	-1.5E-05	4.9E-05	0.288	0.497	0.596
Shore	-1.6E-05	-5.8E-05	2.7E-05	0.475	0.311	0.655
Shelf	-7.1E-05	-1.4E-04	-4.5E-06	0.036	0.124	0.237
Open Sea	-2.2E-06	-2.6E-05	2.1E-05	0.855	0.236	0.317
First trimester						
All probes	-2.0E-06	-3.2E-05	2.8E-05	0.894	0.690	0.949
CpG islands	7.9E-06	-3.4E-05	5.0E-05	0.708	0.636	0.560
Shore	-7.8E-06	-6.3E-05	4.8E-05	0.782	1.000	1.000
Shelf	-6.8E-05	-1.5E-04	1.8E-05	0.120	0.445	0.852
Open Sea	4.5E-06	-2.6E-05	3.5E-05	0.771	0.626	1.000
Second trimester						
All probes	1.2E-05	-1.7E-05	4.2E-05	0.411	0.500	0.425
CpG islands	-4.5E-05	-8.6E-05	-4.6E-06	0.029	0.056	0.105
Shore	1.8E-05	-3.7E-05	7.3E-05	0.515	0.599	0.235
Shelf	2.0E-05	-6.5E-05	1.1E-04	0.641	0.464	0.290
Open Sea	2.9E-05	-6.7E-07	6.0E-05	0.055	0.444	0.244
Third trimester						
All probes	2.0E-05	-1.0E-05	4.9E-05	0.194	0.347	0.659
CpG islands	-1.3E-05	-5.4E-05	2.7E-05	0.520	0.766	0.881
Shore	2.2E-05	-3.3E-05	7.6E-05	0.431	0.420	0.658
Shelf	7.4E-05	-1.1E-05	1.6E-04	0.088	0.201	0.224
Open Sea	2.3E-05	-6.5E-06	5.3E-05	0.124	0.384	0.407
Last month of pregnancy						
All probes	-3.1E-06	-2.8E-05	2.2E-05	0.807	0.828	1.000
CpG islands	-4.0E-06	-3.9E-05	3.1E-05	0.820	1.000	0.914
Shore	-1.8E-05	-6.4E-05	2.8E-05	0.444	0.592	0.406
Shelf	-9.2E-06	-8.2E-05	6.3E-05	0.802	0.684	0.654
Open Sea	2.7E-06	-2.3E-05	2.8E-05	0.835	0.780	0.712

All analyses were adjusted for technical variables. study center. maternal age. education. parity. pre-pregnancy BMI. smoking in pregnancy and child's sex. The second analysis was additionally adjusted for cell heterogeneity. The confidents represent change in DNA methylation for $10\mu g/m^3$ in PM_{10}.

Table S8. Association between prenatal exposure to PM_{10} and telomere length. the coefficients are adjusted for cell heterogeneity.

	N=365				N=324 (samples with CV>20 were excluded)			
	Coef	95% CI		p value	Coef	95% CI		p value
		Lower	Upper			Lower	Upper	
All pregnancy	-0.0010	-0.0171	0.0150	0.901	0.0067	-0.0096	0.0230	0.420
Pre-implantation	-0.0047	-0.0104	0.0010	0.105	-0.0025	-0.0083	0.0034	0.407
Implantation	-0.0044	-0.0103	0.0015	0.142	-0.0022	-0.0083	0.0040	0.488
Post-implantation	-0.0048	-0.0104	0.0009	0.098	-0.0012	-0.0072	0.0047	0.687
First trimester	-0.0053	-0.0127	0.0021	0.157	-0.0007	-0.0085	0.0070	0.850
Second trimester	0.0020	-0.0053	0.0092	0.593	0.0042	-0.0031	0.0116	0.259
Third trimester	0.0038	-0.0035	0.0111	0.306	0.0020	-0.0055	0.0095	0.598
Last month of pregnancy	0.0009	-0.0053	0.0071	0.771	-0.0010	-0.0076	0.0056	0.765

Adjusted for technical variables, study center, maternal age, education, parity, pre-pregnancy BMI, smoking in pregnancy, child's sex and cell heterogeneity. The confidents represent change in telomere length for $10\mu g/m^3$ in PM_{10}.

Study III Maternal antibiotic use and vaginal infections in the third trimester of pregnancy and the risk of obesity in preschool children

[Pediatric Obesity. 2020;15. e12632. doi: 10.1111/ijpo.12632]

Isaevska E[1], Popovic M[1], Pizzi C[1], Fiano V[1], Rusconi F[2], Merletti F[1], Richiardi L[1], Maule M[1]

1 Department of Medical Sciences, University of Turin, CPO Piemonte, Turin, Italy
2 Unit of Epidemiology, "Anna Meyer" Children's University Hospital, Florence, Italy

Funding: M Maule and L Richiardi received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 774548, STOP project (Science and Technology in childhood Obesity Policy) Bando ex-post di progetti di ricerca di Ateneo - Anno 2018

Abstract

Background: Several exposures during pregnancy are associated with offspring BMI. The objective of this study was to evaluate whether third trimester antibiotic use and vaginal infections are associated with BMI in preschool children.

Methods: The study population included singletons from the NINFEA birth cohort with available anthropometric measurements at the age of 4 (3151 born with vaginal and 1111 born with cesarean delivery). Self-reported use of antibiotics and the presence of vaginal infection in the third trimester were analyzed in association with the child's BMI, classified into 3 categories: thinness, normal and overweight/obesity, using both the International Obesity Task Force (IOTF) and the World Health Organization (WHO) recommended cut-offs.

Results: Maternal vaginal infections in the third trimester of pregnancy were associated with higher relative risk ratios (RRR) for overweight/obesity at age of 4 in children delivered vaginally: 1.92 (95% confidence interval (CI): 1.37 to 2.70). This association appeared stronger for children born to women with pre-pregnancy BMI>25 kg/m2 (RRR: 4.78; 95% CI 2.45 to 9.35), and was robust when different obesity cut-offs were used. The results regarding third trimester antibiotic use in vaginal deliveries were less conclusive (RRRs for overweight/obesity: 1.43 (0.92 to 2.21) and 1.11 (0.57 to 2.20), for the IOTF and WHO cut-offs, respectively). Third trimester vaginal infections were not associated with BMI in children delivered by cesarean section.

Conclusion: Maternal third trimester vaginal infections are associated with an increased overweight/obesity risk in children born by vaginal delivery, and especially in children of mothers with pre-pregnancy overweight/obesity.

Introduction

Obesity in children is an important public health challenge due to its high prevalence, complex etiology and long-term consequences.[1,2] According to the Developmental Origins of Health and Disease hypothesis (DOHaD) the risk for many metabolic diseases, including obesity, can originate during early development.[3] Exposures such as high pre-pregnancy BMI, excess gestational weight gain and exposure to tobacco smoke during pregnancy have already been identified as early risk factors for obesity in children.[4,5]

Due to the emerging role of altered human gut microbiota in many metabolic and inflammatory diseases, especially obesity, we focused on two common gestational exposures closely linked[6] with the maternal and initial neonatal microbiota: antibiotics and vaginal infections in the third trimester of pregnancy. The first major microbial colonization of the newborn happens during the birthing process and maternal antibiotic use or vaginal infections in pregnancy may not only disturb the maternal microbiota but also the initial colonizing microbiota of the newborn. Antibiotics are among the most commonly prescribed drugs in pregnancy[7,8] and their administration in many circumstances is both effective and life-saving. However, undesired effects on the long-term health of the offspring are also linked with their use. [9] The association between gestational antibiotic exposure and obesity in children is not well understood, possibly due to the complex interplay between many important factors that exert their effect in this period, with roles difficult to disentangle.[10] Some of the studies on this topic reported positive associations between antibiotic exposure in some trimesters, or certain types of antibiotics and overweight or obesity, while others studies reported null associations. [11–17] Most of the studies on antibiotics either used BMI as a continuous variable or included children who are underweight in the reference category, potentially attenuating the association with overweight/obesity. Vaginal dysbiosis in pregnancy is linked with maternal and fetal morbidity, in particular with increased risk for chorionamnionitis and premature rupture of the membranes[18] and preterm birth[19,20], a risk factors for later metabolic diseases.[21] To date, there are no studies that explore the possible association between vaginal infections during pregnancy and childhood BMI outcomes.

Since the identification of early modifiable risk factors for metabolic diseases can help promote childhood health as early as from pregnancy, our aim was to evaluate, within the framework of a mother-child cohort study, whether antibiotic use and vaginal infections in the last trimester of pregnancy are associated with BMI in pre-school children. To address

possible gaps in the literature and explore the association with both sides of the BMI spectrum, we modeled child's BMI as categorical variable with three categories (thinness, normal weight and overweight/obesity). To account for the fact that all mothers undergoing cesarean delivery are pretreated with antibiotics, and therefore, all children could be considered as exposed, our main analyses are focused only on children born by vaginal delivery.

We focused on exposures occurring during the third trimester of pregnancy because, from a theoretical point of view, we hypothesized that exposures that alter the maternal microbiome closer to the time of delivery might be the most relevant ones for obesity risk, and from a technical point of view, information on the exposures in the third trimester was the most complete, as detailed in the following Methods section.

Methods

1. Study population

The NINFEA study (Nascita e Infanzia: gli Effetti dell'Ambiente; Birth and Childhood: Effects of the Environment) is an Italian internet-based mother-child cohort (https://www.progettoninfea.it/index_en - n1) set up to investigate exposures during pre-natal and early post-natal life that may affect health later in life.[22,23] Members of the cohort are children born to women who have access to the Internet and have enough knowledge of the Italian language to complete the online questionnaires. Since 2005 approximately 7500 pregnant women were recruited and completed the baseline questionnaire at any time during the pregnancy. The follow-up questionnaires are completed at 6 and 18 months after delivery and when the child turns 4, 7, 10 and 13 years. For this study we used the 2019.11 version of the NINFEA database, in which 4841 children were followed up at age of 4.

The study population included 4262 singletons with available height and weight measurements at the age of 4. Of them 3151 were born with vaginal and 1111 with cesarean delivery (Supplementary Figure S1). The NINFEA study was approved by the Ethical Committee of the San Giovanni Battista Hospital and CTO/CRF/Maria Adelaide Hospital of Turin (approval N.0048362 and following amendments) and all the participants gave written informed consent at enrolment.

2. Exposure, outcome and confounding variables

We used the data from the second questionnaire completed 6 months after delivery, covering third trimester exposures, in order to obtain information on antibiotic use and the occurrence of vaginal infections in the first and third trimester of pregnancy. We focused on third trimester exposures due to the larger sample size. The first questionnaire covering first and secondtrimester exposures can be completed at any point in the pregnancy and thus would require exclusion of all participants who completed the questionnaire before the end of the corresponding trimesters, leading to reduction in sample size. In the second questionnaire, mothers were asked to complete two separate pre-specified checklists, one for medication use and one for pathological conditions. Therefore, antibiotic use was defined as any antibiotic taken by the mother independently of indication, while vaginal infection was defined as any vaginal infection.

Child weight and height measurements at approximately 4 years of age were reported by mothers at the 4-year follow-up questionnaire. The offspring body mass index (BMI) at age of 4 years was calculated using the standard formula, weight in kilograms divided by height in meters squared (kg/m^2). Since we modeled BMI as a categorical variable with three categories (thinness, normal weight and overweight/obesity) and differences in prevalence are known to exists when different definitions are used, we decided to use two cut-offs: the ones proposed by the International Obesity Task Force (IOTF)[24] and the World Health Organization (WHO).[25] The IOTF cut-offs are based on and linked to the corresponding adult BMI cut-offs at age 18 of the IOTF reference population and classify children into six categories thinness grade 3, 2 and 1, normal BMI, overweight and obesity. The WHO Child Growth Standards (0-5 years) on the other hand, are based on traditional z-scores. The z scores recommended to define thinness and overweight/obesity for children under the age of five are -2 and 2. The WHO z score of -2 closely corresponds to the cut-off for thinness level 2 in IOTF. Therefore, for comparability reasons for the main IOTF analyses we used the age and sex-specific cut-off for thinness grade 2 to define thinness and included thinness grade 1 in the normal BMI category. We also conducted sensitivity analyses classifying thinness grade 1 together with thinness grade 2 and 3.

Maternal age (continuous), maternal education (primary school or less, secondary education, and university degree or higher), parity (yes, no), maternal pre-pregnancy BMI (continuous), smoking during pregnancy (ever smoked, never) and gestational diabetes were pre-selected as

potential confounders. We also adjusted for gestational age to account for different pregnancy length and, therefore, the period in which the child could have been exposed prenatally. Although antibiotics can be used to treat vaginal infections, we not to adjust for antibiotic use when studying vaginal infections, because in this case the antibiotic would act as a mediator in the association, and not a confounder.[26]

Children born with cesarean delivery are pretreated with intrapartum antibiotic prophylaxis and therefore for our main analyses we focused only on vaginal deliveries and additionally performed the analyses regarding vaginal infections in cesarean deliveries.

3. Statistical analyses

We used a complete case analyses approach, excluding all subjects with missing data in any of the variables of interest. For the main statistical analyses, we used multinomial logistic regression, an extension of logistic regression, used when the dependent variable is nominal with more than two categories. The exponentiated coefficients provide an estimate of relative risk and are presented as relative risk ratios (RRRs) and 95% confidence intervals (CIs). RRR represents the relative risk of thinness or overweight at age 4 (compared to normal BMI) among children of exposed vs unexposed mothers. We further investigated whether the studied associations are modified by maternal pre-pregnancy BMI (categorized as ≤25 and >25) and sex of the child. All the analyses were performed using STATA version 15 (STATA Corp., Texas, USA).

Results

In our population of children born with vaginal delivery, the prevalence of antibiotic use in the third trimester was close to 6%, while the prevalence of vaginal infections was slightly higher, around 9%. We observed notable differences in the prevalence of thinness and overweight (that includes obesity) when different cut-offs were used, Figure 1.

According to the WHO z-score cut-offs, 4.5% of the NINFEA children were classified in the thinness category, while 7.7% were classified as thinness grade 3 or 2 according to the IOTF cut-offs. Around 15% of the NINFEA children were classified as IOTF thinness grade 1, but fell within the normal BMI limits according to WHO cut-offs. The prevalence of overweight according to WHO and IOTF was 5.3% and 12%, respectively. Other characteristics of the population are summarized in Table 1. The corresponding percentages regarding cesarean deliveries are reported in Supplementary Table S1.

Figure 1. Comparison between the IOTF and WHO classification in NINFEA children aged 4 born with vaginal delivery.

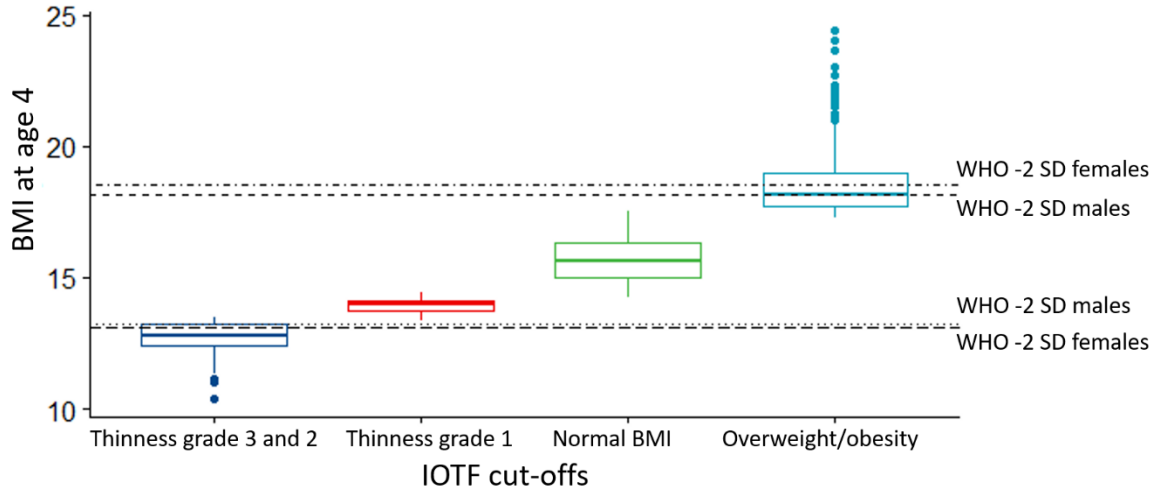


Table 1. Characteristics of mother-child dyads born with vaginal delivery and with available height and weight measures at age 4 (N=3151)

	N	Mean (SD) or %
Maternal age at delivery	3151	33.5 (4.1)
Maternal pre-pregnancy BMI	3100	22.2 (3.6)
Missing values	51	–
Maternal education		
Primary school or less	112	3.6%
Secondary school	985	31.5%
University degree or higher	2035	64.8%
Missing values	19	–
Gestational diabetes		
No	2707	91.5%
Yes	253	8.6%
Missing values	191	–
Smoked in pregnancy		
No	2876	92.4%
Yes	237	7.6%
Missing values	38	–
First pregnancy		
No	858	27.8%
Yes	2227	72.8%
Missing values	66	–
Antibiotic use in third trimester		
No	2820	94.1%
Yes	177	5.9%
Missing values	154	–
Vaginal infections in third trimester		
No	2733	91.2%
Yes	264	8.8%
Missing values	151	–
Child's BMI (IOTF cut-offs)		
Thinness grade 3 and 2	242	7.7%
Thinness grade 1	478	15.2%
Normal BMI	2053	65.2%

Overweight/obesity	378	12%
Child's BMI (WHO cut-offs)		
Thinness	142	4.5%
Normal BMI	2843	90.2%
Overweight/obesity	166	5.3%

For this study, we used the complete case analysis approach and excluded 10% of the subjects due to missing data, leaving a total of 2837 children in our main analyses. The percentage of missing data was below 6% for all explanatory variables. There were no differences in the baseline characteristics between included and excluded subjects, with the exception of maternal smoking in pregnancy, which was more common in the excluded group (7.2% vs. 12.0 %, p-value 0.004). The characteristics of the included and excluded subjects are summarized in Supplementary Table S2.

The prevalence of both thinness and overweight/obesity was smaller by WHO standards and thus the analyses were more limited in power, in comparison to those by IOTF. The relative risk of being overweight/obese at 4 years of age in children whose mothers had vaginal infection during the third trimester *vs* those did not was 1.92 (1.37 to 2.70) (Table 2). The strength of association did not change when different overweight/obesity cut-off was used, Supplementary Table S3.

Table 2. Associations of antibiotic use and vaginal infections in the third trimester with BMI outcomes at age of 4 in NINFEA children born with vaginal delivery. International Obesity Task Force cut-offs were used to define thinness and overweight/obesity. Thinness grade 1 is classified as normal BMI.

	Cases/exposed cases (%)	RRRcrude (95% CI)	RRRadj* (95% CI)
Antibiotic use	170/2837 (6.0)		
Thinness	14/223 (6.3)	1.11 (0.63 to 1.96)	1.10 (0.62 to 1.95)
Overweight/obesity	26/328 (7.9%)	1.43 (0.92 to 2.21)	1.40 (0.90 to 2.16)
Vaginal infections	251/2837 (8.8)		
Thinness	19/223 (8.5)	1.07 (0.65 to 1.76)	1.06 (0.64 to 1.74)
Overweight/obesity	49/328 (14.9)	2.01 (1.43 to 2.85)	1.92 (1.37 to 2.70)

Data are presented as n, n(%) and Relative Risk Ratios (RRR) and 95% confidence intervals (CI). *Adjusted for: maternal age, education, parity, pre-pregnancy BMI, smoking during pregnancy, gestational diabetes and gestational age (only for the third trimester analyses)

The association between third trimester antibiotic use and overweight/obesity was less robust. The corresponding RRRs for antibiotic use were 1.43 (0.92 to 2.21) and 1.11 (0.57 to 2.20), for the IOTF (Table 2) and WHO analysis (Supplementary Table S3), respectively.

We did not observe an association between antibiotic use or vaginal infections with thinness at age 4. The estimates only slightly changed when thinness grade 1 (15% of all children) was included in the thinness category in the IOTF analyses (data not shown).

We found that maternal pre-pregnancy BMI modified the association of vaginal infections with childhood overweight and obesity (p-value for multiplicative interaction 0.002). The stratified analysis showed that the association was particularly pronounced in the category of overweight/obese mothers with RRR of 4.78 (2.45 to 9.35) (Table 3). There was no evidence of effect modification for child sex.

Table 3. Pre-pregnancy BMI as a potential effect modifier in the association between vaginal infections in the third trimester and overweight/obesity at age 4 in NINFEA children born with vaginal delivery

	RRRs (95% CI) for each stratum of pre-pregnancy BMI and vaginal infection status with a single reference category				RRRs (95% CI) for vaginal infection in the strata of pre-pregnancy BMI
	Vaginal infection=0		Vaginal infection=1		
Pre-pregnancy BMI	N	RRR (95% CI)	N	RRR (95% CI)	
≤ 25	208/234	1.00	26/234	1.34 (0.86 to 2.08)	1.33 (0.86 to 2.07)
> 25	71/94	1.64 (1.22 to 2.22)	23/94	7.5 (4.01 to 13.97)	4.78 (2.45 to 9.35)

All RRRs are adjusted for maternal age, education, parity, pre-pregnancy BMI, smoking during pregnancy, gestational diabetes and gestational age

Due to the slightly high power, the analyses regarding maternal vaginal infections and BMI outcomes in cesarean deliveries was performed only based on IOTF cut-offs. We found no evidence of association between vaginal infections during the third trimester and child's overweight and obesity (Supplementary Table S4). However, these analyses were performed on smaller number of subjects (N=1111) in comparison with those born with vaginal delivery (N=3151), the number of exposed cases was small and the confidence intervals quite wide.

Discussion

In the NINFEA cohort, vaginal infections in the third trimester of pregnancy were associated with overweight/obesity at age of 4 in children delivered vaginally. Maternal pre-pregnancy BMI modified the association between vaginal infections and child BMI. The findings regarding antibiotic use were less conclusive, although it seemed to be some indication of an association in our IOTF analysis.

We did not find a similar study in the literature to compare our findings on vaginal infections. They are one of the most common gynecological conditions and vaginal dysbiosis is known to be a risk factor for preterm birth and premature rupture of the membranes[18–20], indicating how unbalanced vaginal microbial communities can influence intrauterine growth and duration of gestation and increasing the risk for some conditions related to metabolic

syndrome.[21,27] However, in our cohort, the effect of vaginal infections is not likely to be explained by prematurity, since the percentage of preterm births is very small (3%).

It is known that altered human gut microbiota has a role in many metabolic, inflammatory pro-obesogenic pathways,[28] The first major microbial colonization of the newborn happens during the birthing process and maternal antibiotic use or vaginal infections in pregnancy may not only disturb the maternal microbiota but also the initial colonizing microbiota of the newborn. The skin, oral and nasopharyngeal orifices and gut of vaginally delivered infants are initially enriched in microbial communities that resembles the maternal vaginal microbiota [6,29] The vaginal microbiome in pregnant women is characterized by more stable, less rich and diverse communities dominated by Lactobacillus spp. that benefit both the mother and the child[30–32] and having maternal vaginal microbiota enriched with health promoting bacteria, could give competitive advantage to the composition of healthy non-obesogenic neonatal intestinal microflora. Due to the novelty of the findings, further research is needed to replicate our findings and explore possible mechanisms that underlie this association.

Furthermore, our analyses indicate that maternal pre-pregnancy BMI modifies the association between vaginal infections in the third trimester and the child's risk for overweight and obesity at age 4. It is known that maternal BMI increases the risk for adverse maternal and birth outcomes[4] as well as childhood obesity risk.[5] A recent study[33] that analyzed the first neonatal stool of neonates found that among the neonates born vaginally, those born to overweight and obese mothers (BMI≥25) had different gut microbiota structure compared with children of normal weight mothers, that ultimately could lead to metabolic differences later in childhood. As the same associations were not found in children born by Cesarean deliveries, the authors argued that the observed differences are likely the result of mother-to-offspring transmission of microbiota during childbirth. We also did not observe an association between vaginal infections and overweight/obesity in cesarean deliveries. However, our effect estimates are based on small number of cases and additional studies are needed with larger statistical power to provide evidence for lack of association.

Several studies explored the association between antibiotic use in pregnancy and elevated offspring BMI and while some found an increased risk of overweight and obesity, others did not observe an association. Some of the studies modeled BMI z scores and others used percentiles or IBTF cut-offs and the inconsistent strength of the effect estimates could be,

therefore, at least in part explained by different overweight and obesity cut-offs used. Muller et al.[11] reported a higher risk for obesity at age 7 after antibiotic use in the second and third trimester. Mor et al.[12] reported higher prevalence ratios for overweight and obesity in school-aged children (7-16 years). Cassidy-Bushrow et al.[14] found an increased risk of overweight/obesity and higher BMI z-score at age 2 after antibiotic use in the first two trimesters. Poulsen et al.[13] and Sejersen et al.[34] did not find an association between prenatal antibiotics and BMI z-scores at ages 1, 3 and 6 years. In addition, three recent large studies[15–17] (based on more than 40,000 mother-child dyads) presented reassuring results with effects going towards the null for most of the associations even when different trimester of exposure, dose-response, spectrum and type of antibiotic were considered. However, association with repeated use of antibiotics during pregnancy and exposure to broad spectrum antibiotics and obesity at age 7 was reported by Wang et al.[15] and Jess et al.[17], respectively.

The found associations could be due to underlying alterations in the infant gut microbiome[35] or could be confounded by maternal susceptibility to infection. Future studies should address other possible mechanisms such as immunologic, metabolomic or epigenetic pathways that might start their effect even before birth.

The results from previous studies are not directly comparable to ours, since only some of them performed stratified analyses by delivery mode, and none reported estimates for third trimester antibiotics use and overweight/obesity in vaginal deliveries. The effect estimate and 95% CIs for vaginal deliveries for ever antibiotic use in pregnancy and obesity in children aged 7-16 reported by Mor et al. was 1.34 (1.08 to 1.75). The estimate, however, was not adjusted for maternal BMI. The effect estimates for vaginal deliveries by Wang et al. and Jess et al. were 1.18 (1.02 to 1.38) for repeated use of antibiotics in pregnancy and obesity at age 7 and 1.08 (0.99 to 1.17) for ever antibiotic use in pregnancy and overweight/obesity at age of 7, respectively. Our study is conducted in a smaller study population with lower prevalence of antibiotic use and therefore our failure to reach more conclusive and robust results may be due to the limited power combined with the present, but relatively smaller, effect on antibiotics on overweight/obesity compared to vaginal infections.

Both exposures are self-reported, but this is common in epidemiological studies particularly with exposures such as antibiotic use/infections. The prevalence of antibiotics is only slightly lower in our study population (6% and 7% for vaginal and caesarean deliveries, respectively), compared to the 9% prevalence of antibiotic use in the third trimester for vaginal and

caesarean deliveries combined in the study by Jess et al.[17] based on more than 40 000 subjects and with antibiotic data coming from the Danish National Prescription Registry, that could be due to the fact that NINFEA, as in many other cohorts, participants mainly originate from a population with high educational attainment. However, it has been extensively shown that baseline selection does not imply biased estimates of the exposure-disease associations.[36,37] We also acknowledge the possibility of missclassification of exposure driven by under-reporting rather than over-reporting, that would translate into high specificity, such as reporting well-known antibiotics and symptomatic vaginal infections, rather that asymptomatic bacterial vaginosis.

Our outcome was measured prospectively, at 4 years of age, and we used parentally-reported height and weight measurements to calculate the BMI. NINFEA is an internet-based cohort and therefore only has data on parentally-reported weight and height. NINFEA birth weight data was previously linked with the Piedmont Birth Registry and the comparison showed a very high validity of the NINFEA birth weight data.[38] The questionnaire at age 4, uses the fact that children in Italy are regularly measured by health professionals and their measures are recorded in the children's personal health booklets. Mothers during the compilation of anthropometric measures were asked how the height and weight measurements were recalled, specifically whether they reported the measurements written in their child's booklets. At age 4, around 40% of the mothers reported written weight and height measures. The association with vaginal infections did not change when we restricted the analysis only on children with anthropometric measures reported from the booklets (RRR for vaginal infections 2.05, 95% CI 1.16 to 3.57 in vaginal deliveries). Several articles on the agreement between parentally-reported and measured height and weight in younger children (less than 10 years) were published that reported specificity ≥95% to accurately classify obesity and variable, but relatively lower, percentages of sensitivity,[39–42] that might lead to underestimation of obesity and produce estimates that bias towards the null. it is generally considered that non-differential exposure and disease misclassification is a greater concern in interpreting studies that seem to indicate the absence of an effect.[43]

Our study has several limitations. Unfortunately, we were not able to explore further the role of factors such as the type of vaginal infection, whether it was ever treated and the type of antibiotic used, nor the composition of the maternal vaginal or offspring's gut microbiota that might underlie the found associations. We also were not able to assess first and second trimester exposures due to smaller sample size.

We used the detailed information available in NINFEA to control for a number of carefully chosen covariates, including maternal BMI, maternal education and age, smoking, parity and gestational diabetes, factors that may strongly influence the associations of interest, and we found very little evidence of confounding. We modelled BMI as categorical variable including the thinness as a separate category and we carried out sensitivity analyses that took into account different BMI cut-offs. Furthermore, we assessed the presence of interaction by sex and maternal pre-pregnancy BMI, and found that the latter modified the effect of vaginal infections with offspring overweight/obesity. We were also able to examine the association between vaginal infections and BMI in children from our cohort born *via* cesarean section, although with lower power.

Conclusion

We reported a novel association regarding vaginal infections in pregnancy and later childhood BMI in children delivered vaginally. Further studies should replicate these findings. Maternal pre-pregnancy BMI seems to modify the association between vaginal infection in the third trimester and overweight/obesity at age 4 which might be an important point for preconception counseling. [4] Our results regarding antibiotics were less conclusive.

Obesity has a complex etiology and is influenced by variety of genetic, dietary and environmental factors. Identifying potential modifiable risk factors for childhood obesity as early as in preconception or pregnancy might have long term beneficial effect on the child's metabolic health. Future studies should take into careful consideration the complex interplay between maternal pre-pregnancy BMI, antibiotic use and vaginal infections.

Bibliography:

1. Abarca-Gómez L, Abdeen ZA, Hamid ZA, et al. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *Lancet*. 2017;390(10113):2627-2642. doi:10.1016/S0140-6736(17)32129-3
2. WHO. *Report of the Commision on Ending Childhood Obesity*.; 2016. http://www.who.int/end-childhood-obesity/news/launch-final-report/en/.
3. 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. doi:10.1056/NEJMra0708473
4. Voerman E, Santos S, Inskip H, et al. Association of Gestational Weight Gain With Adverse Maternal and Infant Outcomes. *JAMA*. 2019;321(17):1702. doi:10.1001/jama.2019.3820
5. Woo Baidal JA, Locks LM, Cheng ER, Blake-Lamb TL, Perkins ME, Taveras EM. Risk Factors for Childhood Obesity in the First 1,000 Days: A Systematic Review. *Am J Prev Med*.

2016;50(6):761-779. doi:10.1016/j.amepre.2015.11.012

6. Stiemsma LT, Michels KB. The Role of the microbiome in the developmental origins of health and disease. *Pediatrics*. 2018;141(4). doi:10.1542/peds.2017-2437

7. Palmsten K, Hernández-Díaz S, Chambers CD, et al. The Most Commonly Dispensed Prescription Medications Among Pregnant Women Enrolled in the U.S. Medicaid Program. *Obstet Gynecol*. 2015;126(3):465-473. doi:10.1097/AOG.0000000000000982

8. Stephansson O, Granath F, Svensson T, Haglund B, Ekbom A, Kieler H. Drug use during pregnancy in Sweden - assessed by the Prescribed Drug Register and the Medical Birth Register. *Clin Epidemiol*. 2011;3:43-50. doi:10.2147/CLEP.S16305

9. Popovic M, Rusconi F, Zugna D, et al. Prenatal exposure to antibiotics and wheezing in infancy: A birth cohort study. *Eur Respir J*. 2016;47(3):810-817. doi:10.1183/13993003.00315-2015

10. Cox LM, Blaser MJ. Antibiotics in early life and obesity. *Nat Rev Endocrinol*. 2015;11(3):182-190. doi:10.1038/nrendo.2014.210

11. Mueller NT, Whyatt R, Hoepner L, et al. Prenatal exposure to antibiotics, cesarean section and risk of childhood obesity. *Int J Obes*. 2015;39(4):665-670. doi:10.1038/ijo.2014.180.Prenatal

12. Mor A, Antonsen S, Kahlert J, et al. Prenatal exposure to systemic antibacterials and overweight and obesity in Danish schoolchildren: A prevalence study. *Int J Obes*. 2015;39(10):1450-1455. doi:10.1038/ijo.2015.129

13. Poulsen MN, Pollak J, Bailey-Davis L, Hirsch AG, Glass TA, Schwartz BS. Associations of prenatal and childhood antibiotic use with child body mass index at age 3 years. *Obesity*. 2017;25(2):438-444. doi:10.1002/oby.21719

14. Cassidy-Bushrow AE, Burmeister C, Havstad S, et al. Prenatal antimicrobial use and early-childhood body mass index. *Int J Obes*. 2018;42(1):1-7. doi:10.1038/ijo.2017.205

15. Wang B, Liu J, Zhang Y, et al. Prenatal Exposure to Antibiotics and Risk of Childhood Obesity in a Multicenter Cohort Study. *Am J Epidemiol*. 2018;187(10):2159-2167. doi:10.1093/aje/kwy122

16. Heerman WJ, Daley MF, Boone-Heinonen J, et al. Maternal antibiotic use during pregnancy and childhood obesity at age 5 years. *Int J Obes*. January 2019:1. doi:10.1038/s41366-018-0316-6

17. Jess T, Morgen CS, Harpsøe MC, et al. Antibiotic use during pregnancy and childhood overweight: A population-based nationwide cohort study. *Sci Rep*. 2019;9(1). doi:10.1038/s41598-019-48065-9

18. Brown RG, Al-Memar M, Marchesi JR, et al. Establishment of vaginal microbiota composition in early pregnancy and its association with subsequent preterm prelabor rupture of the fetal membranes. *Transl Res*. 2019;207:30-43. doi:10.1016/J.TRSL.2018.12.005

19. Fettweis JM, Serrano MG, Brooks JP, et al. The vaginal microbiome and preterm birth. *Nat Med*. 2019;25(6):1012-1021. doi:10.1038/s41591-019-0450-2

20. Hočevar K, Maver A, Vidmar Šimic M, et al. Vaginal Microbiome Signature Is Associated With Spontaneous Preterm Delivery. *Front Med*. 2019;6. doi:10.3389/fmed.2019.00201

21. Markopoulou P, Papanikolaou E, Analytis A, Zoumakis E, Siahanidou T. Preterm Birth as a Risk Factor for Metabolic Syndrome and Cardiovascular Disease in Adult Life: A Systematic Review and Meta-Analysis. *J Pediatr*. 2019;210:69-80.e5. doi:10.1016/j.jpeds.2019.02.041

22. Firestone R, Cheng S, Pearce N, et al. Internet-Based Birth-Cohort Studies: Is This the Future

for Epidemiology? *JMIR Res Protoc*. 2015;4(2):e71. doi:10.2196/resprot.3873

23. Blumenberg C, Zugna D, Popovic M, Pizzi C, Barros AJD, Richiardi L. Questionnaire Breakoff and Item Nonresponse in Web-Based Questionnaires: Multilevel Analysis of Person-Level and Item Design Factors in a Birth Cohort. *J Med Internet Res*. 2018;20(12):e11046. doi:10.2196/11046

24. Breton C V., Gao L, Yao J, Siegmund KD, Lurmann F, Gilliland F. Particulate matter, the newborn methylome, and cardio-respiratory health outcomes in childhood. *Environ Epigenetics*. 2016. doi:10.1093/eep/dvw005

25. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva: World Health Organization, 2006.

26. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: Methods, interpretation and bias. *Int J Epidemiol*. 2013;42(5):1511-1519. doi:10.1093/ije/dyt127

27. Gritz EC, Bhandari V. The Human Neonatal Gut Microbiome: A Brief Review. *Front Pediatr*. 2015;3. doi:10.3389/fped.2015.00017

28. Lynch S V., Pedersen O. The Human Intestinal Microbiome in Health and Disease. Phimister EG, ed. *N Engl J Med*. 2016;375(24):2369-2379. doi:10.1056/NEJMra1600266

29. Robertson RC, Manges AR, Finlay BB, Prendergast AJ. The Human Microbiome and Child Growth – First 1000 Days and Beyond. *Trends Microbiol*. 2019;27(2):131-147. doi:10.1016/j.tim.2018.09.008

30. Freitas AC, Chaban B, Bocking A, et al. The vaginal microbiome of pregnant women is less rich and diverse, with lower prevalence of Mollicutes, compared to non-pregnant women. *Sci Rep*. 2017;7(1). doi:10.1038/s41598-017-07790-9

31. MacIntyre DA, Chandiramani M, Lee YS, et al. The vaginal microbiome during pregnancy and the postpartum period in a European population. *Sci Rep*. 2015;5. doi:10.1038/srep08988

32. Romero R, Hassan SS, Gajer P, et al. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome*. 2014;2(1). doi:10.1186/2049-2618-2-4

33. Mueller NT, Shin H, Pizoni A, et al. Birth mode-dependent association between pre-pregnancy maternal weight status and the neonatal intestinal microbiome. *Sci Rep*. 2016;6(April):1-9. doi:10.1038/srep23133

34. Sejersen TS, Vinding RK, Stokholm J, et al. Antibiotic exposure in infancy and development of BMI and body composition in childhood. *EClinicalMedicine*. December 2019. doi:10.1016/j.eclinm.2019.10.020

35. Zhang M, Differding MK, Benjamin-Neelon SE, Østbye T, Hoyo C, Mueller NT. Association of prenatal antibiotics with measures of infant adiposity and the gut microbiome. *Ann Clin Microbiol Antimicrob*. 2019;18(1):18. doi:10.1186/s12941-019-0318-9

36. Pizzi C, De Stavola B, Merletti F, et al. Sample selection and validity of exposureedisease association estimates in cohort studies. *J Epidemiol Community Health*. 2011;65(5):407-411. doi:10.1136/jech.2009.107185

37. Pizzi C, De Stavola BL, Pearce N, et al. Selection bias and patterns of confounding in cohort studies: The case of the NINFEA web-based birth cohort. *J Epidemiol Community Health*. 2012;66(11):976-981. doi:10.1136/jech-2011-200065

38. Pizzi C. Validation of the NINFEA birth data. https://www.progettoninfea.it/attachments/27%0D.

39. Huybrechts I, Himes JH, Ottevaere C, et al. Validity of parent-reported weight and height of preschool children measured at home or estimated without home measurement: a validation study. *BMC Pediatr*. 2011;11(1):63. doi:10.1186/1471-2431-11-63

40. Sekine M, Yamagami T, Hamanishi S, Kagamimori S. Accuracy of the Estimated Prevalence of Childhood Obesity from Height and Weight Values Reported by Parents: Results of the Toyama Birth Cohort Study. *J Epidemiol*. 2002;12(1):9-13. doi:10.2188/jea.12.9

41. Garcia-Marcos L, Valverde-Molina J, Sanchez-Solis M, et al. Validity of parent-reported height and weight for defining obesity among asthmatic and nonasthmatic schoolchildren. *Int Arch Allergy Immunol*. 2006;139(2):139-145. doi:10.1159/000090389

42. O'Connor DP, Gugenheim JJ. Comparison of measured and parents' reported height and weight in children and adolescents. *Obesity*. 2011;19(5):1040-1046. doi:10.1038/oby.2010.278

43. Rothman KJ, Greenland S. Validity and Generalizability in Epidemiologic Studies. In: *Wiley StatsRef: Statistics Reference Online*. Chichester, UK: John Wiley & Sons, Ltd; 2014. doi:10.1002/9781118445112.stat05223

Supplementary Material: Study III

Figure S1. Flow chart describing the selection of the study population

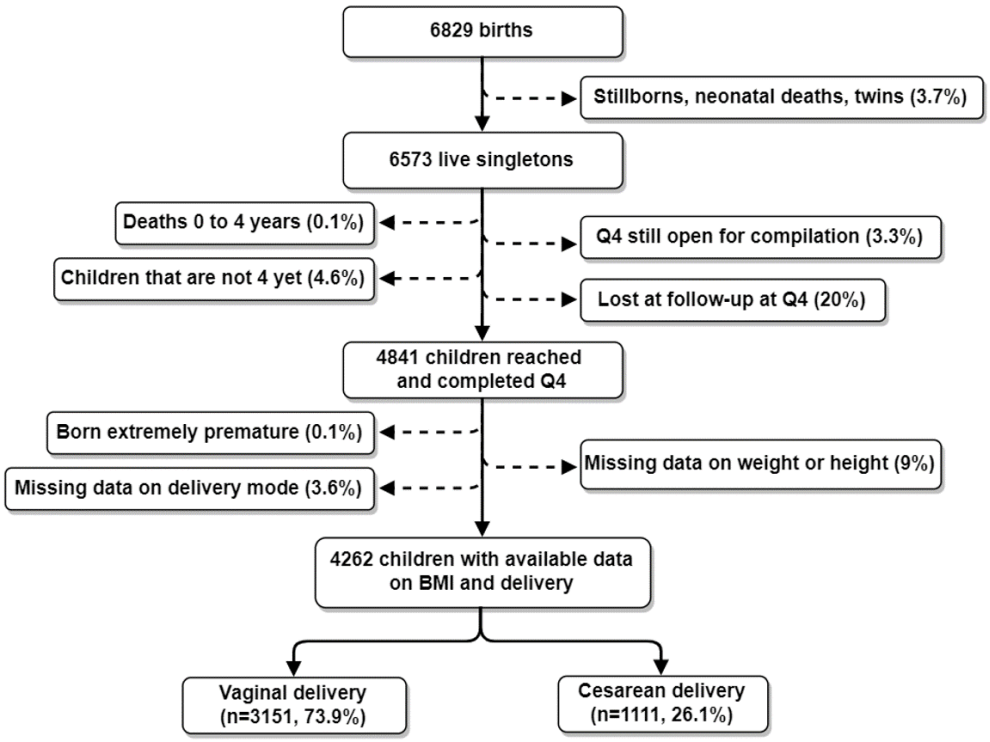


Table S1. Characteristics of the study population born with cesarean delivery

	N	Mean (SD) or %
Maternal age at delivery	1111	34.5 (4.2)
Maternal pre-pregnancy BMI	1089	23.1 (4.3)
Missing values	22	–
Maternal education		
Primary school or less	45	4.1%
Secondary school	343	31.1%
University degree or higher	715	64.8%
Missing values	8	–
Smoked in pregnancy		
No	1019	92.6%
Yes	81	7.4%
Missing values	11	–
First pregnancy		
No	292	27.1%
Yes	786	72.9%
Missing values	33	–
Antibiotic use in third trimester		
No	975	92.9%
Yes	75	7.1%
Missing values	61	–
Vaginal infections in third trimester		
No	917	87.3%
Yes	133	12.7%
Missing values	61	–
Child's BMI (IOTF cut-offs)		
Thinness grade 3 and 2	97	8.7%
Thinness grade 1	178	16.0%
Normal BMI	696	62.7%
Overweight/obesity	140	12.6%
Child's BMI (WHO cut-offs)		
Thinness	55	5.0%
Normal BMI	997	89.7%
Overweight/obesity	59	5.3%

Table S2. Characteristics between the included and excluded subjects in the complete case analysis in vaginal deliveries.

	Included	Excluded	p-value
	N(%) or mean (SD)	N(%) or mean (SD)	
Number of subjects	2837 (90%)	314 (10%)	
Maternal age at delivery	33.5 (4.0)	33.7 (4.4)	0.216
Maternal pre-pregnancy BMI	22.2 (3.6)	22.3 (3.2)	0.701
Maternal education			
Primary school or less	99 (3.5%)	13 (4.4%)	
Secondary school	888 (31.3%)	97 (32.9%)	0.578
University degree or higher	1850 (65.2%)	185 (62.7%)	
Gestational diabetes			
No	2599 (91.6%)	108 (87.8%)	0.139
Yes	238 (8.4%)	15 (12.2%)	
Smoked in pregnancy			
No	2633 (92.8%)	243 (88.0%)	
Yes	204 (7.2%)	33 (12.0%)	0.004
First pregnancy			
No	778 (27.4%)	80 (32.3%)	
Yes	2059 (72.6%)	168 (67.7%)	0.103
Antibiotic use in third trimester			
No	2667 (94.0%)	153 (95.6%)	
Yes	170 (6.0%)	7 (4.4%)	0.399
Vaginal infections in third trimester			
No	2586 (91.2%)	147 (91.9%)	
Yes	251 (8.9%)	13 (8.13%)	0.754
Child's BMI (IOTF cut-offs)			
Thinness grade 3 and 2	223 (7.9%)	19 (6.1%)	
Thinness grade 1	432 (15.2%)	46 (14.7%)	
Normal BMI	1854 (65.4%)	199 (63.4%)	
Overweight/obesity	328 (11.6)	50 (15.9%)	0.116
Child's BMI (WHO cut-offs)			
Thinness	134 (4.7%)	8 (2.6%)	
Normal BMI	2560 (90.2%)	283 (90.1%)	
Overweight/obesity	143 (5.0%)	23 (7.3%)	0.056

Table S3. Associations of antibiotic use and vaginal infections in the third trimester with BMI outcomes at age of 4 in NINFEA children born with vaginal delivery. World Health Organization Growth Standards were used.

	Cases/exposed cases (%)	RRRcrude (95% CI)	RRRadj* (95% CI)
Antibiotic use	170/2837 (6.0)		
Thinness	14/223 (6.3)	1.11 (0.63 to 1.96)	1.10 (0.62 to 1.95)
Overweight/obesity	26/328 (7.9%)	1.43 (0.92 to 2.21)	1.40 (0.90 to 2.16)
Vaginal infections	251/2837 (8.8)		
Thinness	19/223 (8.5)	1.07 (0.65 to 1.76)	1.06 (0.64 to 1.74)
Overweight/obesity	49/328 (14.9)	2.01 (1.43 to 2.85)	1.92 (1.37 to 2.70)

Data are presented as n, n(%) and Relative Risk Ratios (RRR) and 95% confidence intervals (CI). *Adjusted for: maternal age, education, parity, pre-pregnancy BMI, smoking during pregnancy, gestational diabetes and gestational age (only for the third trimester analyses)

Table S4. Associations between vaginal infections and BMI outcomes at age 4 in children born with cesarean delivery

	Cases/exposed cases (%)	RRRcrude (95% CI)	RRRadj* (95% CI)
Vaginal infections			
Thinness	7/90 (7.8)	0.56 (0.25 to 1.25)	0.60 (0.26 to 1.35)
Overweight/obesity	17/123 (13.8)	1.07 (0.61 to 1.86)	1.07 (0.61 to 1.88)

Data are presented as n, n(%) and Relative Risk Ratios (RRR) and 95% confidence intervals (CI). *Adjusted for: maternal age, education, parity, pre-pregnancy BMI, smoking during pregnancy, gestational diabetes and gestational age (only for the third trimester analyses)

Overall interpretation, implications and final conclusions

In this thesis we focused on common exposures and their associations with newborn molecular markers and childhood outcomes. The first 1000 days of life are particularly vulnerable period, when individuals' circumstances and their surrounding environment may leave them with health risks that they will carry on well into childhood and adulthood.

Two thirds of this thesis are focused on air pollution exposure in early-life and newborn molecular markers. The systematic review provided evidence that air pollutants have the potential to cause alterations in DNA methylation, both on global scale and locus specific, and to unstabilise the delicate telomere maintenance system. These findings were supported by our own analysis in cord blood, in particular those relative to locus specific methylation, where we observed alterations in CpGs mapped to genes with roles in cell replication, differentiation and response to oxidative stress. In particular, we observed hypermethylation in CpGs located in gene promoters, that are known to lead to long-term gene silencing, and gene body hypomethylation, often linked with synthesis of aberrant proteins.

These findings are relevant first because ambient air pollution is a ubiquitous exposure with concentrations that often go beyond the recommended limits and second, because pregnant women and their families are exposed to pollutants that are virtually unseparable from the environment where they live and work. Findings like these can also address some of the criticism directed towards DOHaD studies, such as that they erroneously perpetuate the notion that pregnant women, and only them, are responsible for the health of their children. Today, the majority of governments and international health organizations recognize the effects of air pollution and make efforts to lower gas emissions from vehicles and industries to improve overall air quality. However, since pregnant women and the developing fetus seem particularly vulnerable to the effects of air pollution, effort to lower the exposure itself are needed, as well as specific public health policies that support and promote health in vulnerable groups living in highly polluted places, especially in those with multiple risk factors, such as mothers with acute or chronic health issues or those that live in socio-economic adversity.

The analysis across different windows of exposure during pregnancy, outlined the beginning of the pregnancy as a particularly sensitive period for exposure to air pollution, when DNA methylation patterns are actively remodeling. These findings, might be of particular interest for research groups that study the mechanisms behind air pollution exposure and long-term health outcomes, considering that the major part of the methylation patterns is established around the beginning of the pregnancy and are believed to remain relatively stable throughout life.

The last, third part of the thesis was dedicated to common, non-environmental, gestational exposures and their link to pre-school obesity. We observed that vaginal infections and, to smaller extent, antibiotic use in the third trimester was associated with elevated BMI in early childhood. These observations, point to a potential role of the microbiome in childhood obesity. Further studies should assess whether there are specific microbial communities in the gut and/or study specific metabolites that may help unravel the mechanisms behind early adiposity rebound and postnatal growth, as well as identify children with multiple risk factors for developing obesity.

It seems that maternal pre-pregnancy BMI modifies the association between vaginal infections in the third trimester and obesity at the age of 4, pointing out to the importance of pre-conception counselling regarding maternal pre-pregnancy BMI, already known to negatively influence pregnancy and birth outcomes and increase the risk for childhood obesity. Maternal pre-pregnancy BMI might be considered as overall indicator for mother's health, since its intertwined with multiple clinical, nutritional, lifestyle, and socio-economic factors, and as such, it requires multidisciplinary approach for targeted health interventions, both on individual and societal level.

The thesis is based on two approaches that provide high quality scientific evidence: a systematic review and prospective cohort study design. The systematic review allowed me to summarize the current knowledge on the topic in the light of the accelerated growth of scientific information.

There are advantages and some limitations regarding the methodology used in the two cohort-based studies. We used an indirect measure of exposure assessment to study air pollution exposure in pregnancy, by applying an advanced machine learning method, based on daily satellite PM_{10} data to provide cost-effective assessment of PM_{10} exposure based on maternal residential address, that has good predictive performance and small prediction

errors. The outcome was two family of molecular markers: DNA methylation, measured by the latest Illumina methylation array, targeting 850 000 CpGs; and telomere length, by applying the method by Cawthon already used in similar studies. A major advantage was the possibility to calculate PM_{10} concentration across different windows of exposure during pregnancy and the possibility to replicate the findings in a smaller subset of the cohort. Nevertheless, there is a possibility for misclassification of exposure, since the exposure assessment does not take into account daily activity pattern, the outcome was measured only at one-time point, birth, and therefore implications regarding the stability of the markers at different time points cannot be made.

The study based on NINFEA data used self-reported questionnaire data to assess the exposure: antibiotic use and vaginal infections in the third trimester. We discussed the possibility that the slightly lower percentage of reported antibiotic, use in comparison to other studies might be due to the higher percentage of mothers with high educational attainment in the NINFEA cohort and how the baseline selection does not imply biased estimates of the exposure-disease associations. One of the limitations of the study is that we do not have data on the indication for the antibiotic and if the reported vaginal infection was pharmacologically treated, weather antibiotics were used. The outcome (child's BMI at age 4) based on children's height and weight measurements was measured prospectively and reported by the parents. We also confirmed that parental reporting of the child's measurements did not bias our estimates, after we restricted the analyses on children whose parents reported the height and weight measurements from the child's health booklets.

In conclusion, findings from this thesis may not only serve to elucidate disease ethology, but also to convey some public health messages, underlining the need to support and strengthen health since conception. Future studies should be able to address some of the issues underlined in this thesis and pinpoint some of the mechanisms between early risk factors with specific childhood outcomes by joining forces in novel multidisciplinary approaches and collaborative projects.

Future perspectives

Based on the findings of this thesis, these are possible continuations that might expand the knowledge within the field.

First, environmental exposures can have effects that are not limited to the period when they are observed, but they can be delayed in time. Therefore, our air pollution analyses may benefit from the use of statistical models that add a time dimension to the exposure–response relationship. In order to deal with delayed effects, we could estimate mean PM_{10} concentrations by gestational week and use lag-distributed models to model the association between PM_{10} exposure across pregnancy and molecular markers in cord blood.

Second, we have seen that air pollution influences locus-specific methylation in gene promotors and bodies, with key functions in cell replication, differentiation and oxidative stress. As mentioned in the Introduction section, apart from these two regions, there are several other places where DNA methylation is essential for the growth and the development of the fetus, such as methylation at imprinting genes. It is believed that there are around 200 imprinting genes that influence, not just intrauterine growth and development, but also later metabolic programing and neurocognitive pathways. By specifically studying whether air pollution has an influence on the methylation of these genes, it may help elucidate some of the effects on air pollution on intrauterine growth restriction, being born small or large-for-gestational age, or later metabolic and neurodevelopmental outcomes, such as susceptibility to obesity or autism.

Third, many studies reported an association between air pollution exposure in pregnancy and global methylation loss, indicating a possible involvement of DNA methylation machinery. Possible research hypotheses could include studying genotypes for DNA methytransferases responsible for de-novo methylation or maintenance of already established methylation patterns and/or one-carbon metabolism responsible for providing the methyl groups needed for methylation. Other possibility is to conduct candidate gene methylation analysis to see wheather air pollution silences genes that responsible for the expression of DNA methylation enzymes.

Fourth, advanced machine learning approached already found a way to study LINE1 methylation patterns by using high-thoughoutput platforms such as Illumina. Future studies

could assess if the findings for air pollution exposure seen in studies using traditional approaches for measuring LINE1 or Alu, are be comparable to studies that use a measure of LINE1/Alu derived from Illumina platforms.

Fifth, it was seen that DNA methylation at specific CpG sites is related to age. Steve Horvath constructed the first "epigenetic clock", a predictor based on the methylation status of hundreds of age-dependent CpGs that can estimate human age with high accuracy using DNA methylation. The recent development of gestational age predictors based on cord blood methylation made possible to study gestational age acceleration/deceleration (GAA/GAD) in newborns. While age acceleration (defined as the difference between DNA methylation age and chronological age or as the residuals from regressing DNAm age on chronological age) in adults was consistently linked with adverse exposures and age-related diseases, both GAA and GAD were previously linked to gestational exposures and birth outcomes. It is unknown if air pollution exposure during pregnancy influences epigenetic age at birth and if so, in what direction.

Sixth, it was seen that methylation at single CpG level is difficult to replicate. But it may well be that DNA methylation does not leave single CpG marks, but rather creates characteristic signatures in CpGs dispersed throughout the genome. Several prediction scores based on hundreds of CpGs were created with the help of advanced statistical methods to predict gestational exposure to tobacco. It could be argued that epigenetic clocks function similarly. Considering that air pollution exposures creates generalized oxidative stress and inflammation, it is plausible that it can leave a characteristic signature in multiple CpGs dispersed throughout the genome. Large sample size and harmonized pipeline should probably be needed in order to detect such signatures.

Seventh, stochastic epigenetic mutations (SEMs) are defined as an individual having an extreme methylation value at a single CpG locus when compared to the rest of the population. The number of stochastic mutations is known to increase exponentially with human aging and with exposure to toxic substances. Until now, there are no studies that analyzed SEMs in neonatal tissues in relation to adverse prenatal exposures, such as air pollution.

Eighth, in the NINFEA study we only included BMI at one specific time point (4 years). Future studies could use weight and height data available at several time points from birth to childhood to estimate individual weight trajectories using methods such as the SITAR: Super

Imposition by Translation and Rotation model.

Lastly, DNA methylation is another plausible mechanism though which exposures such as antibiotics or vaginal infections can influence childhood obesity. Many birth cohorts, apart from cord blood, collect other types of easily assessable samples, such as nasal epithelia or saliva. The NINFEA cohort has samples from both maternal and child's saliva, and studying epigenetic patterns in oral epithelial cells might be more relevant for exposures that influence the establishment of the newborn gut microbiome.

Acknowledgments

Many people have contributed to the completion of this thesis, and to my personal and professional growth in the past years.

Professor Lorenzo Richiardi, my supervisor, for believing in me and giving me the opportunity to work on this project. For letting me make independent decisions and providing me structure and guidance when I needed support. For his thorough and systematic approach in everything he does. For his enthusiasm for scientific research and his availability to discuss and explain complex issues with ease and simplicity.

Professor Milena Maule, for being my first tutor. For guiding me with endless patience through my first steps in epidemiologic research, for picking me up when I was down and letting me know when I was on the right path. For her availability, warm support and never-ending encouragement.

Professor Franco Merletti, for bringing together a group of extraordinary people, that function and respect each other as a big family, and for introducing me to them. For his good heart, kind spirit and excellent sense of humor.

Maja Popovic, for giving daily advice on study design and data analysis. For her bright spirit and sharp mind, but mostly for always being there when I was feeling stuck or overwhelmed.

Valentina Fiano, Morena Trevisan, Laura DeMarco and Chiara Grasso for introducing me to the world of molecular markers and laboratory analyses, and for always leaving the door of their lab open for me.

Federica Asta, Daniela Porta, Maria Antonieta Stazi, Lorenza Nistico, Sonia Brescianini, Franca Rusconi, Luigi Gagliardi and Luca Ronfani for their hands on approach, for their passionate participation and all stimulating meetings.

Costanza Pizzi and Daniela Zugna, for creating an atmosphere of mutual support in the office and for approaching with enthusiasm every problem that might arise.

Giovenale Moirano and Chiara Moccia for brightening up the spirit in the office. For all our chats and coffee breaks.

Fulvio Lazzarato for providing me assistance while running the complex analysis on our super computer.

Milena, Preeta, Slavica, and Aleksandar for being friends for life.

My parents, Srecko and Snezhana, and my brother Gorgi, for being an impeccable support system. For their endeless love and for providing me a solid base to grow and develop as an individual. ***My aunt, uncle and cousin*** for their constant encouragement and support.

For all the families that participate in Piccolipiù and NINFEA, that devoted their precious time to carefully compile the questionnaires, and ***for all the Piccolipiù and NINFEA staff***. A great part of this thesis was financially supported by the Italian Ministry of Health – CCM. NINFEA was partially funded by the Compagnia di San Paolo.

Turin, February 15th, 2021 *Elena Isaevska*

List of other related papers

Taylor K, Elhakeem A, Nader JLT, Yang T, **Isaevska E**, Richiardi L, Vrijkotte T, Pinot de Moira A, Murray DM, Finn D, Mason D, Wright J, Oddie S, Roeleveld N, Harris JR, Nybo Andersen AM, Caputo M, Lawlor DA. *The effect of maternal pre-/early-pregnancy BMI and pregnancy smoking and alcohol on congenital heart diseases: a parental negative control study*. (submitted; preprint: https://doi.org/10.1101/2020.09.29.20203786)

Popovic M, Moccia C, **Isaevska E**, Moirano G, Pizzi C, Zugna D, Rusconi F, Merletti F, Maule M, Richiardi L. *COVID-19-like symptoms and their relation to SARS-CoV-2 epidemic in children and adults of an Italian birth cohort*. (submitted; https://doi.org/10.21203/rs.3.rs-34027/v1)

Jaddoe V.W.V, Felix J.F, Andersen A.-M.N, Charles M.-A, Chatzi L, Corpeleijn E, Donner N, Elhakeem A, Eriksson J.G, Foong R, Grote V, Haakma S, Hanson M, Harris J.R, Heude B, Huang R.-C, Inskip H, Jarvelin M.-R, Koletzko B, Lawlor D.A, Lindeboom M, McEachan R.R.C, Mikkola T.M, Nader J.L.T, de Moira A.P, Pizzi C, Richiardi L, Sebert S, Schwalber A, Sunyer J, Swertz M.A, Vafeiadi M, Vrijheid M, Wright J, Duijts L, El Marroun H, Gaillard R, Santos S, Geurtsen M.L, Kooijman M.N, Mensink-Bout S.M, Vehmeijer F.O.L, Voerman E, Nieuwenhuijsen M, Basagana X, Bustamante M, Casas M, de Castro M, Cirugeda L.E, Fernandez-Barres S, Fossati S, Garcia R, Julvez J, Lertxundi A.C, Lertxundi N, Llop S, Lopez-Vicente M, Lopez-Espinosa M.-J.B, Maitre L, Murcia M, Lea J, Urquiza H, Warembourg C, Zugna D, Popovic M, **Isaevska E**, Maule M, Moccia C, Moirano G, Rasella D, Hanson M.A, Inskip H.M, Jacob C.M, Salika T, Cadman T, Strandberg-Larsen K.M, Pedersen M, Vinther J.L, Wilson P, Mason D, Yang T.C, Cardol M, van Enckevoort E, Hyde E, Scholtens S, Snieder H, Thio C.H.L, Chatzi L, Margetaki K.C.A, Roumeliotaki T, Nader J.L, Knudsen G.P, Magnus P, Panico L, Ichou M, de Lauzon-Guillain B, Dargent-Molina P, Cornet M, Florian S.M, Harrar F, Lepeule J, Lioret S, Melchior M, Plancoulaine S, Mannikko M, Parmar P, Rautio N, Ronkainen J, Tolvanen M, Mikkola T.M, Aumuller N, Closa-Monasterolo R, Escribano J, Ferre N, Gruszfeld D, Gurlich K, Langhendries J.-P, Luque V, Riva E, Schwarzfischer P, Totzauer M, Verduci E, Xhonneux A, Zaragoza-Jordana M, Schwalber A, Foong R.E, Hall G.L, Lin A, Carson J, Melton P, Rauschert S. *The LifeCycle Project-EU Child Cohort Network: a federated analysis infrastructure and harmonized data of more than 250,000 children and parents.* Eur J Epidemiol. 2020 Jul 1 ;35(7):709–24. https://doi.org/10.1007/s10654-020-00662-z

Moccia C, Popovic M, **Isaevska E,** Moirano G, Pizzi C, Rusconi F, Merletti F, Maule MM, Richiardi L. *Factors associated with self-perceived knowledge of COVID-19: a study among women from the NINFEA birth cohort*. Epidemiol Prev. 2020 (accepted, in print)

Popovic M, Fiano V, **Isaevska E,** Moccia C, Trevisan M, Rusconi F, De Marco L, Polidoro S, Merletti F, Pizzi C, Richiardi L. *Determination of saliva epigenetic age in infancy, and its association with parental socio-economic characteristics and pregnancy outcomes.* J Dev Orig Health Dis. 2020;1-9.