A REVIEW OF THE EPIGENETIC CLOCK: EMERGING BIOMARKERS FOR ASTHMA AND ALLERGIC DISEASE

Denise Daley[1], Denitsa Vasileva[1], and Celia Greenwood M T[2]

[1]The University of British Columbia Centre for Heart Lung Innovation
[2]Lady Davis Institute for Medical Research

March 23, 2023

Abstract

DNA methylation (DNAm) is a dynamic, age-dependent, epigenetic modification that can be used to study interactions between genetic and environmental factors. Environmental exposures during critical periods of growth and development may alter DNAm patterns, leading to increased susceptibility to diseases such as asthma and allergy. One method to study the role of DNAm is the epigenetic clock – an algorithm that uses DNAm levels at select age informative Cytosine-phosphate-Guanine (CpG) dinucleotides to predict epigenetic age (EA). The difference between EA and calendar age (CA) is termed epigenetic age acceleration (EAA) and reveals information about the biological capacity of an individual. Associations between EAA and disease susceptibility have been demonstrated for a variety of age-related conditions and, more recently, phenotypes such asthma and allergic diseases which often begin in childhood and progress throughout the lifespan. In this review, we explore different epigenetic clocks and how they have been applied, particularly as related to childhood asthma and *in utero* and early life exposures (e.g., smoking, diet, and drugs) that result in methylation changes.

A REVIEW OF THE EPIGENETIC CLOCK: EMERGING BIOMARKERS FOR ASTHMA AND ALLERGIC DISEASE

POTENTIAL AUTHOR LIST:

Denitsa Vasileva[1] (ORCID: 0000-0002-8778-6442)

Celia M T Greenwood[2] (ORCID: 0000-0002-2427-5696)

Denise Daley[1,3*](ORCID: 0000-0001-9624-3565)

AFFILIATIONS:

[1]Centre for Heart Lung Innovation, University of British Columbia and Saint Paul's Hospital, Vancouver BC, Canada

[2] Lady Davis Institute for Medical Research, Montreal, Canada

[4] McGill University: Gerald Bronfman Department of Oncology, Department of Epidemiology & Biostatistics, and Department of Human Genetics

[3]Department of Medicine, Respiratory Division, University of British Columbia, Vancouver, BC, Canada

*Correspondence should be addressed to Dr. Denise Daley, Centre for Heart Lung Innovation, St. Paul's Hospital, 1081 Burrard Street Room 166, Vancouver Canada V6J1Y6.

E-mail address: denise.daley@hli.ubc.ca

Funding Acknowledgements: Canadian Epigenetics, Environment, and Health Research Consortium (CEEHRC) Team Grant funding reference number 128091 (PI DD and COI CMTG). Four Year Fellowship (DV).

CONFLICT OF INTEREST: The authors declare no conflict of interest in relation to this work.

N=170 words,

Limit:200

Abstract

DNA methylation (DNAm) is a dynamic, age-dependent, epigenetic modification that can be used to study interactions between genetic and environmental factors. Environmental exposures during critical periods of growth and development may alter DNAm patterns, leading to increased susceptibility to diseases such as asthma and allergy. One method to study the role of DNAm is the epigenetic clock – an algorithm that uses DNAm levels at select age informative Cytosine-phosphate-Guanine (CpG) dinucleotides to predict epigenetic age (EA). The difference between EA and calendar age (CA) is termed epigenetic age acceleration (EAA) and reveals information about the biological capacity of an individual. Associations between EAA and disease susceptibility have been demonstrated for a variety of age-related conditions and, more recently, phenotypes such asthma and allergic diseases which often begin in childhood and progress throughout the lifespan. In this review, we explore different epigenetic clocks and how they have been applied, particularly as related to childhood asthma and *in utero* and early life exposures (e.g., smoking, diet, and drugs) that result in methylation changes.

Key Words (limit 5). Allergy. Asthma. Biomarker. Epigenetic Clock.

Limit=4500

Current Word Count: 3606

Introduction:

DNA methylation (DNAm) is a dynamic epigenetic modification that refers to the bonding of a methyl (CH_3) group to the 5^{th} carbon of a Cytosine base to form 5-methyl-Cytosine (5mC)[1]. This process primarily occurs at Cytosine-phosphate-Guanine (CpG) dinucleotides[2]. DNAm is altered by factors such as environmental exposures (e.g. smoking[3], pesticide exposures[4]), disease (including asthma[5]), cell type[6], sex[7] and age[2]. As such, the study of DNAm has great potential to help characterize the impact of gene-environment interactions on the development of disease.

There is growing interest in considering aging as more than just the passing of calendar years but, rather, as a life-long process beginning at birth and influenced by an accumulation of environmental exposures and disease. Consistent patterns of age-associated DNAm changes have been identified[2,8-11], resulting in the development of epigenetic clocks[12] that reflect biological aging (e.g. susceptibility to disease, fragility and early mortality). These 'clocks' refer to mathematical algorithms that use DNAm levels at select CpG sites to calculate epigenetic age (EA)[13]. A variety of disease phenotypes[14-18], including asthma and allergy[19,20], have demonstrated associations with epigenetic age acceleration (EAA), where epigenetic age is different than chronological, or actual, age (CA).

EAA captures biological aging which may be faster or slower than calendar time. A key question is whether EAA is established during childhood and impacts disease risk in later life. This aligns with the Developmental Origins of Health and Disease (DOHaD)[21] hypothesis that prenatal and early life exposures during the first 1000 days cause DNAm changes that influence the likelihood of disease in later life. This premise has led to a focus on the study of epigenetic aging in relation to prenatal and childhood exposures (**Fig.1A**). There is evidence of differential methylation in cord blood due to *in utero* exposures, including elevated maternal BMI[22], air pollution[23], and the widely reported and replicated effects of maternal smoking[24](**Table 1**). In addition, maternal smoking has been linked to increased EAA well into childhood (years 6-11)[4].

Asthma is a complex phenotype influenced by genetic and environmental factors, that demonstrates age and sex specific prevalence patterns[25]. Childhood asthmatics are predominately male (65%), while 65% of adult asthmatics are female (**Fig.1B**)[25]. There are two main asthma subtypes – allergic and non-allergic. Allergic asthmatics have an additional allergic disease (e.g., atopy or eczema), whereas non-allergic asthmatics have asthma without additional allergic disease. Understanding the relationship between EAA and asthma[26] may clarify the mechanisms whereby early life exposures affect methylation and disease susceptibility.

Purpose

The purpose of this review is to provide an overview of current epigenetic clocks, explore their application in early childhood, and highlight applications to asthma and allergic disease.

Overview of DNA Methylation (DNAm)

About 70% of CpG sites in the human genome are methylated[27]. CpGs are concentrated in CpG islands (CGIs) – regions at least 200 base pairs (bp) in length – where C-G dinucleotides make up more than 50% of the sequence[1,28]. CGIs house the promoter regions of ˜ 70% of human genes[29,28]. The effect of DNAm on gene expression is influenced by CpG density in these promoter regions[30].

Different technologies have been developed for assessing DNAm, but arrays and sequencing protocols form the basis of the literature. Both rely on the bisulfite conversion of DNA. Arrays compare signal intensities between methylated and unmethylated probes at specific sites while in sequencing, the proportion of methylated Cytosines is calculated. Three arrays have been most commonly used to study DNAm in humans: the legacy IlluminaHumanMethylation27 BeadChip[31] , the Illumina HumanMethylation450 BeadChip[32] and finally the Illumina MethylationEPIC BeadChip array[33]. Each arrays features progressive expansion of CpG coverage and increased representation of different regions of the genome. The EPIC array covers ˜30x more CpGs compared to the Illumina 27K array and puts greater focus on CpGs outside of CGIs as these regions are important for gene regulation[33].

DNAm and Aging

Consistent changes in DNAm over the lifespan across individuals have been identified by multiple studies[9,34-37]. Cord blood tends to show low levels of DNAm[10,38]followed by a rapid increase in the early years of life[38,39], and then a gradual loss of methylation with aging[2]. In particular, CpG sites linked to embryonic developmental genes gain methylation during childhood, while regions related to immune processes lose methylation[40]. Genes located in the Major Histocompatibility Complex (MHC) class I and II[40] – in particular*HLA-B* , *HLA-C* , *HLA-DMA* , *HLA-DPB1* – become demethylated with age. This is significant as MHC I and II genes have previously been associated with asthma and allergic disease[41-44,40,45,46].

The immune system has two key mechanisms: innate immunity and adaptive immunity. MHC complexes are part of the adaptive immune system and are involved in recognizing and destroying pathogens[47]. The innate system is present in the fetus and at birth, but is subdued to tolerate the stress of fetal development[48]. The adaptive immune response develops throughout the lifespan, with T cells playing a key role; the helper T cells can be further divided into Th1 and Th2 cells. Th2 cells stimulate the production of antibodies and have been linked to increased IgE response in atopy[49], and to asthma and other allergic diseases[50]. Fetal and neonatal T cells differ significantly from adult cells. Environmental exposures may activate fetal/neonatal T-cells, resulting in a Th2 immune response[48]. For example, Zhang *et al* . found that changes in DNAm within Th2 pathway genes between the ages of 10 and 18 increase the risk for acquisition of asthma in girls[51].

The rate of DNAm fluctuation during childhood is three to four times greater than in adulthood[52]. Studies have suggested that early life changes might follow a logarithmic, rather than linear, pattern[52-54]. In older age, there is a pattern of increased DNAm discordance across individuals (the so-called "epigenetic drift" phenomenon). Epigenetic drift[2] is defined by the accumulation of random changes in DNAm over time, leading to increased inter-individual variance in methylation patterns with age[36,37,55-57], a difference that is more pronounced in later life[36,37,55]. The increase in epigenetic drift in older age has led researchers to

postulate that aging may be a process of 'memorizing' life-time environmental exposures[58]. Understanding the effect of these exposures on DNAm is essential to the study of differential rates of aging.

Epigenetic Clocks

The epigenetic clock exploits the reproducible relationship between DNAm at specific CpGs and age, to predict epigenetic age (EA). The difference between EA and CA highlights changes in cell or tissue function[13]. Positive EAA (EA > CA) in adults has been implicated in increased susceptibility to disease (e.g. Alzheimer's disease, B cell lymphoma)[15,18]and increased mortality[16,59]. In children, positive EAA has been found in connection with maternal smoking[60] and alcohol use[61], as well as with diseases such as allergy and asthma[19]. There have also been exposures (e.g. exercise, consumption of fish and fruits and vegetables) associated with negative EAA[62,63](epigenetic age < chronological age).

CpGs in most epigenetic clocks were selected using penalized linear regression methods[13]such as elastic net, which protect against overfitting in models containing many predictor variables[64]. However, these linear methods do not account for any non-additive interactions between CpG sites, and may not fully capture the complexity of DNAm in the aging process[65]. This limitation can be addressed by using non-linear methods in the development of the epigenetic clocks.

Epigenetic Clock Training Metrics

There are two types of epigenetic clocks – first and second-generation. First-generation clocks use raw or log-transformed CA[53] as the dependent variable, whereas a composite measure of aging is the dependent variable in second generation clocks. For the composite measure, "proxies" of biological aging, such as markers of heart and kidney function, are used in addition to CA[66]. These clocks are described in detail below.

First-Generation Epigenetic Clocks

First generation clocks have been the most widely studied to date in the studies of epigenetic aging. They are broadly generalizable to different populations, and their accuracy can be easily assessed. Despite only being trained on CA, the EAA calculated by these clocks has been implicated in the incidence of many diseases[14]. First generation clocks can be further classified into single and multi-tissue clocks. In this review, we will discuss in detail the first-generation clocks that have broad utility (e.g. the Hannum clock[67], the Horvath pan-tissue[53] and Skin & Blood clocks[68], the PedBE clock[69]). The features of other clocks tailored to narrower use cases (e.g., those aimed at studying specific tissues or primarily focusing on an age group such as neonates), are summarized in **Table 2** . The accuracy of first-generation epigenetic clocks is assessed in relation to CA, usually using Absolute Error (AE=|epigenetic-chronological age|) or Pearson's correlation coefficient (r) between EA and CA. The EA calculated by these clocks is correlated with CA but the deviation between the two has been shown to be informative of 'biological capacity' (e.g. physical fragility, disease susceptibility) in adults[65]. First-generation clocks represent facets of both chronological and biological aging; separating these components remains a major challenge[65].

Bocklandt *et al* developed the first epigenetic clock in 2011, marking a milestone in the field. However, this clock was aimed specifically at saliva samples and due to the tissue specificity of DNAm has not been generalized to other sample types (**Table 2)** [57]. Shortly thereafter, the Hannum[67] clock – a broadly used blood epigenetic clock – was published. Using Illumina 450K array data from 656 samples (482 training set and 174 testing set) of whole blood (age range: 19-101years), the Hannum clock was developed in stages[67]. In the first step, ˜70,000 age associated autosomal CpG sites were identified. Then, elastic net regression with bootstrapping was performed to build an epigenetic clock consisting of 71 CpGs, most of which were located close to genes implicated in age-related conditions[67](**Table 2**). When applied to pediatric samples (CA <18years) the Hannum clock demonstrated low accuracy[61], potentially due to a lack of pediatric samples in the dataset used to develop this clock[61].

First-Generation Multi-Tissue Epigenetic Clocks

The next major milestone in the field of epigenetic aging was the development of multi-tissue epigenetic clocks. Teschendorff *et al.* , 2010[70] described a set of 69 CpGs with age-associated increases in methylation in both blood and epithelial tissue, demonstrating a pan-tissue signature of aging. Koch *et al.* used Illumina 27K data from four different tissue types (**Table 2**) to develop a multi-tissue clock with mean AE of 11 years[71]. This clock was quickly followed by two multi-tissue clocks developed by Horvath *et al.* : the pan-tissue[53] and Skin & Blood clocks[68].

Since its development, the pan-tissue Horvath clock[53] has been the backbone of epigenetic aging studies. It was developed using 8,000 samples (from 51 healthy tissues) of Illumina 27K and 450K data, divided into training and validation cohorts. Elastic net with 10-fold cross validation was performed on the DNAm values of 21,369 CpG sites with a log- transformed version of CA as the dependent variable. This regression yielded a clock of 353 CpG sites with a median AE of 3.6 years in the validation cohort[53]. The pan-tissue Horvath clock has been repeatedly validated and has shown high accuracy, even when applied to data from the Illumina 850K EPIC array (missing 19/353 CpG sites)[72], and robustness to changes in cell type composition[61]. However, the pan-tissue clock underestimates epigenetic age in older individuals[73]. In addition, while this clock is more accurate in pediatric samples than the Hannum clock[61], it was predominantly developed using adult samples and may not contain the CpG sites associated with early developmental processes.

The Skin & Blood epigenetic clock is another multi-tissue clock developed by Horvath *et al* . It aims to improve the accuracy of the Horvath pan-tissue clock in fibroblasts[68]. The Skin & Blood clock consists of 391 CpG sites and was developed through a similar process as the pan-tissue clock, but using Illumina 450K or 850K array data from buccal cells, fibroblasts, keratinocytes, endothelial cells, blood, and saliva[68]. In addition to better performance in fibroblasts, this clock is more accurate than both the Horvath pan-tissue clock and Hannum clock in blood samples (median AE=2.5 years vs. 3.7 and 5.1 years)[68].

Pediatric Epigenetic Clock

As the field of epigenetic aging has broadened to study the effects of childhood exposures on aging, pediatric epigenetic clocks have been developed. The most prominent childhood clock is the Pediatric-Buccal-Epigenetic clock[69] (PedBE) - a 94-CpG buccal epithelial cell clock developed using exclusively pediatric samples (n=1,032, age range: 0.17-19.47 years) and elastic net regression. PedBE's performance was evaluated in an independent set of 689 buccal samples (age range: 0.01-19.96 years)[69] where it had median AE=0.35 years and r=0.98, demonstrating greater accuracy for that age group compared to the pan-tissue Horvath clock[69]. However, when applied to an independent set of blood samples (n=134), the PedBE clock was not as accurate (median AE =3.26 years) as the Horvath pan-tissue clock (median AE= 0.57 years)[69]. This performance discrepancy (blood vs. buccal samples) was expected due to the tissue and cell-type specificity of DNAm. DNAm patterns are highly influenced by tissue types as well as by cell type proportions in whole blood.

Gestational Age Clocks

The cell type specificity of DNAm has complicated the use of the epigenetic clock to study pre-natal environmental exposures[74] as the composition of cord blood is distinct from venous blood, thereby rendering clocks developed using venous blood unsuitable for gestational epigenetic age (GEA) predictions. The Horvath pan-tissue clock incorporated cord blood samples in its training set but set their CA at "0"[75]. This may lead to lower accuracy in neonatal blood samples as it does not account for their actual gestational age. Since then, gestational epigenetic clocks have been developed as summarized in **Table 2** .

The study of the relationship between maternal exposures and DNAm has expanded rapidly. However, some studies use gestational age (as estimated through either last menstrual period or ultrasound methods)[24] while others use GEA. Much like the use of EAA in childhood and adult studies, GEA can provide insight into the role of DNA methylation in traits in infancy.

Second-Generation Epigenetic Clocks

First-generation epigenetic clocks are useful in the study of phenotypes and healthy aging, but may not

select health informative CpG sites because they were trained exclusively on CA (rather than variables more tightly linked to health status[12]. Second-generation epigenetic clocks are trained on proxy variables (e.g., five plasma proteins, and smoking status) for biological aging in addition to CA. These clocks aim to improve on the performance of the first-generation clocks in predicting disease development and mortality. PhenoAge[66] and GrimAge[76](**Table 3**) are second-generation clocks and predict mortality more accurately than first generation clocks. However, there are several key limitations. The clinical markers used to generate a composite "biological age" value makes these clocks difficult to implement in many contexts, as such detailed health data may not always be available. Clinical variables used in the development of these clocks may be relevant to aging in older adults but may not be informative of biological aging in pediatric samples. Pediatric-specific second-generation clocks may be needed to assess the accuracy of second-generation clocks in predicting childhood phenotypes.

Metrics of EAA

The cell and tissue specificity of DNAm also has implications for the calculation of EAA.

In blood data, there are two prominent measures of EAA: Intrinsic Epigenetic Age Acceleration (IEAA) and Extrinsic Epigenetic Age Acceleration (EEAA). IEAA is used in conjunction with the Horvath pan-tissue clock and treats cell type proportions as confounders, highlighting "intrinsic aging processes", which are not influenced by cell type composition changes. EEAA is used to assess EAA from Hannum clock age estimates. Since this clock was developed solely using blood samples and its results have been shown to reflect changes in cell type compositions, EEAA represents changes in the immune system.[77]. Other commonly used methods for assessing EAA are the difference between EA and CA and the residual of the regression of EA on CA.

Applications of the First-Generation Epigenetic Clocks to Asthma and Allergic Disease Research

The asthma phenotype, and in particular, allergic asthma, presents a test case for the utility of the epigenetic clock in studying prenatal and childhood traits and exposures over the lifespan. Asthma often begins in the early years and may be transient. The clinical presentation is heterogeneous, especially in early childhood, making timely diagnosis difficult[19]. Wheeze in young children may resolve without intervention or may persist and transition into asthma. In addition, there is an unexplained switch in the sex-specific prevalence of asthma between childhood and adolescence[25]. According to the atopic march theory[78], asthma and allergic diseases begin in infancy, with the first presenting symptoms being eczema/atopic dermatitis, progressing to infant food allergies, then asthma and allergic rhinitis. The atopic march is thought to be initiated by environmental exposures, and it has been proposed that the impact of environmental exposures may be reflected in DNAm, explaining the heterogeneity of asthma[19,20].

Genetic variants have been shown to account for ˜61-75% of susceptibility to asthma[79],[80]. Numerous genome-wide association studies have demonstrated associations between the HLA region and asthma and allergic disease. The remaining ˜25-40% of the risk is thought to be due to environmental factors. Epigenome Wide Association Studies (EWAS) have repeatedly implicated DNAm in both childhood and adult asthma[5,81,82].

One exposure of particular interest is viral infection, which may have either a punitive or protective effect, depending upon the age when infected and viral subtype. Infection may skew the immune response towards the Th2 pattern observed in allergy[50,83,84]. For example, infection with the respiratory syncytial virus (RSV) in infancy is associated with increased risk for asthma, and this effect appears to be mediated through changes in DNAm[83-85]. Differential methylation at three CpG sites, mapped to airway and immune response genes, can separate (with area under the curve (AUC) =1) children who will develop recurrent wheeze and asthma subsequent to an RSV infection from those who recover normally[85].

The few available studies show a positive association between allergy and asthma and EAA[19,20]. Peng *et al.* found that EEAA was linked to asthma and allergic disease (i.e. atopy, food allergy) in Project Viva – a longitudinal birth cohort which included 408 mother-child pairs with blood DNAm data at mid-childhood (mean age: 7.8 years, range: 6.7-10.2 years)[19]. Epigenetic age as predicted by the pan-tissue Horvath clock

and IEAA have also been associated with allergic disease[19]: the Peng study found a 1.21 increase in the odds of developing atopic disease and food allergy and 1.16 increase in the odds for asthma at mid-childhood for every one year increase in EA[19]. These results were independently replicated. A study examining DNAm of nasal epithelium cells of 547 Project Viva participants (mean age=12.9, range:11.9-15.3 years) showed increased EAA in children with asthma (0.74 years) and allergic asthma (1.30 years)[20].

The epigenetic clock may provide a greater understanding of the sex-specific asthma prevalence between childhood and adulthood[25,86]. It is believed that hormonal fluctuations during puberty, menstruation, pregnancy and menopause are associated with asthma pathogenesis, exacerbations and disease severity[25](**Fig.1B&C**). Sex hormones are key to determining immune response[87]; thus, an epigenetic clock that captures early development and puberty may be crucial to understanding the relationship between EAA and asthma (**Fig. 1C**). EAA has been shown to be affected by biological sex with males having higher acceleration rates compared to females[77]. This pattern is seen at mid-childhood (˜year 7) with persistence into adolescence (˜year 17) and adulthood[61].

A study by Patel *et al.* identified 13 CpG sites with sex-specific methylation that were associated with the acquisition of asthma between the ages of 10 and 18[88]. Ten of these sites were replicated in an independent cohort. Epigenetic clocks may unravel the relationships between DNAm, asthma and sex, and how they are associated with the mechanisms of sex-reversal in asthma prevalence.

EAA and the Developmental Origins of Health and Disease

Development is a highly complex essential process (**Fig. 1A**). Factors such as environmental exposures may perturb DNAm during development, and thereby lead to long-term changes that influence susceptibility to disease. In early life, the majority of CpG sites across the genome are hypomethylated, leaving them vulnerable to aberrant methylation caused by environmental factors (**Fig. 1A**). Subsequent DNAm patterns may reflect exposures during these stages of life and influence the epigenetic clock and EAA. Previous studies on development have primarily focused on the pre-natal and early life periods, but development also encompasses other life stages of rapid change such as puberty and even menopause. Studies have shown that exposures such as smoking in adolescence can alter DNAm during adolescence[89]. Whether this leads to impacts on epigenetic aging which continue in later life remains under-explored.

It has been suggested that an individual's EAA trajectory is established in childhood and continues at the same rate throughout adulthood[90]. However, more work examining this question is needed. In addition, it is necessary to study whether the rate of EAA is influenced by environmental exposures during critical developmental periods such as puberty, pregnancy, and menopause. Most current epigenetic clocks, except the PedBE clock and gestational age clocks, were developed using mainly adult samples. This means that CpG sites involved in early growth and developmental processes may not be captured in these clocks.

Conclusion

In this review, we have summarized the most frequently used epigenetic clocks as well as their applicability to childhood phenotypes, with a particular focus on asthma and allergic disease. EAA has been successfully used to link DNA methylation with disease onset. We focused on the study of epigenetic aging in asthma, and allergic disease and the potential of epigenetic clocks as a biomarker in asthma diagnosis. This review has shown the potential of epigenetic clocks as a tool in the study of aging beginning at birth and has identified areas for continued work. Novel clocks incorporating pediatric longitudinal data can help further characterize the connection between DNAm, asthma and epigenetic aging.

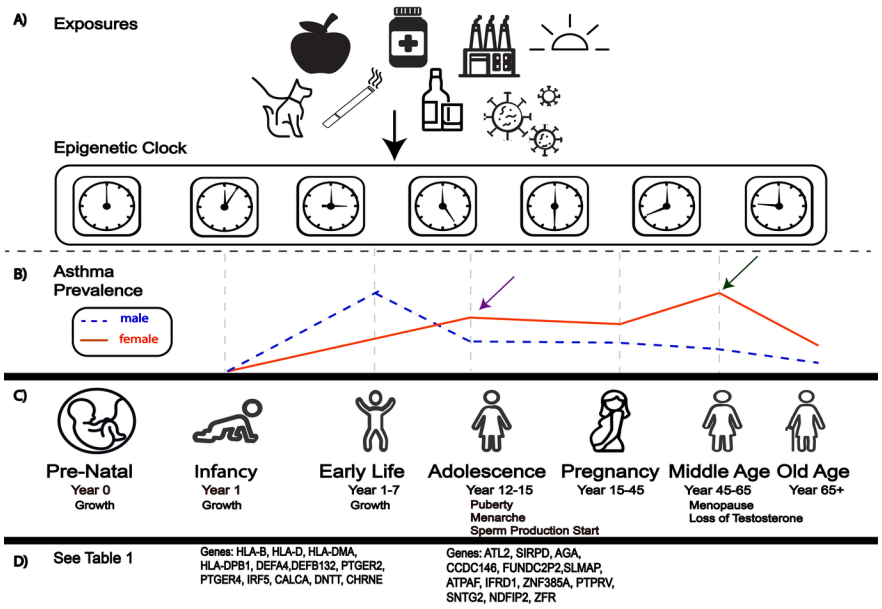

Figure 1. The epigenetic clock and Asthma Across the Life Course . Panel A illustrates different environmental exposures which may alter the rate of the epigenetic clock. Panel B is an illustrative representation of the proportion of male vs. female asthmatics over the life stages. Arrows show key changes in sex-specific patterns of asthma prevalence as extrapolated from[25,91,92]and are intended for illustrative purposes only. Panel C shows the different life stages. DNAm during periods of significant hormone changes (e.g., puberty, pregnancy, and menopause) DNA is particularly vulnerable to alteration by environmental exposures. Panel D shows the key genes undergoing changes in methylation during development ([39,40,88]).

Publication	CpG Site	CHR	Position (hg38)	Gene	Associated Exposure
[23]	cg14547404	10	48653753	ARHGAP22	Air Pollution
[23]	cg06517429	10	113679876	CASP7	Air Pollution
[24]	cg26995690	13	35772239	DCLK1	Birthweight
[24]	cg00637745	2	120739758		Birthweight
[24]	cg07133097	2	120739962		Birthweight
[22]	cg10593758	5	76952917	CRHBP	Elevated Maternal BMI
[22]	cg07621682	19	41321853	*CCDC97*	Elevated Maternal BMI
[24]	cg11932158	3	155704340	*PLCH1*	Gestational Age
[24]	cg18623216	3	155704181	*PLCH1*	Gestational Age
[24]	*cg16103712*	8	98011641	*MATN2*	Gestational Age
[24]	*cg17133774*	1	6138607	*CHD5*	Gestational Age
[24]	*cg12713583*	19	940724	*ARID3A*	Gestational Age
[24]	cg04347477	12	124517461	*NCOR2*	Gestational Age
[24]	*cg08817867*	17	19753241		Gestational Age
[24]	*cg02001279*	19	940967	*ARID3A*	Gestational Age
[24]	cg08412913	16	85395916	*DOCK6*	Gestational Age
[24]	cg06870470	19	11205091		Gestational Age
[24]	cg05549655	15	74726802	*CYP1A1*	Maternal Smoking
[24]	cg11924019	15	74726942	*CYP1A1*	Maternal Smoking
[24]	cg22549041	15	74726910	*CYP1A1*	Maternal Smoking
[24]	cg23067299	5	323791	*AHRR*	Maternal Smoking

Publication	**CpG Site**	**CHR**	**Position (hg38)**	**Gene**	**Associated Exposure**
[24]	cg22132788	7	44962886	*MYO1G*	Maternal Smoking
[24]	cg18092474	15	74726961	*CYP1A1*	Maternal Smoking
[24]	cg12803068	7	44963320	*MYO1G*	Maternal Smoking
[24]	cg12101586	15	74726862	*CYP1A1*	Maternal Smoking

Type	**Epigenetic Clock**	**Tissue Type**	**Methodology Used**	**Methylation Technology**	**Strengths**	**Limitations**
Single Tissue	[57]	Single Tissue: Saliva	Association Analysis	Illumina 27K array	First epigenetic clock	Low accuracy (mean AE:5.2 years)
	[67]	Single Tissue: Whole Blood	Elastic Net with bootstrapping	Illumina 450K Array	Accurate in blood. Extensively used	Limited age range of training samples: 19-101 years
	[93]	Single Tissue: Whole Blood	Multivariate Linear Regression	Pyrosequencing	Consists of only three CpG sites	Low accuracy (mean AE:5.4 years)
	[94]	Single Tissue: Breast Tissue	Elastic Net Regression with cross-validation	TruSeq Methyl Capture EPIC library	Improved accuracy in breast tissue	TruSeq Methyl Capture not yet widely used
Multi-Tissue	[71]	Multi-Tissue: Epidermis, dermis, T-cells, cervical smear, monocytes	Pearson Correlation	Illumina 27K array	First multi-tissue clock	Relative low accuracy (mean AE:11 years)
	[75]	Multi-Tissue: 51 tissue and cell types	Elastic Net with ten-fold cross-validation	Illumina 27K array & Illumina 450K Array	Accurate across tissues; Extensively used	Mostly adult samples Age of neonate samples set at "0"
	[68]	Multi-Tissue including blood	Elastic Net with ten-fold cross-validation	Illumina 450K & Illumina EPIC array	Accurate (mean AE: 2.5 years)	Not widely used yet
Pediatric Single Tissue	[69]	Single Tissue: Buccal Cells	Elastic Net with cross-validation	Illumina EPIC Array	Pediatric-only clock	Low accuracy in blood

Gestational Age	[74]	Cord Blood	Elastic Net Regression with cross-validation	Illumina 27K array & Illumina 450K Array	Median error :1.24 weeks	Gestational Age Only
	[95]	Cord Blood	Lasso Regression with cross-validation	Illumina EPIC array	Uses EPIC array	Gestational Age Only

Epigenetic Clock Citation	**Tissue Type**	**Methodology Used**	**Platform**	**Strengths**	**Limitations**
[66]	Elastic Net with Cross Validation	Phenotypic Age	Illumina EPIC Array	Composite of aging; well correlated with morbidity	Utility in childhood samples unknown
[76]	Elastic Net with Cross Validation	Time-to-Death	Illumina 450K & EPIC arrays	Well correlated with mortality	Utility in childhood samples unknown

Table 3. ***Second-Generation Epigenetic Clocks***

Reference:

1. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics.* 2012;13(7):484-492.

2. Jones MJ, Goodman SJ, Kobor MS. DNA methylation and healthy human aging. *Aging cell.*2015;14(6):924-932.

3. Zeilinger S, Kühnel B, Klopp N, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PloS one.* 2013;8(5):e63812.

4. de Prado-Bert P, Ruiz-Arenas C, Vives-Usano M, et al. The early-life exposome and epigenetic age acceleration in children. *Environment International.*2021;155:106683.

5. Legaki E, Arsenis C, Taka S, Papadopoulos NG. DNA methylation biomarkers in asthma and rhinitis: Are we there yet? *Clinical and translational allergy.*2022;12(3):e12131.

6. Titus AJ, Gallimore RM, Salas LA, Christensen BC. Cell-type deconvolution from DNA methylation: a review of recent applications. *Hum Mol Genet.* 2017;26(R2):R216-r224.

7. Grant OA, Wang Y, Kumari M, Zabet NR, Schalkwyk L. Characterising sex differences of autosomal DNA methylation in whole blood using the Illumina EPIC array. *Clin Epigenetics.* 2022;14(1):62.

8. Bjornsson HT, Sigurdsson MI, Fallin MD, et al. Intra-individual change over time in DNA methylation with familial clustering. *Jama.* 2008;299(24):2877-2883.

9. Rakyan VK, Down TA, Maslau S, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res.* 2010;20(4):434-439.

10. Martino DJ, Tulic MK, Gordon L, et al. Evidence for age-related and individual-specific changes in DNA methylation profile of mononuclear cells during early immune development in humans. *Epigenetics.* 2011;6(9):1085-1094.

11. Florath I, Butterbach K, Müller H, Bewerunge-Hudler M, Brenner H. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. *Hum Mol Genet.*2014;23(5):1186-1201.

h12. Bell CG, Lowe R, Adams PD, et al. DNA methylation aging clocks: challenges and recommendations.*Genome biology.* 2019;20(1):249.

13. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing.*Nature Reviews Genetics.* 2018;19(6):371-384.

14. Fransquet PD, Wrigglesworth J, Woods RL, Ernst ME, Ryan J. The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis.*Clin Epigenetics.* 2019;11(1):62.

15. Dugué PA, Bassett JK, Joo JE, et al. DNA methylation-based biological aging and cancer risk and survival: Pooled analysis of seven prospective studies. *International journal of cancer.* 2018;142(8):1611-1619.

16. Chen BH, Marioni RE, Colicino E, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging.* 2016;8(9):1844-1865.

17. Verschoor CP, Lin DTS, Kobor MS, et al. Epigenetic age is associated with baseline and 3-year change in frailty in the Canadian Longitudinal Study on Aging. *Clin Epigenetics.* 2021;13(1):163-163.

18. Grodstein F, Lemos B, Yu L, et al. The association of epigenetic clocks in brain tissue with brain pathologies and common aging phenotypes. *Neurobiology of disease.*2021;157:105428.

19. Peng C, Cardenas A, Rifas-Shiman SL, et al. Epigenetic age acceleration is associated with allergy and asthma in children in Project Viva. *J Allergy Clin Immunol.*2019;143(6):2263-2270.e2214.

20. Cardenas A, Sordillo JE, Rifas-Shiman SL, et al. The nasal methylome as a biomarker of asthma and airway inflammation in children. *Nature Communications.*2019;10(1):3095.

21. Barker DJ. The origins of the developmental origins theory. *Journal of internal medicine.*2007;261(5):412-417.

22. Ghildayal N, Fore R, Lutz SM, et al. Early-pregnancy maternal body mass index is associated with common DNA methylation markers in cord blood and placenta: a paired-tissue epigenome-wide association study. *Epigenetics.*2022;17(7):808-818.

23. Park J, Kim WJ, Kim J, et al. Prenatal Exposure to Traffic-Related Air Pollution and the DNA Methylation in Cord Blood Cells: MOCEH Study. *International journal of environmental research and public health.* 2022;19(6).

24. Akhabir L, Stringer R, Desai D, et al. DNA methylation changes in cord blood and the developmental origins of health and disease - a systematic review and replication study. *BMC genomics.* 2022;23(1):221.

25. Chowdhury NU, Guntur VP, Newcomb DC, Wechsler ME. Sex and gender in asthma. *European Respiratory Review.* 2021;30(162):210067.

26. DeVries A, Vercelli D. Epigenetic Mechanisms in Asthma. *Annals of the American Thoracic Society.*2016;13 Suppl 1(Suppl 1):S48-50.

27. Strichman-Almashanu LZ, Lee RS, Onyango PO, et al. A genome-wide screen for normally methylated human CpG islands that can identify novel imprinted genes. *Genome Res.*2002;12(4):543-554.

28. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. *Journal of molecular biology.*1987;196(2):261-282.

29. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences of the United*

*States of America.*2006;103(5):1412-1417.

30. Weber M, Hellmann I, Stadler MB, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature genetics.*2007;39(4):457-466.

31. Bibikova M, Le J, Barnes B, et al. Genome-wide DNA methylation profiling using Infinium® assay.*Epigenomics.* 2009;1(1):177-200.

32. Bibikova M, Barnes B, Tsan C, et al. High density DNA methylation array with single CpG site resolution.*Genomics.* 2011;98(4):288-295.

33. Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome biology.* 2016;17(1):208.

34. Bell JT, Tsai PC, Yang TP, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS genetics.* 2012;8(4):e1002629.

35. Day K, Waite LL, Thalacker-Mercer A, et al. Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. *Genome biology.* 2013;14(9):R102.

36. Heyn H, Li N, Ferreira HJ, et al. Distinct DNA methylomes of newborns and centenarians. *Proceedings of the National Academy of Sciences.* 2012;109(26):10522-10527.

37. Talens RP, Christensen K, Putter H, et al. Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs.*Aging cell.* 2012;11(4):694-703.

38. Herbstman JB, Wang S, Perera FP, et al. Predictors and consequences of global DNA methylation in cord blood and at three years. *PloS one.* 2013;8(9):e72824.

39. 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 biology.*2013;14(5):R42.

40. Acevedo N, Reinius LE, Vitezic M, et al. Age-associated DNA methylation changes in immune genes, histone modifiers and chromatin remodeling factors within 5 years after birth in human blood leukocytes. *Clin Epigenetics.* 2015;7(1):34.

41. Lasky-Su J, Himes BE, Raby BA, et al. HLA-DQ strikes again: genome-wide association study further confirms HLA-DQ in the diagnosis of asthma among adults. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology.* 2012;42(12):1724-1733.

42. Waage J, Standl M, Curtin JA, et al. Genome-wide association and HLA fine-mapping studies identify risk loci and genetic pathways underlying allergic rhinitis. *Nature genetics.* 2018;50(8):1072-1080.

43. Daya M, Cox C, Acevedo N, et al. Multiethnic genome-wide and HLA association study of total serum IgE level. *J Allergy Clin Immunol.* 2021;148(6):1589-1595.

44. Ramasamy A, Kuokkanen M, Vedantam S, et al. Genome-wide association studies of asthma in population-based cohorts confirm known and suggested loci and identify an additional association near HLA. *PloS one.* 2012;7(9):e44008.

45. Asai Y, Eslami A, van Ginkel CD, et al. A Canadian genome-wide association study and meta-analysis confirm HLA as a risk factor for peanut allergy independent of asthma.*J Allergy Clin Immunol.* 2018;141(4):1513-1516.

46. Madore AM, Vaillancourt VT, Asai Y, et al. HLA-DQB1*02 and DQB1*06:03P are associated with peanut allergy. *Eur J Hum Genet.* 2013;21(10):1181-1184.

47. Wieczorek M, Abualrous ET, Sticht J, et al. Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation.*Frontiers in immunology.* 2017;8:292.

48. Simon AK, Hollander GA, McMichael A. Evolution of the immune system in humans from infancy to old age.*Proceedings Biological sciences.* 2015;282(1821):20143085.

49. Berger A. Th1 and Th2 responses: what are they? *BMJ (Clinical research ed).* 2000;321(7258):424.

50. Holgate ST. Innate and adaptive immune responses in asthma. *Nature medicine.* 2012;18(5):673-683.

51. Zhang H, Tong X, Holloway JW, et al. The interplay of DNA methylation over time with Th2 pathway genetic variants on asthma risk and temporal asthma transition. *Clin Epigenetics.* 2014;6(1):8.

52. Alisch RS, Barwick BG, Chopra P, et al. Age-associated DNA methylation in pediatric populations.*Genome Res.* 2012;22(4):623-632.

53. Horvath S. DNA methylation age of human tissues and cell types. *Genome biology.* 2013;14(10):R115.

54. Snir S, Farrell C, Pellegrini M. Human epigenetic ageing is logarithmic with time across the entire lifespan. *Epigenetics.* 2019;14(9):912-926.

55. Fraga MF, Ballestar E, Paz MF, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences.*2005;102(30):10604-10609.

56. Poulsen P, Esteller M, Vaag A, Fraga MF. The Epigenetic Basis of Twin Discordance in Age-Related Diseases. *Pediatric Research.* 2007;61(7):38-42.

57. Bocklandt S, Lin W, Sehl ME, et al. Epigenetic predictor of age. *PloS one.* 2011;6(6):e14821.

58. Vaiserman A. Developmental Tuning of Epigenetic Clock. *Front Genet.* 2018;9.

59. Perna L, Zhang Y, Mons U, Holleczek B, Saum K-U, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort.*Clin Epigenetics.* 2016;8(1):64.

60. de Prado-Bert P, Ruiz-Arenas C, Vives-Usano M, et al. The early-life exposome and epigenetic age acceleration in children. *Environ Int.* 2021;155:106683.

61. Simpkin AJ, Hemani G, Suderman M, et al. Prenatal and early life influences on epigenetic age in children: a study of mother-offspring pairs from two cohort studies. *Hum Mol Genet.* 2016;25(1):191-201.

62. Quach A, Levine ME, Tanaka T, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. *Aging.* 2017;9(2):419-446.

63. Song AY, Feinberg JI, Bakulski KM, et al. Prenatal Exposure to Ambient Air Pollution and Epigenetic Aging at Birth in Newborns. *Front Genet.* 2022;13:929416.

64. Zou H, Hastie T. Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society Series B (Statistical Methodology).*2005;67(2):301-320.

65. Field AE, Robertson NA, Wang T, Havas A, Ideker T, Adams PD. DNA Methylation Clocks in Aging: Categories, Causes, and Consequences. *Molecular cell.*2018;71(6):882-895.

66. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan.*Aging.* 2018;10(4):573-591.

67. Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular cell.* 2013;49(2):359-367.

68. Horvath S, Oshima J, Martin GM, et al. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. *Aging.*2018;10(7):1758-1775.

69. McEwen LM, O'Donnell KJ, McGill MG, et al. The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proceedings of the National Academy of Sciences.* 2020;117(38):23329-23335.

70. Teschendorff AE, Menon U, Gentry-Maharaj A, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res.*2010;20(4):440-446.

71. Koch CM, Wagner W. Epigenetic-aging-signature to determine age in different tissues.*Aging.* 2011;3(10):1018-1027.

72. McEwen LM, Jones MJ, Lin DTS, et al. Systematic evaluation of DNA methylation age estimation with common preprocessing methods and the Infinium MethylationEPIC BeadChip array.*Clin Epigenetics.* 2018;10(1):123.

73. El Khoury LY, Gorrie-Stone T, Smart M, et al. Systematic underestimation of the epigenetic clock and age acceleration in older subjects. *Genome biology.*2019;20(1):283-283.

74. Knight AK, Craig JM, Theda C, et al. An epigenetic clock for gestational age at birth based on blood methylation data. *Genome biology.* 2016;17(1):206.

75. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biology.* 2013;14(10):3156.

76. Lu AT, Quach A, Wilson JG, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan.*Aging.* 2019;11(2):303-327.

77. Horvath S, Gurven M, Levine ME, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome biology.* 2016;17(1):171.

78. Bantz SK, Zhu Z, Zheng T. The Atopic March: Progression from Atopic Dermatitis to Allergic Rhinitis and Asthma. *J Clin Cell Immunol.* 2014;5(2):202.

79. McGeachie MJ, Stahl EA, Himes BE, et al. Polygenic heritability estimates in pharmacogenetics: focus on asthma and related phenotypes. *Pharmacogenetics and genomics.*2013;23(6):324-328.

80. Thomsen SF. Genetics of asthma: an introduction for the clinician. *European clinical respiratory journal.* 2015;2.

81. Xu CJ, Söderhäll C, Bustamante M, et al. DNA methylation in childhood asthma: an epigenome-wide meta-analysis. *The Lancet Respiratory medicine.*2018;6(5):379-388.

82. Forno E, Wang T, Qi C, et al. DNA methylation in nasal epithelium, atopy, and atopic asthma in children: a genome-wide study. *The Lancet Respiratory medicine.*2019;7(4):336-346.

83. Paller AS, Spergel JM, Mina-Osorio P, Irvine AD. The atopic march and atopic multimorbidity: Many trajectories, many pathways. *J Allergy Clin Immunol.*2019;143(1):46-55.

84. Daley D. The evolution of the hygiene hypothesis: the role of early-life exposures to viruses and microbes and their relationship to asthma and allergic diseases.*Curr Opin Allergy Clin Immunol.* 2014;14(5):390-396.

85. Pischedda S, Rivero-Calle I, Gómez-Carballa A, et al. Role and Diagnostic Performance of Host Epigenome in Respiratory Morbidity after RSV Infection: The EPIRESVi Study. *Frontiers in immunology.* 2022;13:875691.

86. Chlamydas S, Markouli M, Strepkos D, Piperi C. Epigenetic mechanisms regulate sex-specific bias in disease manifestations. *Journal of molecular medicine (Berlin, Germany).*2022;100(8):1111-1123.

87. Taneja V. Sex Hormones Determine Immune Response. *Frontiers in immunology.* 2018;9:1931.

88. Patel R, Solatikia F, Zhang H, et al. Sex-specific associations of asthma acquisition with changes in DNA methylation during adolescence. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology.*2021;51(2):318-328.

89. Han L, Zhang H, Kaushal A, et al. Changes in DNA methylation from pre- to post-adolescence are associated with pubertal exposures. *Clinical Epigenetics.* 2019;11(1):176.

90. Kananen L, Marttila S, Nevalainen T, et al. The trajectory of the blood DNA methylome ageing rate is largely set before adulthood: evidence from two longitudinal studies.*AGE.* 2016;38(3):65.

91. Fuseini H, Newcomb DC. Mechanisms Driving Gender Differences in Asthma. *Curr Allergy Asthma Rep.*2017;17(3):19.

92. Shah R, Newcomb DC. Sex Bias in Asthma Prevalence and Pathogenesis. *Frontiers in immunology.*2018;9:2997.

93. Weidner CI, Lin Q, Koch CM, et al. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome biology.* 2014;15(2):R24.

94. Castle JR, Lin N, Liu J, et al. Estimating breast tissue-specific DNA methylation age using next-generation sequencing data. *Clin Epigenetics.* 2020;12(1):45.

95. Haftorn KL, Lee Y, Denault WRP, et al. An EPIC predictor of gestational age and its application to newborns conceived by assisted reproductive technologies. *Clin Epigenetics.* 2021;13(1):82.