

Research Progress on Human Age Estimation from the DNA Methylation of Skeletal Materials

Tianle Lu

Nanjing Medical University, Nanjing, Jiangsu, China

ntzxltl@163.com

Abstract

Within forensic science, age estimation supports the identification of unknown individuals and carries considerable significance for the investigation and resolution of both criminal and civil cases. Epigenetic clocks, which are age prediction models built on DNA methylation, have become a focus of research in forensic age estimation. Most existing work constructs epigenetic age clocks from biological materials such as blood, saliva, buccal epithelial cells, or semen and attains a mean absolute error of roughly five years, whereas research and application centered on skeletal materials remain comparatively limited. This review first traces the three generations through which epigenetic clocks have evolved, moving from single tissue linear regression models to composite models that integrate phenotypic indicators and then to models that assess the pace of aging. It then gives a systematic account of progress in human age estimation based on DNA methylation in skeletal materials and offers an integrated analysis across sample type and target sites, detection methods and technologies, and data processing and model construction.

Keywords

Epigenetic Clock; DNA Methylation; Age Estimation; Skeleton.

1. Introduction

Age estimation holds considerable value in both clinical medicine and forensic science, and its central aim is to infer the chronological and the biological age of an individual. Chronological age (CA) refers to the number of years a person has survived counted from the day of birth, and it is absolute and irreversible. Biological age (BA) reflects the functional state of the body and gauges, from a medical and biological standpoint, how well an individual matches the expectation set by chronological age. Being relative and reversible, it offers a more accurate assessment of the degree of aging [1]. In forensic practice age estimation falls into two categories, namely estimation in the deceased and estimation in the living [2]. At the scenes of criminal cases or disasters, certain biological materials cannot be matched against a DNA database, which leaves the identity of an individual unresolved. An accurate determination of age in such circumstances assists the recognition of unknown suspects and unidentified victims, supplies investigative leads, narrows the scope of inquiry, and accelerates the identification of unknown remains. Age estimation also helps settle a range of civil disputes, among them the capacity assessment of minors and the matching of missing persons. Individual age estimation has therefore long remained both a priority and a challenge within forensic science.

Blood, saliva, semen, and similar biological materials are widely distributed and readily collected and analyzed, which makes them common samples in forensic biological evidence work. Under many real conditions, however, these materials are vulnerable to temperature, humidity, environmental pH, microorganisms, and related factors, and the DNA they contain

undergoes degradation and fragmentation that hinders precise detection [3]. Bone and teeth are more stable by comparison. Components such as hydroxyapatite and tooth enamel protect DNA at the physical level against environmental erosion and suppress the activity of enzymes and microorganisms. Even in bodies subjected to extreme environmental exposure or in skeletonized remains at a late stage, the DNA preserved within bone and teeth remains an important target for analysis [4]. Bone and teeth are therefore regarded as ideal materials for individual age estimation.

Traditional age estimation has relied chiefly on the combination of morphological observation with imaging techniques. Research shows that the morphology of bone and teeth changes markedly with age [5], and imaging has accordingly been used to observe and measure these morphological patterns. For children and adolescents the available approaches include the pattern of replacement between deciduous and permanent teeth [6] and the quantification of epiphyseal development in the wrist [7]. For adults the indicators include the pubic symphyseal surface, the auricular surface of the ilium, calcification of the costal cartilage, closure of the cranial sutures, and tooth wear [5]. These approaches place high demands on the completeness of the material and are readily swayed by the subjective judgment of the assessor. Adult estimation is also less accurate than the relatively mature methods used for minors, with errors reaching 3.5 to 8.5 years [5]. As molecular biology advanced rapidly, age estimation gradually moved from morphological methods toward molecular ones. Among the molecular options, the measurement of noncoding RNA expression, telomere length, skewed X chromosome inactivation, and chromatin remodeling is easily disturbed by the environment, storage conditions, pathological state, and other factors, so the precision is limited and broad application has been constrained [8, 9]. DNA methylation, as one form of epigenetic modification, offers high stability and a large information content together with a strong correlation with age, and it has therefore been studied and applied widely [8].

Studies of human age estimation based on DNA methylation have so far concentrated on materials such as blood, saliva, buccal epithelial cells, and semen, and the use of bone and teeth has been comparatively rare. This review surveys the literature of the past decade that has applied DNA methylation in skeletal materials to human age estimation, examines the problems present in current work, and considers future directions, with the aim of offering a reference for related research and practical application.

2. Overview of Epigenetic Clocks

2.1. Theoretical Basis of Epigenetic Clocks

Epigenetics examines the stable and heritable changes in gene expression or biological phenotype that occur without any alteration of the DNA sequence, and the principal forms include DNA methylation, change in telomere length, X chromosome inactivation, and chromatin remodeling [8]. DNA methylation (DNAm) denotes the covalent addition of a methyl group to the fifth carbon of the cytosine within a CpG dinucleotide under the catalysis of DNA methyltransferase. As one important form of epigenetic modification, it participates in the regulation of cell proliferation and differentiation as well as numerous physiological and pathological processes [8].

Compared with other epigenetic modifications, DNA methylation carries three advantages. It holds a larger quantity of information, since the methylation patterns across different CpG sites can store a great deal of genetic information. It shows higher stability, because short term changes in the internal and external environment usually do not noticeably alter the methylation map of the cellular genome. It also displays a marked association with age together with spatial and temporal specificity. As age increases, the activity of methyltransferase declines and the overall level of DNA methylation in various tissues falls

accordingly [10], while certain genes have been found to gain methylation with advancing age [11].

Epigenetic clocks are chronological age prediction models grounded in DNA methylation, and they have supplied an important tool for the study of aging and of age related biology [12]. The methylation level at many CpG sites in the human genome changes in a regular manner over time, which makes it possible to select a subset of these sites to measure the pace of aging in an individual [13]. Building on this property, machine learning is used to screen the methylation sites most strongly correlated with age and to establish a mathematical model linking methylation data to an estimate of biological age, and the result reflects the physiological state, the health risk, and the degree of aging of the individual.

2.2. Evolution of Epigenetic Clocks

Epigenetic age clocks based on DNA methylation have passed through three generations of theoretical and technical change. Working with saliva samples, Sven Bocklandt and colleagues in 2011 found that the methylation level at 88 CpG sites correlated significantly with age, and the regression model they built on this basis produced predicted ages that aligned closely with actual ages at a mean absolute error of about 5.2 years. That study was the first to propose a method of age estimation from DNA methylation [14] and laid a foundation for the later development of epigenetic age clocks.

The first generation of epigenetic age clocks depended on a specific tissue sample and a limited number of CpG sites, estimated biological age through linear regression, and drew most of its samples from whole blood or a particular organ. The best known among them is HorvathClock2013, which was the first epigenetic clock to achieve age prediction across tissues. It covered 51 types of tissue and cell, analyzed methylation at 353 characteristic CpG sites, of which 193 became more methylated and 160 less methylated with rising age, and in doing so revealed that different tissues age at different rates [15]. HannumClock2013 instead carried out a methylation analysis across the whole genome using whole blood samples, and its optimal model incorporated 71 methylation markers strongly correlated with age, yielding a correlation coefficient of 0.96 between predicted and actual age and an error of roughly 3.9 years, which represented a high level of accuracy [16]. These early models offered a scientifically workable line of thought and method for the studies of epigenetic clocks that followed.

In contrast with the first generation, the second generation of epigenetic age clocks no longer confined itself to predicting chronological age. It integrated composite indicators of phenotypic age that capture differences in lifespan and healthspan, and it covered a larger and broader set of CpG sites. The DNAmPhenoAge2018 model developed by Levine and colleagues in 2018 built a predictor of phenotypic age from 513 weighted CpG sites in whole blood, and the resulting values correlated closely not only with chronological age but also with the risk of death and with the outcomes of several aging related diseases [17]. In 2019 Ake T. Lu and colleagues proposed seven methods for estimating plasma proteins from DNA methylation and used them to develop the lifespan predictor DNAm GrimAge, which markedly strengthened the prediction of time to death and of the onset of coronary heart disease and cancer [18]. In 2022 the same team added two new plasma protein indicators to the original version and constructed GrimAge2, which further raised the accuracy of prediction for mortality and age related disease across several racial and ethnic groups [19]. The second generation thus improved the precision of chronological age prediction and demonstrated the potential of epigenetics in the assessment of health risk and of mortality outcomes.

The third generation of epigenetic age clocks comprises chiefly DunedinPoAm2020, DunedinPACE2022, and Systems Age. Compared with the previous two generations, the third generation adopts a variety of machine learning methods and can monitor and assess the

aging of different physiological systems of the body at the same time. DunedinPoAm2020 [20] and DunedinPACE2022 [21], both developed by Belsky and colleagues, rest on blood methylation data spanning the whole genome and incorporate 46 and 173 CpG sites respectively, which allows them to track the change of multiple organ systems with age. Drawing likewise on blood DNA methylation, Raghav Sehgal and colleagues in 2025 monitored the aging of 11 physiological systems at once and constructed Systems Age, an indicator able to predict a range of diseases and aging phenomena in an integrated way [22]. The arrival of the third generation has further improved the accuracy of biological age prediction and has widened the view of research on multi organ disease and aging.

Table 1 summarizes the basic features of the mainstream age prediction models across the three generations of epigenetic clocks, so that their core characteristics and line of development appear clearly.

Table 1. The three generation evolution of classical epigenetic clocks

Clock name	Researcher	Year	Sample type	No. of CpG sites	Core feature
HorvathClock2013 [15]	Horvath	2013	Multiple tissues	353	First age prediction model to achieve detection across tissues
HannumClock2013 [16]	Hannum	2013	Blood	71	Carried out methylation analysis across the whole genome, with an age correlation of 0.96 and an error of about 3.9 years in the optimal model
DNAmPhenoAge2018 [17]	Levine	2018	Blood	513	Able to predict the risk of death and the outcomes of several aging related diseases
DNAm GrimAge [18]	Ake T. Lu	2019	Blood	1030	New lifespan predictor with strong prediction of time to death and of the onset of coronary heart disease and cancer
DNAm GrimAge2 [19]	Ake T. Lu	2022	Blood	1030	Improved accuracy of prediction for mortality and aging related disease across several racial and ethnic groups
DunedinPoAm2020 [20]	Belsky	2020	Blood	150	Precisely tracks the pace and degree of aging of multiple organ systems
DunedinPACE2022 [21]	Belsky	2022	Blood	231	New aging rate methylation biomarker with accurate prediction of disease, disability, and death and higher test retest reliability
Systems Age [22]	Raghav Sehgal	2025	Blood	/	Monitors 11 organ systems at once and accurately predicts various diseases and aging phenomena

The epigenetic clocks of every generation carry broad value for the estimation of individual biological age, for the assessment of aging, and for the management of health risk. They have exposed the deep connection between DNA methylation and aging, have furnished comparatively precise tools for estimating phenotypic age with an error that usually stays within five years, and have provided quantitative means for predicting the risk of death, inferring disease outcomes, and evaluating the effect of anti aging measures. Clear limitations remain when these classical models are applied to forensic science. The samples used to build them come mostly from living clinical subjects, whereas forensic materials are drawn mostly from bodies after death. The number of CpG sites involved is also large, which makes the models difficult to apply to degraded or damaged DNA materials. There is therefore a pressing need to develop epigenetic clocks better suited to the scenarios and conditions of forensic science so that accurate age estimation becomes attainable.

3. Forensic Application of Epigenetic Clocks

3.1. Small Epigenetic Clocks in Forensic Science

The classical epigenetic clocks of every generation detect a comparatively large number of CpG sites, ranging from several dozen to more than five hundred [16,17,20,21]. A model with many sites raises the accuracy of age prediction, yet it places a higher demand on the integrity of the DNA and brings greater cost in time and money. In forensic practice the DNA of materials recovered under varying conditions is often prone to degradation and of uncertain quality, which cannot support the detection of a large number of sites. The construction of a more accurate epigenetic clock from fewer sites has therefore become a central concern and a difficulty for forensic age estimation in the years ahead.

Working with the EpiTYPER system, Lei Feng and colleagues in 2018 selected the nine best sites from 153 previously reported age related CpG sites and built an optimal age estimation model through linear regression with a mean absolute error of about 2.89 years [23], which carries considerable potential for forensic practice. Drawing on the same EpiTYPER platform, Li Shanfei and colleagues in 2019 screened eight age related sites from the blood of Han Chinese males and constructed a model with a mean absolute deviation of about 2.69 years [24]. Differing from these two studies, Hao Ting and colleagues in 2021 mainly used hair samples, screened ten CpG sites strongly correlated with age through SNaPshot technology, and incorporated sex as a factor, so that the optimal model reached a mean absolute deviation of about 3.68 years [25]. Han Xueli and colleagues in 2022 then built a methylation SNaPshot multiplex detection system from eight age related CpG sites and established two age estimation models with mean absolute deviations of 3.52 years and 2.88 years [26], an approach that proved simpler and more effective while improving both sensitivity and accuracy. In the most recent work, Zhang Tao and colleagues in 2026 adopted a domestically produced droplet digital PCR (ddPCR) technique to establish an age estimation model suited to oral swabs that incorporated seven characteristic CpG sites with a mean absolute deviation of about 3.45 to 4.30 years [27], a convenient and efficient method with sound prospects for forensic application.

The age related CpG sites employed in the models above never exceed ten, a sharp reduction from the classical clocks that span several hundred sites. This streamlined design not only lowers cost markedly but also shortens detection time and so raises efficiency considerably. Fewer sites at the same time mean a lower demand on the quality of the material and on the integrity of the DNA, which suits the trace, degraded, or mixed materials commonly encountered in forensic science and gives these models stronger general applicability and greater promise for wider use.

3.2. Epigenetic Clocks Based on Skeletal Materials

With the exception of HorvathClock2013 [15], which is a multi tissue age prediction model, the classical models rely for the most part on materials such as blood and saliva [16-21], and the small clocks described above likewise draw mainly on blood, hair, and oral swabs [23-27]. Such materials are easy to collect, extract, and analyze, and their application is very broad, whereas bone and teeth have been used comparatively little. Across the varied scenarios of forensic science, however, body fluid and soft tissue materials are readily affected by environmental factors such as temperature, humidity, ultraviolet light, and microorganisms as well as by the physiological and pathological state of the individual, and the DNA within them often breaks and degrades in a way that resists accurate detection, which limits the applicability of the models above. The hydroxyapatite inside bone and the enamel on the surface of teeth provide sound physical protection for DNA [4] and keep it intact even in late-stage bodies or in damaged and degraded materials, which makes bone and teeth ideal materials for forensic examination. Research has further shown that DNA methylation carries marked tissue specificity [28], so a clock model built on blood or saliva cannot be transferred directly to skeletal materials. The construction of epigenetic clocks suited to the many scenarios of forensic science from the DNA methylation of bone or teeth is for these reasons especially urgent and important.

Scholars in China and abroad have already used a small number of specific CpG sites on bone and tooth materials to construct several age estimation models, which suit the forensic field better than the three generations of classical clocks.

As early as 2015 Bram Bekaert and colleagues identified four age related genes, namely ASPA, PDE4C, ELOVL2, and EDARADD. They built an age estimation model from blood samples supplemented by third molar dentine samples, and the prediction error for the dentine samples was about 4.86 years [29], which confirmed the stability and reproducibility of age prediction from tooth DNA methylation markers. Also working with teeth, Cristina Giuliani and colleagues in 2016 collected DNA methylation data from dentine, cementum, and pulp tissue and used MALDI-TOF mass spectrometry to analyze the methylation level at particular CpG sites within the ELOVL2, FHL2, and PENK genes, obtaining a model error of about 2.25 to 7.07 years with the highest prediction accuracy in pulp tissue [30]. Together these two studies displayed the considerable potential of teeth, and of pulp tissue in particular, for the construction of DNA methylation clocks. Ana Belen Marquez Ruiz and colleagues in 2020 compared the age prediction capacity of these common genes and found that several CpG sites on ELOVL2 and PDE4C, together with telomere length, predicted age more strongly than the ASPA gene, and they noted that DNA methylation carries more information than telomere length for the purpose of age estimation [31].

In 2021 Helena Correia Dias and colleagues combined blood, rib, and molar samples and built two multi tissue age estimation models, one with Sanger sequencing and one with SNaPshot detection. The first incorporated seven CpG sites within the ELOVL2, EDARADD, PDE4C, FHL2, and C1orf132 genes and the second incorporated three CpG sites within the ELOVL2, KLF14, and C1orf132 genes, and the errors of the two models were about 6.06 to 6.49 years [32]. In the same year Sara C. Zapico and colleagues analyzed the DNA methylation level in the pulp tissue of adult third molars, verified the value of CpG sites linked to ELOVL2 and FHL2 for age estimation, and went on to assess the three new marker sites NPTX2, KLF14, and SCGN, with the four multiple regression models they built reaching a mean absolute error of only 1.5 to 2.13 years [33]. Anna Wozniak and colleagues in 2021 used six CpG sites within the ELOVL2, KLF14, PDE4C, and ASPA genes in bone tissue, and the model they built reached a mean absolute error of 3.4 years [34].

Table 2. Epigenetic clock models constructed from the DNA methylation of skeletal materials

Researcher	Bone type	No. of CpG sites	Detection technology	Modeling method	Error range
Bekaert [29]	Teeth	19	Pyrosequencing	Linear, logarithmic, and quadratic regression models	plus or minus 4.86 years
Giuliani [30]	Teeth	5 to 13	MALDI-TOF mass spectrometry	Multiple linear regression model	plus or minus 2.25 to 7.07 years
Marquez Ruiz [31]	Teeth	9	Pyrosequencing	Quantile regression model	plus or minus 4.80 to 5.08 years
Correia Dias [32]	Blood, rib, molar	3 to 7	Sanger sequencing and SNaPshot detection	Simple and multiple linear regression models	plus or minus 6.06 to 6.49 years
Sara C. Zapico [33]	Third molar	12 to 14	Pyrosequencing	Multiple linear regression model	plus or minus 1.5 to 2.13 years
Anna Wozniak [34]	Blood, buccal mucosa cells, bone	5 to 6	MPS detection	Multiple linear regression model	plus or minus 3.2 to 3.7 years
Jung MS [35]	Costal cartilage	9	MPS detection	Multiple linear regression model	plus or minus 4.97 years
Siqi Chen [36]	Rib	8	Pyrosequencing	Multiple linear regression model	plus or minus 4.813 years
Charlotte Sutter [37]	Rib, femur, clavicle, iliac crest, petrous bone	44	MPS detection	Linear transformation model	plus or minus 4.37 to 8.10 years

In 2024 Ju Yeon Jung and colleagues developed an age prediction tool based on costal cartilage, constructing a regression model from nine CpG sites within the MIR29B2CHG, FHL2, TRIM59, ELOVL2, KLF14, and ASPA genes with a mean absolute error of about 4.97 years [35]. Siqi Chen and colleagues in 2025 turned to rib samples, analyzed the methylation level at eight CpG sites within the ELOVL2, FHL2, KLF14, and FAM123C genes, and developed a rib specific age prediction model. Comparing the sternum, rib, and frontal bone samples of twelve deceased individuals, they found no significant statistical difference in methylation level or in estimated age among the different bone types [36]. In 2026 Charlotte Sutter and colleagues analyzed five bone types, namely rib, femur, clavicle, iliac crest, and petrous bone, and found that the accuracy of age estimation varied across bone type, with the largest error in the petrous bone. By simulating a range of environmental conditions, that study also revealed the

influence of the freshness of the material on the result of age estimation [37]. These studies developed further new detection genes and CpG sites for bone and tooth materials and widened the range over which skeletal DNA methylation markers can serve in epigenetic clocks. They also examined in a systematic way how different bone types and the freshness of the material affect the error of age estimation and so improved the accuracy of the several models.

Taken together, the studies show that gene sites performing well in blood models do not apply in full to bone models, which is exactly an expression of the tissue specificity of DNA methylation. The genes used in age estimation models based on bone samples cluster around ELOVL2, EDARADD, PDE4C, FHL2, ASPA, and C1orf132, and the number of specific detection sites generally stays within ten, which suits the trace and degraded materials met under various forensic conditions and keeps the detection process simpler and more efficient. The errors of these models mostly hover around five years, with higher errors for certain bone types or detection tools, and overall, they come close to the models built on blood, saliva, and similar materials, displaying relatively high prediction accuracy. Table 2 collects the studies that have so far constructed epigenetic clocks from the DNA methylation of skeletal materials.

Systematic research into age estimation from the DNA methylation of skeletal materials nonetheless remains comparatively scarce, a systematic evaluation and screening of skeletal detection sites is still lacking, and only one or two related papers exist at home and abroad. Hwan Young Lee and colleagues in 2020 carried out methylation profiling across the whole genome on 32 identified skeletal remains, obtained more than eight hundred thousand CpG methylation data points with the Infinium MethylationEPIC BeadChip, and screened from them 19 candidate age related CpG sites, some of which were verified in an independent sample [38]. That study drew its samples only from a Korean population and worked with a small sample size, so the correlation of these sites with age and their practical value still await further verification. The study populations in existing work are concentrated in Europe and a few parts of Asia and lack analysis and verification in other ethnic groups, and the racial bias arising from this single source of population has become a bottleneck that constrains the extension of current models across regions and ethnicities. Beyond differences among populations, factors such as sex, physiological and pathological state, and epigenetic age acceleration are often neglected and readily lead to prediction error. Epigenetic age acceleration (EAA) refers to the difference between the biological age estimated for an individual by an epigenetic clock and the actual age, and it reflects the pace at which the body of a given individual ages [39].

4. Detection Technologies for DNA Methylation

The assessment of the pattern and the level of methylation forms the core link of an epigenetic clock and has always stood at the center of work in this field. A variety of methylation detection and analysis techniques have arisen in response, the scope of detection has gradually expanded from a particular locus to the level of the whole genome, and the limits of throughput and efficiency have been pushed back step by step. In the work that builds clock models from the DNA methylation of skeletal materials, the detection means used are likewise several mainstream sequencing methods in wide current use.

Appropriate pretreatment of the DNA material is a key step before methylation detection, and the classical strategies fall into two kinds, namely bisulfite conversion and treatment with methylation sensitive restriction enzymes. The great majority of detection techniques now take bisulfite conversion as the basis for subsequent analysis. Bisulfite conversion specifically transforms unmethylated cytosine into uracil without altering 5-methylcytosine, which allows the methylation state to be distinguished at the level of a single nucleotide in later sequencing

and carries both qualitative and quantitative capacity [40]. The method offers comparatively high sensitivity and accuracy and completes qualitative and quantitative analysis at the same time with good efficiency, yet its reaction time is long, the environment of strong acid below pH 5 and high temperature between 64 and 98 degrees Celsius readily causes damage and degradation of DNA, and regions of high CG content often convert incompletely and so produce bias [41], which limits its use in degraded materials or in samples of low input. To overcome these shortcomings Qing Dai and colleagues in 2025 proposed an ultra mild bisulfite conversion method known as UMBS-seq [42], which markedly lowers the influence of DNA degradation and background noise and can serve samples with input as low as 10 picograms, an ideal choice for the treatment of degraded materials such as skeletal remains. Treatment with methylation sensitive restriction enzymes instead uses a range of restriction enzymes to recognize and cut unmethylated sequences, which eases the later design of specific primers and the completion of PCR amplification [43]. Compared with bisulfite conversion it is simpler and more stable and lower in both cost and time, though its detection error is larger and its demand on sample quantity is higher [44].

Once the cytosine in genomic DNA has been chemically modified into a distinguishable sequence, methylation analysis can be completed with the help of various DNA detection techniques (10). According to the purpose of detection the work divides into single site analysis and scanning across the whole genome [45]. The construction of skeletal clock models at present relies mainly on platforms such as Sanger sequencing, SNaPshot detection, pyrosequencing, massively parallel sequencing (MPS), and EpiTYPER mass spectrometry. Because the results of different platforms carry an inherent systematic bias, researchers commonly apply statistical means such as Z-score transformation to correct the data and remove the influence of technical heterogeneity [46].

Pyrosequencing is the most widely used method for the detection of DNA methylation, is regarded in the forensic community as the gold standard for methylation detection, and served as the main platform in the skeletal clock studies above (29,31,33,36). It is in essence a technique of enzymatic synthesis sequencing that can accurately quantify the methylation level on the basis of bisulfite conversion and PCR amplification [47], and it carries the advantages of simple operation, high sensitivity, and good reproducibility, although its low throughput and limited capacity for multiplex detection restrict its application in large scale site screening. EpiTYPER mass spectrometry, also known as MassARRAY, is a method of bisulfite sequencing built on the MALDI-TOF mass spectrometry platform that can carry out quantitative and high throughput methylation detection over a particular region [48]. Compared with pyrosequencing, which rests on the sequencing of a single fragment, it shows a clear gain in detection throughput and in the breadth of gene coverage and is better suited to large scale quantitative analysis. Ling Wang and colleagues in 2020 compared the performance of pyrosequencing and MassARRAY in the quantitative detection of DNA, found a marked systematic batch effect between the two platforms, and confirmed that Z-score transformation can effectively remove such differences, which offered a reliable statistical correction for the integration and comparison of data across platforms [46].

Massively parallel sequencing, founded on the principle of second generation sequencing, also holds an important place in methylation detection through its strength in high throughput and coverage across the whole genome, and the studies of Anna Wozniak [34], Ju Yeon Jung [35], and Charlotte Sutter [37] all rested on a multiplex MPS platform. The method shows good compatibility with DNA samples of low abundance or low quality and offers comparatively high resolution, yet its high infrastructure cost and the complex process of data analysis are the main challenges in practical application.

Sanger sequencing and the SNaPshot detection method derived from it are likewise in wide use in forensic science. Sanger sequencing is the classic representative of first generation

sequencing technology and resolves a methylation sequence chiefly through single base extension (SBE) together with capillary electrophoresis. It is simple and convenient but demands high specificity of the primers and does not suit the simultaneous detection of multiple sites. The SNaPshot method shares the same principle as Sanger sequencing and can be used for semi quantitative analysis of the degree of methylation, and its strength lies in the simultaneous detection of multiple sites through a combination of multiplex primers, which widens the scope of analysis. The construction of a multiplex system nonetheless adds to the cost in time and to the complexity of operation, and the semi quantitative character also limits the precision of the result [49]. Helena Correia Dias and colleagues in 2021 integrated these two platforms and built separate age prediction models from the data of each, with mean absolute deviations of 6.06 years and 6.49 years [32].

5. Data Processing and Model Construction

The modeling methods now used to build epigenetic clocks from DNA methylation data include chiefly multiple linear regression, multiple quantile regression, support vector machines, artificial neural networks, and random forests [50]. Among the studies above the skeletal clock models adopted for the most part multiple linear regression and multiple quantile regression [29-37].

Multiple linear regression is the modeling method most commonly used at present and describes the relationship between one dependent variable and one or more independent variables [50], which here means the linear relationship between the methylation level at several CpG sites and age. Helena Correia Dias and colleagues [32] first built a simple linear regression model from specific CpG sites and then used stepwise regression to select the best combination of sites before constructing a multiple linear regression model. The model is concise and intuitive with strong interpretability, can portray the approximately linear association between methylation level and age, and holds general applicability for most studies of skeletal clocks.

Multiple quantile regression can estimate the value of the dependent variable at different quantiles given the independent variables, which in this field means the inference of the median age from the methylation level at particular sites [50]. Ana Belen Marquez Ruiz and colleagues [31] built a series of quantile age prediction models from the methylation level at CpG sites within the ELOVL2, ASPA, and PDE4C genes in 65 tooth samples, with a mean absolute deviation of about 4.80 to 5.08 years. The method has seen wide use beyond skeletal clocks in studies of blood and saliva clocks as well, with X. Guan [51] and Ambroa Conde [52] and colleagues both adopting it to obtain prediction errors of 6.243 years and 3.54 years respectively. Quantile regression handles heteroscedasticity, in which the variance is not constant, and outliers effectively, and its prediction interval widens gradually as age increases, which accords better with the pattern of natural human aging [50].

The support vector machine is a powerful method of supervised learning in wide use for tasks of classification and regression that gains excellent generalization by maximizing the margin between samples of different classes [53]. Several studies show that the support vector machine performs markedly better than multiple linear regression and other models in age estimation [54, 55]. Its advantages lie chiefly in the capture of the complex relationship between DNA methylation and age through a variety of kernel functions so that prediction precision rises, in the improvement of accuracy through the maximum margin hyperplane, and in good adaptation to high dimensional methylation data [50]. Xueli Han and colleagues [26] built both a multiple linear regression model and a support vector machine model from eight age related CpG sites in blood samples, with errors of 3.52 years and 2.88 years respectively.

The artificial neural network can learn and express highly complex nonlinear relationships and on that basis construct complex decision boundaries that are piecewise linear, and it shows a high tolerance for missing data, so its overall performance again surpasses that of traditional linear regression [50]. The single site artificial neural network model that Thong and colleagues [56] built from blood samples reached an error of about 3.7 years, lower than that of the linear regression model, and could raise prediction accuracy while reducing the amount of DNA required.

The random forest carries equally clear advantages in the construction of epigenetic clocks and suits the age estimation of degraded samples in particular, which makes it one of the ideal choices for the modeling of forensic skeletal clocks. Haoliang Fan and colleagues [57] built age prediction models from 34 CpG sites in blood samples with four machine learning algorithms, among which the random forest gave the smallest mean absolute error and so displayed the full promise of the method for application.

6. Summary and Outlook

DNA methylation is a comparatively stable biological marker within epigenetics that carries marked tissue specificity and a strong correlation with age, and it can both assess the degree of aging well and predict disease outcome and the risk of death with reasonable accuracy. Since its capacity for age estimation came to light, the epigenetic clock has passed through three important stages of development. Most of these models nonetheless rest on samples of blood, saliva, hair, or other soft tissue and cover several dozen and at times several hundred CpG sites, which leaves them unsuited to the degraded and trace DNA samples of forensic science.

Bone, teeth, and similar materials are highly stable and resistant to degradation, and the construction of epigenetic clocks from their DNA methylation has gradually grown into an important and active direction of research. A reasonable choice of material type and of CpG sites, together with optimized detection technology and modeling strategy, will help improve the precision and the accuracy of age estimation. Scholars in China and abroad have already carried out several studies on the basis of skeletal DNA methylation, and the models built by Cristina Giuliani, Sara C. Zapico, Ju Yeon Jung, and others all attained comparatively high accuracy with an error around five years, which opened a new vista for the application of epigenetics in forensic age estimation.

The studies above nonetheless carry several prominent problems that have held back their translation from scientific research toward practical application. One problem is the insufficient representativeness of the study populations. Existing samples cluster in Europe and a few parts of Asia and lack large scale validation in populations from other regions, and because patterns of DNA methylation differ markedly with racial and genetic background, this bias causes the models to show prediction deviation in populations on which they were not trained and limits their global applicability. Future work needs to widen the coverage of populations, broaden the sources of data, integrate samples across many regions and ethnic groups, and build a general age prediction model with stability across populations. A second problem is the scarcity of systematic research into skeletal specific clocks. CpG sites of high correlation with age have not yet been screened and validated in a systematic way across the whole genome in bone tissue, and most studies still borrow the sites of blood or other tissues, which leaves the accuracy and the specificity of skeletal clocks wanting. Effort should advance the screening of specific sites across the whole genome in bone samples and mine in a systematic way the methylation markers that are specific to bone tissue, little disturbed by the environment, and strongly correlated with age, so that the accuracy and the reliability of the models improve. A third problem is the single biological dimension of the models. Most

studies fit age on the basis of methylation level alone and do not bring sex, physiological and pathological state, or epigenetic age acceleration into the model, which weakens the biological explanatory power of the model and its capacity to portray individual characteristics. Sex, physiological rhythm, disease state, and epigenetic age acceleration should accordingly be taken into account as variables in future data collection and modeling.

This review has traced in a systematic way the three generation evolution of the epigenetic clock together with the research of the past decade on epigenetic clocks built from the DNA methylation of skeletal materials, and it has analyzed the advantages and the limitations of the various models across the choice and number of gene sites, the detection methods and technologies, and the strategies of model construction. Compared with clocks based on body fluids or soft tissue such as blood and saliva, research that takes skeletal remains as the material still appears insufficient and awaits more systematic and deeper exploration, so that epigenetics may make further substantial contributions to forensic age estimation.

References

- [1] Chen, C. (2021). Research progress of biological age in healthy aging [in Chinese]. *Chinese Journal of Geriatrics*, 40(6), 798. <https://doi.org/10.3760/cma.j.issn.0254-9026.2021.06.023>.
- [2] Wang, Y. H. (2024). Frontiers and challenges in forensic age estimation of the living [in Chinese]. *Journal of Forensic Medicine*, (0), 1. <https://doi.org/10.12116/j.issn.1004-5619.2024.240410>.
- [3] Al-Munim, M. F. A., & Al-Rashedi, N. A. M. (2022). Stability of DNA methylation of blood evidence under temperature and humidity conditions. *MJPS*, 9(1). <https://doi.org/10.52113/2/09.01.2022/59-69>.
- [4] Yudianto, A. (2024). Bone and dental DNA damage due to extreme high temperature exposure through STR-CODIS, Y-STR and mtDNA examinations. *Gaceta Médica de Caracas*, 132(2). <https://doi.org/10.47307/GMC.2024.132.2.7>.
- [5] Fan, F. (2020). Research progress on imaging methods for adult age estimation [in Chinese]. *Journal of Forensic Medicine*, 36(5), 605. <https://doi.org/10.12116/j.issn.1004-5619.2020.05.002>.
- [6] Liu, F. (2017). Research progress in forensic odontology [in Chinese]. *Journal of Forensic Medicine*, 33(2), 175. <https://doi.org/10.3969/j.issn.1004-5619.2017.02.015>.
- [7] Zhou, Z. L. (2025). Quantification of wrist MRI epiphyseal development and bone age assessment in Han Chinese male adolescents of East China [in Chinese]. *Journal of Forensic Medicine*, (0), 1. <https://doi.org/10.12116/j.issn.1004-5619.2023.231203>.
- [8] Zhang, B., Yuan, Y., Di, P. X., Xu, F. L., Lin, Z. Q., & Zhang, L. (2024). Current research status of epigenetics in forensic biological evidence [in Chinese]. *Life Science Research*, 28(5), 395–401. <https://doi.org/10.16605/j.cnki.1007-7847.2023.12.0224>.
- [9] Zhang, J. Y. (2021). Research progress of RNA in forensic genetics [in Chinese]. *Chinese Journal of Forensic Sciences*, (3), 36. <https://doi.org/10.3969/j.issn.1671-2072.2021.03.004>.
- [10] Ao, J. H., Dong, W. J., Lü, M. R., Wu, Z. L., Jing, D., Hou, Z. P., & Yang, G. M. (2023). Research progress on the functional features of DNA methylation and its forensic application [in Chinese]. *Chinese Journal of Birth Health & Heredity*, 31(10), 2183–2188. <https://doi.org/10.13405/j.cnki.issn.1003-9406.2023.10.0050>.
- [11] Nie, Y. C. (2017). Research progress on detection methods of DNA methylation and their forensic application [in Chinese]. *Journal of Forensic Medicine*, 33(3), 293. <https://doi.org/10.3969/j.issn.1004-5619.2017.03.017>.
- [12] Farrell, C., Hu, C., Lapborisuth, K., Pu, K., Snir, S., & Pellegrini, M. (2024). Identifying epigenetic aging moderators using the epigenetic pacemaker. *Frontiers in Bioinformatics*, 3, 1308680. <https://doi.org/10.3389/fbinf.2023.1308680>.
- [13] Li, A., Koch, Z., & Ideker, T. (2022). Epigenetic aging: Biological age prediction and informing a mechanistic theory of aging. *Journal of Internal Medicine*, 292(5), 733–744. <https://doi.org/10.1111/joim.13533>.

- [14] Bocklandt, S., Lin, W., Sehl, M. E., Sanchez, F. J., Sinsheimer, J. S., & Horvath, S., et al. (2011). Epigenetic predictor of age. *PLOS ONE*, 6(6), e14821. <https://doi.org/10.1371/journal.pone.0014821>.
- [15] Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, 14(10), R115. <https://doi.org/10.1186/gb-2013-14-10-r115>.
- [16] Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., & Sadda, S., et al. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular Cell*, 49(2), 359–367. <https://doi.org/10.1016/j.molcel.2012.10.016>.
- [17] Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., & Hou, L., et al. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging*, 10(4), 573–591. <https://doi.org/10.18632/aging.101414>.
- [18] Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., & Raj, K., et al. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*, 11(2), 303–327. <https://doi.org/10.18632/aging.101684>.
- [19] Lu, A. T., Binder, A. M., Zhang, J., Yan, Q., Reiner, A. P., & Cox, S. R., et al. (2022). DNA methylation GrimAge version 2. *Aging*, 14(23), 9484–9549. <https://doi.org/10.18632/aging.204434>.
- [20] Belsky, D. W., Caspi, A., Arseneault, L., Baccarelli, A., Corcoran, D. L., & Gao, X., et al. (2020). Quantification of the pace of biological aging in humans through a blood test, the DunedinPoAm DNA methylation algorithm. *eLife*, 9, e54870. <https://doi.org/10.7554/eLife.54870>.
- [21] Belsky, D. W., Caspi, A., Corcoran, D. L., Sugden, K., Poulton, R., & Arseneault, L., et al. (2022). DunedinPACE, a DNA methylation biomarker of the pace of aging. *eLife*, 11, e73420. <https://doi.org/10.7554/eLife.73420>.
- [22] Sehgal, R., Markov, Y., Qin, C., et al. (2025). Systems Age: A single blood methylation test to quantify aging heterogeneity across 11 physiological systems. *Nature Aging*, 5, 1880–1896. <https://doi.org/10.1038/s43587-025-00958-3>.
- [23] Feng, L., Peng, F., Li, S., Jiang, L., Sun, H., & Ji, A., et al. (2018). Systematic feature selection improves accuracy of methylation-based forensic age estimation in Han Chinese males. *Forensic Science International: Genetics*, 35, 38–45. <https://doi.org/10.1016/j.fsigen.2018.03.009>.
- [24] Li, S. F., Peng, F. D., Wang, J. N., et al. (2019). Construction and performance evaluation of a methylation based age estimation model [in Chinese]. *Journal of Forensic Medicine*, 35(1), 17. <https://doi.org/10.12116/j.issn.1004-5619.2019.01.004>.
- [25] Hao, T., Guo, J. L., Liu, J. D., et al. (2021). Predicting human age by detecting DNA methylation status in hair [in Chinese]. *Electrophoresis*, 42(6), 752–760. <https://doi.org/10.1002/elps.202000349>.
- [26] Han, X., Xiao, C., Yi, S., Li, Y., Chen, M., & Huang, D. (2022). Accurate age estimation from blood samples of Han Chinese individuals using eight high-performance age-related CpG sites. *International Journal of Legal Medicine*, 136(6), 1655–1665. <https://doi.org/10.1007/s00414-022-02865-3>.
- [27] Zhang, T., Zhang, J., Ma, M., et al. (2026). A seven CpG age estimation model for oral swabs based on ddPCR technology [in Chinese]. *Criminal Technology*, (2), 145–151. <https://doi.org/10.16467/j.1008-3650.2025.0089>.
- [28] Slieker, R. C., Relton, C. L., Gaunt, T. R., Slagboom, P. E., & Heijmans, B. T. (2018). Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception. *Epigenetics & Chromatin*, 11(1), 25. <https://doi.org/10.1186/s13072-018-0191-3>.
- [29] Bekaert, B., Kamalandua, A., Zapico, S. C., Van De Voorde, W., & Decorte, R. (2015). Improved age determination of blood and teeth samples using a selected set of DNA methylation markers. *Epigenetics*, 10(10), 922–930. <https://doi.org/10.1080/15592294.2015.1080413>.
- [30] Giuliani, C., Cilli, E., Bacalini, M. G., Pirazzini, C., Sazzini, M., Gruppioni, G., et al. (2016). Inferring chronological age from DNA methylation patterns of human teeth. *American Journal of Physical Anthropology*, 159(4), 585–595. <https://doi.org/10.1002/ajpa.22921>.

- [31] Marquez-Ruiz, A. B., Gonzalez-Herrera, L., Luna, J. D., & Valenzuela, A. (2020). DNA methylation levels and telomere length in human teeth: usefulness for age estimation. *International Journal of Legal Medicine*, 134(2), 451–459. <https://doi.org/10.1007/s00414-019-02242-7>.
- [32] Correia Dias, H., Manco, L., Corte Real, F., & Cunha, E. (2021). A blood-bone-tooth model for age prediction in forensic contexts. *Biology*, 10(12), 1312. <https://doi.org/10.3390/biology10121312>.
- [33] Zapico, S. C., Gauthier, Q., Antevska, A., & McCord, B. R. (2021). Identifying methylation patterns in dental pulp aging: application to age-at-death estimation in forensic anthropology. *International Journal of Molecular Sciences*, 22(7), 3717. <https://doi.org/10.3390/ijms22073717>.
- [34] Wozniak, A., Heidegger, A., Piniewska-Rog, D., Pospiech, E., Xavier, C., & the VISAGE Consortium. (2021). Development of the VISAGE enhanced tool and statistical models for epigenetic age estimation in blood, buccal cells and bones. *Aging*, 13(5), 6459–6484. <https://doi.org/10.18632/aging.202783>.
- [35] Jung, J. Y., So, M. H., Jeong, K., Kim, S., Kim, E. J., Park, J. H., et al. (2024). Epigenetic age prediction using costal cartilage for the investigation of disaster victims and missing persons. *Journal of Forensic Sciences*, 69(5), 1578–1586. <https://doi.org/10.1111/1556-4029.15470>.
- [36] Chen, S., Zhang, C., Wen, D., Wang, C., Tang, X., Li, X., et al. (2025). Epigenetic age signatures in postmortem rib samples. *Electrophoresis*, 46(18), 1397–1406. <https://doi.org/10.1002/elps.70014>.
- [37] Sutter, C., Haas, C., & Neubauer, J. (2026). Assessing the impact of bone type on the accuracy of methylation-based forensic age estimation. *Forensic Science International: Genetics*, 84, 103459. <https://doi.org/10.1016/j.fsigen.2026.103459>.
- [38] Lee, H. Y., Hong, S. R., Lee, J. E., Hwang, I. K., Kim, N. Y., Lee, J. M., et al. (2020). Epigenetic age signatures in bones. *Forensic Science International: Genetics*, 46, 102261. <https://doi.org/10.1016/j.fsigen.2020.102261>.
- [39] Chervova, O., Panteleeva, K., Chernysheva, E., Widayati, T. A., Baronik, Z. F., Hrbkova, N., et al. (2024). Breaking new ground on human health and well-being with epigenetic clocks: a systematic review and meta-analysis of epigenetic age acceleration associations. *Ageing Research Reviews*, 102, 102552. <https://doi.org/10.1016/j.arr.2024.102552>.
- [40] Leti, F., Llaci, L., Malenica, I., & DiStefano, J. K. (2018). Methods for CpG methylation array profiling via bisulfite conversion. In J. K. DiStefano (Ed.), *Disease Gene Identification* (pp. 233–254). Springer. https://doi.org/10.1007/978-1-4939-7471-9_13.
- [41] Dai, Q., Ye, C., Irklienko, I., Wang, Y., Sun, H. L., Gao, Y., et al. (2024). Ultrafast bisulfite sequencing detection of 5-methylcytosine in DNA and RNA. *Nature Biotechnology*, 42(10), 1559–1570. <https://doi.org/10.1038/s41587-023-02034-w>.
- [42] Dai, Q., Baldwin, T., Lyu, R., Daniels, B., Ye, C., Cao, C., et al. (2025). Ultra-mild bisulfite outperforms existing methods for 5-methylcytosine detection with low input DNA. *Nature Communications*, 16(1), 9939. <https://doi.org/10.1038/s41467-025-66033-y>.
- [43] Khulan, B., Thompson, R. F., Ye, K., Fazzari, M. J., Suzuki, M., Stasiak, E., et al. (2006). Comparative isoschizomer profiling of cytosine methylation: The HELP assay. *Genome Research*, 16(8), 1046–1055. <https://doi.org/10.1101/gr.5273806>.
- [44] Wang, J. Y., & Li, S. J. (2025). Research progress of aging related biomolecular markers in forensic age estimation [in Chinese]. *Chinese Journal of Forensic Medicine*, 40(2), 189–195. <https://doi.org/10.13618/j.issn.1001-5728.2025.02.006>.
- [45] Laird, P. W. (2010). Principles and challenges of genome-wide DNA methylation analysis. *Nature Reviews Genetics*, 11(3), 191–203. <https://doi.org/10.1038/nrg2732>.
- [46] Wang, L. (2020). Differences between pyrosequencing and MassARRAY in the quantitative detection of DNA methylation for age estimation [in Chinese]. *Journal of Southern Medical University*, 40(12), 1831. <https://doi.org/10.12122/j.issn.1673-4254.2020.12.21>.
- [47] Harrington, C. T., Lin, E. I., Olson, M. T., & Eshleman, J. R. (2013). Fundamentals of pyrosequencing. *Archives of Pathology & Laboratory Medicine*, 137(9), 1296–1303. <https://doi.org/10.5858/arpa.2012-0463-RA>.

- [48] Suchiman, H. E. D., Sliker, R. C., Kremer, D., Slagboom, P. E., Heijmans, B. T., & Tobi, E. W. (2015). Design, measurement and processing of region-specific DNA methylation assays: The mass spectrometry-based method EpiTYPER. *Frontiers in Genetics*, 6, 287. <https://doi.org/10.3389/fgene.2015.00287>.
- [49] Castagnola, M. J., Medina-Paz, F., & Zapico, S. C. (2024). Uncovering forensic evidence: A path to age estimation through DNA methylation. *International Journal of Molecular Sciences*, 25(9), 4917. <https://doi.org/10.3390/ijms25094917>.
- [50] Huang, Y. H. (2023). Modeling methods and influencing factors of age estimation based on DNA methylation [in Chinese]. *Journal of Forensic Medicine*, (0), 1. <https://doi.org/10.12116/j.issn.1004-5619.2023.530106>.
- [51] Guan, X., Ohuchi, T., Hashiyada, M., & Funayama, M. (2021). Age-related DNA methylation analysis for forensic age estimation using post-mortem blood samples from Japanese individuals. *Legal Medicine*, 53, 101917. <https://doi.org/10.1016/j.legalmed.2021.101917>.
- [52] Ambroa-Conde, A., Giron-Santamaria, L., Mosquera-Miguel, A., Phillips, C., Casares de Cal, M. A., Gomez-Tato, A., et al. (2022). Epigenetic age estimation in saliva and in buccal cells. *Forensic Science International: Genetics*, 61, 102770. <https://doi.org/10.1016/j.fsigen.2022.102770>.
- [53] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>.
- [54] Xu, C., Qu, H., Wang, G., Xie, B., Shi, Y., Yang, Y., et al. (2015). A novel strategy for forensic age prediction by DNA methylation and support vector regression model. *Scientific Reports*, 5(1), 17788. <https://doi.org/10.1038/srep17788>.
- [55] Vidaki, A., Gonzalez, D. M., Jimenez, B. P., & Kayser, M. (2021). Male-specific age estimation based on Y-chromosomal DNA methylation. *Aging*, 13(5), 6442–6458. <https://doi.org/10.18632/aging.202775>.
- [56] Thong, Z., Tan, J. Y. Y., Loo, E. S., Phua, Y. W., Chan, X. L. S., & Syn, C. K. C. (2021). Artificial neural network, predictor variables and sensitivity threshold for DNA methylation-based age prediction using blood samples. *Scientific Reports*, 11(1), 1744. <https://doi.org/10.1038/s41598-021-81556-2>.
- [57] Fan, H., Xie, Q., Zhang, Z., Wang, J., Chen, X., & Qiu, P. (2022). Chronological age prediction: Developmental evaluation of DNA methylation-based machine learning models. *Frontiers in Bioengineering and Biotechnology*, 9, 819991. <https://doi.org/10.3389/fbioe.2021.819991>.