

**Document Version**

Final published version

**Licence**

CC BY

**Citation (APA)**

Bizzarri, D., Reinders, M. J. T., Kuiper, L., Beekman, M., Deelen, J., van Meurs, J. B. J., van Dongen, J., Pool, R., van den Akker, E. B., & More Authors (2024). NMR metabolomics-guided DNA methylation mortality predictors. *EBioMedicine*, 107, Article 105279. <https://doi.org/10.1016/j.ebiom.2024.105279>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.  
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

**Sharing and reuse**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# NMR metabolomics-guided DNA methylation mortality predictors



Daniele Bizzari,<sup>a,b,c</sup> Marcel J. T. Reinders,<sup>b,c</sup> Lieke Kuiper,<sup>d,e</sup> Marian Beekman,<sup>a</sup> Joris Deelen,<sup>a,f,g</sup> Joyce B. J. van Meurs,<sup>d,h</sup> Jenny van Dongen,<sup>i,j,k</sup> René Pool,<sup>l,k</sup> Dorret I. Boomsma,<sup>l,j,k</sup> Mohsen Ghanbari,<sup>l</sup> Lude Franke,<sup>m</sup> BBMRI Consortium,<sup>n,o</sup> Pieterella E. Slagboom,<sup>a,f</sup> and Erik B. van den Akker<sup>a,b,c,\*</sup>



<sup>a</sup>Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands

<sup>b</sup>Leiden Computational Biology Center, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands

<sup>c</sup>Delft Bioinformatics Lab, TU Delft, Delft, the Netherlands

<sup>d</sup>Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands

<sup>e</sup>Center for Nutrition, Prevention and Health Services, National Institute for Public Health and Environment (RIVM), Bilthoven, the Netherlands

<sup>f</sup>Max Planck Institute for the Biology of Ageing, Cologne, Germany

<sup>g</sup>Cologne Excellence Cluster on Cellular Stress Responses in Aging Associated Diseases, University of Cologne, Cologne, Germany

<sup>h</sup>Department of Orthopaedics & Sports, Erasmus Medical Center, Rotterdam, the Netherlands

<sup>i</sup>Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

<sup>j</sup>Amsterdam Reproduction and Development (AR&D) Research Institute, Amsterdam, the Netherlands

<sup>k</sup>Amsterdam Public Health Research Institute, Amsterdam, the Netherlands

<sup>l</sup>Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands

<sup>m</sup>Department of Genetics, University Medical Center Groningen, Groningen, the Netherlands

## Summary

**Background** <sup>1</sup>H-NMR metabolomics and DNA methylation in blood are widely known biomarkers predicting age-related physiological decline and mortality yet exert mutually independent mortality and frailty signals.

**Methods** Leveraging multi-omics data in four Dutch population studies (N = 5238, ~40% of which male) we investigated whether the mortality signal captured by <sup>1</sup>H-NMR metabolomics could guide the construction of DNA methylation-based mortality predictors.

**Findings** We trained DNA methylation-based surrogates for 64 metabolomic analytes and found that analytes marking inflammation, fluid balance, or HDL/VLDL metabolism could be accurately reconstructed using DNA-methylation assays. Interestingly, a previously reported multi-analyte score indicating mortality risk (MetaboHealth) could also be accurately reconstructed. Sixteen of our derived surrogates, including the MetaboHealth surrogate, showed significant associations with mortality, independent of relevant covariates.

**Interpretation** The addition of our metabolic analyte-derived surrogates to the well-established epigenetic clock GrimAge demonstrates that our surrogates potentially represent valuable mortality signal.

**Funding** BBMRI-NL, X-omics, VOILA, Medical Delta, NWO, ERC.

**Copyright** © 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

**Keywords:** DNA methylation predictors; NMR metabolomics; Ageing biomarkers; Epigenetic clock; Metabolic risk score; Epidemiology

eBioMedicine  
2024;107: 105279  
Published Online xxx  
<https://doi.org/10.1016/j.ebiom.2024.105279>

\*Corresponding author. Leiden Computational Biology Center, Leiden University Medical Center, Einthovenweg 20, 2333, ZC, Leiden, the Netherlands.

E-mail address: [e.b.van\\_den\\_akker@lumc.nl](mailto:e.b.van_den_akker@lumc.nl) (E.B. van den Akker).

<sup>b</sup>BIOS Consortium, see Consortium Banner [Supplementary Material](#).

<sup>o</sup>BBMRI-NL Consortium, see Consortium Banner [Supplementary Material](#).

### Research in context

#### Evidence before this study

The development of a comprehensive “biomarker of ageing” capable of identifying frailty and functional decline in the global ageing population would represent a significant leap forward in geroscience and healthcare prevention. Omics clocks, sophisticated prediction models condensing extensive molecular data into manageable multi-biomarker scores, hold immense promise in elucidating the biological trajectory of ageing. Focusing our attention on second generation of clocks, which draw predictive value from longitudinal endpoints, <sup>1</sup>H-NMR metabolomics and DNA methylation are emerging as pivotal omics-layers, with prominent scores such as MetaboHealth, PhenoAge, and GrimAge. DNA methylation emerges as a remarkably precise estimator of chronological age, with additional potential to encode for risk factors like as smoking, consequently also predicting mortality and disease susceptibility. Nonetheless, also <sup>1</sup>H-NMR metabolomics exhibits robust predictive capabilities in tracking various risk factors and endpoints, particularly when concerning cardiovascular mortality. Recent investigations underscore the distinct and complementary nature of mortality-oriented scores derived from both -omics layers. They appear to independently contribute to the prediction of frailty and mortality among older adults, highlighting the significance of integrating two omics layers.

#### Added value of this study

In this study, we investigate the potential of DNA methylation profiles to serve as epigenetic predictors for 64 metabolomics features and one metabolomics mortality score

(MetaboHealth). Leveraging a dataset comprising of four large Dutch epidemiological cohorts, and designing rigorous calibration and feature selection techniques, we achieved a successful quantification of metabolomics markers related to inflammation, fluid balance, HDL-, and VLDL-related markers, and MetaboHealth. These markers exhibited distinct signals independent of previously trained methylation-based surrogates, both quantitatively and as mortality indicators, as validated in an independent test set. Consequently, a mortality model incorporating nine of our DNAm-metabolomics features demonstrated superior predictive capability compared to current state-of-the-art models.

#### Implications of all the available evidence

While previous research has successfully harnessed models to leverage clinical health variables, cell counts, and proteomics markers from methylation profiles, our study unveil the added value of integrating them with faithful metabolomics surrogates. These additional DNAm-based features not only expand the scope of epigenetic clocks to encompass cardiometabolic signal, but also demonstrate complementary signal to GrimAge and its surrogates in predicting mortality. Moreover, through a meticulous evaluation of the epigenetic content within our DNAm-based metabolomics models, we observe an optimised selection of highly informative CpG sites. These sites resonate with metabolic and developmental processes, cell differentiation dynamics, and the intricate landscape of cardiometabolic diseases, thus enriching our understanding of ageing-related pathways and facilitating more precise predictive modelling.

## Introduction

A common goal in geroscience is to identify mechanisms that drive ageing and design interventions that might slow down or even reverse the rate of ageing.<sup>1</sup> For this purpose, it is essential to have indicators not only quantifying ageing, but simultaneously marking the trajectory of overall health decline.<sup>2</sup> While calendar age is a core risk factor for almost any common disease, it has many limitations for capturing the variability in health-span. Crucially, calendar age does not capture the effects of an individual’s lifestyle, nor incorporates readouts of functional decline. Instead, faithful markers of biological age would allow to quantify the vulnerability to acute and chronic diseases irrespective of an individual’s calendar age, and to develop and monitor effective healthy lifestyle advice and anti-ageing interventions. The earliest approaches to construct such markers of biological age relied on clinical measures of physiological capacity.<sup>3</sup> Later molecular and -omics approaches gained popularity, initially including markers such as leukocyte telomere length,<sup>4</sup> followed by multi-marker algorithms based on high-throughput platforms, such as DNA methylation,<sup>5</sup>

transcriptomics,<sup>6</sup> metabolomics,<sup>7</sup> and proteomics.<sup>8</sup> Importantly, these algorithms were trained to estimate cross-sectional chronological age. Of these omics approaches, particularly DNA methylation-based algorithms exhibited remarkably high accuracies in predicting calendar age,<sup>5</sup> and were named ‘DNA methylation clocks’. Nonetheless, while interesting by itself, this observation highlighted a fundamental limitation in this design of markers of biological age. Since nearly-perfect age predictors would arrive to similar observations as chronological age, they would lose their characteristics as age-independent health status indicators.<sup>9</sup>

Concomitantly, a second generation of -omics markers was introduced, which instead were trained to predict the mortality risk. Prominent examples of these mortality-trained multivariate markers include the DNA methylation-based PhenoAge<sup>10</sup> and GrimAge<sup>11</sup> and the <sup>1</sup>H-NMR metabolomics-based MetaboHealth.<sup>12</sup> These predictors were trained quite differently. The wide availability of the Nightingale Health <sup>1</sup>H-NMR metabolomics in large prospective population studies, in combination with its relatively narrow though

informative content (~250 analytes), allowed for a more classic and direct approach. Deelen et al. trained MetaboHealth as a linear combination of 14 metabolic features, showing a strong predictive value, not only for mortality risk, but also for other outcomes, including pneumonia,<sup>13</sup> and frailty.<sup>14</sup> Conversely, the DNA methylation platform by Illumina contains hundreds of thousands of features, and thus requires additional guidance to robustly capture the mortality signal. Hence, the PhenoAge and GrimAge were trained using the so-called two-stage approaches, in which more widely-available markers associated with mortality were leveraged to help extract the mortality signal.<sup>10,11</sup> PhenoAge achieved this by first training an all-cause mortality predictor based on clinical measures (e.g., glucose, C-reactive-protein), which was then re-estimated using DNA methylation. Similarly, DNAm-GrimAge is composed by a combination of DNA methylation-based surrogates for molecular or phenotypic markers known to associate with mortality. Interestingly, both two-step training strategies yielded DNA methylation-based scores that can associate not only with mortality, but also with a wide diversity of disease outcomes. These developments indicate that mortality-trained predictors for biological age can be trained using different omics platforms, and moreover, that DNA-methylation might serve as a platform to integrate these signals captured by different data sources. This latter concept was recently further substantiated by the work of Gadd et al., who systematically developed DNA-methylation-based predictors for 109 plasma proteins, called EpiScores. Their findings demonstrated significant associations with various incident morbidities over a span of 14-years.<sup>15</sup> Moreover, these surrogates were afterwards employed, in an attempt to refine the GrimAge score in a clock known as bAge.<sup>16</sup>

In a recent study we demonstrated that mortality-based predictors such as MetaboHealth and GrimAge are instrumental in predicting frailty in studies of middle-aged and elderly individuals.<sup>14</sup> Importantly, we also showed that these scores confer mutually independent information for predicting five frailty indicators and mortality (specific details in Kuiper et al.<sup>14</sup>). Viewing these developments in the field, we thus pose the question to what extent the mortality signal captured by <sup>1</sup>H-NMR metabolomics could be transferred and integrated with the mortality signals captured by the DNA-methylation platform. For this purpose, we will evaluate both strategies for training two-stage DNA-methylation based mortality predictors. On one hand, we will train a DNA methylation-based predictor re-estimating directly MetaboHealth, akin the strategy of PhenoAge. On the other hand, we will train DNA-methylation surrogates for single metabolomics features, and combine these in an overall score, akin GrimAge. Moreover, we will evaluate to what extent DNA-methylation surrogates features from different

origins capture mutually independent signals, also with respect to predicting mortality risk.

## Methods

### Dataset description

#### Cohorts

This study was performed using DNA methylation data (DNAm, Illumina 450 k array) and <sup>1</sup>H-NMR metabolomics (Nightingale Health, platform version 2020) from 4 Dutch cohorts: LifeLines-Deep (LIFELINES), Leiden Longevity Study (LLS-PAROFFS), Netherlands Twin Register (NTR) and Rotterdam Study (RS-I), all part of the BIOS consortium.<sup>17,18</sup> For the current study, the BIOS multi-omics compendium was further extended with 1145 samples from the NTR for which the entire process of array measurement to quality control and normalization was done together with the other BIOS-NTR samples,<sup>19</sup> and 904 samples from the Rotterdam Study (RS-II).<sup>14</sup> A thorough description of all cohorts and their ethics statement are provided in the [Supplementary Materials](#). The datasets were realised by the Dutch part of the Biobanking and BioMolecular Resources and Research Infrastructure (BBMRI-NL). The final dataset contained data for 5238 individuals.

#### Metabolomics data

The metabolomics data were generated by the BBMRI-NL Metabolomics Consortium for the cohorts LIFELINES, LLS-PAROFFS, NTR and RS-I. The current study employs a total of 4334 EDTA plasma samples quantified with the high-throughput proton Nuclear Magnetic Resonance (<sup>1</sup>H-NMR) platform made available by Nightingale Health Ltd., Helsinki, Finland (platform version 2020). This technique can quantify over 250 metabolic features, including also ratios and derived features.<sup>20,21</sup>

#### DNA methylation data

DNA methylation data were generated by the subsection of BBMRI-NL named Biobank-based Integrative Omics Study (BIOS) Consortium for the cohorts LIFELINES, LLS-PAROFFS, NTR, RS-I, and RS-II (total of 5238 samples). The DNAm was assessed from whole blood samples with an Illumina iScan BeadChip according to the manufacturer's protocol: the Illumina HumanMethylation450 BeadChip (450 k array). For compatibility with the following versions of the Illumina array, we only considered CpG sites which are available in the Illumina HumanMethylation450 BeadChip and the MethylationEPIC BeadChip. We analysed the DNAm  $\beta$  values, which range from 0 to 1, to indicate the proportion of methylated sites at a specific CpG in a sample.

#### Mortality data

We evaluated the associations of the DNAm-based features with all-cause mortality in a subsample of the

Rotterdam Study (RS-I and RS-II) comprising a total of 1544 samples, of which a subset of 640 (RS-I, 104 deaths) had also Illumina 450 k and Nightingale Health metabolomics. The information on the vital status of the participants in RS was last updated on the 20th of October 2022. The dataset comprehends 1544 samples, 285 of which are deceased. All the DNAm-based features were z-scaled within the RS.

### Pre-processing

#### Quality control of the metabolomics dataset

To ensure the quality of our data, we applied standardised quality control processes, which have been described in previous publications (summarised in [Figure S1](#)).<sup>7,22</sup> First, we limited our analyses to a subset of 65 features (out of 250), previously selected to be a mutually independent subset.<sup>7,12,22</sup> This selection includes fatty acids, routine lipid concentrations, lipoprotein subclasses and low molecular weight metabolites. A complete list of the variables can be found in the [Supplementary Materials Table S3](#). In addition, pyruvate was excluded due to its high missingness in NTR (80%). Despite a small percentage of values under detection limit for acetoacetate (8% in NTR), and an even smaller percentage of outliers in glucose and xl\_hdl\_c (less than 0.15%), we decided to retain all other variables ([Figure S1C–E](#)). Samples with more than 1 outlier (2 from Lifelines and 1 from RS-I) were further removed. We then used *nipals* (from the package *pcaMethods*) to impute the 584 missing values, which accounted for 0.211% of the remaining values. The final dataset included 4334 samples and 64 metabolic measures.

#### Quality control of the DNA methylation dataset

The quality control and normalization of the DNA methylation (DNAm) was performed using a workflow developed by the BIOS Consortium for each cohort and thoroughly described in DNAmArray ([https://molepi.github.io/DNAmArray\\_workflow/](https://molepi.github.io/DNAmArray_workflow/)). In brief, sample-level QC was performed with the R package *MethylAid*.<sup>23</sup> Probes were set to missing based on the number of available beads ( $\leq 2$ ), intensity equal to zero, or the detection p value ( $p < 0.01$ ). Probes with more than 5% missing were excluded from all samples. The remaining missingness was imputed using *impute.knn* from the R package *impute*.<sup>24</sup> Functional normalization was then applied as implemented in *minfi*. Finally, we removed an ulterior set of ~60,000 underperforming probes as suggested by Zhou et al.<sup>25</sup>

#### Calibration of <sup>1</sup>H-NMR-metabolomics

To minimise any bias that may arise from batch effects among the four cohorts included in our study, we performed a cross-cohort calibration. We followed the assumption that similar phenotypic characteristics

result in similar metabolomics profiles.<sup>26</sup> We used sex, age, and BMI as matching characteristics, given their well-known association with the metabolomic features in the Nightingale Health Platform.<sup>7,22,26–28</sup> We considered LIFELINES as our reference cohort, as it spanned a broad range of age, and BMI and had an equal representation amongst sexes. To further minimise the impact of sex on our results, we selected the subset of samples used for cross-cohort matching to have equal numbers of men and women. Following this strategy, we identified the following subsets of participants used for matching: 73 men and 73 women in LLS-PAROFFS; 140 men and 140 women in NTR; 37 men and 37 women in RS ([Figure S2](#)).

Based on these matching samples between cohorts, we calculated the shift in mean and standard deviation for each metabolic feature required to transform the distribution of values observed in a cohort to match the distribution in the reference cohort. We then applied this transformation to all samples of each cohort (see [Supplementary Materials](#)). The final dataset was log-transformed and standard normalised (zero mean and unit standard deviation) across all samples to obtain normally distributed concentration with comparable ranges across all metabolic features.

T-distributed stochastic neighbour embedding (tSNE, R package *Rtsne*) was used to visually inspect the effect of this calibration, comparing the sample similarities before and after calibration. Moreover, K-nearest neighbour batch effect test (kBET, R package *kBET*) was applied to the matching samples of each biobank before and after calibration to quantitatively evaluate the mixing of the samples.<sup>29</sup> Finally, we used principal variance Component Analysis (PVCA, R package *pvca*), to determine if the calibration disrupted the sources of variability of the dataset.<sup>30</sup>

### Application of previously trained multivariate models

#### MetaboHealth

The MetaboHealth model is a mortality predictor based on Nightingale Health metabolomics concentration.<sup>12</sup> We applied this model both on the uncalibrated and calibrated version of the <sup>1</sup>H-NMR metabolomics dataset using the R-package *MiMIR*.<sup>31</sup>

#### Epigenetic clocks

We projected the Horvath, Hannum, DNAm PhenoAge in our data using the R package *methylclock* and the DNAm GrimAge clocks using Python scripts provided by Lu et al.<sup>11,32</sup> About 1000 out of 30,000 CpG sites required for calculating these biological age scores were missing or removed during QC and were imputed using the “datMiniAnnotation3\_GOLD.csv” file,<sup>33</sup> as outlined by the authors of the original papers.<sup>11</sup> The accuracy of the epigenetic clocks projected in BIOS is indeed preserved ([Figure S13](#)).

### EpiScores

We projected the EpiScores and bAge using the code available from the work of Bernabeu et al.<sup>16</sup>

### Estimation and evaluation of the epigenetic-based metabolic features

We derived prediction models for the 64 metabolomics features and the MetaboHealth score using blood methylation data. NTR, and LIFELINES, respectively the largest cohort and the cohort used as a reference for calibration, were employed for model development using the internal loop of a nested 5-Fold-Cross-Validation (5-Fold CV; [Supplementary Material SM2](#)). We employed ElasticNET regression from the R package *glmnet* to train the models ([Supplementary Material SM3](#)).

Other studies show the benefit of pre-selecting the features before using ElasticNET regression.<sup>16,34</sup> For this reason, we performed Epigenome Wide Association Studies (EWASes) to identify CpG sites showing linear association with each feature separately in NTR and LIFELINES (*metabolic feature ~ CpG site*) and pre-selected the CpG probes with a consistent association sign (positive or negative in both cohorts) and nominal significance (p value <0.05) to enrich for potentially predictive CpGs, while minimizing the risk of prematurely excluding valuable signal. Subsequent training of the ElasticNET regression models determined the final set of CpG sites included in each model ([Supplementary Material SM2](#)). An overview of the number of CpGs selected during the EWAS pre-selection and ElasticNET regression phases can be found in [Supplementary Material Table S4](#). CpG selection and optimization of the penalization parameter  $\lambda$  were performed in the internal loop of the nested 5-Fold-Cross-Validation. The mixing parameter alpha was fixed at 0.5, based on previous work.<sup>11,22</sup> Model performances in NTR and LIFELINES are reported using the outer loop of the nested 5-Fold-Cross-Validation and should give an unbiased impression of the performances of models created with our training procedure. The final models were obtained by training the ElasticNET models with the optimised parameters on the whole of NTR and LIFELINES. Resulting models were then applied to the held-out LLS-PAROFFS and RS datasets. Finally, reported measures for model performance are the Pearson correlation (R) and the root mean square error (RMSE) of the predicted DNAm *metabolic features* with their measured concentrations.

### CpG sites characterization

To gain more insight into the biological phenomena that characterise our DNAm metabolomics models we evaluated their fully data-driven selection of 22,145 CpG sites.

#### EWAS enrichment analysis

We utilised the MRC-IEU EWAS Catalog<sup>35</sup> and the EWAS Atlas<sup>36</sup> to assess the previously known phenotypic annotation (traits) of the CpG sites selected by our

models. Both the EWAS Catalog and EWAS Atlas are online databases that compile results from Epigenome Wide Association Study results. By merging these two resources, both downloaded on March 13th, 2023, we aimed to gather a comprehensive list of previous EWASs. We gathered only the associations conducted with Illumina 450 k and accompanied with a PMID. We then unified redundant or possible synonymous traits between the two catalogues (EWAS Atlas and EWAS Catalog) and selected only the trait-to-CpG associations significant after Bonferroni correction (using the number of CpGs in the Illumina 450 k ~480,000). This process yielded 742,635 CpG-trait associations.

Next, we employed Fisher's exact test to assess the enrichments of each of the CpG sites selected by our DNAm models. To account for multiple testing, we used the Benjamini-Hochberg correction.

#### Annotation of the genomic position of the CpG sites

We used the R package *annotatr* to annotate the genomic features. CpG sites from the 450 k array were annotated using CpG Island (CGI) centric categories. The annotations we utilised are as follows: CGI (annotated in the R package *AnnotationHub*), shores (2 Kb upstream or downstream the CGI), shelves (2 kb flanking the CpG shores), interCGI (the rest of the CpGs). As for the genic annotations we considered regions 1–5 Kb upstream of the transcription starting site (TSS), promoters (<1 Kb upstream of the TSS), 5' UTR, 3' UTR, exons, introns, boundaries between introns and exons, and intergenic regions. Additionally, we report the annotations of active enhancers determined by Anderson et al.<sup>37</sup>

For the enrichment analyses of the annotations described above we calculated odds ratios (OR) of the CpGs included in each model compared to the rest of the 450 k array. Statistical significance was evaluated using the Fisher's exact test. The significance of the associations was established with an FDR <0.05.

#### Gene ontology enrichment analyses

To gain further insights into the genetic context of the set of CpG sites selected, we investigated the genes in cis, considering a maximum distance of 100 kb of distance.

Next, we utilised the genes associated with each CpG selections from our models to perform a functional enrichment using Gene Ontology. We employed GOfuncR package to explore the Biological *Processes* and *Molecular Functions*. The significance of the associations was established with an FDR <0.05. This analysis resulted in 2365 significant associations between CpG sites and genes.

### Associations with mortality in the Rotterdam Study (RS-I and RS-II)

#### Univariate mortality associations

We used Cox Proportional hazard to univariately associate our 65 DNAm metabolomics features, the

pre-trained DNAm clocks (e.g., PhenoAge, GrimAge), and 109 protein EpiScores with mortality (see [Supplementary Materials](#)). All models were corrected for age at blood sampling and sex. Additionally, we evaluated the association with mortality of our DNAm metabolomics features when correcting for sex, age and GrimAge. All p values were corrected using Benjamini Hochberg and considered significant if the FDR <0.05. We used the R-package *survival* to calculate the Cox regressions.

#### Multivariate mortality models

We then combined the DNAm features with sex and age in 4 different stepwise Cox regression models (see [Supplementary Materials](#)). The first base model included our DNAm metabolomics features. The second and third model added to the first model respectively DNAm-GrimAge and the DNAm-based components of the GrimAge model. Finally, the fourth model is based on a combination of our DNAm metabolomics features, the DNAm-based components of the GrimAge and the DNAm-based protein EpiScores.

To select the interesting DNAm surrogate, we used a stepwise (backward/forward) procedure for each Cox regression model. For each of the above-described selections, we started from a model containing the full set of variables and we removed or added an unselected metabolic surrogate at each round based on the improvement on the model calculated from the C-index, taking also into account the significance of the p value of each variable included in the model.

To compare the performances of the Cox regression models we used the R package *survcomp* within the Rotterdam Study.<sup>38</sup> We compared the C-indices of the newly developed models with baseline (GrimAge) using a Student t-test as described in Haibe Kans et al.<sup>39</sup> Moreover, we plotted the ROC curves at 5 and 10 years of all the models.

#### Role of funders

BBMRI-NL and BIOS contributed to the generation of the metabolomics data and the data sharing and computational resource infrastructure. The funding sources had no role in the design of this study, and did not have any role during its execution, analyses, interpretation of the data, or decision to submit results.

#### Ethics

The complete ethical statements for each cohort are available in “[Supplementary Materials: BIOS Consortium, Ethics Statements](#)”.

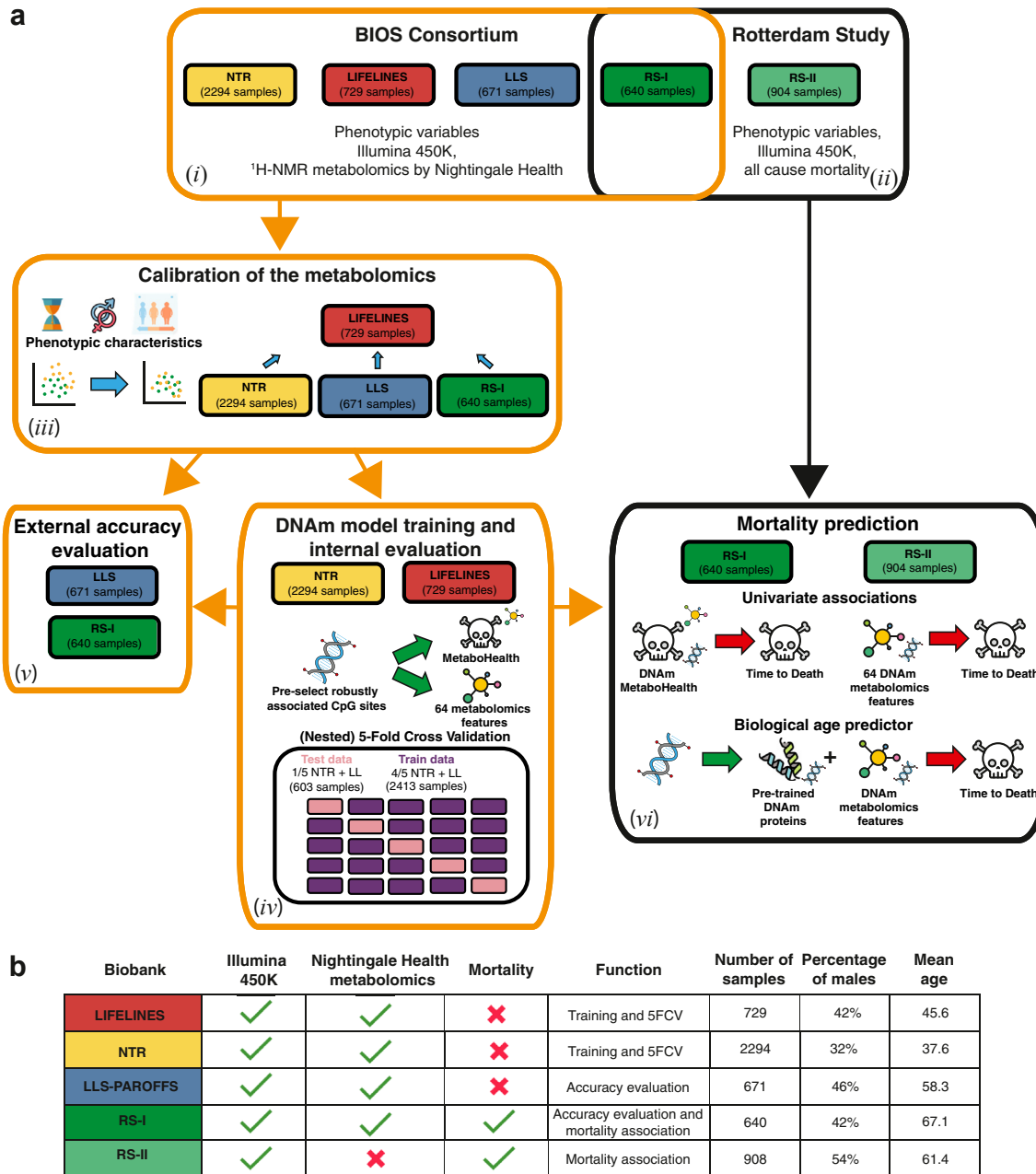
## Results

### Cross-cohort calibration of <sup>1</sup>H-NMR metabolomics data

To derive DNA methylation-based models predicting metabolic features we analysed data gathered by

partners of the BIOS consortium,<sup>17,18</sup> totalling 4334 individuals for whom both DNA methylation (Illumina 450 k) and <sup>1</sup>H-NMR Metabolomics (Nightingale Health Plc) data have been assayed. The resulting dataset had contributions of four independently collected population studies: LIFELINES-DEEP (LIFELINES), Leiden Longevity Study Partners-Offspring (LLS-PAROFFS), Rotterdam Study (RS), and Netherlands Twin Register (NTR), each with their own inclusion criteria, as reflected by differences in subject characteristics that range from the younger and leaner population of NTR (mean age = 37.57 years and mean BMI = 24.32 cm/kg<sup>2</sup>) to the older and heavier population of RS (mean age = 67.15 years and mean BMI = 27.71 cm/kg<sup>2</sup>) (Fig. 1, Table S1). A reduced dimensionality projection using a t-distributed neighbour embedding (tSNE) suggested that the interindividual variance in metabolomics data could not only be attributed to interindividual phenotypic variability but was also capturing some systematic differences between studies (Figure S3a–c). Following Mäkinen et al., we implemented a calibration technique suitable for cross-cohort harmonization, which starts with the assumption that individuals with similar phenotypic characteristics should on average exhibit similar metabolomics profiles.<sup>26</sup> For this purpose, we identified pairs of samples across cohorts with matching age, sex, and BMI, and used LIFELINES as a common reference to calibrate the other studies (Figure S2, more details in methods). A t-SNE projection of the calibrated data revealed a substantial reduction of the systematic differences between studies, as also quantified by k-BET (k-nearest neighbour Batch Effect Test) (Fig. 2a and b, Figure S3). Principal Variance Component Analyses (PVCAs) further confirmed this observation indicating that the variation attributable to study differences was attenuated, while maintaining the variation attributed to relevant biological variability (Fig. 2c).

The construction of the MetaboHealth score as published by Deelen et al.<sup>12</sup> does not include a cross-cohort calibration, but instead standardises the individual metabolic features per study prior to computation of the score (Figure S4). While this does make the MetaboHealth score more comparable across cohorts, and satisfactory for most meta-analysis purposes, it does also negate any real biological differences that may exist between studies. Conversely, when computing the MetaboHealth score on the calibrated data, i.e., after removing unwanted study differences and supplying all data as one dataset, it produced scores with interpretable differences and consistent trends across cohorts (Fig. 2d–f). For instance, consistent with our expectation, the calibrated MetaboHealth scores now tend to be higher among the studies with the older individuals RS and LLS-PAROFFS (Figure S4a). In addition, it showed a more pronounced age-associated increase in men than



**Fig. 1:** Study and methods overview. a) Study overview. (i) We employed 4334 samples, from 4 cohort of the BIOS Consortium, DNAm methylation and metabolomics to train and test our surrogates (orange border). (ii) Coupled with 1544 samples from the Rotterdam Study (black border) to evaluate their associations with mortality. (iii) We applied a calibration to harmonise all the metabolomics data, using Lifelines as a reference dataset. (iv) We then train ElasticNET models, on LIFELINES and NTR. Using the DNA methylation data, we predict two types of outcomes: 1) the pre-trained metabolomics mortality predictor (MetaboHealth), and 2) the 64 metabolic features. (v) The DNAm models are evaluated using 1) the hold-out valuation sets (LLS and RS) and 2) a 5-Fold Cross Validation on the training sets (NTR and LIFELINES). (vi) Finally, we use the DNAm models to generate surrogate metabolomics features in the RS dataset (1544 samples) and 1) evaluate their univariate associations to mortality (while correcting for age, sex), and 2) trained a complete Cox regression combining our DNAm metabolomics features and the pre-trained DNAm surrogates. b) Data availability, usage, and main phenotypic characteristics of the individuals in each cohort.



metabolomics-based mortality predictor (Fig. 1). To enforce the selection of consistent signal in different cohorts, we implement an output-specific pre-selection of consistent DNA methylation sites in the studies reserved for model development, NTR and LIFELINES (methods, Supplementary Material SM2). This Epigenome Wide Association Study (EWAS) yielded 17,705 CpG sites showing a consistent univariate association with MetaboHealth, both in direction of association and nominal significance ( $p$  value  $< 0.05$ ). Pre-selected sites were then used as input for the ElasticNET regression model predicting the MetaboHealth values (Supplementary Material SM2, Table S4). The resulting model, indicated as “DNAm-MetaboHealth” comprised ~1000 sites and showed good accuracy in the 5-Fold Cross Validation test sets (5-FCV) (median  $R \sim 0.52$ ,  $RMSE \sim 0.43$ ), which was slightly lower, but stable, in the replication sets (LLS-PAROFF:  $R \sim 0.34$ ,  $RMSE \sim 0.38$ ; RS:  $R \sim 0.33$ ,  $RMSE \sim 0.5$ ) (Fig. 3).

In parallel, we built distinct predictors for 64 metabolic features from Nightingale Health Plc, following the same training design as for DNAm MetaboHealth (Fig. 1a and Supplementary Material SM2). The resulting DNAm-based surrogates for the metabolomic features showed diverse mean accuracies over the different test sets (5-FCV, LLS-PAROFFS and RS), with 23 models being accurate (mean  $R$  across test sets  $> 0.35$ ), 20 mildly accurate ( $0.2 > \text{mean } R \text{ across test sets} \leq 0.35$ ), and 21 low accuracy models (mean  $R$  across test sets  $< 0.2$ ) (Fig. 3 and Figure S5). In the latter group we find 5 out of 8 amino acids, several LDL-related variables, all the ketone bodies and all the glycolysis related markers. The middle group is enriched with IDL related markers, 6 out of 14 fatty acids, and 2 out of 3 glycolysis related metabolites. The accurate group of DNAm-metabolomic features included, 8 out of 10 HDL-related markers, 4 out of 8 VLDL-related molecules, glycoprotein acetyls, creatinine, 3 out of 8 amino acids, and several fluid balance markers (e.g., MUFA%, Omega6%). Higher accuracies are often accompanied by a higher correlation with age (e.g., DNAm *Leucine* and *Isoleucine*) or sex (e.g., DNAm *Creatinine*) (Fig. 3, inner circles), similar to what is observed for the GrimAge DNAm based components.<sup>11</sup> Notably, only 7 of the most accurate surrogate markers were part of the 14 original metabolomic features composing the MetaboHealth score. Nevertheless, for 18 of the 23 most accurate surrogate markers, it was previously shown that the respective metabolic features significantly associated with mortality.<sup>12</sup>

### DNAm metabolomics surrogates confer a unique and relevant signal

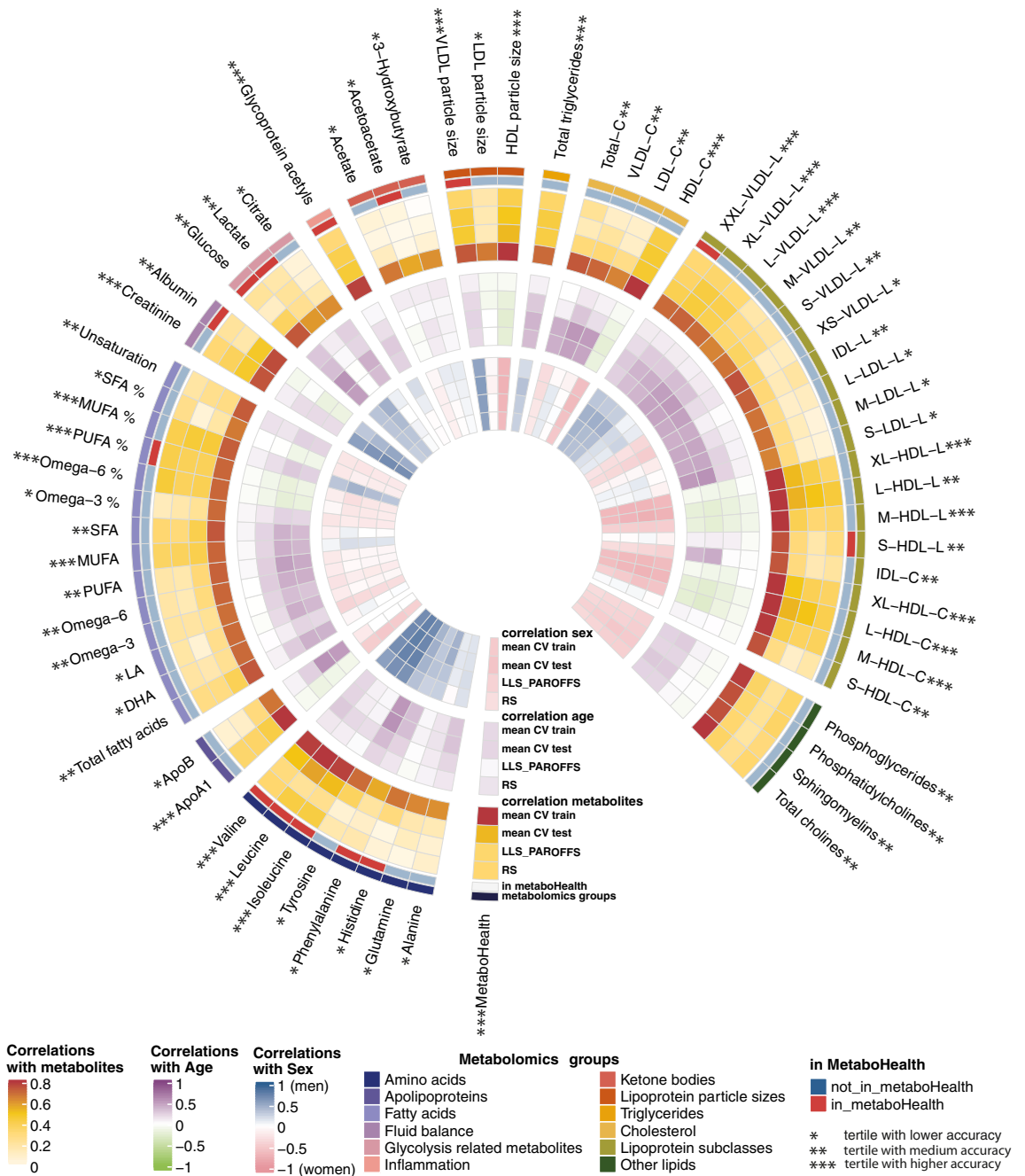
To foster the concept that DNA-methylation measurements might potentially serve as a platform to integrate biomarker signals captured from various data sources, we conducted two types of experiments. First, we ensured that the signals conveyed by our DNAm

surrogates of metabolomic features, constitute mutually independent markers, and not a multitude of highly similar signals (Figure S6a). Then, comparing the correlations between our surrogates with the original metabolomic features (Figure S6a, upper triangle), we observed a structure remarkably congruent with the correlation structure observed between the original metabolites (Figure S6a, lower triangle), albeit overall at slightly lower magnitude. This indicates that, apart from the correlation structure between the original markers, no systematic high inter correlations are observed, which thus suggests that the information in DNA methylation measurements are sufficiently rich to reconstitute many closely related biomarker signals, without introducing artificial interdependency.

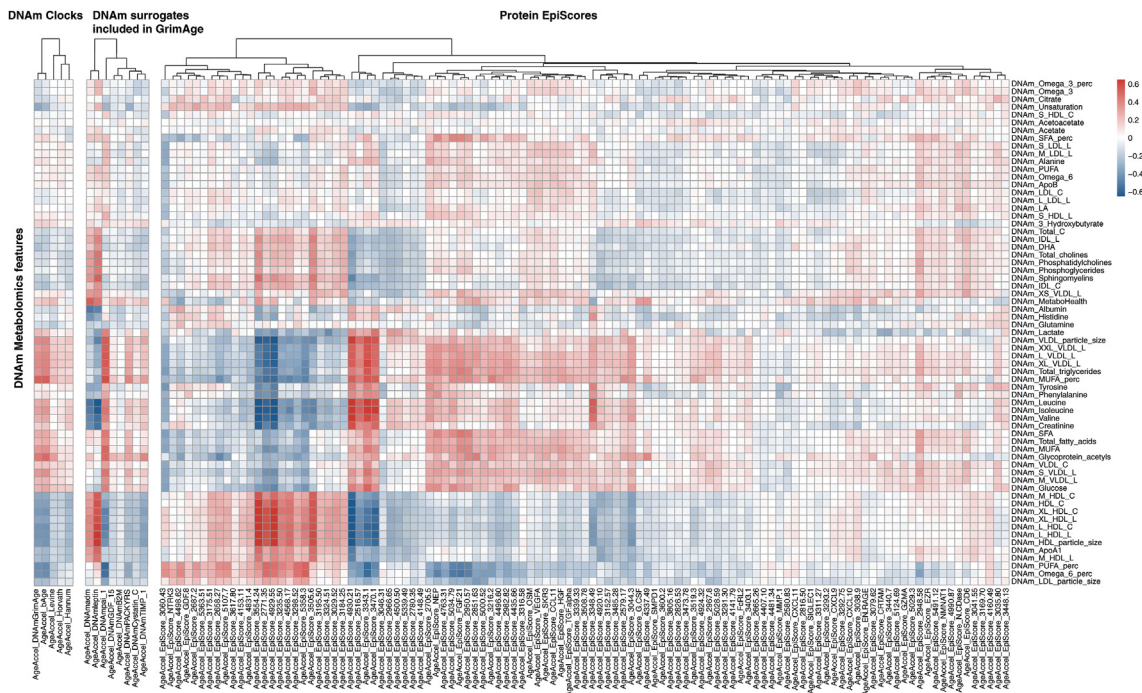
Secondly, we explored to what extent our DNAm surrogates of metabolomic features report signal compared to previously constructed DNA-methylation estimates, after regressing out age. Overall, modest correlations (mean  $|R| \sim 0.13$ , min  $|R| \sim 0.01$ , max  $|R| \sim 0.5$ ) are observed between our DNAm-based metabolomics surrogates and DNAm-based multivariate clocks (Horvath, Hannum, PhenoAge, GrimAge, and bAge) (Fig. 4). Furthermore, correlations with other pre-trained DNAm-based surrogate molecular markers (GrimAge components and the 109 protein EpiScores) are generally modest, with a few notable exceptions. Particularly, the GrimAge surrogates DNAm-leptin and DNAm-adm, and 4 EpiScores (2771.35 [Gene: IGFBP1], 4929.55 [Gene: SHBG], 3505.6 [Gene: LT $\alpha$ ], CD6), present a relatively high positive correlation with HDL related surrogate markers and a relatively high negative correlation with the amino acids (DNAm-Leucine, DNAm-Isoleucine, and DNAm-Valine). We observe the inverse pattern for the GrimAge surrogate DNAm-PAI-1, and 4 other EpiScores (4930.21 [Gene: STC1], 2516.57 [Gene: CCL21], 3343.1 [Gene: ACY1], 3470.1 [Gene: SELE]) which also show a high correlation with VLDL surrogate markers. Notably, these correlations might suggest a link between the metabolome and protein markers related to immune signalling (LT $\alpha$ , CD6, CCL21, SELE), energy balance and metabolism-related hormones (IGFBP1, SHBG, ADM, leptin, and STC1), and atherosclerosis/thrombosis inhibitor (PAI1). Nonetheless, most of the markers exhibit limited correlations with their predecessors (Fig. 4), implying the presence of previously unexplored information in DNA methylation patterns.

### Metabolomic surrogates improve the mortality predictions of GrimAge

Next, we evaluated the DNAm-metabolomics features for their predictive value for all-cause mortality in the Rotterdam Study (RS-I and RS-II). We emphasise that this whole cohort was solely employed for evaluation purposes through the analyses. For this purpose, we utilised a total of 1544 samples from this cohort (mean



**Fig. 3:** DNAm metabolites accuracies. Circular heatmap representing the accuracies of the DNAm-based models for 64 1H-NMR metabolite features by Nightingale Health and MetaboHealth. The outer ring shows the correlation between measured and DNAm-based metabolomics features, while correlation between DNAm surrogates with age and sex are shown in the middle and inner ring, respectively. Mean CV states for mean cross validation results in the cohorts LL and NTR together in the training (train) and test (test) sets, while results in the left-out set are indicated with RS and LLS. Moreover, the metabolomics features are annotated for their metabolomics group type (e.g., amino acids, fatty acids etc.) and if they were or were not included in MetaboHealth. Finally, we indicated with asterisks the tertiles of mean correlations with the measured metabolites over the test sets.



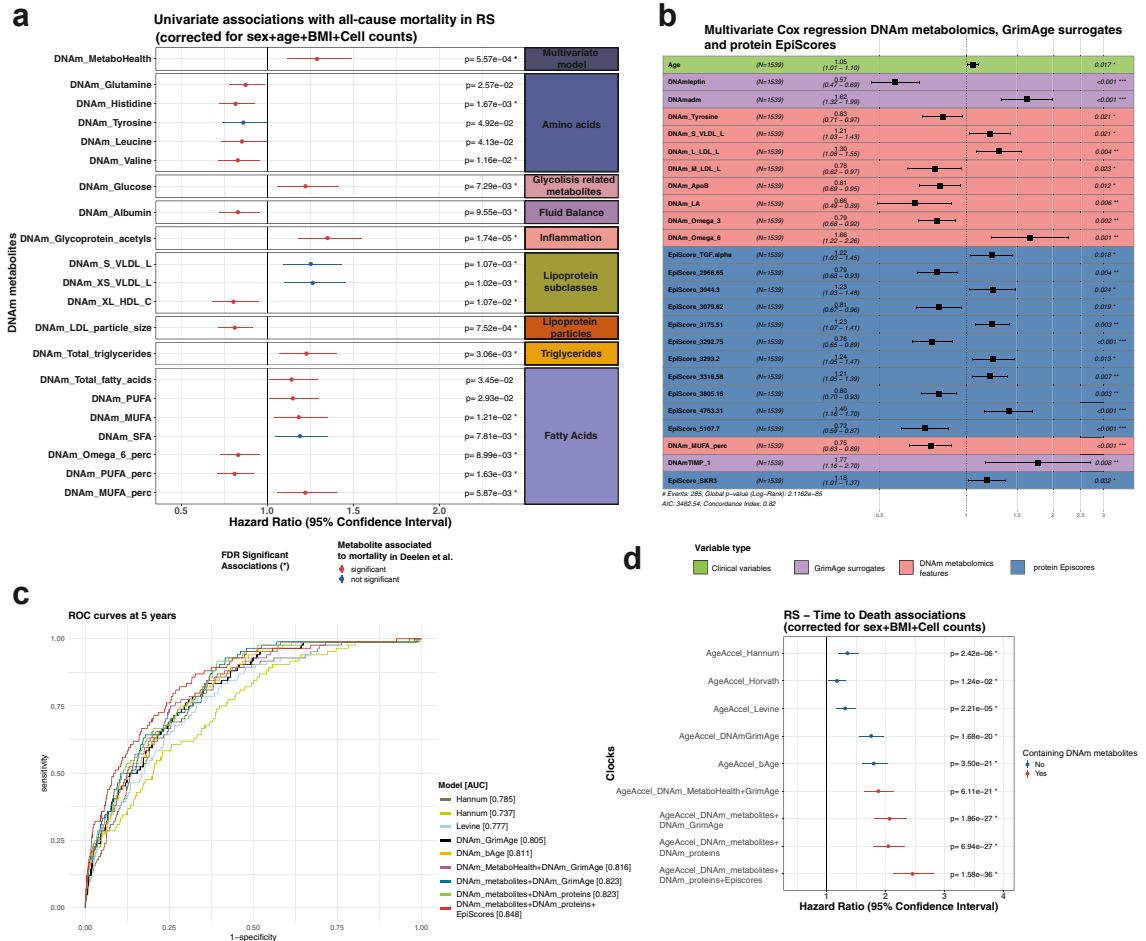
**Fig. 4:** Correlations with pre-trained DNAm scores after regressing out the ageing signal. Correlations between our DNAm metabolic features and previously trained clocks (Hannum, Horvath, PhenoAge, bAge, and GrimAge), the DNAm surrogates included in GrimAge and the 109 DNAm-based surrogates for proteins (EpiScores) by Gall et al. These correlations were calculated on the age regressed DNA methylation-based features (indicated as “AgeAccel”).

age at baseline of 64 years, 251 deceased, and a median follow-up of 11 years; Fig. 1), by incorporating an additional 863 samples with available Illumina 450 k and mortality information, but not <sup>1</sup>H-NMR metabolomics. In accordance with previous studies,<sup>11,14,15</sup> we assess univariate Cox proportional hazard models (in years of follow up) adjusted for relevant covariates, specifically sex, together with age, BMI and cell counts levels at blood sampling (Fig. 5A and Figure S7A).

First, we evaluated our *DNAm-MetaboHealth* predictor, which showed a significant association with all-cause mortality (HR = 1.29, p = 5.57 × 10<sup>-04</sup> [cox regression]) (Fig. 5a), in line with the original metabolomics-based *MetaboHealth* score (Figure S8C). Next, we evaluated the individual DNAm metabolomics features and observed significant associations with mortality for 15 out of the 64 surrogate metabolites, of which 6 (out of 14) metabolomics features were included in the original *MetaboHealth* score. In addition, our estimated effects were overall consistent with those found by the study performed by Deelen et al. which employed a considerably larger dataset of 44,168 individuals (Supplemental S7c), with our most significant findings being amongst their strongest effects. We observed an increased risk for higher estimates of 8 DNAm-based features, with the strongest being *DNAm-Glucose* (HR = 1.22, p = 7.29 × 10<sup>-03</sup> [cox regression]),

*DNAm-Glycoprotein Acetyls* (HR = 1.35, p = 1.74 × 10<sup>-05</sup> [cox regression]). Conversely, we observed protective effects for 7 features, such as *DNAm-PUFA%* (HR = 0.81, p = 1.63 × 10<sup>-03</sup> [cox regression]), *DNAm-Histidine* (HR = 0.82, p = 1.67 × 10<sup>-03</sup> [cox regression]), *DNAm-Valine* (HR = 0.83, p = 1.6 × 10<sup>-02</sup> [cox regression]), and *DNAm-Albumin* (HR = 0.83, p = 9.5 × 10<sup>-03</sup> [cox regression]). In addition, 5 nominal significant additional metabolites showed discordant mortality associations between sexes. Explicitly, for males we observed mortality associations with DNAm-Glutamine, whereas DNAm-Tyrosine, DNAm-Leucine, DNAm-Total Fatty acids, and DNAm-PUFA associated with mortality in women (Figure S7B).

Almost all the pre-trained DNAm clocks that we considered (Hannum, Horvath, PhenoAge, GrimAge) and some of their intermediate surrogates exert mortality associations in RS (Figure S9a–c). Next, we attempted to refine the current standard for biological age estimation, specifically GrimAge (CI = 0.79, p = 4.6 × 10<sup>-77</sup> [cox regression]), by training a multivariate all-cause mortality predictor including our DNAm metabolomics features. As a first exploration, we trained a Cox regression model with age at blood sampling, sex, DNAm-GrimAge and DNAm-MetaboHealth, which only showed minor, but significant, improvements in the C-index (CI = 0.8, p = 4.6 × 10<sup>-77</sup> [cox regression]) (Figure S10b). As



**Fig. 5:** Associations with time to death. a) Significant univariate associations of the DNAm metabolomics features with time to all-cause mortality in RS (N = 1542 with 285 reported deaths). The associations are grouped based on the metabolomics groups coloured by the significant associations or the metabolites with mortality in Deelen et al. The asterisks (\*) separates nominal significant DNAm metabolomics features from the FDR significant ones [cox regression]. b) Stepwise Cox regression predicting of time to all-cause mortality optimised in RS, composed combining age, 3 DNAm surrogates included in GrimAge, and 9 DNAm metabolic models and 12 protein EpiScores. Finally, c) presents the ROC curves and the accuracies (AUC) at 5-years mortality for our newly developed clocks as compared to previously trained scores. Similarly, d) reports the Hazard Ratios for each of the Age accelerated multivariate scores, regressing also for sex, BMI and cell counts. In this last figure the scores as split for containing our DNAm metabolomics features (red) or not (blue).

a second exploration, we performed a stepwise (backward/forward) Cox regression model to identify a minimal set of features including age, sex, our 64 DNAm-metabolomic features and DNAm-GrimAge (CI = 0.81,  $p = 1.7 \times 10^{-83}$  [cox regression]) (Figure S10d). Nonetheless, the best performing model was obtained when including age, sex, 3 out of 8 GrimAge components (predicting Leptin and ADM and TIMP\_1), 12 of the 109 protein EpiScores, and 9 out of 64 DNAm metabolites (CI = 0.82,  $p = 1 \times 10^{-85}$  [cox regression]) (Fig. 5b). The selected DNAm metabolic features included *DNAm-Tyrosine*, *DNAm-S-VLDL-L*, *DNAm-S-LDL-L*, *DNAm-M-LDL-L*, *DNAm-APOB*, *DNAm-LA*, *DNAm-omega3* and *DNAm-omega6*, and *DNAm-MUFA*. Analyses of Schoenfeld residuals highlighted no violation

of the proportional hazard assumption (Figure S14). In any case, all the newly introduced scores exhibited a significantly improved C-index (Figure S10e)<sup>39</sup> and a higher AUC at 5 and 10 years compared to the GrimAge (Fig. 5c and d). Overall, this indicates that DNAm-surrogates from different origin, phenotypic, proteomic, or metabolomic, might confer mutually independent information for mortality prediction.

**DNAm metabolomics models introduce relevant CpG selections**

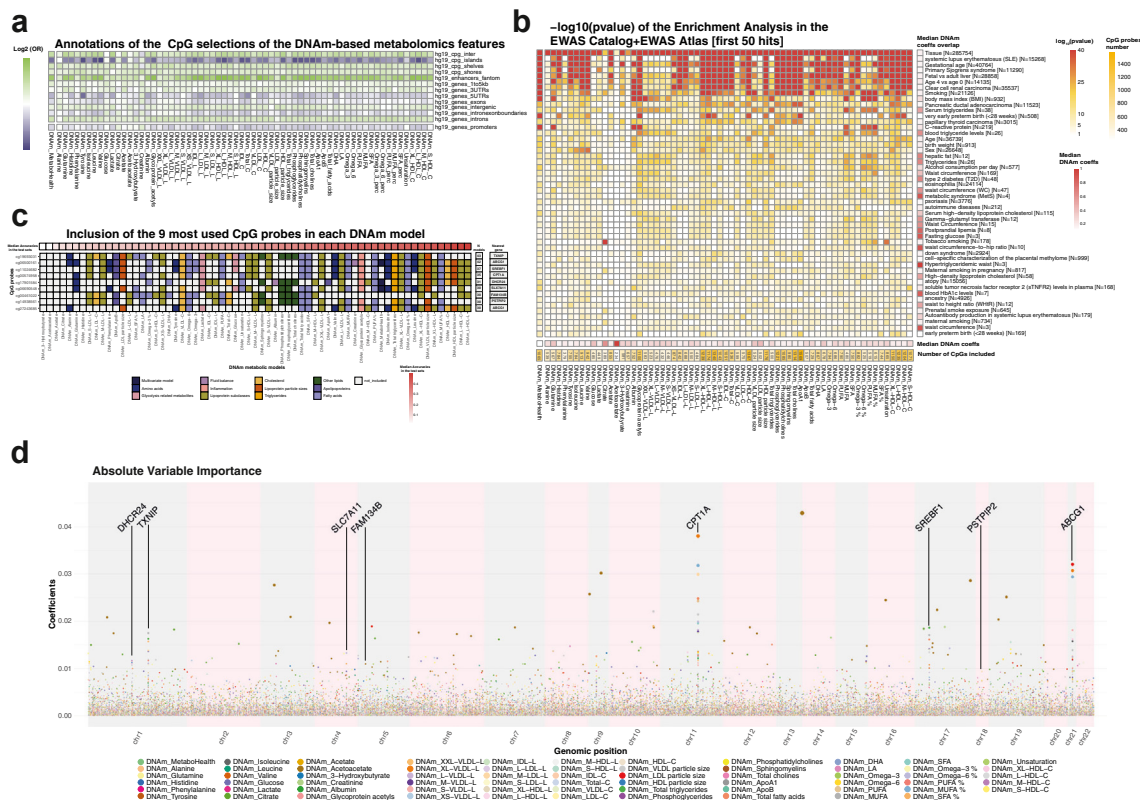
After establishing the value of our DNAm metabolomics features in predicting mortality, we explored the nature of the signal included in our models by investigating the

CpG sites picked by the ElasticNET regression, which can shrink contributions of unnecessary features to zero. Predictors selected a median of ~750 CpG-sites, with a minimum of 234 CpG-sites for *DNAm-Acetoacetate* and a maximum of 1569 for *DNAm-ApoA1* (Fig. 6B, and Table S4). A total of 22,145 probes were included in at least 1 model. Comparison of the genomic positions of the selected CpG-sites with the rest of the 450 k array highlighted an underrepresentation of probes positioned in CpG Islands, and a preferential selection for CpG shelves and shores, known to be more dynamic areas (Fig. 6a).<sup>40,41</sup> Noteworthy is the higher tendency to select CpGs co-locating with enhancers, cis-acting short regions of DNA that control the temporal and cell-specific activation of gene expression (Fig. 6a).<sup>37</sup>

Functional enrichment analyses using the most proximal genes to the selected CpG-sites highlighted pathways associated to “developmental processes”, “cell differentiation”, and “regulation of metabolic processes”

from Biological Processes in Gene Ontology (Figure S11c). Concomitantly, enrichment analyses of phenotypic annotations in the EWAS Catalog and EWAS Atlas (Fig. 6b), indicated that the CpG-sites are known to be largely related to peripheral tissue differentiation,<sup>42</sup> foetal brain development<sup>43</sup> and gestational age.<sup>44</sup> Nonetheless, the CpG sites with the highest median coefficients across all our models were the ones annotated for metabolite-related traits, such as “Triglycerides”, and “Fasting Glucose” (Fig. 6b, Figure S11a and b). In total, 203 traits exhibited a significant enrichment for the CpG selections made by our models. Notably we find also highly significant associations with “Ageing”, and “all-cause mortality”, indicating that we do identify CpGs related to age-related processes.

Despite their interesting overarching signal, the DNAm-metabolic models show little overlap with each other in their CpG selections, with the majority showing overlaps well below 15%, apart for a few



**Fig. 6:** CpG selections of the ElasticNET models. a) Log2 Odds ratio indicating the enriched in annotations of the CpG by our ElasticNET models. b) The central heatmap reports the log 10 p values [Fisher test] of the enrichments CpG sites selected by our models (rows) and the 50 most significant traits in the EWAS Catalog and Atlas enriched (rows). Bottom: the median coefficients in each DNAm model, and the number of CpGs per model. Right: the median coefficients given by our DNAm model to the overlapping CpGs with each trait. c) The nine most used probes (rows) over the 65 ElasticNET models (columns), coloured by metabolic groups. Top: The models were ordered by the mean accuracy over the test sets (CV, LLS, and RS). Right: The number of models which include each CpG and their nearest genes. d) Manhattan plot-like figure indicating the Variable importance of the single CpG probes in the DNAm metabolic models.

exceptions of highly correlated metabolites (e.g., 83% between *DNAm-Total\_cholines* and *DNAm-phosphoglycerides*; Figures S6 and S12). Nonetheless, a handful of CpG probes were chosen in more than 30 models with largely consistent coefficient signs (Fig. 6c). Interestingly, while some of these 9 features have a higher importance weight on the DNAm metabolomics models (e.g., cg00574958, or cg06500161), others only exert a more minor influence (e.g., cg14938561, cg00461022). The 9 CpG sites with higher importance weight don't favour one specific metabolic group but seems to be relevant to many metabolic markers (Fig. 6d). Not surprisingly, also the nearest genes to these 9 probes are noteworthy. For instance, *TXNIP*, which includes cg19693031 (chosen in 43 DNAm-metabolomics models), was previously associated to hyperglycaemia and insulin resistance, and *ABCG1*, nearby cg06500161 (in 42 DNAm-metabolomics models) was associated to plasma lipid levels and stroke (Fig. 6d).

## Discussion

A comprehensive quantification of biological ageing, as a way to assess the overall, holistic health status and disease susceptibility of individuals,<sup>14</sup> would constitute a major advance for healthcare and preventive research. A diversity of molecular markers has been proposed as indicators of biological age relating to health- and lifespan. Here we integrated well-established DNA methylation-based and <sup>1</sup>H-NMR metabolomics resources for biological age prediction with mortality as a primary endpoint. To our knowledge, the potential synergistic effects arising from combining these two molecular sources remained thus far largely unexplored, and we believe that a collection of models predicting metabolomics features may be relevant within the rapidly growing repertoire of DNA methylation-based estimates.<sup>10,11,15,45</sup> A structured training and evaluation design aided us to demonstrate the robustness of our features. We highlighted the distinct signal expressed by our models and their feature selection. Finally, we explored the use of our DNAm-based surrogates of metabolomics features in combination with previously trained DNAm-based surrogates (e.g., GrimAge constituents) suggesting that these confer complementary information.

We applied ElasticNET regression, a widely utilised algorithm to train epigenetic clocks, to data of four large population cohorts to derive DNAm-based surrogates for a previously derived multi-analyte score indicating mortality (MetaboHealth), and for 64 individual metabolomics features. The direct estimation of metabolomics-based mortality by constructing a DNAm surrogate for the MetaboHealth score showed promising results (mean R in test-sets = 0.397). Moreover, we were able to construct DNAm surrogates for many, but

not all, metabolomics features with good replication accuracies (mean R in test-sets >0.35), including health markers for HDL and VLDL metabolism, inflammation, and fluid balance. Less accurate were the DNAm surrogates for amino-acids, ketone bodies, glycolysis, and LDL-related markers (mean R in test-sets <0.2). Nevertheless, considering the limited number of available markers and the low accuracy thresholds previously used for DNAm scores (R >0.1 in test sets),<sup>15,45</sup> we continued evaluating all 65 models. This decision was further corroborated by a previous report by Stevenson et al. who suggested that their DNAm surrogate for CRP was a more reliable indication of chronic inflammation than its measured counterpart, even when considering the modest correlation between CRP and its surrogate.<sup>46</sup> Overall, our DNAm metabolomic features conveyed a signal coherent with the quantified metabolomics variables and independent from most of the previously reported DNA methylation-based clocks and molecular surrogates.

Great emphasis was given to the harmonization of metabolomic data collected across different cohorts, prior to training our DNAm-based models for individual metabolites or the MetaboHealth score. Non-biological variability that may originate from inter cohort differences in sample collection, storage, or handling could confound model training. Typically, this challenge in epidemiology is addressed by applying a z-scaling per cohort prior to conducting a meta-analysis, which in effect discards all differences, both technical and biological, between cohorts. In other words, while allowing to draw conclusions on the similarities in associations with endpoints between cohorts, this strategy does not allow for a direct comparison of the underlying molecular profiles between cohorts. To address this issue we applied a calibration technique, which we developed adapting methodologies previously applied in longitudinal studies.<sup>26</sup> This calibration technique showed its merit in harmonizing the metabolomics profiles, while preserving the natural biological heterogeneity within and between the different study populations. Importantly, this approach allowed for an evaluation of the MetaboHealth score across cohorts, showing consistent age and sex specific trends per study, and global predictive power for established clinical variables, such as hsCRP and diabetes.

Previous studies have shown advantages of pre-selecting CpGs when training ElasticNET regression models.<sup>16,47–49</sup> Following this example, we implemented a pre-selection of CpG sites showing a high variability and consistent association with the outcome of interest during the training phase of our 5-Fold Cross Validation procedure. During this selection process, we avoided imposing stringent criteria (it was based on nominal p value) and refrained from correcting for covariates to prevent an excessive loss of potentially meaningful signal. Approximately 22,000 CpG sites were included in at least one DNAm-based models. Enrichment

analyses showed that the selected CpGs are more likely to be enhancers in CpG shelves and shores and are in the proximity of genes enriched for regulation of metabolic and developmental processes, or cell differentiation. This finding resonates with a longstanding hypothesis, that the ageing methylome reflects processes underlying intricate cellular and molecular changes linked with development and differentiation.<sup>50</sup> Furthermore, CpG sites selected for our surrogates were also previously associated to age (e.g., Ageing, all-cause mortality), inflammatory (C-reactive proteins), or metabolically related traits (e.g., triglycerides and metabolic syndrome). Strikingly, we found a highly recurrent selection of 9 CpGs in at least 30 distinct DNAm surrogate models, suggesting that these CpGs form a fundamental link between the blood metabolome and DNA methylome. All these loci have been previously found associated with metabolic traits and processes,<sup>51</sup> and most of these 9 CpGs and their nearest genes are considered powerful classifiers for diabetes stratification.<sup>52–54</sup> Remarkably, 3 of these 9 CpG probes showed significant univariate association with mortality within the Rotterdam Study (Figure S8D). This reassures over the valuable cardiometabolic content latent in our DNAm models.

Besides, our main intent was to evaluate the possibility to extrapolate the mortality signal from the metabolome to DNA methylation. To do so, we tested which of our surrogates might be indicative of all-cause mortality in a subset of the Rotterdam Study (1544 persons, 285 deaths). Notably, we observed a successful detection, albeit partial, of the mortality signal exerted by the metabolomics platform. We could successfully derive a DNAm-based version of MetaboHealth, which significantly associates with all-cause mortality, although it showed a lower hazard ratio than the original score.<sup>14</sup> This might in part be explained by the fact that only 6 of the 14 DNAm surrogates for the metabolites constituting the MetaboHealth showed associations with all-cause mortality. Overall, we observed significant associations with mortality for 15 out of 64 DNAm-based metabolites. The detected effects are consistent with the results previously reported by Deelen et al. in a large study using the original metabolomic features measured in 44,168 individuals. This consistency further underpins that DNAm surrogates for metabolomic features could potentially be leveraged as epigenetic markers of biological ageing.

To further explore this concept, we trained a multivariate model for all-cause mortality, that was allowed to select from all available DNAm surrogates using a stepwise forward/backward regression. This final model included 9 DNAm metabolomic features together with the competing covariates age, 3 GrimAge components and 12 plasma protein EpiScores. The resulting model combining DNAm surrogates from different origin showed a significantly improved mortality prediction

(C-index = 0.82) compared to the GrimAge score (C-index = 0.79) (Fig. 5 and Figure S10). Our composite scores showed a substantial refinement of the AUC at 5 and 10 compared to the original GrimAge (Figure S10G and H). Overall, this suggests that a broader collection of DNAm-surrogates of independent origin, such as proteomics, phenotypes, and now also metabolomics, might confer a more comprehensive indication on epigenetic-based biological ageing.

An important limitation of the current study for leveraging mortality signals is its limited sample size, which is modest when compared to the large dataset that Deelen et al. employed to evaluate the mortality associations of the metabolomics features and to build a multi-analyte predictor for mortality. Despite the limited power, we found significant associations with mortality for the DNAm surrogates of the multi-analyte score MetaboHealth and 15 individual metabolic features, which were consistent with those observed by Deelen et al. A second limitation consists in the inclusion of only Dutch population cohorts, which does not ensure a correct replication in other populations. A third limitation is the usage of a single endpoint, mortality, for evaluating the potential applications of our DNAm surrogates as marker for biological age. We acknowledge that ageing and its associated decline in overall health is a complex multi-factorial process, that is only partially captured by mortality risk. Previous work reported the merits of the <sup>1</sup>H-NMR metabolomics in estimating several different types of endpoints,<sup>7,13,55,56</sup> or even end-of-life related-phenotypes such as frailty,<sup>14</sup> leading us to speculate that our DNAm surrogates for metabolomic features might also be instrumental for capturing these ageing endophenotypes.

In conclusion, we have demonstrated that metabolite markers previously associated with mortality could be leveraged to help extract the mortality signal captured by the DNA methylation platforms. Moreover, we showed that our DNAm surrogates capture mortality signal that is independent of the mortality signal captured by previous DNAm scores, such as GrimAge or its separate DNAm surrogate constituents. Overall, this does suggest that even more mortality signal could be extracted given the availability of proper mortality-associated biomarkers.

#### Contributors

EBvdA, DB, MJTR and PES conceived and wrote the manuscript. DB performed the analyses. EBvdA and MJTR verified and supervised the analyses. PES, MB, JBJvM, JvD, DIB, RP, MG, LF were involved in data acquisition of the cohort data. All authors (DB, MJTR, LMK, MB, JD, JBJvM, JvD, RP, DIB, MG, LF, PES, EBvdA) discussed the results read, and approved the final version of the manuscript. Data used in this manuscript was mostly funded by the BBMRI-NL Metabolomics Consortium.

#### Data sharing statement

BBMRI-nl and BIOS-nl data are available upon request at <https://www.bbMRI.nl/services/samples-images-data>. All DNAm metabolomics scores can be obtained with a script at: [https://github.com/DanieleBizzarri/DNAM\\_metabolomics\\_scores](https://github.com/DanieleBizzarri/DNAM_metabolomics_scores).

## Declaration of interests

Authors declare no competing interests.

## Acknowledgements

This work was performed within the BBMRI Metabolomics Consortium funded by: BBMRI-NL (financed by NWO 184.021.007 and 184.033.111), X-omics (NWO 184.034.019), VOILA (ZonMW 457001001) and Medical Delta (METABODELTA: Metabolomics for clinical advances in the Medical Delta). EBvdA is funded by a personal grant of the Dutch Research Council (NWO; VENI: 09150161810095). Acknowledgements for all contributing studies can be found in the Supplementary Material-BIOS Consortium. Additional NTR samples were funded by the European Research Council (ERC-230374) project Genetics of Mental Illness (DIB).

## Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.ebiom.2024.105279>.

## References

- López-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of aging: an expanding universe. *Cell*. 2023;186. <https://doi.org/10.1016/j.cell.2022.11.001>.
- Partridge L, Deelen J, Slagboom PE. Facing up to the global challenges of ageing. *Nature*. 2018;561:45–56.
- Comfort A. Test-Battery to measure ageing-rate in man. *Lancet*. 1969;294:1411–1415.
- Blackburn EH, Greider CW, Szostak JW. Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. *Nat Med*. 2006;12:1133–1138.
- Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14:R115.
- Peters MJ, Joehanes R, Pilling LC, et al. The transcriptional landscape of age in human peripheral blood. *Nat Commun*. 2015;6:8570.
- van den Akker EB, Trompet S, Barkey Wolf JJH, et al. Metabolic age based on the BBMRI-NL 1H-nmr metabolomics repository as biomarker of age-related disease. *Circ Genom Precis Med*. 2020;13(5): 541–547. <https://doi.org/10.1161/CIRCGEN.119.002610>.
- Menni C, Kiddle SJ, Mangino M, et al. Circulating proteomic signatures of chronological age. *J Gerontol A Biol Sci Med Sci*. 2015;70:809–816.
- Zhang Q, Vallerga CL, Walker RM, et al. Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. *Genome Med*. 2019;11:54.
- Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Ageing (Albany NY)*. 2018;10:573–591.
- Lu AT, Quach A, Wilson JG, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Ageing (Albany NY)*. 2019;11:303–327.
- Deelen J, Kettunen J, Fischer K, et al. A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. *Nat Commun*. 2019;10:1–8.
- Nightingale Health UK Biobank Initiative, Julkunen H, Cichońska A, Slagboom PE, Würtz P. Metabolic biomarker profiling for identification of susceptibility to severe pneumonia and COVID-19 in the general population. *Elife*. 2021;10:e63033.
- Kuiper LM, Polinder-Bos HA, Bizzarri D, et al. Epigenetic and metabolomic biomarkers for biological age: a comparative analysis of mortality and frailty risk. *J Gerontol A Biol Sci Med Sci*. 2023;78:1753–1762.
- Gadd DA, Hillary RF, McCartney DL, et al. Epigenetic scores for the circulating proteome as tools for disease prediction. *Elife*. 2022;11:e71802.
- Bernabeu E, McCartney DL, Gadd DA, et al. Refining epigenetic prediction of chronological and biological age. *Genome Med*. 2023;15:12.
- Zhernakova DV, Deelen P, Vermaat M, et al. Identification of context-dependent expression quantitative trait loci in whole blood. *Nat Genet*. 2017;49:139–145.
- Bonder MJ, Luijk R, Zhernakova DV, et al. Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet*. 2017;49:131–138.
- van Dongen J, Gordon SD, McRae AF, et al. Identical twins carry a persistent epigenetic signature of early genome programming. *Nat Commun*. 2021;12:5618.
- Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiovasc Genet*. 2015;8:192–206.
- Würtz P, Kangas AJ, Soininen P, Lawlor DA, Davey Smith G, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in large-scale epidemiology: a primer on -omic technologies. *Am J Epidemiol*. 2017;186:1084–1096.
- Bizzarri D, Reinders MJT, Beekman M, Slagboom PE, Bbmri-NL null, van den Akker EB. 1H-NMR metabolomics-based surrogates to impute common clinical risk factors and endpoints. *eBioMedicine*. 2022;75:103764.
- van Iterson M, Tobi EW, Sliker RC, et al. MethylAid: visual and interactive quality control of large Illumina 450k datasets. *Bioinformatics*. 2014;30:3435–3437.
- Hastie T, Tibshirani R, Narasimhan B, Chu G. *impute: impute: imputation for microarray data*. 2023. <https://doi.org/10.18129/B9.bioc.impute>.
- Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res*. 2017;45:e22.
- Mäkinen V-P, Karsikas M, Kettunen J, et al. Longitudinal profiling of metabolic ageing trends in two population cohorts of young adults. *Int J Epidemiol*. 2022;51:1970–1983.
- Telle-Hansen VH, Christensen JJ, Formo GA, Holven KB, Ulven SM. A comprehensive metabolic profiling of the metabolically healthy obesity phenotype. *Lipids Health Dis*. 2020;19:90.
- Ala-Korpela M, Lehtimäki T, Kähönen M, et al. Cross-sectionally calculated metabolic aging does not relate to longitudinal metabolic changes-support for stratified aging models. *J Clin Endocrinol Metab*. 2023;108:2099–2104.
- Büttner M, Miao Z, Wolf FA, Teichmann SA, Theis FJ. A test metric for assessing single-cell RNA-seq batch correction. *Nat Methods*. 2019;16:43–49.
- Li J, Bushel PR, Chu T-M, Wolfinger RD. Principal variance components analysis: estimating batch effects in microarray gene expression data. In: *Batch effects and noise in microarray experiments*. John Wiley & Sons, Ltd; 2009:141–154.
- Bizzarri D, Reinders MJT, Beekman M, Slagboom PE, van den Akker EB. MiMIR: R-shiny application to infer risk factors and endpoints from Nightingale Health's 1H-NMR metabolomics data. *Bioinformatics*. 2022;38:3847–3849.
- Pelegi-Sisó D, de Prado P, Ronkainen J, Bustamante M, González JR. methylclock: a bioconductor package to estimate DNA methylation age. *Bioinformatics*. 2021;37:1759–1760.
- Mukherjee S, Stamatis D, Bertsch J, et al. Genomes OnLine Database (GOLD) v.8: overview and updates. *Nucleic Acids Res*. 2021;49:D723–D733.
- Higgins-Chen AT, Thrush KL, Wang Y, et al. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. *Nat Aging*. 2022;2:644–661.
- Batram T, Yousefi P, Crawford G, et al. The EWAS catalog: a database of epigenome-wide association studies. *Wellcome Open Res*. 2022;7:41.
- Xiong Z, Yang F, Li M, et al. EWAS open platform: integrated data, knowledge and toolkit for epigenome-wide association study. *Nucleic Acids Res*. 2022;50:D1004–D1009.
- Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. *Nature*. 2014;507: 455–461.
- Schröder MS, Culhane AC, Quackenbush J, Haibe-Kains B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics*. 2011;27:3206–3208.
- Haibe-Kains B, Desmedt C, Sotiriou C, Bontempo G. A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all? *Bioinformatics*. 2008;24:2200–2208.
- Ziller MJ, Gu H, Müller F, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature*. 2013;500: 477–481.
- Irizarry RA, Ladd-Acosta C, Wen B, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet*. 2009;41:178–186.

- 42 Islam SA, Goodman SJ, MacIsaac JL, et al. Integration of DNA methylation patterns and genetic variation in human pediatric tissues help inform EWAS design and interpretation. *Epigenetics Chromatin*. 2019;12:1.
- 43 Spiers H, Hannon E, Schalkwyk LC, et al. Methylomic trajectories across human fetal brain development. *Genome Res*. 2015;25:338–352.
- 44 Bohlin J, Håberg SE, Magnus P, et al. Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biol*. 2016;17:207.
- 45 Chen Q, Dwaraka VB, Carreras-Gallo N, et al. OMICmAge: an integrative multi-omics approach to quantify biological age with electronic medical records. *bioRxiv*. 2023. <https://doi.org/10.1101/2023.10.16.562114>.
- 46 Stevenson AJ, McCartney DL, Hillary RF, et al. Characterisation of an inflammation-related epigenetic score and its association with cognitive ability. *Clin Epigenetics*. 2020;12:113.
- 47 Choi H, Joe S, Nam H. Development of tissue-specific age predictors using DNA methylation data. *Genes (Basel)*. 2019;10:888.
- 48 Bergersen LC, Ahmed I, Frigessi A, Glad IK, Richardson S. Pre-selection in lasso-type analysis for ultra-high dimensional genomic exploration. In: Frigessi A, Bühlmann P, Glad IK, Langaas M, Richardson S, Vannucci M, eds. *Statistical analysis for high-dimensional data*. Cham: Springer International Publishing; 2016:37–66.
- 49 Croiseau P, Legarra A, Guillaume F, et al. Fine tuning genomic evaluations in dairy cattle through SNP pre-selection with the Elastic-Net algorithm. *Genet Res*. 2011;93:409–417.
- 50 Seale K, Horvath S, Teschendorff A, Eynon N, Voisin S. Making sense of the ageing methylome. *Nat Rev Genet*. 2022;23:585–605.
- 51 Gomez-Alonso MDC, Kretschmer A, Wilson R, et al. DNA methylation and lipid metabolism: an EWAS of 226 metabolic measures. *Clin Epigenetics*. 2021;13:7.
- 52 Soriano-Tárraga C, Jiménez-Conde J, Giralt-Steinhauer E, et al. Epigenome-wide association study identifies TXNIP gene associated with type 2 diabetes mellitus and sustained hyperglycemia. *Hum Mol Genet*. 2016;25:609–619.
- 53 Krause C, Sievert H, Geißler C, et al. Critical evaluation of the DNA-methylation markers ABCG1 and SREBF1 for Type 2 diabetes stratification. *Epigenomics*. 2019;11:885–897.
- 54 Lai C-Q, Parnell LD, Smith CE, et al. Carbohydrate and fat intake associated with risk of metabolic diseases through epigenetics of CPT1A. *Am J Clin Nutr*. 2020;112:1200–1211.
- 55 Buergel T, Steinfeldt J, Ruyoga G, et al. Metabolomic profiles predict individual multidisease outcomes. *Nat Med*. 2022;28:2309–2320.
- 56 Ahola-Olli AV, Mustelin L, Kalimeri M, et al. Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four finnish cohorts. *Diabetologia*. 2019;62:2298–2309.