

Population DNA methylation studies in the DOHaD framework

Janine F. Felix^{1,2,3}, Charlotte A. M. Cecil^{1,2,4,5}

SELF-ARCHIVING VERSION

1. The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
2. Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
3. Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
4. Department of Child and Adolescent Psychiatry, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
5. Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

Corresponding author: Janine F. Felix MD PhD, The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, P.O. Box 2040, 3000 CA Rotterdam, the Netherlands, j.felix@erasmusmc.nl, Ph.+31 10 70 43997.

Citation: Felix JF & Cecil CAM (2018). Population DNA methylation studies in the DOHaD framework. *Journal of Developmental Origins of Health and Disease*.

ABSTRACT

Epigenetic changes represent a potential mechanism underlying associations of early-life exposures and later-life health outcomes. Population-based cohort studies starting in early life are an attractive framework to study the role of such changes. DNA methylation is the most studied epigenetic mechanism in population research. We discuss the application of DNA methylation in early-life population studies, some recent findings, key challenges and recommendations for future research. Studies into DNA methylation within the DOHaD framework generally either explore associations between prenatal exposures and offspring DNA methylation or associations between offspring DNA methylation in early life and later health outcomes. Only few studies to date have integrated prospective exposure, epigenetic and phenotypic data in order to explicitly test the role of DNA methylation as a potential biological mediator of environmental effects on health outcomes. Population epigenetics is an emerging field which has challenges in terms of methodology and interpretation of the data. Key challenges include tissue specificity, cell type adjustment, issues of power and comparability of findings, genetic influences, and exploring causality and functional consequences. Ongoing studies are working on addressing these issues. Large collaborative efforts of prospective cohorts are emerging, with clear benefits in terms of optimizing power and use of resources, and in advancing methodology. In the future, multidisciplinary approaches, within and beyond longitudinal birth and preconception cohorts will advance this complex, but highly promising, field of research.

Keywords: Epigenetics; Cohort Studies; DNA-methylation; Epidemiology, Molecular

Introduction

In the Developmental Origins of Health and Disease framework^{1,2}, many associations between the early-life environment (e.g. prenatal exposure to maternal obesity, smoking and suboptimal diet) and later health outcomes have been described³⁻⁶. However, the underlying molecular mechanisms are often not known. Epigenetic changes have recently emerged as a potential mechanism that may at least partially explain these associations.

What is epigenetics

The ‘epigenome’ refers to a collection of molecular mechanisms, which do not affect the DNA sequence, but can affect gene regulation by changing the structure and accessibility of the DNA. Three main types of epigenetic mechanisms can be distinguished: DNA methylation, histone modifications and non-coding RNAs. These mechanisms are not independent, but rather function together as part of the larger, complex epigenetic machinery⁷. Here, we focus specifically on DNA methylation, as this is currently the most studied epigenetic mark in large population studies. DNA methylation refers to the binding of a methyl group to the DNA, mainly at positions where a cytosine is located next to a guanine, a cytosine-phosphate-guanine (CpG) site⁸. Typically (but not always), higher levels of DNA methylation at CpG sites in promoter or enhancer regions of a gene impede transcription factors from binding to the DNA, thereby inhibiting gene expression. On the contrary, methylation of CpG sites in a gene body generally leads to increased expression⁸. Because DNA methylation patterns can be passed on mitotically during cell division, they can lead to long-term alterations in gene activity and related phenotypes⁹. At the same time, DNA methylation patterns can also show a considerable degree of flexibility over time, enabling cells to respond to changing internal and external inputs.

DNA methylation within the DOHaD framework

Interest in DNA methylation within the DOHaD framework stems from the fact that (a) it has been shown to respond to both genetic and environmental influences, beginning *in utero*, and (b) it plays a fundamental role in development^{10,11}. Consequently, DNA methylation shows promise as a mechanism for understanding how environmental and genetic influences shape trajectories of health and disease across the lifespan.

Early life, in particular, may be a critical period for epigenetic effects on health outcomes. Extensive reprogramming of DNA methylation patterns takes place *in utero*¹². Thus, aberrant modifications that occur at this early stage, before tissue differentiation, could result in alterations in gene expression across many tissue types. In addition, environmental and genetic effects on organ maturation during critical periods in pregnancy and postnatally may be mediated through the epigenome. Evidence for this has come primarily from animal experimental studies, which enable careful manipulation of environmental exposures¹³. In humans, an early example is the work performed in the Dutch Famine Study, which has shown that exposure to famine in the first trimester of pregnancy, but not the last, is associated with DNA methylation changes at the imprinted *IGF2* gene, a key modulator of fetal development¹⁴. Timing-specific effects were later found for further loci, some of which also associated with cardiometabolic phenotypes, such as birth weight and LDL cholesterol levels^{15,16}. Together, this work provided preliminary evidence that (very) early life may be a critical period for epigenetically-mediated effects of environmental exposures on health outcomes. Also, twin studies have been

instrumental in disentangling the relative contributions of genetic, environmental and stochastic factors to DNA methylation and phenotypic variability^{11,17}.

Population-based cohort studies starting in early life offer an attractive platform to build on this early work and characterize associations between early-life exposures, epigenetic mechanisms and later-life health. Below, we discuss the application of DNA methylation in early-life population research, including some recent findings, key challenges and recommendations for future research.

DNA methylation in early-life population-based cohort studies

In population-based studies, there is increasing interest in the role of DNA methylation changes in observational associations of adverse exposures in the preconception period, during pregnancy and in early infancy with later-life common diseases. Through the development of high-throughput cost-effective arrays which analyze hundreds of thousands of DNA methylation markers across the genome, it has become possible to measure DNA methylation in (relatively) large samples. Levels of DNA methylation at these markers are then associated with determinants and outcomes of health and disease in “epigenome-wide association studies” or EWASs. In addition, many long-running population-based studies have longitudinal collection of data and biological samples in extensive biobanks. This makes it possible to study DNA methylation over time in relation to repeated measurements of exposures and outcomes and to test key mediational hypotheses.

So far, population-based research examining epigenetics within the DOHaD framework has fallen mainly into two categories – (i) studies investigating associations between prenatal exposures and offspring DNA methylation patterns, e.g. dietary, environmental, and psychosocial exposures (Fig 1a: *a* path); and (ii) studies investigating associations between offspring DNA methylation patterns in early life and later health outcomes, e.g. neurodevelopment or cardiometabolic health (Fig 1a: *b* path). In both cases, DNA methylation has typically been examined at birth, either in cord blood or in placental tissue, given its role as a barrier and interface between maternal and fetal environments.

Early-life exposures and DNA methylation

To date, the prenatal exposure that has been most robustly linked to DNA methylation changes in the offspring is maternal smoking¹⁸⁻²³. In a recent large study based on data from 13 cohorts ($n = 6,685$), Joubert *et al* found that sustained smoking during pregnancy associated with altered DNA methylation patterns across thousands of CpG sites, an effect that was still observable – albeit to a lesser extent – in older children, although the sample size was around half that of the newborn analysis¹⁸. This study’s top finding was a CpG in *AHRR*, which has been associated with both personal and maternal smoking in multiple other studies^{19,24,25}. Pathway analyses found neurological and developmental pathways, which is in line with the maternal smoking-related child outcomes. Alterations in DNA methylation at birth have also been documented in relation to other chemicals, such as prenatal exposure to arsenic²⁶, mercury²⁷ and air pollution^{28,29}.

Maternal pre-pregnancy body mass index (BMI) has also been associated with offspring DNA methylation, with findings showing associations with both lower and higher maternal BMI in a recent meta-analysis of 9340 newborns³⁰. Effects were generally small, with a median increase of 0.03% in methylation level per unit increase in maternal BMI, and strong evidence for a direct causal intra-uterine

effect was lacking for most loci³⁰. Furthermore, dietary factors have been examined, with a particular focus on folate intake, due to its established role in both epigenetic regulation and fetal brain development³¹. The largest of these studies, based on data from two birth-cohorts (n = 1,988), reported associations between maternal plasma folate levels during pregnancy and differential DNA methylation in cord blood across multiple genes involved in folate biology and neurodevelopmental processes³². Interestingly, this study found reduced methylation with higher levels of folate, a methyl donor, at the vast majority of CpG sites, a finding which underlines the complex associations of the one-carbon metabolism pathway with methylation³². In contrast, other known risk factors for poorer offspring health have not been found to be strongly associated with neonatal DNA methylation. For example, a recent meta-analysis on maternal alcohol use found no evidence for an association with cord blood DNA methylation³³.³³ Similarly, a meta-analysis of prenatal maternal stress did not identify associations with DNA methylation in cord blood, although an enrichment for genes related to methyltransferase activity was identified³⁴. Besides a true null effect, the lack of a strong signal in these studies may reflect low levels of more severe forms of exposure in the general population, relatively small sample sizes, resulting in decreased power to detect effects, or heterogeneity between the included studies in terms of population, exposure definition or confounding structure.

DNA methylation and childhood health outcomes

Compared to prenatal exposures, fewer population-based studies have examined the relationship between DNA methylation patterns and childhood health outcomes. We highlight here two important areas of progress in the field of DNA methylation in population-based studies: neurodevelopment and adiposity.

For neurodevelopmental outcomes, most of the evidence to date on this topic has been drawn from cross-sectional data in smaller, selected samples, e.g. high-risk/clinical samples, although efforts to characterize these associations in the wider population are fast increasing. One such study, based on data from 537 newborns, found that epigenetic patterning of glucocorticoid genes in the placenta prospectively associated with infant neurodevelopmental profiles³⁵. Specifically, infant membership to a reactive, poorly regulated profile associated with increased placental methylation of the *NR3C1* glucocorticoid receptor gene, and decreased methylation of the *HSD11B2* gene involved in limiting fetal exposure to maternal circulating cortisol. Other population studies have examined DNA methylation in relation to risk for neurodevelopmental disorders – particularly ADHD, but also other psychiatric problems with early developmental origins, such as oppositional defiant and conduct problem behavior³⁶⁻³⁹. For example, a longitudinal epigenome-wide study based on over 800 children found that DNA methylation patterns at birth (but not at age 7) prospectively associated with trajectories of ADHD symptoms from childhood to adolescence³⁹. Of the 13 genome-wide significant loci identified at birth, several were annotated to genes involved in fatty acid metabolism and neural tube development. Importantly, one of the loci was annotated to a gene (*ST4GAL3*) that has been recently identified as the top genetic hit for ADHD in the largest GWAS to date. Studies are also beginning to integrate brain imaging data to investigate the links between DNA methylation, brain development and neuropsychiatric outcomes. For example, a recent EWAS study found that *SP6* methylation at birth prospectively associated with lower amygdala-hippocampal volume and higher psychotic symptoms in young adulthood – an association that was replicated in an independent sample of patients with schizophrenia⁴⁰.

Most of the evidence regarding DNA methylation and adiposity stems from adult studies⁴¹⁻⁴⁶. The largest of these identified and replicated cross-sectional associations between DNA methylation at 187 loci and BMI in over 10,000 samples, showing that these are mostly a consequence rather than a cause of

adiposity⁴⁶. Less is known about DNA methylation and childhood adiposity. Several studies have described associations of DNA methylation with birth weight, with a limited number of CpG sites identified in more than one study⁴⁷⁻⁵¹. No strong evidence was found for persistence of associations into later life in the two studies that explored longitudinal relationships^{47,51}. In targeted analyses, a replicable association of *RXRA* methylation in umbilical cord tissue with childhood fat mass was shown⁵². Several differentially methylated regions in newborn blood, such as the promoter of the long non-coding RNA ANRIL, have been associated with childhood adiposity measures based on epigenome-wide scans^{53,54}. In cross-sectional analyses in children and adolescents, DNA methylation in several regions has been associated with adiposity measures, including the *SOCS3* locus⁵⁵⁻⁵⁷. It remains unclear whether DNA methylation patterns are a *risk factor for* or *consequence of* adiposity. Detailed associations in the large adult EWAS show that changes in DNA methylation are mostly a consequence of obesity⁴⁶. This is supported by smaller studies in children and individuals before and after bariatric surgery⁵⁸⁻⁶¹.

DNA methylation as a mediator

Despite promising findings, very few population-based studies to date have integrated prospective exposure, epigenetic and phenotypic data in order to explicitly test the role of DNA methylation as a potential biological *mediator* of environmental effects on health outcomes. Notable exceptions include an EWAS study based on 321 cord blood samples, which showed that, in males, DNA methylation of the *PONI* gene at birth associated with prenatal mercury levels and predicted cognitive performance scores during childhood²⁷. Using repeated measures of DNA methylation, the authors found that this effect persisted in early childhood (3-5 years) and was attenuated in mid-childhood (6-10 years). Furthermore, DNA methylation of the *PONI* locus at birth was shown to associate with gene expression levels in an independent set of cord blood samples. Another exception is a longitudinal EWAS study examining epigenetic correlates of adolescent substance use (tobacco, alcohol and cannabis use; n = 244; age 14–18). This study found that neonatal DNA methylation levels in a tightly interconnected network of genes enriched for neurodevelopmental processes (n = 65 epigenome-wide corrected loci) partially mediated the effect of prenatal tobacco smoking on later substance use risk in adolescence⁶². Relatedly, maternal prenatal smoking has been found by other studies to associate with lower birth weight partially through alterations in cord blood DNA methylation using both targeted and epigenome-wide approaches^{24,63}.

Current challenges and future directions

The field of population epigenetics is still in its infancy. Although great strides have already been taken, there is still much to discover and learn. Here we present some of the key challenges that current research in this area is facing, and how these may be addressed.

Tissue specificity

DNA methylation patterns are largely tissue-specific. In large cohort studies, blood is typically the main source of DNA collected. In some cases there may be DNA from buccal swabs, umbilical cord or placental tissue, but other tissues are generally inaccessible for population researchers. This poses challenges, as the extent to which findings may be generalized to the target tissue of interest is often unclear^{64,65}. One way to address this limitation is to refer to publicly available datasets that show cross-tissue concordance in DNA methylation patterns, based on selected samples (e.g. comparison of blood

and brain in postmortem samples or blood and adipose tissue in patients undergoing surgery), for example: <http://epigenetics.iop.kcl.ac.uk/bloodbrain>, and <https://redgar598.shinyapps.io/BECOn/>⁶⁶⁻⁶⁸. However, it is not just the concordance in DNA methylation levels between tissues, but also, and maybe more so, the variability of those DNA methylation levels in association with the exposure or outcome of interest that matters. Concordance between tissues in associations with specific phenotypes has been shown, for example for BMI between blood, adipose and liver tissue, but overall evidence is limited^{46,64,69}. The NIH Roadmap Epigenomics Project can be used as a reference for epigenetic marks in a large number of tissues (www.roadmapepigenomics.org)⁷⁰.

Cell type adjustment

DNA methylation patterns vary not only between, but also within tissues. In blood samples, DNA is collected from leukocytes, which represent a heterogeneous population characterized by cell-type specific DNA methylation. As such, the relative proportions of cell types from which the DNA is extracted may influence epigenetic results. Ideally, detailed measurements of leukocyte subtypes would be available for correction in all study participants, but this is often not feasible. As an alternative, the use of reference-based approaches for estimating cell type proportions has become popular, such as methods developed for whole blood^{71,72} and, more recently, cord blood⁷³⁻⁷⁵. Reference-free methods are also available⁷⁶⁻⁷⁸. These methods, however, are still under development and population-based studies with specific cell type measurements can play a pivotal role in developing and validating new analytic approaches.

Methodological considerations

DNA methylation data are multifactorial, high dimensional and inter-correlated, raising questions about how best they should be analyzed⁷⁹. So far, studies have varied widely in study characteristics, methodology and analytic routines, limiting comparability of findings. Furthermore, many studies have been underpowered to identify effects, especially when examining exposures or outcomes with low prevalence in the general population. To address these challenges, large collaborative efforts of prospective cohorts, such as the Pregnancy And Childhood Epigenetics (PACE) Consortium, have emerged, similar to the field of genome-wide association studies although numbers are still much smaller⁸⁰. Such collaborations have shown enormous potential in increasing statistical power and optimizing the use of resources⁸¹⁻⁸⁵, as well as allowing replication and comparison of effects between studies with potentially heterogeneous confounding structures, thus strengthening the evidence for reported associations. In addition, they are a valuable platform for standardization of study design and analysis protocols to maximize comparability and help to delineate best practices in the field. Multidisciplinary teams, bringing together basic, population and clinical scientists, will enable easier exchange of knowledge and ideas, translation between bench, population and bedside, and strengthening of scientific conclusions by bringing together multiple lines of evidence.

Role of genetics

It has been reported that a substantial proportion of the variation in DNA methylation between individuals is due to genetic variants, acting in *cis* or *trans* (mQTLs)⁸⁶⁻⁸⁸. As such, measured DNA methylation levels may reflect genotype and mediate associations of genetic variants and health outcomes, or, if genotype drives both DNA methylation and the outcome under study, changes in DNA methylation may be an epiphenomenon of the genetic risk rather than part of the causal pathway. Furthermore, genetic factors

may moderate associations between environmental exposures, DNA methylation and outcomes (Fig1a). Thus, genetic variability should always be considered when examining DNA methylation – ideally directly, or if unavailable or underpowered, by referring to online mQTL databases (e.g. www.mqtl.org⁸⁷).

Causal inference

DNA methylation is a dynamic process that may change under the influence of genetic and environmental factors, but also with disease states. Thus, DNA methylation may be a cause, a consequence, or an epiphenomenon of the phenotype under study (Fig1b). This issue of reverse causality is particularly problematic when using cross-sectional designs. While the use of prospective, longitudinal data partially addresses directionality - especially when DNA methylation is measured before the onset of disease - observational studies in general are subject to confounding. The application of advanced causal inference methods, such as two-step Mendelian randomization offers a valuable tool for probing causal pathways⁸⁹. For example, a recent study applied this method to support a small, causal effect of prenatal maternal B₁₂ levels on childhood IQ levels, via DNA methylation at birth⁹⁰. The tissue specificity of DNA methylation is an additional challenge to causal inference analyses.

Determining whether DNA methylation plays a causal role in disease is important for understanding the etiology of the disease as well as potentially developing new therapies in future. Yet, even if DNA methylation changes are not causal to a disease, but rather a consequence or an epiphenomenon, they may still serve as biomarkers of the disease or its progression (Fig1b)⁹¹. In addition, if an association of DNA methylation with a phenotype is non-causal, the DNA methylation level may be used as a proxy for a specific exposure. This may be especially helpful for exposures that are difficult to measure (reliably). Previous studies have shown the potential of using DNA methylation as a biomarker in a number of these instances, including for alcohol consumption, maternal smoking during pregnancy and gestational age⁹²⁻⁹⁵.

Functional characterization

The buck does not stop at DNA methylation. To understand pathways leading to disease, we need to study the functional effects in terms of gene expression as well as the roles of other epigenetic marks. Effect sizes in EWAS are often small, for example in an EWAS of maternal BMI, the largest effect estimates for the significantly associated CpG sites were an 0.1% decrease or increase in methylation level per unit increase in maternal BMI (kg/m²)³⁰. For a strong exposure such as sustained maternal smoking, the maximum effect size was a 10% decrease in methylation levels for sustained smokers compared to non-smokers¹⁸. It is currently unclear to what extent statistical significance parallels biological meaningfulness. In future, studies that incorporate both DNA methylation and expression data, preferably within the same population and at the same time point, can shed more light on this question of functionality. Such data are not widely available in longitudinal birth cohorts yet, with some noticeable exceptions (INMA Study, Gambia Study)^{96,97}. In addition, the ENCODE database can be used to explore whether identified DNA methylation sites are located in potential regulatory regions³⁶. Intensifying collaborations and expanding them to researchers in neighboring fields, such as clinical and basic researchers, will make it easier to build chains of evidence across species, across research fields and across the life course, leading to strong and innovative research.

Starting at the very beginning

Most early-life cohorts begin in pregnancy and thus generally do not have prospective and hands-on data about the preconception and early pregnancy periods. Some preconception cohorts do exist (Southampton Women's Survey: www.mrc.soton.ac.uk/sws⁹⁸, Generation R Next Study: www.generationr.nl/next) and they, together with longitudinal birth cohorts with follow-up long enough to include the next generation, will provide valuable information on exposures in these very early life stages and their epigenetic associations. Furthermore, the vast majority of epigenetic research in birth cohorts is currently focused on maternal exposures during pregnancy. In contrast, the contribution of paternal epigenetics and exposures is currently understudied, but has the potential to lend important new insights in the development of the offspring epigenome and the pathways leading from early life exposures to child and adolescent health.

Conclusion

Population epigenetics is an emerging field with strong potential to shed light on mechanisms underlying associations of early-life adverse exposures and later life health. Yet, the field is still under development and poses challenges in terms of methodology and interpretation of the data. The use of multidisciplinary approaches, combining various “omics” studies linking epigenetic changes to functional readouts and phenotypes, within and beyond longitudinal birth and preconception cohorts are needed to advance the field and shed further light on this complex, but highly promising, field of research.

Acknowledgements: None

Financial support

JFF received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 633595 (DynaHEALTH) and No 733206 (LifeCycle). CC is supported by the Economic and Social Research Council (grant ref: ES/N001273/1).

Conflicts of interest: None

References

1. Barker DJ. Fetal origins of coronary heart disease. *BMJ*. 1995;311,171-174.
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;359,61-73.
3. Godfrey KM, Barker DJ. Fetal nutrition and adult disease. *Am J Clin Nutr*. 2000;71,1344S-1352S.
4. Moller SE, Ajslev TA, Andersen CS, Dalgard C, Sorensen TI. Risk of childhood overweight after exposure to tobacco smoking in prenatal and early postnatal life. *PLoS One*. 2014;9,e109184.
5. Roseboom TJ, Painter RC, van Abeelen AF, Veenendaal MV, de Rooij SR. Hungry in the womb: what are the consequences? Lessons from the Dutch famine. *Maturitas*. 2011;70,141-145.
6. Yu Z, Han S, Zhu J, et al. Pre-pregnancy body mass index in relation to infant birth weight and offspring overweight/obesity: a systematic review and meta-analysis. *PLoS One*. 2013;8,e61627.
7. Yamada L, Chong S. Epigenetic studies in Developmental Origins of Health and Disease: pitfalls and key considerations for study design and interpretation. *J Dev Orig Health Dis*. 2017;8,30-43.
8. Jang HS, Shin WJ, Lee JE, Do JT. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes (Basel)*. 2017;8.
9. Ziller MJ, Gu H, Muller F, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature*. 2013;500,477-481.
10. Bestor TH, Edwards JR, Boulard M. Notes on the role of dynamic DNA methylation in mammalian development. *Proc Natl Acad Sci U S A*. 2015;112,6796-6799.
11. Gordon L, Joo JE, Powell JE, et al. Neonatal DNA methylation profile in human twins is specified by a complex interplay between intrauterine environmental and genetic factors, subject to tissue-specific influence. *Genome Res*. 2012;22,1395-1406.
12. Marsit CJ. Influence of environmental exposure on human epigenetic regulation. *J Exp Biol*. 2015;218,71-79.
13. Kundakovic M, Jaric I. The Epigenetic Link between Prenatal Adverse Environments and Neurodevelopmental Disorders. *Genes (Basel)*. 2017;8.
14. Heijmans BT, Tobi EW, Stein AD, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci U S A*. 2008;105,17046-17049.
15. Tobi EW, Lumey LH, Talens RP, et al. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum Mol Genet*. 2009;18,4046-4053.
16. Tobi EW, Slieker RC, Stein AD, et al. Early gestation as the critical time-window for changes in the prenatal environment to affect the adult human blood methylome. *Int J Epidemiol*. 2015;44,1211-1223.
17. Martino D, Loke YJ, Gordon L, et al. Longitudinal, genome-scale analysis of DNA methylation in twins from birth to 18 months of age reveals rapid epigenetic change in early life and pair-specific effects of discordance. *Genome Biol*. 2013;14,R42.
18. 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;98,680-696.
19. Joubert BR, Haberg SE, Nilsen RM, et al. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect*. 2012;120,1425-1431.

20. Lee KW, Richmond R, Hu P, et al. Prenatal exposure to maternal cigarette smoking and DNA methylation: epigenome-wide association in a discovery sample of adolescents and replication in an independent cohort at birth through 17 years of age. *Environ Health Perspect.* 2015;123,193-199.
21. Markunas CA, Xu Z, Harlid S, et al. Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environ Health Perspect.* 2014;122,1147-1153.
22. Richmond RC, Simpkin AJ, Woodward G, et al. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum Mol Genet.* 2015;24,2201-2217.
23. Rzehak P, Saffery R, Reischl E, et al. Maternal Smoking during Pregnancy and DNA-Methylation in Children at Age 5.5 Years: Epigenome-Wide-Analysis in the European Childhood Obesity Project (CHOP)-Study. *PLoS One.* 2016;11,e0155554.
24. Kupers 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;44,1224-1237.
25. Zeilinger S, Kuhnel B, Klopp N, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One.* 2013;8,e63812.
26. Green BB, Karagas MR, Punshon T, et al. Epigenome-Wide Assessment of DNA Methylation in the Placenta and Arsenic Exposure in the New Hampshire Birth Cohort Study (USA). *Environ Health Perspect.* 2016;124,1253-1260.
27. Cardenas A, Rifas-Shiman SL, Agha G, et al. Persistent DNA methylation changes associated with prenatal mercury exposure and cognitive performance during childhood. *Sci Rep.* 2017;7,288.
28. Gruzjeva O, Xu CJ, Breton CV, et al. Epigenome-Wide Meta-Analysis of Methylation in Children Related to Prenatal NO₂ Air Pollution Exposure. *Environ Health Perspect.* 2017;125,104-110.
29. Breton CV, Gao L, Yao J, et al. Particulate matter, the newborn methylome, and cardio-respiratory health outcomes in childhood. *Environ Epigenet.* 2016;2,dvw005.
30. 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;26,4067-4085.
31. Irwin RE, Pentieva K, Cassidy T, et al. The interplay between DNA methylation, folate and neurocognitive development. *Epigenomics.* 2016;8,863-879.
32. 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;7,10577.
33. Sharp GC, Arathimos R, Reese SE, et al. Maternal alcohol consumption and offspring DNA methylation: findings from six general population-based birth cohorts. *Epigenomics.* 2018;10,27-42.
34. Rijlaarsdam J, Pappa I, Walton E, et al. An epigenome-wide association meta-analysis of prenatal maternal stress in neonates: A model approach for replication. *Epigenetics.* 2016;11,140-149.
35. Paquette AG, Lester BM, Lesueur C, et al. Placental epigenetic patterning of glucocorticoid response genes is associated with infant neurodevelopment. *Epigenomics.* 2015;7,767-779.
36. Barker ED, Walton E, Cecil CAM. Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *J Child Psychol Psychiatry.* 2017.

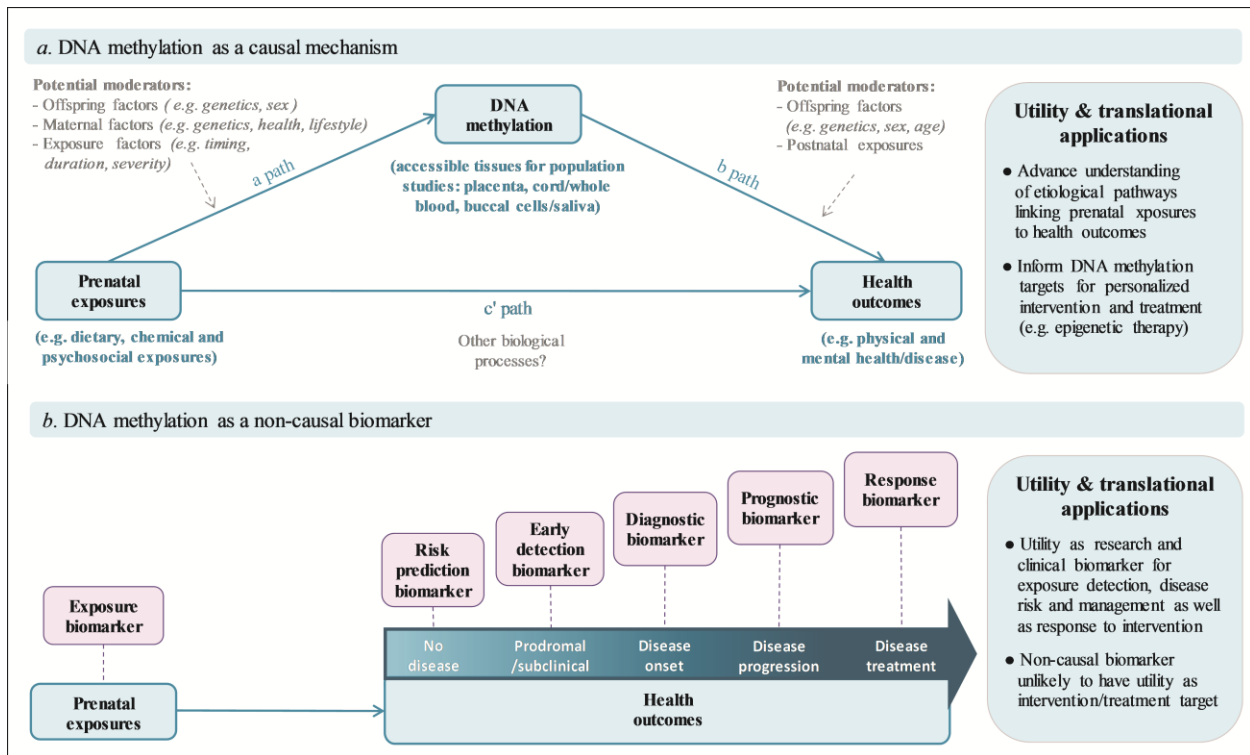
37. Cecil CAM, Walton E, Jaffee SR, et al. Neonatal DNA methylation and early-onset conduct problems: A genome-wide, prospective study. *Dev Psychopathol.* 2017;1-15.
38. van Mil NH, Steegers-Theunissen RP, Bouwland-Both MI, et al. DNA methylation profiles at birth and child ADHD symptoms. *J Psychiatr Res.* 2014;49,51-59.
39. Walton E, Pingault JB, Cecil CA, et al. Epigenetic profiling of ADHD symptoms trajectories: a prospective, methylome-wide study. *Mol Psychiatry.* 2017;22,250-256.
40. Walton E, Cecil CAM, Suderman M, et al. Longitudinal epigenetic predictors of amygdala:hippocampus volume ratio. *J Child Psychol Psychiatry.* 2017;58,1341-1350.
41. Agha G, Houseman EA, Kelsey KT, et al. Adiposity is associated with DNA methylation profile in adipose tissue. *Int J Epidemiol.* 2015;44,1277-1287.
42. Aslibekyan S, Demerath EW, Mendelson M, et al. Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference. *Obesity (Silver Spring).* 2015;23,1493-1501.
43. Demerath EW, Guan W, Grove ML, et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Hum Mol Genet.* 2015;24,4464-4479.
44. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet.* 2014;383,1990-1998.
45. Meeks KAC, Henneman P, Venema A, et al. An epigenome-wide association study in whole blood of measures of adiposity among Ghanaians: the RODAM study. *Clin Epigenetics.* 2017;9,103.
46. Wahl S, Drong A, Lehne B, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature.* 2017;541,81-86.
47. Agha G, Hajj H, Rifas-Shiman SL, et al. Birth weight-for-gestational age is associated with DNA methylation at birth and in childhood. *Clin Epigenetics.* 2016;8,118.
48. Engel SM, Joubert BR, Wu MC, et al. Neonatal genome-wide methylation patterns in relation to birth weight in the Norwegian Mother and Child Cohort. *Am J Epidemiol.* 2014;179,834-842.
49. Haworth KE, Farrell WE, Emes RD, et al. Methylation of the FGFR2 gene is associated with high birth weight centile in humans. *Epigenomics.* 2014;6,477-491.
50. Pan H, Lin X, Wu Y, et al. HIF3A association with adiposity: the story begins before birth. *Epigenomics.* 2015;7,937-950.
51. Simpkin AJ, Suderman M, Gaunt TR, et al. Longitudinal analysis of DNA methylation associated with birth weight and gestational age. *Hum Mol Genet.* 2015;24,3752-3763.
52. Godfrey KM, Sheppard A, Gluckman PD, et al. Epigenetic gene promoter methylation at birth is associated with child's later adiposity. *Diabetes.* 2011;60,1528-1534.
53. Lillycrop K, Murray R, Cheong C, et al. ANRIL Promoter DNA Methylation: A Perinatal Marker for Later Adiposity. *EBioMedicine.* 2017;19,60-72.
54. van Dijk SJ, Peters TJ, Buckley M, et al. DNA methylation in blood from neonatal screening cards and the association with BMI and insulin sensitivity in early childhood. *Int J Obes (Lond).* 2018;42,28-35.
55. Ali O, Cerjak D, Kent JW, Jr., et al. Methylation of SOCS3 is inversely associated with metabolic syndrome in an epigenome-wide association study of obesity. *Epigenetics.* 2016;11,699-707.

56. Rzehak P, Covic M, Saffery R, et al. DNA-Methylation and Body Composition in Preschool Children: Epigenome-Wide-Analysis in the European Childhood Obesity Project (CHOP)-Study. *Sci Rep*. 2017;7,14349.
57. Wang S, Song J, Yang Y, et al. HIF3A DNA Methylation Is Associated with Childhood Obesity and ALT. *PLoS One*. 2015;10,e0145944.
58. Barres R, Kirchner H, Rasmussen M, et al. Weight loss after gastric bypass surgery in human obesity remodels promoter methylation. *Cell Rep*. 2013;3,1020-1027.
59. Benton MC, Johnstone A, Eccles D, et al. An analysis of DNA methylation in human adipose tissue reveals differential modification of obesity genes before and after gastric bypass and weight loss. *Genome Biol*. 2015;16,8.
60. Richmond RC, Sharp GC, Ward ME, et al. DNA Methylation and BMI: Investigating Identified Methylation Sites at HIF3A in a Causal Framework. *Diabetes*. 2016;65,1231-1244.
61. Guenard F, Deshaies Y, Cianflone K, et al. Differential methylation in glucoregulatory genes of offspring born before vs. after maternal gastrointestinal bypass surgery. *Proc Natl Acad Sci U S A*. 2013;110,11439-11444.
62. Cecil CA, Walton E, Smith RG, et al. DNA methylation and substance-use risk: a prospective, genome-wide study spanning gestation to adolescence. *Transl Psychiatry*. 2016;6,e976.
63. Bouwland-Both MI, van Mil NH, Tolhoek CP, et al. Prenatal parental tobacco smoking, gene specific DNA methylation, and newborns size: the Generation R study. *Clin Epigenetics*. 2015;7,83.
64. Jiang R, Jones MJ, Chen E, et al. Discordance of DNA methylation variance between two accessible human tissues. *Sci Rep*. 2015;5,8257.
65. Davies MN, Volta M, Pidsley R, et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol*. 2012;13,R43.
66. Edgar RD, Jones MJ, Meaney MJ, Turecki G, Kobor MS. BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. *Transl Psychiatry*. 2017;7,e1187.
67. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*. 2015;10,1024-1032.
68. Huang YT, Chu S, Loucks EB, et al. Epigenome-wide profiling of DNA methylation in paired samples of adipose tissue and blood. *Epigenetics*. 2016;11,227-236.
69. Stueve TR, Li WQ, Shi J, et al. Epigenome-wide analysis of DNA methylation in lung tissue shows concordance with blood studies and identifies tobacco smoke-inducible enhancers. *Hum Mol Genet*. 2017;26,3014-3027.
70. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518,317-330.
71. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13,86.
72. Reinius LE, Acevedo N, Joerink M, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One*. 2012;7,e41361.

73. Bakulski KM, Feinberg JI, Andrews SV, et al. DNA methylation of cord blood cell types: Applications for mixed cell birth studies. *Epigenetics*. 2016;11,354-362.
74. de Goede OM, Razzaghi HR, Price EM, et al. Nucleated red blood cells impact DNA methylation and expression analyses of cord blood hematopoietic cells. *Clin Epigenetics*. 2015;7,95.
75. Gervin K, Page CM, Aass HC, et al. Cell type specific DNA methylation in cord blood: A 450K-reference data set and cell count-based validation of estimated cell type composition. *Epigenetics*. 2016;11,690-698.
76. Kaushal A, Zhang H, Karmaus WJJ, et al. Comparison of different cell type correction methods for genome-scale epigenetics studies. *BMC Bioinformatics*. 2017;18,216.
77. McGregor K, Bernatsky S, Colmegna I, et al. An evaluation of methods correcting for cell-type heterogeneity in DNA methylation studies. *Genome Biol*. 2016;17,84.
78. Teschendorff AE, Zheng SC. Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. *Epigenomics*. 2017;9,757-768.
79. Almouzni G, Altucci L, Amati B, et al. Relationship between genome and epigenome--challenges and requirements for future research. *BMC Genomics*. 2014;15,487.
80. Felix JF, Joubert BR, Baccarelli AA, et al. Cohort Profile: Pregnancy And Childhood Epigenetics (PACE) Consortium. *Int J Epidemiol*. 2017.
81. Benke KS, Nivard MG, Velders FP, et al. A genome-wide association meta-analysis of preschool internalizing problems. *J Am Acad Child Adolesc Psychiatry*. 2014;53,667-676 e667.
82. Bradfield JP, Taal HR, Timpson NJ, et al. A genome-wide association meta-analysis identifies new childhood obesity loci. *Nat Genet*. 2012;44,526-531.
83. Felix JF, Bradfield JP, Monnereau C, et al. Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Hum Mol Genet*. 2016;25,389-403.
84. Horikoshi M, Beaumont RN, Day FR, et al. Genome-wide associations for birth weight and correlations with adult disease. *Nature*. 2016;538,248-252.
85. Pappa I, St Pourcain B, Benke K, et al. A genome-wide approach to children's aggressive behavior: The EAGLE consortium. *Am J Med Genet B Neuropsychiatr Genet*. 2015.
86. Bell JT, Pai AA, Pickrell JK, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol*. 2011;12,R10.
87. Gaunt TR, Shihab HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biol*. 2016;17,61.
88. Grundberg E, Meduri E, Sandling JK, et al. Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. *Am J Hum Genet*. 2013;93,876-890.
89. Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol*. 2012;41,161-176.
90. Caramaschi D, Sharp GC, Nohr EA, et al. Exploring a causal role of DNA methylation in the relationship between maternal vitamin B12 during pregnancy and child's IQ at age 8, cognitive performance and educational attainment: a two-step Mendelian randomization study. *Hum Mol Genet*. 2017;26,3001-3013.

91. Ladd-Acosta C, Fallin MD. The role of epigenetics in genetic and environmental epidemiology. *Epigenomics*. 2016;8,271-283.
92. Bohlin J, Haberg SE, Magnus P, et al. Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biol*. 2016;17,207.
93. Knight AK, Craig JM, Theda C, et al. An epigenetic clock for gestational age at birth based on blood methylation data. *Genome Biol*. 2016;17,206.
94. Liu C, Marioni RE, Hedman AK, et al. A DNA methylation biomarker of alcohol consumption. *Mol Psychiatry*. 2016.
95. Reese SE, Zhao S, Wu MC, et al. DNA Methylation Score as a Biomarker in Newborns for Sustained Maternal Smoking during Pregnancy. *Environ Health Perspect*. 2017;125,760-766.
96. Guxens M, Ballester F, Espada M, et al. Cohort Profile: the INMA--Infancia y Medio Ambiente-- (Environment and Childhood) Project. *Int J Epidemiol*. 2012;41,930-940.
97. Moore SE, Fulford AJ, Darboe MK, et al. A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: the ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth*. 2012;12,107.
98. Inskip HM, Godfrey KM, Robinson SM, et al. Cohort profile: The Southampton Women's Survey. *Int J Epidemiol*. 2006;35,42-48.

Figure 1



a. The causal mediation model, whereby prenatal exposures (independent variable) partly influence health outcomes in the offspring (dependent variable) via changes in DNA methylation (mediator variable). Of note, both the *a* path and the *b* path are hypothesised to be moderated by genetic effects, as well as additional factors. Furthermore, DNA methylation may also mediate genetic (as well as environmental) effects. *b.* The alternative non-causal model, whereby DNA methylation can serve as a biomarker of, but not causal mechanism in, exposure-outcome associations. Note that we present here the two models that are most relevant to the DOHaD framework; however, it is important to note that other models have also been proposed. For example, DNA methylation may function as a moderator of genetic and environmental influences on outcomes or as a mediator of genetic influences on outcomes. Also, stochastic changes may influence DNA methylation. More complex models are also possible [see for a more detailed discussion:⁹¹].