How to Translate DNA Methylation Biomarkers Into Clinical Practice

Wolfgang Wagner[1,2]*

[1]Helmholtz-Institute for Biomedical Engineering, Stem Cell Biology and Cellular Engineering, RWTH Aachen University Medical School, Aachen, Germany, [2]Center for Integrated Oncology Aachen Bonn Cologne Düsseldorf (CIO ABCD), Aachen, Germany

Recent advances in sequencing technologies provide unprecedented opportunities for epigenetic biomarker development. Particularly the DNA methylation pattern—which is modified at specific sites in the genome during cellular differentiation, aging, and disease—holds high hopes for a wide variety of diagnostic applications. While many epigenetic biomarkers have been described, only very few of them have so far been successfully translated into clinical practice and almost exclusively in the field of oncology. This discrepancy might be attributed to the different demands of either publishing a new finding or establishing a standardized and approved diagnostic procedure. This is exemplified for epigenetic leukocyte counts and epigenetic age-predictions. To ease later clinical translation, the following hallmarks should already be taken into consideration when designing epigenetic biomarkers: 1) Identification of best genomic regions, 2) pre-analytical processing, 3) accuracy of DNA methylation measurements, 4) identification of confounding parameters, 5) accreditation as diagnostic procedure, 6) standardized data analysis, 7) turnaround time, and 8) costs and customer requirements. While the initial selection of relevant genomic regions is usually performed on genome wide DNA methylation profiles, it might be advantageous to subsequently establish targeted assays that focus on specific genomic regions. Development of an epigenetic biomarker for clinical application is a long and cumbersome process that is only initiated with the identification of an epigenetic signature.

Keywords: epigenetic, biomarker, hallmarks, DNA methylation, aging, blood counts, methods, IVDD

INTRODUCTION

Epigenetics is a fascinating branch of research that will increasingly find its way into clinical diagnostics. There are different types of epigenetic modifications, including DNA methylation (DNAm), histone modifications, and higher order chromatin conformation (Zhu et al., 2013). For epigenetic diagnostics DNAm appears to be best suited, as it can be precisely determined at specific cytosine residues of CG dinucleotides (CpG sites) (Smith and Meissner, 2013). The enzymes that catalyze *de novo* methylation and maintenance of DNAm levels have long been known (Okano et al., 1999; Reik et al., 2001). In contrast, despite the impressive scientific descriptions of the epigenetic landscape in the last years, it remains unclear how these enzymes are directed to specific sites in the genome and how the complex epigenetic network is orchestrated. In the future, a better understanding of the underlying regulatory mechanisms that direct DNAm can help to derive even more focused and functionally relevant biomarkers. Nonetheless, the currently rather descriptive analysis of DNAm alterations can already provide accurate insights into cellular and developmental processes and thus provide valuable diagnostic tools.

FIGURE 1 | Aspects to be considered when designing diagnostic epigenetic biomarkers.

FIELDS OF APPLICATION FOR DNA METHYLATION BIOMARKERS

An epigenetic biomarker can be defined as any epigenetic mark, which is stable and reproducible during sample processing and can be measured in body fluids or primary tissue. They provide tools for diagnosis, prognosis, monitoring of disease evolution, and can support clinical decision-making (Garcia-Gimenez et al., 2017). The scope of DNAm biomarkers can be classified in four categories:

1) Cellular composition. Since DNAm is fundamentally linked to cellular development, it is very well suited to provide insights into cellular composition of a given tissue. Each of our cell types has a characteristic DNAm pattern (Moss et al., 2018). Tissue composition can then be reliably estimated using deconvolution algorithms, as each cell only has two copies of chromosomal DNA (Houseman et al., 2012; Schmidt et al., 2020). With such assays it is for instance possible to investigate the origin of cell-free DNA in blood plasma as a liquid biopsy of potentially affected tissues (Moss et al., 2018; Neuberger et al., 2022).

2) Environmental influences and lifestyle. For example, cigarette smoking affects the DNAm levels at multiple genomic loci (Tsaprouni et al., 2014) and this effect can be reliably tracked in blood and brain specimen (Gadd et al., 2021). There are also DNAm biomarkers for alcohol consumption (Liu et al., 2018) and other parameters, such as physical activity, exercise and body weight, which can impact on the epigenome.

3) Aging. It is remarkable how well DNAm correlates at specific CpGs with chronological age (Horvath, 2013; Weidner and Wagner, 2014). Corresponding epigenetic signatures, so-called "epigenetic clocks", can support forensic investigation of

blood traces or of donors with allegedly unknown age. On the other hand, there is increasing evidence that epigenetic age more closely reflects biological age than chronological age (Bell et al., 2019). In fact, accelerated epigenetic age is associated with higher all-cause mortality in later life (Marioni et al., 2015). Whether pure epigenetic clocks can be trained to better capture aspects of biological age—independent from epigenetic changes by other confounding parameters of life-style, cellular composition, and diseases—still requires further validation. Either way, a biomarker for biological age has great potential for individualized medicine to evaluate therapeutic options. In addition, such studies can help to uncover factors that influence aging to adjust life-style for healthy aging.

4) Diseases. So far, epigenetic diagnostic biomarkers have almost exclusively been established for applications in oncology (Locke et al., 2019). Epigenetic aberrations can mimic genomic mutations, e.g., by DNAm aberrations in the gene *DNMT3A* in acute myeloid leukemia (Jost et al., 2014). In fact, next to driver mutations it seems to be particularly epigenetic aberrations that cause malignant transformation (Schoofs et al., 2014; Wagner et al., 2015), and such epimutations may either arise in the absence of DNA sequence changes (primary epimutations), or secondary to genetic variants (secondary epimutation) (Cerrato et al., 2020). In contrast, episignatures rather resemble complex aberrant DNAm patterns that are not functionally restricted to specific sites in the genome (Bozic et al., 2022). Epimutations as well as episignatures can provide insight into the malignant clone and can be used for initial diagnostics and disease stratification. Furthermore, DNAm patterns can be used to predict response to a specific therapeutic regimen and to track measurable residual disease after treatment (Bozic et al., 2022). Although there are currently no approved *in vitro* diagnostic (IVD) tests targeting methylation outside of oncology, there is clear evidence that many other diseases, including imprinting disorders, neurodegenerative and psychiatric disorders, involve epigenetic aberrations that may be addressed by epigenetic signatures (Beltran-Garcia et al., 2019; Taryma-Lesniak et al., 2020).

In view of the rapid development in DNAm studies on the one hand, and the increasing regulatory requirements on the other hand, it is to be expected that the gap between potential applications and actual implementations will continue to widen. Two examples of potential applications are further highlighted below.

Epigenetic Leukocyte Counts

Leukocyte counts in blood is one of the most common diagnostic tests, which is conventionally performed with automated cell counting devices and particularly for stratification of lymphocyte subsets with flow cytometry (Pitoiset et al., 2018). Yet, deconvolution of leukocyte subsets based on DNAm may have several advantages as compared to the conventional regimen (Sontag et al., 2022): 1) It is applicable to very small volumes of blood that can be harvested by a finger-prick; 2) DNAm analysis is possible with frozen blood; 3) it might be applied to coagulated samples or specimen with ineffective antibody binding; and 4) the precise measurement of DNAm levels might provide less variability in inter-laboratory comparison. Deconvolution models were initially developed based on cell-type specific hypo- or

hypermethylation that was analyzed in purified leukocyte subsets (Houseman et al., 2012). Derivation of such signatures for granulocytes, lymphocytes, B cells, T cells (CD4 or CD8), NK cells, and monocytes was particularly based on Infinium BeadChip mircoarrays (Accomando et al., 2014; Salas et al., 2018).

It has been demonstrated that leukocyte counts can also be determined by targeted assays at specific CpGs. These assays are based on quantitative PCR (qPCR), pyrosequencing or digital droplet PCR (ddPCR) (Baron et al., 2018; Frobel et al., 2018; Malic et al., 2019) and appear to be applicable for clinical use. We have recently further optimized and validated our targeted pyrosequencing and ddPCR assays for leukocyte deconvolution using 332 venous and 122 capillary blood samples from healthy donors (Sontag et al., 2022). In addition, we tested 36 samples from ring trials and venous blood from 266 patients diagnosed with different hematological diseases. Overall, the targeted DNAm analysis by pyrosequencing or ddPCR is a valid alternative to quantify leukocyte subsets (Sontag et al., 2022). However, much research remains necessary for further optimization, validation, and approval as an *in vitro* diagnostic device (IVDD) before epigenetic blood counts can ultimately find their way into clinical application.

Epigenetic Clocks

As already mentioned above, a multitude of epigenetic signatures have been described to estimate either chronological age or biological age. Ten years ago, our group was one of the first to describe such epigenetic clocks (Bocklandt et al., 2011; Koch and Wagner, 2011). Since then, these signatures significantly improved by the rapidly growing number of available Infinium BeadChip datasets and more elegant bioinformatic considerations (Horvath, 2013). Integration of Infinium BeadChip measurements facilitated human biomarker development, that was unparalleled in other species (Wagner, 2017). Only recently, the Infinium Mouse Methylation BeadChip became available, as well as the mammalian methylation array that covers highly conserved CpGs (Arenson et al., 2022).

In the future, epigenetic aging signatures might move from measuring DNAm levels at individual CpGs to a probabilistic analysis of the binary sequel of methylated and non-methylated CpGs on individual reads. This method was first described for barcoded amplicon sequencing and may reflect heterogeneity of epigenetic aging within a sample (Han et al., 2020a; Han et al., 2020b). It was further developed within the scAge framework that is applicable with few bisulfite sequencing reads (Trapp et al., 2021). Probabilistic epigenetic age-predictions could be performed with tagmentation-based indexing for methylation sequencing (TIME-Seq) that is applicable for Methyl-ATAC-seq with low-cost shallow sequencing (Griffin et al., 2021). It might eventually even be used for single-cell methyl-ATAC-seq data to ultimately elucidate heterogeneity of epigenetic aging within a given sample. Other sequencing technologies with longer reads, such as nanopore sequencing or PacBio sequencing, will further strengthen such probabilistic pattern-based approaches. While the shallow-sequencing may significantly reduce the sequencing costs, the bioinformatic demands are relatively high and further validation is still elusive. So far, these

methods can hardly be standardized to be accredited as diagnostic procedure (e.g., according to ISO13485).

For a clinical or forensic application of epigenetic clocks, it might be advantageous to rather focus on a few individual age-associated CpGs with targeted methods (Weidner et al., 2014). In fact, some individual CpGs reveal very high correlation with chronological age and they also seem to capture aspects of biological aging (Lin et al., 2016). Targeted epigenetic clocks for pyrosequencing, MassArray, ddPCR, or barcoded amplicon sequencing can reach almost similar precision as described for Infinium BeadChip clocks (Han et al., 2020a). They provide a tradeoff between the less CpGs to be integrated into a robust network for epigenetic age-predictions, and the more accurate DNAm measurements at individual CpGs. Lastly, it is worth mentioning that aging is *per se* not a disease and therefore, depending on the application, an accreditation for clinical use may not be necessary.

HALLMARKS FOR DNA METHYLATION BIOMARKERS

To successfully establish an IVDD it is important to benchmark accuracy and robustness in comparison to conventional markers. The sensitivity (ability of the test to correctly identify patients with a disease) and specificity (ability to correctly identify a subject without the disease) are dependent on a multitude of variables in the analytical process (Taryma-Lesniak et al., 2020). Many methods for DNAm analysis have evolved and can be classified in genome-wide approaches or targeted measurements at specific genomic regions (Locke et al., 2019) (**Table 1**). Each of these methods has advantages and disadvantages that can only be briefly touched in the context of this mini-review. It is not possible to give general applicable guidelines for development of epigenetic biomarkers due to the very different fields of application, molecular biological features, and clinical requirements. However, the following parameters should be considered when designing such studies (**Figure 1**).

Identification of the Best Genomic Regions

There are approximately 28 million CpG sites in the human genome and selection of the best suited regions is the first challenge for biomarker development. Whole Genome Bisulfite Sequencing (WGBS) provides insight into the entire methylome based on unbiased genome wide DNAm analysis, but this necessitates high coverage of deep-sequencing reads, which is relatively costly. Reduced Representation Bisulfite Sequencing (RRBS) can significantly reduce the sequencing effort, but coverage at specific locations of the genome sometimes deviates significantly, reducing transferability. Within the last decade human biomarker development has been revolutionized by Infinium BeadChip technology. These microarray platforms covered initially approximately 27,000 CpGs (27k), later 450,000 CpGs (450k), and currently, with the human EPIC BeadChip, approximately 850,000 CpGs (Pidsley et al., 2016). A major advantage of this method is that many DNA methylation profiles of public databases can be directly compared and

TABLE 1 | Comparison of selected methods for DNAm analysis.

Method	Mechanism for DNAm detection	Genome wide analysis	Targeted DNAm analysis	Accuracy of DNAm at individual CpGs	Integrative analysis with other datasets	Scalability for many samples	CE accredited instrumentation [a]	Bioinformatic requirement	Usual turn around time	Costs [b]
Whole Genome Bisulfite Sequencing (WGBS)	High-throughput sequencing of bisulfite or enzym. converted DNA	++	—	+	++	-	+	↑↑	↑↑	↑↑
Reduced Representation Bisulfite Sequencing (RRBS)	High-throughput sequencing of bisulfite or enzym. converted DNA	+	—	+	+	+	—	↑↑	↑↑	↑
Infinium BeadChip Technology (EPIC array)	Microarray analysis of bisulfite-converted DNA	+	—	++	++	+	—	↑	↑	↑
Nanopore sequencing	Detection of current changes during sequencing of non-converted DNA	++	—	-	+	+	+	↑↑	↑	↑
Methylated DNA immunoprecipitation (MeDIP)	Affinity capture of unconverted DNA with antibody and deep sequencing	++	NA	-	+	-	—	↑↑	↑↑	↑
Methyl Binding Domain (MBD) Sequencing	Affinity capture of unconverted DNA with MBD and deep sequencing	++	NA	-	+	-	—	↑↑	↑↑	↑
Probabilistic frameworks for genome wide DNAm pattern analysis	Shallow sequencing data (e.g. RRBS for scAge or methyl-ATAC-seq for TIME-seq)	-		-	+	++	—	↑↑	↑↑	↔
Methylation sensitive restriction enzyme (MSRE) PCR	Unconverted DNA is digested with MRSE and amplified by PCR	NA	+	-	-	+	—	-	↔	↓
Methylation specific PCR	PCR amplicons of bisulfite-converted or enzymatic converted DNA	NA	+	-	-	+	—	-	↔	↓
Quantitative PCR (qPCR)	Allele specific qPCR after bisulfite-conversion	NA	+	+	-	+	—	-	↔	↓
Targeted amplicon sequencing	Deep-sequencing of PCR amplicons of bisulfite converted DNA	NA	+	++	-	++	+	↑	↑	↔
Pyrosequencing	Sequencing of PCR amplicons of bisulfite-converted DNA	NA	+	++	-	+	—	-	↔	↔
Digital droplet PCR (ddPCR)	PCR of bisulfite-converted DNA in droplets	NA	+	++	-	+	+	-	↔	↔
EpiTYPER	Mass-spectrometry of bisulfite-converted DNA	NA	+	++	-	+	—	-	↔	↔

This table shall only exemplify how hallmarks of DNAm biomarkers may be influenced by different methods. It does not claim to represent all available approaches for DNAm analysis. The suggested classifications will vary between applications and laboratories.

[a]Accreditation may also include enzymes and viable components. The regulatory requirements and accreditation may change with time and according to local regulations.

[b]The expenses can only be estimated relatively. They include consumables, instrumentation and personnel costs and depend largely on the number of samples that can be processed in parallel (and on the number of CpGs for the targeted assays). NA, not applicable.

integrated into the analysis to identify and validate suitable signatures. For any epigenetic biomarker development, it is crucial to have a high number of measurements for a training set and for an independent validation cohort.

Pre-Analytical Processing

Sample harvesting, transportation, DNA extraction, treatments such as bisulfite conversion or enzymatic conversion, storage and preservation can have enormous impact on the final results. Since biomarker development requires large training and validation cohorts, biobanks play a crucial role for the initial development of personalized medicine (Peiró-Chova et al., 2016). However, sample processing may be very different in biobanks as compared to a later clinical setting. The samples may even be formalin-fixed paraffin-embedded (FFPE) tissue specimens, which severely affects fragmentation and down-stream analysis of the DNA. It is important to consider the best clinical procedure for sample collection when designing an epigenetic biomarker. For example, blood samples might be taken from venous blood or capillary blood by finger pricks, shipped as dried blood spot or at liquid state, frozen or at room temperature (Sontag et al., 2022). For liquid biopsies, it may also be relevant if DNA is extracted from serum or plasma (Constancio et al., 2020), or if the donor performed physical activity before sample collection (Neuberger et al., 2022). Similar considerations are also required for specimen from urine, stool, airway, or other tissues (Locke et al., 2019). There are many alternative protocols for DNA extraction and conversion that may need to be accredited for the IVDD, too. Depending on the number of samples, barcoding and automation need to be considered.

Accuracy of DNA Methylation Measurement

In contrast to other epigenetic modifications, such as the histone code, the DNAm level can be determined precisely for each CpG, as absolute value ranging from 0 to 100%. For many biomarkers the difference of DNAm levels is relatively low and therefore the accuracy of DNAm measurement is an important parameter. This applies particularly for epigenetic biomarkers that do not simply classify between two DNAm states (yes or no), but rather resemble a continuous variable (as for example in epigenetic leukocyte counts and epigenetic clocks). In this case small deviations in the accuracy of DNAm measurements can have big impact on predictions, even though the biological relevance and the phenotype effect of small changes in DNAm is still under debate. For the deep-sequencing based approaches accuracy depends largely on the sequencing depth. In contrast, Infinium BeadChip microarrays achieve overall relatively high precision, while there is batch-to-batch and inter-laboratory variation, which can be partially compensated for by background correction and normalization procedures (Pidsley et al., 2016). Even higher accuracy and precision of DNAm levels at specific CpGs can be achieved by targeted measurement techniques (Blueprint-consortium, 2016). For instance, by pyrosequencing, mass-spectrometry based EpiTYPER, or bisulfite amplicon sequencing (BA-seq) DNAm differences below 5% are detectable. However, these targeted methods can have an inherent PCR bias if either methylated or non-methylated sequences are preferentially amplified. A possible exception is digital droplet PCR (ddPCR):

The bisulfite-converted DNA is dispersed into small droplets with individual PCR reactions, which are either detected as positive or negative for methylated and non-methylated sequences. The exact DNAm values can then be calculated using Poisson distribution for the different droplets. In fact, there is some evidence that DNAm measurement with ddPCR has even higher accuracy than pyrosequencing, but this may largely depend on the assay design (Han et al., 2020a; Sontag et al., 2022). Either way, inter-laboratory comparison and round robin tests should be considered.

Identification of Confounding Parameters

For clinical applications it is necessary to have a good understanding of confounding factors. Environmental and life-style parameters, aging, cellular composition, and diseases can affect the signatures (Gadd et al., 2021). Furthermore, there are DNAm differences between woman and men, particularly at sex chromosomes. For example, epigenetic age predictions of women are overall under-estimated as compared to men, which might be attributed to their longer life expectancy (Horvath, 2013; Weidner et al., 2014). In addition, epigenetic characteristics can differ between ethnic groups (Becker et al., 2021). Such confounding parameters can be identified by epigenome-wide association studies (EWAS), that have enormous power when utilizing large available datasets. It is important to envisage potential confounding parameters already at study design but they can ultimately only be identified through extensive studies with many patient samples in the validation phase.

Accreditation as a Diagnostic Procedure

An approved diagnostic test needs to be applicable reliably and standardized for many years. For clinical application diagnostic tests are required to have a quality certificate, for example a CE mark from the European Union as an *in vitro* diagnostic device (IVDD) according to IVDR2017/746, which will be mandatory to be accomplished by all diagnostic devices in May 2022, or approval from the Food and Drug Administration (FDA) (Locke et al., 2019). The process to get such approvals is based on validation and benchmarking experiments — it takes years and is very cost-intensive. Also, all instruments need to have certificates for clinical use, which is not always the case. Complex analytical procedures with many suppliers tend to be detrimental to clinical implementation. For example, the changes in Infinium BeadChip platforms (from 27k, to 450k, to EPIC) and their continuous annotation updates are not a major obstacle for basic research, but they are a challenge for an already accredited process. With this in mind, it may be advantageous to keep diagnostic procedures as simple as possible and to rather rely on targeted assays for a higher and consistent throughput.

Standardization of Data Analysis

For an IVDD all processes must be fully standardized, including data analysis. For epigenetic studies the bottleneck is often bioinformatic evaluation. Not only it is becoming increasingly difficult to recruit specially trained bioinformaticians, there are countless integration possibilities of different datasets and the rapidly growing arsenal of tools can hardly be overseen, even by experts. Furthermore, standardized collection of DNAm data and

clinical information is critical. Every program that is used needs to be completely described and approved, even for simple linear regression equations. The programs used for analysis of sequencing data must also be accredited accordingly. Thus, it is also important to simplify and standardize the data analysis procedures.

Turnaround Time

The turnaround time can be crucial for clinical application. The procedures usually comprise DNA isolation and bisulfite conversion, which already require amost one day. Targeted methods, such as pyrosequencing, ddPCR, or EpiTYPER necessitate only a few additional PCR and/or sequencing steps and the results may therefore be available after one or two days (Blueprint-consortium, 2016). In contrast, genome-wide methods and amplicon sequencing are often subject to longer waiting times due to the high frequentation of sequencing instruments. Quality control and bioinformatic evaluation of deep-sequencing data may also require several additional days. For the final decision on the method of the diagnostic procedure, the time aspects should be taken into account — not least because personnel costs represent a significant component of the process costs.

Costs and Customer Requirements

It is advantageous to consider potential customers and marketing already at the time of the initial study design. How would the assay compare with conventional assays? Would it rather be distributed as kit or service? For example, customers in forensics may prefer in-house analysis—in this case a kit for commonly available instrumentation might be advantageous. On the other hand, DNA is relatively stable and can be shipped at room temperature (in contrast to RNA, which is bound to rapid degradation). The samples can be frozen for many years prior to measurement, making retrospective studies possible. Therefore, the samples might alternatively be easily shipped to a service provider for a centralized analysis. Due to the personnel costs the price per assay may largely depend on the number of samples that can be processed in parallel. Automation and parallelization of sample acceptance, DNA isolation, bisulfite conversion (or enzymatic conversion), amplification and measurement, up to the evaluation and the generation of a final report is therefore to be aimed for.

CONCLUSION

There is a "valley of death" between identification of promising epigenetic signatures and translation into clinical practice. Despite the enormous diagnostic potential and rapidly growing numbers of described epigenetic signatures only a small number of epigenetic biomarkers are approved as IVDD (Beltran-Garcia et al., 2019; Locke et al., 2019; Taryma-Lesniak et al., 2020). To bridge this gap, it is necessary to already focus on the clinical demands of a potential epigenetic biomarker at the initial study design. While the genomic regions are usually identified with genome-wide approaches, clinical application may require further translation into a robust and highly standardized targeted assay that rather uses a small number of CpGs. The scientific community needs to acknowledge that optimization, validation, and standardization of existing epigenetic signatures is important research to establish diagnostic DNAm biomarkers.

REFERENCES

Accomando, W. P., Wiencke, J. K., Houseman, E., Nelson, H. H., and Kelsey, K. T. (2014). Quantitative Reconstruction of Leukocyte Subsets Using DNA Methylation. *Genome Biol.* 15, R50. doi:10.1186/gb-2014-15-3-r50

Arenson, A., Haghani, A., Thompson, M. J., Pellegrini, M., Kwon, S. B., Vu, H., et al. (2022). A Mammalian Methylation Array for Profiling Methylation Levels at Conserved Sequences. *Nat. Commun.* 13, 783. doi:10.1038/s41467-022-28355-z

Baron, U., Werner, J., Schildknecht, K., Schulze, J. J., Mulu, A., Liebert, U. G., et al. (2018). Epigenetic Immune Cell Counting in Human Blood Samples for Immunodiagnostics. *Sci. Transl Med.* 10 (452), eaan3508. doi:10.1126/scitranslmed.aan3508

Becker, J., Bohme, P., Reckert, A., Eickhoff, S. B., Koop, B. E., Blum, J., et al. (2021). Evidence for Differences in DNA Methylation between Germans and Japanese. *Int. J. Leg. Med.* Online ahead of print. doi:10.1007/s00414-021-02736-3

Bell, C. G., Lowe, R., Adams, P. D., Baccarelli, A. A., Beck, S., Bell, J. T., et al. (2019). DNA Methylation Aging Clocks: Challenges and Recommendations. *Genome Biol.* 20, 249. doi:10.1186/s13059-019-1824-y

Beltrán-García, J., Osca-Verdegal, R., Mena-Mollá, S., and García-Giménez, J. L. (2019). Epigenetic IVD Tests for Personalized Precision Medicine in Cancer. *Front. Genet.* 10, 621. doi:10.3389/fgene.2019.00621

Blueprint-consortium (2016). Quantitative Comparison of DNA Methylation Assays for Biomarker Development and Clinical Applications. *Nat. Biotechnol.* 34, 726–737. doi:10.1038/nbt.3605

Bocklandt, S., Lin, W., Sehl, M. E., Sánchez, F. J., Sinsheimer, J. S., Horvath, S., et al. (2011). Epigenetic Predictor of Age. *PLoS ONE* 6, e14821. doi:10.1371/journal.pone.0014821

Božić, T., Kuo, C. C., Hapala, J., Franzen, J., Eipel, M., Platzbecker, U., et al. (2022). Investigation of Measurable Residual Disease in Acute Myeloid Leukemia by DNA Methylation Patterns. *Leukemia* 36, 80–89. doi:10.1038/s41375-021-01316-z

Cerrato, F., Sparago, A., Ariani, F., Brugnoletti, F., Calzari, L., Coppedè, F., et al. (2020). DNA Methylation in the Diagnosis of Monogenic Diseases. *Genes (Basel)* 11 (4), 355. doi:10.3390/genes11040355

Constancio, V., Nunes, S. P., Henrique, R., and Jeronimo, C. (2020). DNA Methylation-Based Testing in Liquid Biopsies as Detection and Prognostic Biomarkers for the Four Major Cancer Types. *Cells* 9 (3), 624. doi:10.3390/cells9030624

Frobel, J., Božić, T., Lenz, M., Uciechowski, P., Han, Y., Herwartz, R., et al. (2018). Leukocyte Counts Based on DNA Methylation at Individual Cytosines. *Clin. Chem.* 64, 566–575. doi:10.1373/clinchem.2017.279935

Gadd, D. A., Stevenson, A. J., Hillary, R. F., McCartney, D. L., Wrobel, N., McCafferty, S., et al. (2021). Epigenetic Predictors of Lifestyle Traits Applied to the Blood and Brain. *Brain Commun.* 3, fcab082. doi:10.1093/braincomms/fcab082

García-Giménez, J. L., Seco-Cervera, M., Tollefsbol, T. O., Romá-Mateo, C., Peiró-Chova, L., Lapunzina, P., et al. (2017). Epigenetic Biomarkers: Current Strategies and Future Challenges for Their Use in the Clinical Laboratory. *Crit. Rev. Clin. Lab. Sci.* 54, 529–550. doi:10.1080/10408363.2017.1410520

Griffin, P. T., Kane, A. E., Trapp, A., Li, J., McNamara, M. S., Meer, M. V., et al. (2021). Ultra-cheap and Scalable Epigenetic Age Predictions with TIME-Seq. *BioRxiv*. Preprint. doi:10.1101/2021.10.25.465725

Han, Y., Franzen, J., Stiehl, T., Gobs, M., Kuo, C.-C., Nikolić, M., et al. (2020a). New Targeted Approaches for Epigenetic Age Predictions. *BMC Biol.* 18, 71. doi:10.1186/s12915-020-00807-2

Han, Y., Nikolić, M., Gobs, M., Franzen, J., de Haan, G., Geiger, H., et al. (2020b). Targeted Methods for Epigenetic Age Predictions in Mice. *Sci. Rep.* 10, 22439. doi:10.1038/s41598-020-79509-2

Horvath, S. (2013). DNA Methylation Age of Human Tissues and Cell Types. *Genome Biol.* 14, R115. doi:10.1186/gb-2013-14-10-r115

Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., et al. (2012). DNA Methylation Arrays as Surrogate Measures of Cell Mixture Distribution. *BMC Bioinformatics* 13, 86. doi:10.1186/1471-2105-13-86

Jost, E., Lin, Q., Weidner, C. I., Wilop, S., Hoffmann, M., Walenda, T., et al. (2014). Epimutations Mimic Genomic Mutations of DNMT3A in Acute Myeloid Leukemia. *Leukemia* 28, 1227–1234. doi:10.1038/leu.2013.362

Koch, C. M., and Wagner, W. (2011). Epigenetic-aging-signature to Determine Age in Different Tissues. *Aging* 3, 1018–1027. doi:10.18632/aging.100395

Lin, Q., Weidner, C. I., Costa, I. G., Marioni, R. E., Ferreira, M. R. P., Deary, I. J., et al. (2016). DNA Methylation Levels at Individual Age-Associated CpG Sites Can Be Indicative for Life Expectancy. *Aging* 8, 394–401. doi:10.18632/aging.100908

Liu, C., Marioni, R. E., Hedman, Å. K., Pfeiffer, L., Tsai, P.-C., Reynolds, L. M., et al. (2018). A DNA Methylation Biomarker of Alcohol Consumption. *Mol. Psychiatry* 23, 422–433. doi:10.1038/mp.2016.192

Locke, W. J., Guanzon, D., Ma, C., Liew, Y. J., Duesing, K. R., Fung, K. Y. C., et al. (2019). DNA Methylation Cancer Biomarkers: Translation to the Clinic. *Front. Genet.* 10, 1150. doi:10.3389/fgene.2019.01150

Malic, L., Daoud, J., Geissler, M., Boutin, A., Lukic, L., Janta, M., et al. (2019). Epigenetic Subtyping of white Blood Cells Using a Thermoplastic Elastomer-Based Microfluidic Emulsification Device for Multiplexed, Methylation-specific Digital Droplet PCR. *Analyst* 144, 6541–6553. doi:10.1039/c9an01316d

Marioni, R. E., Shah, S., McRae, A. F., Chen, B. H., Colicino, E., Harris, S. E., et al. (2015). DNA Methylation Age of Blood Predicts All-Cause Mortality in Later Life. *Genome Biol.* 16, 25. doi:10.1186/s13059-015-0584-6

Moss, J., Magenheim, J., Neiman, D., Zemmour, H., Loyfer, N., Korach, A., et al. (2018). Comprehensive Human Cell-type Methylation Atlas Reveals Origins of Circulating Cell-free DNA in Health and Disease. *Nat. Commun.* 9, 5068. doi:10.1038/s41467-018-07466-6

Neuberger, E. W. I., Sontag, S., Brahmer, A., Philippi, K. F. A., Radsak, M. P., Wagner, W., et al. (2022). Physical Activity Specifically Evokes Release of Cell-free DNA from Granulocytes Thereby Affecting Liquid Biopsy. *Clin. Epigenetics*. doi:10.1101/2021.09.01.21262910

Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell* 99, 247–257. doi:10.1016/s0092-8674(00)81656-6

Peiró-Chova, L., Bahamonde Ponce, O., Abril-Tormo, C., Martínez-Santamaría, J., López-Guerrero, J. A., and Riegman, P. H. J. (2016). "Chapter 2 - the Importance of Biobanks in Epigenetic Studies," in *Epigenetic Biomarkers and Diagnostics*. Editor J. L. García-Giménez (Academic Press), 19–35.

Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., et al. (2016). Critical Evaluation of the Illumina MethylationEPIC BeadChip Microarray for Whole-Genome DNA Methylation Profiling. *Genome Biol.* 17, 208. doi:10.1186/s13059-016-1066-1

Pitoiset, F., Cassard, L., El Soufi, K., Boselli, L., Grivel, J., Roux, A., et al. (2018). Deep Phenotyping of Immune Cell Populations by Optimized and Standardized Flow Cytometry Analyses. *Cytometry* 93, 793–802. doi:10.1002/cyto.a.23570

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic Reprogramming in Mammalian Development. *Science* 293, 1089–1093. doi:10.1126/science.1063443

Salas, L. A., Koestler, D. C., Butler, R. A., Hansen, H. M., Wiencke, J. K., Kelsey, K. T., et al. (2018). An Optimized Library for Reference-Based Deconvolution of Whole-Blood Biospecimens Assayed Using the Illumina HumanMethylationEPIC BeadArray. *Genome Biol.* 19, 64. doi:10.1186/s13059-018-1448-7

Schmidt, M., Maié, T., Dahl, E., Costa, I. G., and Wagner, W. (2020). Deconvolution of Cellular Subsets in Human Tissue Based on Targeted DNA Methylation Analysis at Individual CpG Sites. *BMC Biol.* 18, 178. doi:10.1186/s12915-020-00910-4

Schoofs, T., Berdel, W. E., and Müller-Tidow, C. (2014). Origins of Aberrant DNA Methylation in Acute Myeloid Leukemia. *Leukemia* 28, 1–14. doi:10.1038/leu.2013.242

Smith, Z. D., and Meissner, A. (2013). DNA Methylation: Roles in Mammalian Development. *Nat. Rev. Genet.* 14, 204–220. doi:10.1038/nrg3354

Sontag, S., Bocova, L., Hubens, W. H. G., Nüchtern, S., Schnitker, M., Look, T., et al. (2022). Toward Clinical Application of Leukocyte Counts Based on Targeted DNA Methylation Analysis. *Clin. Chem.*. doi:10.1093/clinchem/hvac006

Taryma-Lesniak, O., Sokolowska, K. E., and Wojdacz, T. K. (2020). Current Status of Development of Methylation Biomarkers for *In Vitro* Diagnostic IVD Applications. *Clin. Epigenetics* 12, 100.

Trapp, A., Kerepesi, C., and Gladyshev, V. N. (2021). Profiling Epigenetic Age in Single Cells. *Nat. Aging* 1, 1189–1201. doi:10.1038/s43587-021-00134-3

Tsaprouni, L. G., Yang, T.-P., Bell, J., Dick, K. J., Kanoni, S., Nisbet, J., et al. (2014). Cigarette Smoking Reduces DNA Methylation Levels at Multiple Genomic Loci but the Effect Is Partially Reversible upon Cessation. *Epigenetics* 9, 1382–1396. doi:10.4161/15592294.2014.969637

Wagner, W. (2017). Epigenetic Aging Clocks in Mice and Men. *Genome Biol.* 18, 107. doi:10.1186/s13059-017-1245-8

Wagner, W., Weidner, C. I., and Lin, Q. (2015). Do age-associated DNA Methylation Changes Increase the Risk of Malignant Transformation? *Bioessays* 37, 20–24. doi:10.1002/bies.201400063

Weidner, C. I., and Wagner, W. (2014). The Epigenetic Tracks of Aging. *Biol. Chem.* 395, 1307–1314. doi:10.1515/hsz-2014-0180

Weidner, C., Lin, Q., Koch, C., Eisele, L., Beier, F., Ziegler, P., et al. (2014). Aging of Blood Can Be Tracked by DNA Methylation Changes at Just Three CpG Sites. *Genome Biol.* 15, R24. doi:10.1186/gb-2014-15-2-r24

Zhu, J., Adli, M., Zou, J. Y., Verstappen, G., Coyne, M., Zhang, X., et al. (2013). Genome-wide Chromatin State Transitions Associated with Developmental and Environmental Cues. *Cell* 152, 642–654. doi:10.1016/j.cell.2012.12.033