

1 ¹H-NMR metabolomics-guided DNA methylation mortality 2 predictors

3 D. Bizzarri^{1,2,3}, M.J.T. Reinders^{2,3}, L.M. Kuiper^{4,5}, M. Beekman¹, J. Deelen^{1,6,7}, J.B.J. van
4 Meurs^{4,8}, J. van Dongen^{9,10,11}, R. Pool^{9,11}, D.I. Boomsma^{9,10,11}, M. Ghanbari¹², L. Franke¹³,
5 BIOS Consortium¹⁴, BBMRI-NL Consortium¹⁵, P.E. Slagboom^{1,6}, and E.B. van den
6 Akker^{1,2,3, #}

7 ¹ Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center,
8 Leiden, The Netherlands

9 ² Leiden Computational Biology Center, Department of Biomedical Data Sciences, Leiden University Medical
10 Center, Leiden, The Netherlands

11 ³ Delft Bioinformatics Lab, TU Delft, Delft, The Netherlands

12 ⁴ Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands

13 ⁵ Center for Nutrition, Prevention and Health Services, National Institute for Public Health and Environment
14 (RIVM), Bilthoven, the Netherlands

15 ⁶ Max Planck Institute for the Biology of Ageing, Cologne, Germany

16 ⁷ Cologne Excellence Cluster on Cellular Stress responses in Aging Associated Diseases, University of
17 Cologne,

18 Cologne, Germany

19 ⁸ Department of Orthopaedics & Sports, Erasmus Medical Center, Rotterdam, the Netherlands

20 ⁹ Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.

21 ¹⁰ Amsterdam Reproduction and Development (AR&D) Research Institute, Amsterdam, The Netherlands.

22 ¹¹ Amsterdam Public Health Research Institute, Amsterdam, The Netherlands.

23 ¹² Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands

24 ¹³ Department of Genetics, University Medical Center Groningen, Groningen, The Netherlands

25 ¹⁴ BIOS Consortium, see Consortium Banner Supplement

26 ¹⁵ BBMRI-NL Consortium, see Consortium Banner Supplement

27

28 **Abstract**

29 ¹H-NMR metabolomics and DNA methylation in blood are widely known biomarkers
30 predicting age-related physiological decline and mortality yet exert mutually independent
31 mortality and frailty signals. Leveraging multi-omics data in four Dutch population studies
32 (N=5238) we investigated whether the mortality signal captured by ¹H-NMR metabolomics
33 could guide the construction of novel DNA methylation-based mortality predictors. Hence, we
34 trained DNA methylation-based surrogates for 64 metabolomic analytes and found that
35 analytes marking inflammation, fluid balance, or HDL/VLDL metabolism could be accurately
36 reconstructed using DNA-methylation assays. Interestingly, a previously reported multi-analyte
37 score indicating mortality risk (MetaboHealth) could also be accurately reconstructed. Sixteen
38 of our derived surrogates, including the MetaboHealth surrogate, showed significant
39 associations with mortality, independent of other relevant covariates. Finally, adding our novel
40 surrogates to previously established DNA-methylation markers, such as GrimAge, showed
41 significant improvement for predicting all-cause mortality, indicating that our metabolic
42 analyte-derived surrogates potentially represent novel mortality signal.

43

44 **Introduction**

45 A common goal in geroscience is to identify mechanisms that drive ageing and design
46 interventions that might slow down or even reverse the rate of ageing [1]. For this purpose, it is
47 essential to have indicators not only quantifying ageing, but simultaneously marking the
48 trajectory of overall health decline [2]. While calendar age is a core risk factor for almost any
49 common disease, it has many limitations for capturing the variability in health-span. Crucially,
50 calendar age does not capture the effects of an individual's lifestyle, nor incorporates readouts
51 of functional decline. Instead, faithful markers of biological age would allow to quantify the
52 vulnerability to disease irrespective of an individual's calendar age, and to develop and
53 monitor effective healthy lifestyle advices and anti-aging interventions. The earliest
54 approaches to construct such markers of biological age relied on clinical measures of
55 physiological capacity [3]. Later molecular and -omics approaches gained popularity, initially
56 including markers such as leukocyte telomere length [4], followed by multi-marker algorithms
57 based on high-throughput platforms, such as DNA methylation [5], transcriptomics [6],
58 metabolomics [7], and proteomics [8]. Importantly, these algorithms were trained to estimate
59 cross-sectional chronological age. Of these omics approaches, particularly DNA methylation-
60 based algorithms exhibited remarkably high accuracies in predicting calendar age [5], and were
61 named 'DNA methylation clocks'. Nonetheless, while interesting by itself, this observation
62 highlighted a fundamental limitation in this first design of markers of biological age. Since
63 nearly-perfect age predictors would arrive to similar observations as chronological age, they
64 would lose their characteristics as age-independent health status indicators [9].

65 Concomitantly, a second generation of -omics markers was introduced, which instead were
66 trained to predict the mortality risk. Prominent examples of these mortality-trained
67 multivariate markers include the DNA methylation-based PhenoAge [10] and GrimAge [11] and
68 the ¹H-NMR metabolomics-based MetaboHealth [12]. These predictors were trained quite
69 differently. The wide availability of the Nightingale Health ¹H-NMR metabolomics in large
70 prospective population studies, in combination with its relatively narrow though informative
71 content (~250 analytes), allowed for a more classic and direct approach. Deelen *et al.* trained
72 MetaboHealth as a linear combination of 14 metabolic features, showing a strong predictive
73 value, not only for mortality risk, but also for other outcomes, including pneumonia [13], and
74 frailty [14]. Conversely, the DNA methylation platform by Illumina contains hundreds of
75 thousands of features, and thus requires additional guidance to robustly capture the mortality
76 signal. Hence, the PhenoAge and GrimAge were trained using the so-called two-stage
77 approaches, in which more widely-available markers associated with mortality were leveraged
78 to help extract the mortality signal [10,11]. PhenoAge achieved this by first training an all-cause
79 mortality predictor based on clinical measures (e.g., glucose, C-reactive-protein), which was
80 then re-estimated using DNA methylation. Similarly, DNAm-GrimAge is composed by a
81 combination of DNA methylation-based surrogates for molecular or phenotypic markers known
82 to associate with mortality. Interestingly, both two-step training strategies yielded DNA
83 methylation-based scores that can associate not only with mortality, but also with a wide
84 diversity of disease outcomes. These developments indicate that mortality-trained predictors
85 for biological age can be trained using different omics platforms, and moreover, that DNA-
86 methylation might serve as a platform to integrate these signals captured by different data

87 sources. This latter concept was recently further substantiated by the work of Gadd *et al.*, who
88 systematically trained DNA-methylation-based predictors for 109 plasma proteins showing
89 significant associations with incident morbidities over 14-years [15].

90 In a recent study we demonstrated that mortality-based predictors such as MetaboHealth
91 and GrimAge are instrumental in predicting frailty in studies of middle-aged and elderly
92 individuals [14]. Importantly, we also showed that these scores confer mutually independent
93 information for predicting both frailty and mortality. Viewing these developments in the field,
94 we thus pose the question to what extent the mortality signal captured by ¹H-NMR
95 metabolomics could be transferred and integrated with the mortality signals captured by the
96 DNA-methylation platform. For this purpose, we will evaluate both strategies for training two-
97 stage DNA-methylation based mortality predictors. On one hand, we will train a DNA
98 methylation-based predictor re-estimating directly MetaboHealth, akin the strategy of
99 PhenoAge. On the other hand, we will train DNA-methylation surrogates for single
100 metabolomics features, and combine these in an overall score, akin GrimAge. Moreover, we
101 will evaluate to what extent DNA-methylation surrogates features from different origins
102 capture mutually independent signals, also with respect to predicting mortality risk.

103

104 **Results**

105 **Cross-cohort calibration of ¹H-NMR metabolomics data**

106 To derive DNA methylation-based models predicting metabolic features we analyzed
107 data gathered by partners of the BIOS consortium [16,17], totalling 4,334 individuals for whom
108 both DNA methylation (Illumina 450k) and ¹H-NMR Metabolomics (Nightingale Health Plc) data
109 have been assayed. The resulting dataset had contributions of four independently collected
110 population studies: LIFELINES-DEEP (LIFELINES), Leiden-Longevity-Study Partners-Offspring (LLS-
111 PAROFFS), Rotterdam-Study (RS), and Netherlands-Twin-Register (NTR), each with their own
112 inclusion criteria, as reflected by differences in subject characteristics that range from the
113 younger and leaner population of NTR (mean age=37.57 years and mean BMI=24.32 cm/kg²) to
114 the older and heavier population of RS (mean age=67.15 years and mean BMI=27,71 cm/kg²)
115 (Figure 1, Supplementary Table S1,). A reduced dimensionality projection using a t-distributed
116 neighbour embedding (tSNE) suggested that the interindividual variance in metabolomics data
117 could not only be attributed to interindividual phenotypic variability but was also capturing
118 some systematic differences between studies (Figure S3A-C). Following Makinen *et al.*, we
119 implemented a calibration technique suitable for cross-cohort harmonization, which starts with
120 the assumption that individuals with similar phenotypic characteristics should on average
121 exhibit similar metabolomics profiles [18]. For this purpose, we identified pairs of samples
122 across cohorts with matching age, sex, and BMI, and used LIFELINES as a common reference to
123 calibrate the other studies (Figure S2, more details in **methods**). A t-SNE projection of the
124 calibrated data revealed a substantial reduction of the systematic differences between studies,

125 as also quantified by k-BET (k-nearest neighbor Batch Effect Test) (Figure 2A-B, S3). Principal
126 Variance Component Analyses (PVCA) further confirmed this observation indicating that the
127 variation attributable to study differences was attenuated, while maintaining the variation
128 attributed to relevant biological variability (Figure 2C).

129 The construction of the MetaboHealth score as published by Deelen *et al.* [12] does not
130 include a cross-cohort calibration, but instead standardizes the individual metabolic features
131 per study prior to computation of the score (Figure S4). While this does make the
132 MetaboHealth score more comparable across cohorts, and satisfactory for most meta-analysis
133 purposes, it does also negate any real biological differences that may exist between studies.
134 Conversely, when computing the MetaboHealth score on the calibrated data, i.e., after
135 removing unwanted study differences and supplying all data as one dataset, it produced scores
136 with interpretable differences and consistent trends across cohorts (Figure 2D-F). For instance,
137 consistent with our expectation, the calibrated MetaboHealth scores now tend to be higher
138 among the studies with the older individuals RS and LLS-PAROFFS (Figure S4A). In addition, it
139 showed a more pronounced age-associated increase in men than in women, consistently over
140 all cohorts (Figure 2D). Lastly, higher calibrated MetaboHealth percentiles correlated with
141 increasing age, BMI, high sensitive CRP, and increasing prevalence of diabetes and alcohol
142 usage (Figure 2E-F).

143

144 **DNA methylation-based predictors recapitulate metabolic markers** 145 **previously associated with mortality**

146 Our first objective was to determine if DNA methylation could simulate the MetaboHealth
147 score, our metabolomics-based mortality predictor (Figure 1). To enforce the selection of
148 consistent signal in different cohorts, we implement an output-specific pre-selection of
149 consistent DNA methylation sites in the studies reserved for model development, NTR and
150 LIFELINES (methods). This Epigenome Wide Association (EWAS) yielded 17,705 CpG sites
151 showing a consistent univariate association with MetaboHealth, both in direction of association
152 and nominal significance (p -value <0.05). Pre-selected sites were then used as input for the
153 ElasticNET regression model predicting the MetaboHealth values (Method). The resulting
154 model, indicated as “*DNAm-MetaboHealth*” comprised ~1000 sites and showed good accuracy
155 in the 5-Fold Cross Validation test sets (5-FCV) (median $r\sim 0.52$, $RMSE\sim 0.43$), which was slightly
156 lower, but stable, in the replication sets (LLS-PAROFF: $R\sim 0.34$, $RMSE\sim 0.38$; RS: $R\sim 0.33$,
157 $RMSE\sim 0.5$) (Figure 3).

158 In parallel, we built distinct predictors for 64 metabolic features from Nightingale Health Plc,
159 following the same training design as for DNAm MetaboHealth (Figure 1A). The resulting
160 DNAm-based surrogates for the metabolomic features showed diverse mean accuracies over
161 the different test sets (5-FCV, LLS-PAROFFS and RS), with 23 models being accurate (mean R
162 across test sets >0.35), 20 mildly accurate ($0.2 > \text{mean } R \text{ across test sets} \leq 0.35$), and 21 low
163 accuracy models (mean R across test sets <0.2) (Figure 3 and S5). In the latter group we find 5
164 out of 8 amino acids, several LDL-related variables, all the ketone bodies and all the glycolysis
165 related markers. The middle group is enriched with IDL related markers, 6 out of 14 fatty acids,

166 and 2 out of 3 glycolysis related metabolites. The accurate group of DNAm-metabolomic
167 features included, 8 out of 10 HDL-related markers, 4 out of 8 VLDL-related molecules,
168 glycoprotein acetyls, creatinine, 3 out of 8 amino acids, and several fluid balance markers (e.g.,
169 MUFA%, Omega6%). Higher accuracies are often accompanied by a higher correlation with age
170 (e.g. *DNAm Leucine* and *Isoleucine*) or sex (e.g. *DNAm Creatinine*) (Figure 3, inner circles),
171 similar to what is observed for the GrimAge DNAm based components [11]. Notably, only 7 of
172 the most accurate surrogate markers were part of the 14 original metabolomic features
173 composing the MetaboHealth score. Nevertheless, for 18 of the 23 most accurate surrogate
174 markers, it was previously shown that the respective metabolic features significantly associated
175 with mortality [12].

176

177 **DNAm metabolomics surrogates confer a unique and relevant signal**

178 To foster the concept that DNA-methylation measurements might potentially serve as a
179 platform to integrate biomarker signals captured from various data sources, we conducted two
180 types of experiments. First, we ensured that the signals conveyed by our novel DNAm
181 surrogates of metabolomic features, constitute mutually independent markers, and not a
182 multitude of highly similar signals (Figure S6A). Then, comparing the correlations between our
183 surrogates with the original metabolomic features (Figure S6A, upper triangle), we observed a
184 structure remarkably congruent with the correlation structure observed between the original
185 metabolites (Figure S6A, lower triangle), albeit overall at slightly lower magnitude. This
186 indicates that, apart from the correlation structure between the original markers, no systematic
187 high inter correlations are observed, which thus suggests that the information in DNA

188 methylation measurements are sufficiently rich to reconstitute many closely related biomarker
189 signals, without introducing artificial interdependency.

190 Secondly, we explored to what extent our novel DNAm surrogates of metabolomic features
191 constitute novel signal compared to previously constructed DNA-methylation estimates.
192 Overall, low correlation ($\max |R| \sim 0.4$) are observed between our DNAm-based metabolomics
193 surrogates and DNAm-based multivariate clocks (Horvath, Hannum, PhenoAge, and GrimAge)
194 (Figure 4). Furthermore, correlations with other pre-trained DNAm-based surrogate molecular
195 markers (GrimAge components and the 109 protein EpiScores) are generally modest, with a
196 few notable exceptions. Particularly, the GrimAge surrogates DNAm-leptin and DNAm-adm, and
197 4 EpiScores (2771.35 [Gene: IGFBP1], 4929.55 [Gene: SHBG], 3505.6 [Gene: LT α], CD6), present
198 a relatively high positive correlation with HDL related surrogate markers and a relatively high
199 negative correlation with the amino acids (DNAm-Leucine, DNAm-Isoleucine, and DNAm-
200 Valine). We observe the inverse pattern for the GrimAge surrogate DNAm-PAI-1, and 4 other
201 EpiScores (4930.21 [Gene: STC1], 2516.57 [Gene: CCL21], 3343.1 [Gene: ACY1], 3470.1 [Gene:
202 SELE]) which also show a high correlation with VLDL surrogate markers. Notably, these
203 correlations might suggest a link between the metabolome and protein markers related to
204 immune signaling (LT α , CD6, CCL21, SELE), energy balance and metabolism-related hormones
205 (IGFBP1, SHGB, ADM, leptin, and STC1), and atherosclerosis/thrombosis inhibitor (PAI1).
206 Nonetheless, the majority of the markers exhibit limited correlations with their predecessors
207 (Figure 4), implying the presence of previously unexplored information in DNA methylation
208 patterns.

209

210 **Metabolomic surrogates improve the mortality predictions of** 211 **GrimAge.**

212 Next, we evaluated the DNAm-metabolomics features for their predictive value for all-cause
213 mortality in the Rotterdam Study (RS). For this purpose, we utilized a total of 1544 samples
214 from this cohort (mean age at baseline of 64 years, 251 deceased, and a median follow-up of 11
215 years; Figure 1), by incorporating an additional 863 samples with available Illumina 450k and
216 mortality information, but not ¹H-NMR metabolomics. In accordance with previous studies
217 [11,14,15], we assess univariate Cox proportional hazard models (in years of follow up)
218 adjusted for relevant covariates, specifically sex, together with age, BMI and cell counts levels
219 at blood sampling (Figure 5A and S7A).

220 First, we evaluated our *DNAm-MetaboHealth* predictor, which showed a significant
221 association with all-cause mortality (HR=1.29, $p=5.57 \times 10^{-04}$) (Figure 5A), in line with the original
222 metabolomics-based MetaboHealth score (Figure S8C). Next, we evaluated the individual
223 DNAm metabolomics features and observed significant associations with mortality for 15 out of
224 the 64 surrogate metabolites, of which 6 (out of 14) metabolomics features were included in
225 the original MetaboHealth score. In addition, our estimated effects were overall consistent with
226 those found by the study performed by Deelen *et al.* which employed a considerably larger
227 dataset of 44,168 individuals (Supplemental S7C), with our most significant findings being
228 amongst their strongest effects. We observed an increased risk for higher estimates of 8 DNAm-
229 based features, with the strongest being *DNAm-Glucose* (HR=1.22, $p=7.29 \times 10^{-03}$), *DNAm-*
230 *Glycoprotein Acetyls* (HR=1.35, $p=1.74 \times 10^{-05}$). Conversely, we observed protective effects for 7
231 features, such as *DNAm-PUFA%* (HR=0.81, $p=1.63 \times 10^{-03}$), *DNAm-Histidine* (HR=0.82, $p=1.67 \times 10^{-03}$).

232 ⁰³), *DNAm-Valine* (HR=0.83, $p=1.6 \times 10^{-02}$), and *DNAm-Albumin* (HR=0.83, $p=9.5 \times 10^{-03}$). In
233 addition, 5 nominal significant additional metabolites showed discordant mortality associations
234 between sexes. Explicitly, for males we observed mortality associations with *DNAm-Glutamine*,
235 whereas *DNAm-Tyrosine*, *DNAm-Leucine*, *DNAm-Total Fatty acids*, and *DNAm-PUFA* associated
236 with mortality in women (Figure S7B).

237 Almost all the pre-trained DNAm clocks that we considered (Hannum, PhenoAge, GrimAge
238 and bAge) and some of their intermediate surrogates exert mortality associations in RS (Figure
239 S9A-C). Next, we attempted to refine the current standard for biological age estimation,
240 specifically GrimAge (CI=0.79, $p=4.6 \times 10^{-77}$), by training a multivariate all-cause mortality
241 predictor including our novel DNAm metabolomics features. As a first exploration, we trained a
242 Cox regression model with age at blood sampling, sex, *DNAm-GrimAge* and *DNAm-*
243 *MetaboHealth*, which only showed minor, but significant, improvements in the C-index (CI=0.8,
244 $p=4.6 \times 10^{-77}$) (Figure S10 B). As a second exploration, we performed a stepwise
245 (backward/forward) Cox regression model to identify a minimal set of features including age,
246 sex, our 64 DNAm-metabolomic features and *DNAm-GrimAge* (CI=0.81, $p=1.7 \times 10^{-83}$) (Figure
247 S10D). Nonetheless, the best performing model was obtained when including age, sex, 3 out of
248 8 GrimAge components (predicting Leptin and ADM and TIMP_1), 12 of the 109 protein
249 EpiScores, and 9 out of 64 DNAm metabolites (CI=0.82, $p=1 \times 10^{-85}$) (Figure 5B). The selected
250 DNAm metabolic features included *DNAm-Tyrosine*, *DNAm-S-VLDL-L*, *DNAm-S-LDL-L*, *DNAm-M-*
251 *LDL-L*, *DNAm-APOB*, *DNAm-LA*, *DNAm-omega3* and *DNAm-omega6*, and *DNAm-MUFA*. In any
252 case, all the newly introduced scores exhibited a significantly improved C-index and a higher
253 AUC at 5 and 10 years compared to the GrimAge (Figure S10E-H). Overall, this indicates that

254 DNAm-surrogates from different origin, phenotypic, proteomic, or metabolomic, might confer
255 mutually independent information for mortality prediction.

256

257 **DNAm metabolomics models introduce relevant CpG selections**

258 After establishing the value of our novel DNAm metabolomics features in predicting
259 mortality, we explored the nature of the signal included in our models by investigating the CpG
260 sites picked by the ElasticNET regression, which can shrink contributions of unnecessary
261 features to zero. Predictors selected a median of ~750 CpG-sites, with a minimum of 234 CpG-
262 sites for *DNAm-Acetoacetate* and a maximum of 1,569 for *DNAm-ApoA1* (Figure 6B). A total of
263 22,145 probes were included in at least 1 model. Comparison of the genomic positions of the
264 selected CpG-sites with the rest of the 450k array highlighted an underrepresentation of probes
265 positioned in CpG Islands, and a preferential selection for CpG shelves and shores, known to be
266 more dynamic areas (Figure 6A) [19,20]. Noteworthy is the higher tendency to select CpGs co-
267 locating with enhancers, cis-acting short regions of DNA that control the temporal and cell-
268 specific activation of gene expression (Figure 6A) [21].

269 Functional enrichment analyses using the most proximal genes to the selected CpG-sites
270 highlighted pathways associated to “developmental processes”, “cell differentiation”, and
271 “regulation of metabolic processes” from Biological Processes in Gene Ontology (Figure S11C).
272 Concomitantly, enrichment analyses of phenotypic annotations in the EWAS Catalog and EWAS
273 Atlas (Figure 6B), indicated that the CpG-sites are known to be largely related to peripheral
274 tissue differentiation [22], fetal brain development [23] and gestational age [24]. Nonetheless,
275 the CpG sites with the highest median coefficients across all our models were the ones

276 annotated for metabolite-related traits, such as “Triglycerides”, and “Fasting Glucose” (Figures
277 6B, S11A-B). In total, 203 traits exhibited a significant enrichment for the CpG selections made
278 by our models. Notably we find also highly significant associations with “Ageing”, and “all-cause
279 mortality”, indicating that we do identify CpGs related to age-related processes.

280 Despite their interesting overarching signal, the DNAm-metabolic models show little overlap
281 with each other in their CpG selections, with the majority showing overlaps well below 15%,
282 apart for a few exceptions of highly correlated metabolites (e.g., 83% between *DNAm-*
283 *Total_cholines* and *DNAm-phosphoglycerides*; Figures S6 and S12). Nonetheless, a handful of
284 CpG probes were chosen in more than 30 models with largely consistent coefficient signs
285 (Figure 6C). Interestingly, while some of these 9 features have a higher importance weight on
286 the DNAm metabolomics models (e.g., cg00574958, or cg06500161), others only exert a more
287 minor influence (e.g., cg14938561, cg00461022). The 9 CpG sites with higher importance
288 weight don’t favor one specific metabolic group but seems to be relevant to many metabolic
289 markers (Figure 6D). Not surprisingly, also the nearest genes to these 9 probes are noteworthy.
290 For instance, TXNIP, which includes cg19693031 (chosen in 43 DNAm-metabolomics models),
291 was previously associated to hyperglycemia and insulin resistance, and ABCG1, nearby
292 cg06500161 (in 42 DNAm-metabolomics models) was associated to plasma lipid levels and
293 stroke (Figure 6D).

294

295 **Discussion**

296 A comprehensive quantification of biological ageing, as a way to assess the overall, holistic
297 health status and disease susceptibility of individuals [14], would constitute a major advance for
298 healthcare and preventive research. A diversity of molecular markers has been proposed as

299 indicators of biological age relating to health- and lifespan. Here we integrated well-established
300 DNA methylation-based and ¹H-NMR metabolomics resources for biological age prediction with
301 mortality as a primary endpoint. To our knowledge, the potential synergistic effects arising
302 from combining these two molecular sources remained thus far largely unexplored, and we
303 believe that a collection of models predicting metabolomics features may be relevant within
304 the rapidly growing repertoire of DNA methylation-based estimates [10,11,15]. A structured
305 training and evaluation design aided us to demonstrate the robustness of our features. We
306 highlighted the distinct signal expressed by our novel models and their feature selection.
307 Finally, we explored the use of our novel DNAm-based surrogates of metabolomics features in
308 combination with previously trained DNAm-based surrogates (e.g., Grimage constituents)
309 suggesting that these confer complementary information.

310 We applied ElasticNET regression models to the data of four large population cohorts to
311 derive DNAm-based surrogates for a previously derived multi-analyte score indicating mortality
312 (MetaboHealth), and for 64 individual metabolomics features. The direct estimation of
313 metabolomics-based mortality by constructing a DNAm surrogate for the MetaboHealth score
314 showed promising results (mean R in test-sets=0.397). Moreover, we were able to construct
315 DNAm surrogates for many, but not all, metabolomics features with good replication accuracies
316 (mean R in test-sets>0.35), including health markers for HDL and VLDL metabolism,
317 inflammation, and fluid balance. Less accurate were the DNAm surrogates for amino-acids,
318 ketone bodies, glycolysis, and LDL-related markers (mean R in test-sets <0.2). Nevertheless,
319 considering the limited number of available markers and the low accuracy thresholds previously
320 used for DNAm scores (R>0.1 in test sets) [15], we continued evaluating all 65 models. This

321 decision was further corroborated by a previous report by Stevenson *et al.* who suggested that
322 their DNAm surrogate for CRP was a more reliable indication of chronic inflammation than its
323 measured counterpart, even when considering the modest correlation between CRP and its
324 surrogate [25]. Overall, our DNAm metabolomic features conveyed a signal coherent with the
325 quantified metabolomics variables and independent from most of the previously reported DNA
326 methylation-based clocks and molecular surrogates.

327 Great emphasis was given to the harmonization of metabolomic data collected across
328 different cohorts, prior to training our DNAm-based models for individual metabolites or the
329 MetaboHealth score. Non-biological variability that may originate from inter cohort differences
330 in sample collection, storage, or handling could confound model training. Typically, this
331 challenge in epidemiology is addressed by applying a z-scaling per cohort prior to conducting a
332 meta-analysis, which in effect discards all differences, both technical and biological, between
333 cohorts. In other words, while allowing to draw conclusions on the similarities in associations
334 with endpoints between cohorts, this strategy does not allow for a direct comparison of the
335 underlying molecular profiles between cohorts. To address this issue we applied a calibration
336 technique, which we developed adapting methodologies previously applied in longitudinal
337 studies [18]. This novel calibration technique showed its merit in harmonizing the
338 metabolomics profiles, while preserving the natural biological heterogeneity within and
339 between the different study populations. Importantly, this approach allowed for an evaluation
340 of the MetaboHealth score across cohorts, showing consistent age and sex specific trends per
341 study, and global predictive power for established clinical variables, such as hsCRP and
342 diabetes.

343 Previous studies have shown advantages of pre-selecting CpGs when training ElasticNET
344 regression models [26–29]. Following this example, we implemented a pre-selection of CpG
345 sites showing a high variability and consistent association with the outcome of interest during
346 the training phase of our 5-Fold Cross Validation procedure. Approximately 22,000 CpG sites
347 were included in at least one DNAm-based models. Enrichment analyses showed that the
348 selected CpGs are more likely to be enhancers in CpG shelves and shores and are in the
349 proximity of genes enriched for regulation of metabolic and developmental processes, or cell
350 differentiation. This finding resonates with a longstanding hypothesis, that the ageing
351 methylome reflects processes underlying intricate cellular and molecular changes linked with
352 development and differentiation [30]. Furthermore, CpG sites selected for our surrogates were
353 also previously associated to age (e.g., Ageing, all-cause mortality), inflammatory (C-reactive
354 proteins), or metabolically related traits (e.g., triglycerides and metabolic syndrome). Strikingly,
355 we found a highly recurrent selection of 9 CpGs in at least 30 distinct DNAm surrogate models,
356 suggesting that these CpGs form a fundamental link between the blood metabolome and DNA
357 methylome. All these loci have been previously found associated with metabolic traits and
358 processes [31], and most of these 9 CpGs and their nearest genes are considered powerful
359 classifiers for diabetes stratification [32–34]. Remarkably, 3 of these 9 CpG probes showed
360 significant univariate association with mortality within the Rotterdam Study (Figure S8D). This
361 reassures over the valuable cardiometabolic content latent in our novel DNAm models.

362 Besides, our main intent was to evaluate the possibility to extrapolate the mortality signal
363 from the metabolome to DNA methylation. To do so, we tested which of our surrogates might
364 be indicative of all-cause mortality in a subset of the Rotterdam Study (1544 persons, 285

365 deaths). Notably, we observed a successful detection, albeit partial, of the mortality signal
366 exerted by the metabolomics platform. We could successfully derive a DNAm-based version of
367 MetaboHealth, which significantly associates with all-cause mortality, although it showed a
368 lower hazard ratio than the original score [35]. This might in part be explained by the fact that
369 only 6 of the 14 DNAm surrogates for the metabolites constituting the MetaboHealth showed
370 associations with all-cause mortality. Overall, we observed significant associations with
371 mortality for 15 out of 64 DNAm-based metabolites. The detected effects are consistent with
372 the results previously reported by Deelen *et al.* in a large study using the original metabolomic
373 features measured in 44,168 individuals. This consistency further underpins that DNAm
374 surrogates for metabolomic features could potentially be leveraged as novel epigenetic
375 markers of biological ageing.

376 To further explore this concept, we trained a multivariate model for all-cause mortality, that
377 was allowed to select from all available DNAm surrogates using a stepwise forward/backward
378 regression. This final model included 9 DNAm metabolomic features together with the
379 competing covariates age, 3 GrimAge components and 12 plasma protein EpiScores. The
380 resulting model combining DNAm surrogates from different origin showed a significantly
381 improved mortality prediction (C-index=0.82) compared to the GrimAge score (C-index=0.79)
382 (Figure 5 and S10). Our novel composite scores showed a substantial refinement of the AUC at
383 5 and 10 compared to the original GrimAge (Figure S10G-H). Overall, this suggests that a
384 broader collection of DNAm-surrogates of independent origin, such as proteomics, phenotypes,
385 and now also metabolomics, might confer a more comprehensive indication on epigenetic-
386 based biological ageing.

387 An important limitation of the current study for leveraging mortality signals is its limited
388 sample size, which is modest when compared to the large dataset that Deelen *et al.* employed
389 to evaluate the mortality associations of the metabolomics features and to build a multi-analyte
390 predictor for mortality. Despite the limited power, we found significant associations with
391 mortality for the DNAm surrogates of the multi-analyte score MetaboHealth and 15 individual
392 metabolic features, which were consistent with those observed by Deelen *et al.* A second
393 limitation is the usage of a single endpoint, mortality, for evaluating the potential applications
394 of our DNAm surrogates as novel marker for biological age. We acknowledge that ageing and its
395 associated decline in overall health is a complex multi-factorial process, that is only partially
396 captured by mortality risk. Previous work reported the merits of the ¹H-NMR metabolomics in
397 estimating several different types of endpoints [7,13,36,37], or even end-of-life related-
398 phenotypes such as frailty [35], leading us to speculate that our novel DNAm surrogates for
399 metabolomic features might also be instrumental for capturing these ageing endophenotypes.

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

407 **Materials and Methods**

408 **1. Dataset description**

409 **Cohorts**

410 This study was performed using DNA methylation data (DNAm, Illumina 450k array) and ¹H-
411 NMR metabolomics (Nightingale Health, platform version 2020) from 4 Dutch cohorts:
412 LifeLines-Deep (LL), Leiden Longevity Study (LLS-PARTNER-OFFSPRINGS), Netherlands Twin
413 Register (NTR) and Rotterdam Study (RS), all part of the BIOS consortium [16,17]. For the
414 current study, the BIOS multi-omics compendium was further extended with 1145 samples
415 from the NTR for which the entire process of array measurement to quality control and
416 normalization was done together with the other BIOS-NTR samples [38], and 904 samples from
417 the Rotterdam Study [35]. A thorough description of all cohorts and their ethics statement are
418 provided in the Supplementary Materials. The datasets were realized by the Dutch part of the
419 Biobanking and BioMolecular Resources and Research Infrastructure (BBMRI-NL). The final
420 dataset contained 5,238 samples.

421 **Metabolomics data**

422 The metabolomics data was generated by the BBMRI-NL Metabolomics Consortium. The
423 metabolic features were measured in EDTA plasma samples on the high-throughput proton
424 Nuclear Magnetic Resonance (¹H-NMR) platform made available by Nightingale Health Ltd.,
425 Helsinki, Finland (platform version 2020). This technique can quantify over 250 metabolic
426 features, including also ratios and derived features. [39,40]

427 **DNA methylation data**

428 DNA methylation data for all four cohorts was generated by the subsection of BBMRI-NL
429 named Biobank-based Integrative Omics Study (BIOS) Consortium. The DNAm was assessed

430 from whole blood samples with an Illumina iScan BeadChip according to the manufacturer's
431 protocol: the Illumina HumanMethylation405 BeadChip (450k array). For compatibility with the
432 following versions of the Illumina array, we only considered CpG sites which are available in the
433 Illumina HumanMethylation450 BeadChip and the MethylationEPIC BeadChip. We analysed
434 the DNAm β values, which range from 0 to 1, to indicate the proportion of methylated sites at a
435 specific CpG in a sample.

436 **Mortality data**

437 We evaluated the associations of the DNAm-based features with all-cause mortality in a
438 subsample of the Rotterdam Study (RS) comprising a total of 1544 samples, only 640 of which
439 had also Illumina 450k and Nightingale Health metabolomics. The information on the vital
440 status of the participants in RS was last updated on the 20th of October 2022. The dataset
441 comprehends 1544 samples, 285 of which are deceased. All the DNAm-based features were z-
442 scaled within the RS.

443 **2. Pre-processing**

444 **Quality control of the metabolomics dataset**

445 To ensure the quality of our data, we applied standardized quality control processes, which
446 have been described in previous publications (summarized in Figure S1) [7,41]. First, we limited
447 our analyses to a subset of 65 features (out of 250), previously selected to be a mutually
448 independent subset [7,12,41]. This selection includes fatty acids, routine lipid concentrations,
449 lipoprotein subclasses and low molecular weight metabolites. A complete list of the variables
450 can be found in the Supplementary Materials. In addition, pyruvate was excluded due to its
451 high missingness in NTR (80%). Despite a small percentage of values under detection limit for
452 acetoacetate (8% in NTR), and an even smaller percentage of outliers in glucose and xl_hdl_c

453 (less than 0.15%), we decided to retain all other variables (Figure S1C-E). Samples with more
454 than 1 outlier (2 from Lifelines and 1 from RS) were further removed. We then used *nipals*
455 (from the package *pcaMethods*) to impute the 584 missing values, which accounted for 0.211%
456 of the remaining values. The final dataset included 4,334 samples and 64 metabolic measures.

457 **Quality control of the DNA methylation dataset**

458 The quality control and normalization of the DNA methylation (DNAm) was performed using
459 a workflow developed by the BIOS Consortium for each cohort and thoroughly described in
460 DNAmArray (https://molepi.github.io/DNAmArray_workflow/). In brief, sample-level QC was
461 performed with the R package *MethylAid* [42]. Probes were set to missing based on the number
462 of available beads (≤ 2), intensity equal to zero, or the detection p value ($p < 0.01$). Probes with
463 more than 5% missing were excluded from all samples. The remaining missingness was imputed
464 using *impute.knn* from the R package *impute* [43]. Functional normalization was then applied as
465 implemented in *minfi*. Finally, we removed an ulterior set of ~60,000 underperforming probes
466 as suggested by Zhou et al. [44]

467 **Calibration of ¹H-NMR-metabolomics**

468 To minimize any bias that may arise from batch effects among the four cohorts included in
469 our study, we performed a cross-cohort calibration. We followed the assumption that similar
470 phenotypic characteristics result in similar metabolomics profiles [18]. We used sex, age, and
471 BMI as matching characteristics, given their well-known association with the metabolomic
472 features in the Nightingale Health Platform [7,18,41,45,46]. We considered LIFELINES as our
473 reference cohort, as it spanned a broad range of age, and BMI and had an equal representation
474 amongst sexes. To further minimize the impact of sex on our results, we selected the subset of
475 samples used for cross-cohort matching to have equal numbers of men and women. Following

476 this strategy, we identified the following subsets of participants used for matching: 73 men and
477 73 women in LLS-PAROFFS; 140 men and 140 women in NTR; 37 men and 37 women in RS
478 (Supplementary Figure S2).

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

486 T-distributed stochastic neighbor embedding (tSNE, R package *Rtsne*) was used to visually
487 inspect the effect of this calibration, comparing the sample similarities before and after
488 calibration. Moreover, K-nearest neighbor batch effect test (kBET, R package *kBET*) was applied
489 to the matching samples of each biobank before and after calibration to quantitatively evaluate
490 the mixing of the samples [47]. Finally, we used principal variance Component Analysis (PVCA, R
491 package *pvca*), to determine if the calibration disrupted the sources of variability of the dataset
492 [48].

493 **3. Application of previously trained multivariate models**

494 **MetaboHealth:** The MetaboHealth model is a mortality predictor based on Nightingale
495 Health metabolomics concentration [12]. We applied this model both on the uncalibrated and
496 calibrated version of the ¹H-NMR metabolomics dataset using the R-package *MiMIR* [49].

497 **Epigenetic clocks:** We projected the Horvath, Hannum, DNAm PhenoAge in our data using
498 the R package *methylclock* and the DNAm GrimAge clocks using Python scripts provided by Lu
499 et al. [11,50]. About 1000 CpG sites needed to calculate these biological ages were missing in
500 our cohorts, therefore we imputed them using the “datMiniAnnotation3_GOLD.csv” file,
501 dispatched by the same authors [11].

502 **EpiScores:** We projected the EpiScores and bAge score using the code available from the
503 work of Bernabeu et al. [28].

504 **4. Estimation and evaluation of the epigenetic-based metabolic** 505 **features.**

506 We derived prediction models for the 64 metabolomics features and the MetaboHealth
507 score using blood methylation data. For model development and testing we used NTR, and
508 LIFELINES, respectively the largest cohort and the calibration’s reference cohort. We employed
509 ElasticNET regression from the R package *glmnet* to train the models.

510 Other studies show the benefit of pre-selecting the features before using ElasticNET
511 regression [28,51]. For this reason, we performed Epigenome Wide Association studies (EWAS)
512 to identify CpG sites showing linear association with each feature separately in NTR and
513 LIFELINES (*metabolic feature* ~ *CpG site*). We selected the CpG probes with a consistent
514 association sign (positive or negative in both cohorts) and significant nominal p-value (<0.05), to
515 avoid excluding too much information.

516 We used a nested 5-Fold-Cross-Validation (5-Fold CV), to evaluate the models in the external
517 loop and using the internal loop to optimize the λ parameter, which determines the final set of
518 CpG sites included in each model. The mixing parameter alpha was fixed at 0.5, based on

519 previous work [11,41]. The final ElasticNET models were obtained using both NTR and LIFELINES
520 and the optimized parameters. LLS-PAROFFS and RS were used as replication datasets. Finally,
521 we report Pearson correlations (R) and the root mean square error (RMSE) of the predicted
522 *DNAm metabolic features* with their measured concentrations.

523 **5. CpG sites characterization**

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

526 **EWAS enrichment analysis.**

527 We utilized the MRC-IEU EWAS Catalog [52] and the EWAS Atlas [53] to assess the previously
528 known phenotypic annotation (traits) of the CpG sites selected by our models. Both the EWAS
529 Catalog and EWAS Atlas are online databases that compile results from Epigenome Wide
530 Association Study results. By merging these two resources, both downloaded on March 13th
531 2023, we aimed to gather a comprehensive list of previous EWASs. We gathered only the
532 associations conducted with Illumina 450k and accompanied with a PMID. We then excluded
533 redundancy between the two catalogs and applied Bonferroni correction using the number of
534 CpGs in the Illumina 450k (~480,000). This process yielded 742635 CpG-trait associations.

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

538 **Annotation of the genomic position of the CpG sites**

539 We used the R package *annotatr* to annotate the genomic features. CpG sites from the 450k
540 array were annotated using CpG Island (CGI) centric categories. The annotations we utilized are
541 as follows: CGI (annotated in the R package *AnnotationHub*), shores (2Kb upstream or

542 downstream the CGI), shelves (2kb flanking the CpG shores), interCGI (the rest of the CpGs). As
543 for the genic annotations we considered regions 1-5Kb upstream of the transcription starting
544 site (TSS), promoters (<1Kb upstream of the TSS), 5'UTR, 3'UTR, exons, introns, boundaries
545 between introns and exons, and intergenic regions. Additionally, we report the annotations of
546 active enhancers determined by Anderson et al. [21]

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

551 **Gene Ontology enrichment analyses**

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

554 Next, we utilized the genes associated with each CpG selections from our models to perform
555 a functional enrichment using Gene Ontology. We employed GOfuncR package to explore the
556 Biological *Processes* and *Molecular Functions*. The significance of the associations was
557 established with an $FDR < 0.05$. This analysis resulted in 2,365 significant associations between
558 CpG sites and genes.

559 **6. Associations with mortality in the Rotterdam Study**

560 **Univariate mortality associations:** We used Cox Proportional hazard to univariately
561 associate our 65 DNAm metabolomics features, the pre-trained DNAm clocks (e.g., PhenoAge,
562 GrimAge), and 109 protein EpiScores with mortality (see **Supplementary Materials**). All models
563 were corrected for age at blood sampling and sex. Additionally, we evaluated the association

564 with mortality of our DNAm metabolomics features when correcting for sex, age and GrimAge.
565 All *p-values* were corrected using Benjamini Hochberg and considered significant if the
566 FDR<0.05. We used the R-package *survival* to calculate the Cox regressions.

567 **Multivariate mortality models:** We then combined the DNAm features with sex and age in 4
568 different stepwise Cox regression models (see **Supplementary Materials**). The first base model
569 included our novel DNAm metabolomics features. The second and third model added to the
570 first model respectively DNAm-GrimAge and the DNAm-based components of the GrimAge
571 model. Finally, the fourth model is based on a combination of our DNAm metabolomics
572 features, the DNAm-based components of the GrimAge and the DNAm-based protein
573 EpiScores.

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

580 To compare the performances of the Cox regression models we used the R package
581 *survcomp* within the Rotterdam Study [54]. We compared the C-indices of the newly developed
582 models with baseline (GrimAge) using a Student t-test as described in Haibe Kans et al. [55].
583 Moreover, we plotted the ROC curves at 5 and 10 years of all the models.

584 **Data sharing**
585 BBMRI-nl and BIOS-nl data are available upon request at
586 <https://www.bbmri.nl/services/samples-images-data>. All DNAm metabolomics scores can be
587 obtained with a script at: https://github.com/DanieleBizzarri/DNAm_metabolomics_scores.

588 **Acknowledgements**

589 This work was performed within the BBMRI Metabolomics Consortium funded by: BBMRI-NL
590 (financed by NWO 184.021.007 and 184.033.111), X-omics (NWO 184.034.019), VOILA (ZonMW
591 457001001) and Medical Delta (METABODELTA: Metabolomics for clinical advances in the
592 Medical Delta). EvdA is funded by a personal grant of the Dutch Research Council
593 (NWO;VENI:09150161810095). Acknowledgements for all contributing studies can be found in
594 the Supplementary Material-BIOS Consortium. Additional NTR samples were funded by the
595 European Research Council (ERC-230374) project Genetics of Mental Illness (Boomsma).

596 **Contributors**

597 EbvDA, DB, MJTR and PES conceived and wrote the manuscript. DB performed the analyses.
598 EbvDA and MJTR verified and supervised the analyses. PES, MB, JBJvM, JvD, DIB, RP, MG, LF
599 were involved in data acquisition of the cohort data. All authors discussed the results and
600 contributed to the final manuscript.

601 **Competing Interests**

602 Authors declare no competing interests.

603 **References**

- 604 1. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of aging: An
605 expanding universe. *Cell* [Internet]. 2023 [cited 2023 Jan 10];0. Available from:
606 [https://www.cell.com/cell/abstract/S0092-8674\(22\)01377-0](https://www.cell.com/cell/abstract/S0092-8674(22)01377-0)
- 607 2. Partridge L, Deelen J, Slagboom PE. Facing up to the global challenges of ageing. *Nature*.
608 2018;561:45–56.
- 609 3. Comfort A. TEST-BATTERY TO MEASURE AGEING-RATE IN MAN. *The Lancet*. 1969;294:1411–
610 5.
- 611 4. Blackburn EH, Greider CW, Szostak JW. Telomeres and telomerase: the path from maize,
612 Tetrahymena and yeast to human cancer and aging. *Nat Med*. 2006;12:1133–8.
- 613 5. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*.
614 2013;14:R115.
- 615 6. Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, et al. The
616 transcriptional landscape of age in human peripheral blood. *Nat Commun*. 2015;6:8570.
- 617 7. van den Akker Erik B., Trompet Stella, Barkey Wolf Jurriaan J.H., Beekman Marian, Suchiman
618 H. Eka D., Deelen Joris, et al. Metabolic Age Based on the BBMRI-NL 1H-NMR Metabolomics
619 Repository as Biomarker of Age-related Disease. *Circulation: Genomic and Precision Medicine*
620 [Internet]. [cited 2020 Sep 14];0. Available from:
621 <https://www.ahajournals.org/doi/10.1161/CIRCGEN.119.002610>
- 622 8. Menni C, Kiddle SJ, Mangino M, Viñuela A, Psatha M, Steves C, et al. Circulating Proteomic
623 Signatures of Chronological Age. *J Gerontol A Biol Sci Med Sci*. 2015;70:809–16.
- 624 9. Zhang Q, Vallerga CL, Walker RM, Lin T, Henders AK, Montgomery GW, et al. Improved
625 precision of epigenetic clock estimates across tissues and its implication for biological ageing.
626 *Genome Medicine*. 2019;11:54.
- 627 10. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic
628 biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*. 2018;10:573–91.
- 629 11. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, et al. DNA methylation GrimAge
630 strongly predicts lifespan and healthspan. *Aging (Albany NY)*. 2019;11:303–27.
- 631 12. Deelen J, Kettunen J, Fischer K, van der Spek A, Trompet S, Kastenmüller G, et al. A
632 metabolic profile of all-cause mortality risk identified in an observational study of 44,168
633 individuals. *Nature Communications*. 2019;10:1–8.

- 634 13. Nightingale Health UK Biobank Initiative, Julkunen H, Cichońska A, Slagboom PE, Würtz P.
635 Metabolic biomarker profiling for identification of susceptibility to severe pneumonia and
636 COVID-19 in the general population. Janus ED, editor. *eLife*. 2021;10:e63033.
- 637 14. Kuiper LM, Polinder-Bos HA, Bizzarri D, Vojinovic D, Vallerga CL, Beekman M, et al.
638 Epigenetic and Metabolomic Biomarkers for Biological Age: A Comparative Analysis of Mortality
639 and Frailty Risk. *The Journals of Gerontology: Series A*. 2023;glad137.
- 640 15. Gadd DA, Hillary RF, McCartney DL, Zaghlool SB, Stevenson AJ, Cheng Y, et al. Epigenetic
641 scores for the circulating proteome as tools for disease prediction. Lo YD, Ferrucci L, editors.
642 *eLife*. 2022;11:e71802.
- 643 16. Zhernakova DV, Deelen P, Vermaat M, van Iterson M, van Galen M, Arindrarto W, et al.
644 Identification of context-dependent expression quantitative trait loci in whole blood. *Nature*
645 *Genetics*. 2017;49:139–45.
- 646 17. Bonder MJ, Luijk R, Zhernakova DV, Moed M, Deelen P, Vermaat M, et al. Disease variants
647 alter transcription factor levels and methylation of their binding sites. *Nat Genet*. 2017;49:131–
648 8.
- 649 18. Mäkinen V-P, Karsikas M, Kettunen J, Lehtimäki T, Kähönen M, Viikari J, et al. Longitudinal
650 profiling of metabolic ageing trends in two population cohorts of young adults. *International*
651 *Journal of Epidemiology*. 2022;51:1970–83.
- 652 19. Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LT-Y, Kohlbacher O, et al. Charting a dynamic DNA
653 methylation landscape of the human genome. *Nature*. 2013;500:477–81.
- 654 20. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon
655 cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG
656 island shores. *Nat Genet*. 2009;41:178–86.
- 657 21. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, et al. An atlas of
658 active enhancers across human cell types and tissues. *Nature*. 2014;507:455–61.
- 659 22. Islam SA, Goodman SJ, MacIsaac JL, Obradović J, Barr RG, Boyce WT, et al. Integration of
660 DNA methylation patterns and genetic variation in human pediatric tissues help inform EWAS
661 design and interpretation. *Epigenetics Chromatin*. 2019;12:1.
- 662 23. Spiers H, Hannon E, Schalkwyk LC, Smith R, Wong CCY, O'Donovan MC, et al. Methylomic
663 trajectories across human fetal brain development. *Genome Res*. 2015;25:338–52.
- 664 24. Bohlin J, Håberg SE, Magnus P, Reese SE, Gjessing HK, Magnus MC, et al. Prediction of
665 gestational age based on genome-wide differentially methylated regions. *Genome Biol*.
666 2016;17:207.

- 667 25. Stevenson AJ, McCartney DL, Hillary RF, Campbell A, Morris SW, Bermingham ML, et al.
668 Characterisation of an inflammation-related epigenetic score and its association with cognitive
669 ability. *Clinical Epigenetics*. 2020;12:113.
- 670 26. Choi H, Joe S, Nam H. Development of Tissue-Specific Age Predictors Using DNA
671 Methylation Data. *Genes (Basel)*. 2019;10:888.
- 672 27. Bergersen LC, Ahmed I, Frigessi A, Glad IK, Richardson S. Preselection in Lasso-Type Analysis
673 for Ultra-High Dimensional Genomic Exploration. In: Frigessi A, Bühlmann P, Glad IK, Langaas M,
674 Richardson S, Vannucci M, editors. *Statistical Analysis for High-Dimensional Data*. Cham:
675 Springer International Publishing; 2016. p. 37–66.
- 676 28. Bernabeu E, McCartney DL, Gadd DA, Hillary RF, Lu AT, Murphy L, et al. Refining epigenetic
677 prediction of chronological and biological age. *Genome Medicine*. 2023;15:12.
- 678 29. Croiseau P, Legarra A, Guillaume F, Fritz S, Baur A, Colombani C, et al. Fine tuning genomic
679 evaluations in dairy cattle through SNP pre-selection with the Elastic-Net algorithm. *Genet Res*
680 (Camb). 2011;93:409–17.
- 681 30. Seale K, Horvath S, Teschendorff A, Eynon N, Voisin S. Making sense of the ageing
682 methylome. *Nat Rev Genet*. 2022;23:585–605.
- 683 31. Gomez-Alonso M del C, Kretschmer A, Wilson R, Pfeiffer L, Karhunen V, Seppälä I, et al. DNA
684 methylation and lipid metabolism: an EWAS of 226 metabolic measures. *Clinical Epigenetics*.
685 2021;13:7.
- 686 32. Soriano-Tárraga C, Jiménez-Conde J, Giralt-Steinhauer E, Mola-Caminal M, Vivanco-Hidalgo
687 RM, Ois A, et al. Epigenome-wide association study identifies TXNIP gene associated with type 2
688 diabetes mellitus and sustained hyperglycemia. *Human Molecular Genetics*. 2016;25:609–19.
- 689 33. Krause C, Sievert H, Geißler C, Grohs M, El Gammal AT, Wolter S, et al. Critical evaluation of
690 the DNA-methylation markers ABCG1 and SREBF1 for Type 2 diabetes stratification.
691 *Epigenomics*. 2019;11:885–97.
- 692 34. Lai C-Q, Parnell LD, Smith CE, Guo T, Sayols-Baixeras S, Aslibekyan S, et al. Carbohydrate and
693 fat intake associated with risk of metabolic diseases through epigenetics of CPT1A. *Am J Clin*
694 *Nutr*. 2020;112:1200–11.
- 695 35. Kuiper LM, Polinder-Bos HA, Bizzarri D, Vojinovic D, Vallerga CL, Beekman M, et al.
696 Evaluation of epigenetic and metabolomic biomarkers indicating biological age [Internet].
697 medRxiv; 2022 [cited 2023 Jun 6]. p. 2022.12.05.22282968. Available from:
698 <https://www.medrxiv.org/content/10.1101/2022.12.05.22282968v1>
- 699 36. Buerger T, Steinfeldt J, Ruyoga G, Pietzner M, Bizzarri D, Vojinovic D, et al. Metabolomic
700 profiles predict individual multidisease outcomes. *Nat Med*. 2022;28:2309–20.

- 701 37. Ahola-Olli AV, Mustelin L, Kalimeri M, Kettunen J, Jokelainen J, Auvinen J, et al. Circulating
702 metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from
703 four Finnish cohorts. *Diabetologia*. 2019;62:2298–309.
- 704 38. van Dongen J, Gordon SD, McRae AF, Odintsova VV, Mbarek H, Breeze CE, et al. Identical
705 twins carry a persistent epigenetic signature of early genome programming. *Nat Commun*.
706 2021;12:5618.
- 707 39. Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic
708 resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiovasc Genet*.
709 2015;8:192–206.
- 710 40. Würtz P, Kangas AJ, Soininen P, Lawlor DA, Davey Smith G, Ala-Korpela M. Quantitative
711 Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -
712 Omic Technologies. *Am J Epidemiol*. 2017;186:1084–96.
- 713 41. Bizzarri D, Reinders MJT, Beekman M, Slagboom PE, Bbmri-NI null, van den Akker EB. 1H-
714 NMR metabolomics-based surrogates to impute common clinical risk factors and endpoints.
715 *EBioMedicine*. 2022;75:103764.
- 716 42. van Iterson M, Tobi EW, Slieker RC, den Hollander W, Luijk R, Slagboom PE, et al. MethylAid:
717 visual and interactive quality control of large Illumina 450k datasets. *Bioinformatics*.
718 2014;30:3435–7.
- 719 43. Hastie T, Tibshirani R, Narasimhan B, Chu G. impute: impute: Imputation for microarray data
720 [Internet]. Bioconductor version: Release (3.16); 2023 [cited 2023 Mar 23]. Available from:
721 <https://bioconductor.org/packages/impute/>
- 722 44. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use
723 of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res*. 2017;45:e22.
- 724 45. Telle-Hansen VH, Christensen JJ, Formo GA, Holven KB, Ulven SM. A comprehensive
725 metabolic profiling of the metabolically healthy obesity phenotype. *Lipids Health Dis*.
726 2020;19:90.
- 727 46. Ala-Korpela M, Lehtimäki T, Kähönen M, Viikari J, Perola M, Salomaa V, et al. Cross-
728 sectionally calculated metabolic ageing does not relate to longitudinal metabolic changes -
729 support for stratified ageing models. *J Clin Endocrinol Metab*. 2023;dgad032.
- 730 47. Büttner M, Miao Z, Wolf FA, Teichmann SA, Theis FJ. A test metric for assessing single-cell
731 RNA-seq batch correction. *Nat Methods*. 2019;16:43–9.
- 732 48. Li J, Bushel PR, Chu T-M, Wolfinger RD. Principal Variance Components Analysis: Estimating
733 Batch Effects in Microarray Gene Expression Data. *Batch Effects and Noise in Microarray*
734 *Experiments* [Internet]. John Wiley & Sons, Ltd; 2009 [cited 2023 Mar 23]. p. 141–54. Available
735 from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470685983.ch12>

- 736 49. Bizzarri D, Reinders MJT, Beekman M, Slagboom PE, van den Akker EB. MiMIR: R-shiny
737 application to infer risk factors and endpoints from Nightingale Health’s 1H-NMR metabolomics
738 data. *Bioinformatics*. 2022;38:3847–9.
- 739 50. Pelegí-Sisó D, de Prado P, Ronkainen J, Bustamante M, González JR. methylclock: a
740 Bioconductor package to estimate DNA methylation age. *Bioinformatics*. 2021;37:1759–60.
- 741 51. Higgins-Chen AT, Thrush KL, Wang Y, Minteer CJ, Kuo P-L, Wang M, et al. A computational
742 solution for bolstering reliability of epigenetic clocks: implications for clinical trials and
743 longitudinal tracking. *Nat Aging*. 2022;2:644–61.
- 744 52. Battram T, Yousefi P, Crawford G, Prince C, Sheikhalı Babaei M, Sharp G, et al. The EWAS
745 Catalog: a database of epigenome-wide association studies. *Wellcome Open Res*. 2022;7:41.
- 746 53. Xiong Z, Yang F, Li M, Ma Y, Zhao W, Wang G, et al. EWAS Open Platform: integrated data,
747 knowledge and toolkit for epigenome-wide association study. *Nucleic Acids Research*.
748 2022;50:D1004–9.
- 749 54. Schröder MS, Culhane AC, Quackenbush J, Haibe-Kains B. survcomp: an R/Bioconductor
750 package for performance assessment and comparison of survival models. *Bioinformatics*.
751 2011;27:3206–8.
- 752 55. Haibe-Kains B, Desmedt C, Sotiriou C, Bontempi G. A comparative study of survival models
753 for breast cancer prognostication based on microarray data: does a single gene beat them all?
754 *Bioinformatics*. 2008;24:2200–8.
- 755

756 **Figure descriptions for ¹H-NMR metabolomics guided** 757 **DNA methylation mortality predictors**

758 **Figure 1: Study and methods overview.** A) Study overview. (i) We employed 4334 samples,
759 from 4 cohort of the BIOS Consortium, DNAm methylation and metabolomics to train and test
760 our surrogates. (ii) Coupled with 1,544 samples from the Rotterdam Study to evaluate their
761 associations with mortality. (iii) We applied a calibration to harmonize the metabolomics
762 dataset. (iv) We then train ElasticNET models, on LIFELINES and NTR. Using the DNA
763 methylation data we predict two types of outcomes: 1) the pre-trained metabolomics mortality
764 predictor (MetaboHealth), and 2) the 64 metabolic features. (v) The DNAm models are
765 evaluated using 1) the hold-out valuation sets (LLS and RS) and 2) a 5-Fold Cross Validation on
766 the training sets (NTR and LIFELINES). (vi) Finally, we use the DNAm models to generate
767 surrogate metabolomics features in the RS dataset (1544 samples) and 1) evaluate their
768 univariate associations to mortality (while correcting for age, sex), and 2) trained a complete
769 Cox regression combining our DNAm metabolomics features and the pre-trained DNAm
770 surrogates. B) Availability of data in each cohort and when they are exploited in the study.

771
772 **Figure 2: Harmonization of the metabolomics data.** A) Distribution of glucose in NTR and
773 LIFELINES before (upper figure) and after (lower figure) calibration. B) tSNE of the
774 metabolomics dataset after calibration and colored by the four biobanks (LIFELINES, LLS, RS and
775 NTR). C) Principal Variance Component Analysis (PVCA) before (red) and after (green)
776 calibration, estimating the variance explained in the dataset by available clinical variables (e.g.,
777 sex, age, BMI, diabetes). D) Bar-plots showing the differences in men and women in the
778 calibrated MetaboHealth in the four cohorts. E) Observed mean values of age, BMI, eGFR,
779 hsCRP and pressure ordered following the calibrated MetaboHealth in different percentiles
780 over the entire BIOS population. F) Observed percentage of alcohol consumption, current
781 smoking ordered following the calibrated MetaboHealth in different percentiles over the entire
782 BIOS population.

783
784 **Figure 3: DNAm metabolites accuracies.** Circular heatmap representing the accuracies of the
785 DNAm-based models for 64 ¹H-NMR metabolic features by Nightingale Health and
786 MetaboHealth. The outer ring shows the correlation between measured and DNAm-based
787 metabolomics features, while correlation between DNAm surrogates with age and sex are
788 shown in the middle and inner ring, respectively. Mean CV states for mean cross validation
789 results in the cohorts LL and NTR together in the training (train) and test (test) sets, while
790 results in the left-out set are indicated with RS and LLS. Moreover, the metabolomics features
791 are annotated for their metabolomics group type (e.g., amino acids, fatty acids etc.) and if they
792 were or were not included in MetaboHealth. Finally, we indicated with asterisks the tertiles of
793 mean accuracies over the test sets.

794
795 **Figure 4: Correlations with pre-trained DNAm scores.** Correlations between our DNAm
796 metabolic features and previously trained clocks, the DNAm surrogates included in GrimAge
797 and the 109 DNAm-based surrogates for proteins (EpiScores) by Gall et al.
798

799

800

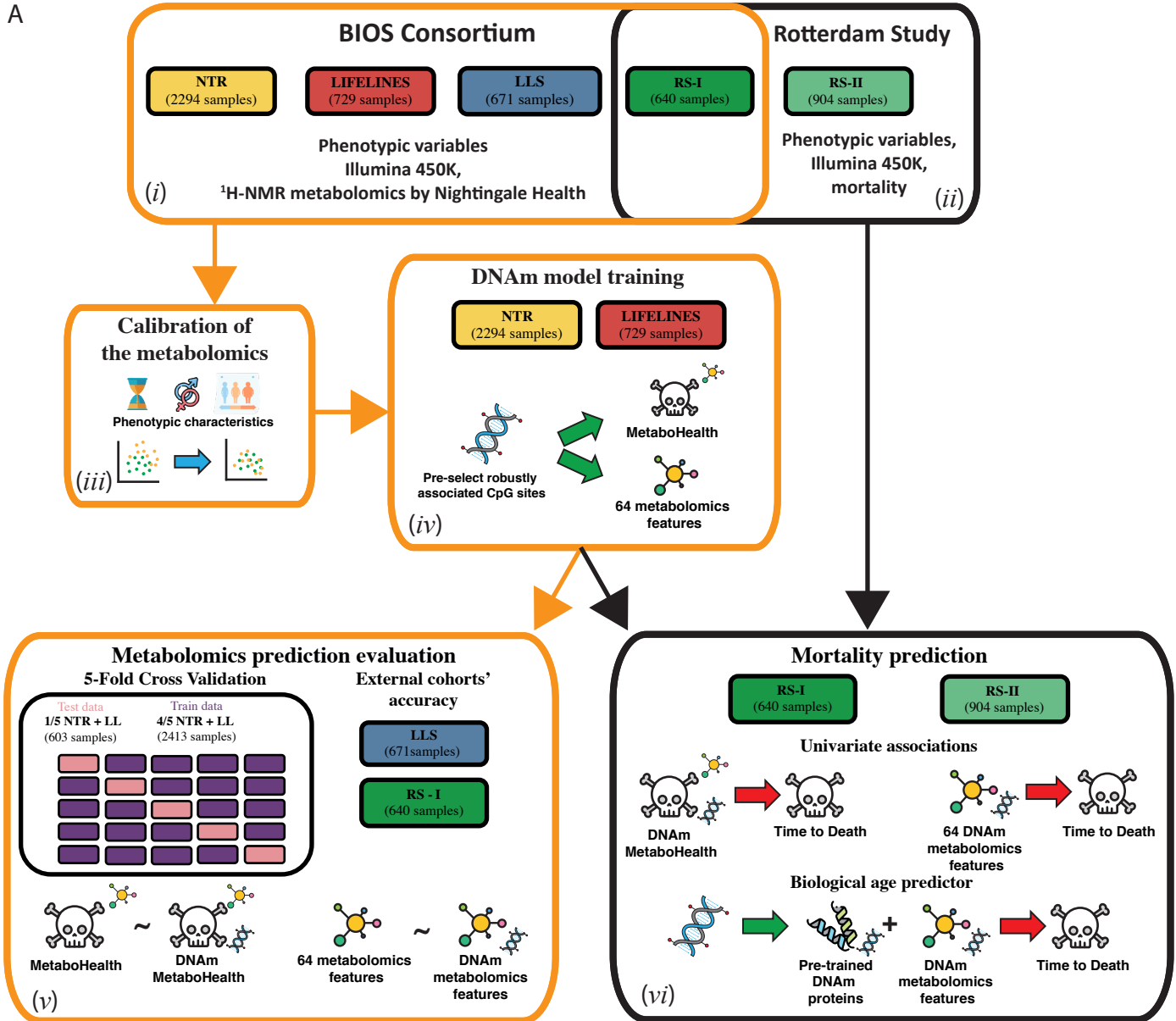
801

802 **Figure 5: Associations with Time to Death.** A) Significant univariate associations of the
803 DNAm metabolomics features with time to all-cause mortality in RS (N = 1542 with 285
804 reported deaths). The associations are grouped based on the metabolomics groups colored by
805 the significant associations or the metabolites with mortality in Deelen et al. The asterisks (*)
806 separates nominal significant DNAm metabolomics features from the FDR significant ones. B)
807 Stepwise Cox regression predicting of time to all-cause mortality optimized in RS, composed
808 combining age, 3 DNAm surrogates included in GrimAge, and 9 DNAm metabolic models and 12
809 protein EpiScores.

810

811 **Figure 6: CpG selections of the ElasticNET models.** A) Log2 Odds ratio indicating the
812 enriched in annotations of the CpG by our ElasticNET models. B) The central heatmap reports
813 the log₁₀ P-values of the enrichments CpG sites selected by our models (rows) and the 50 most
814 significant traits in the EWAS Catalog and Atlas enriched (rows). Bottom: the median
815 coefficients in each DNAm model, and the number of CpGs per model. Right: the median
816 coefficients given by our DNAm model to the overlapping CpGs with each trait. C) The nine
817 most used probes (rows) over the 65 ElasticNET models (columns), colored by metabolic
818 groups. Top: The models were ordered by the mean accuracy over the test sets (CV, LLS, and
819 RS). Right: The number of models which include each CpG and their nearest genes. D)
820 Manhattan plot-like figure indicating the Variable importance of the single CpG probes in the
821 DNAm metabolic models.

Figure 1: Study overview



B

Biobank	Illumina 450K	Nightingale Health metabolomics	Mortality	Function
LIFELINES	✓	✓	✗	Training and 5FCV
NTR	✓	✓	✗	Training and 5FCV
LLS-PAROFFS	✓	✓	✗	Accuracy evaluation
RS-I	✓	✓	✓	Accuracy evaluation and mortality association
RS-II	✓	✗	✓	Mortality association

Figure 2: Harmonization of the metabolomics data and its effect on MetaboHealth

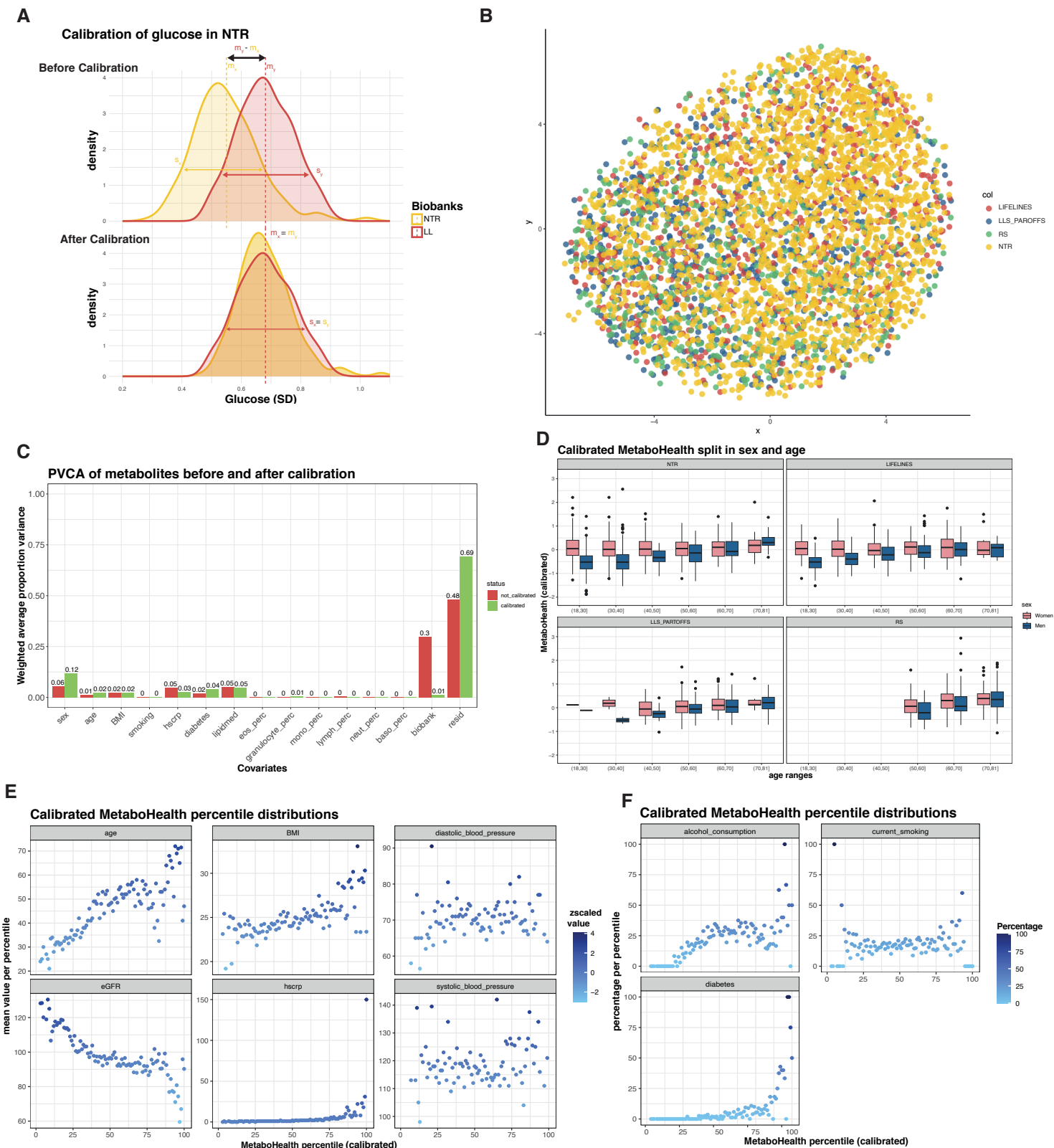


Figure 3: DNAm metabolomics accuracies

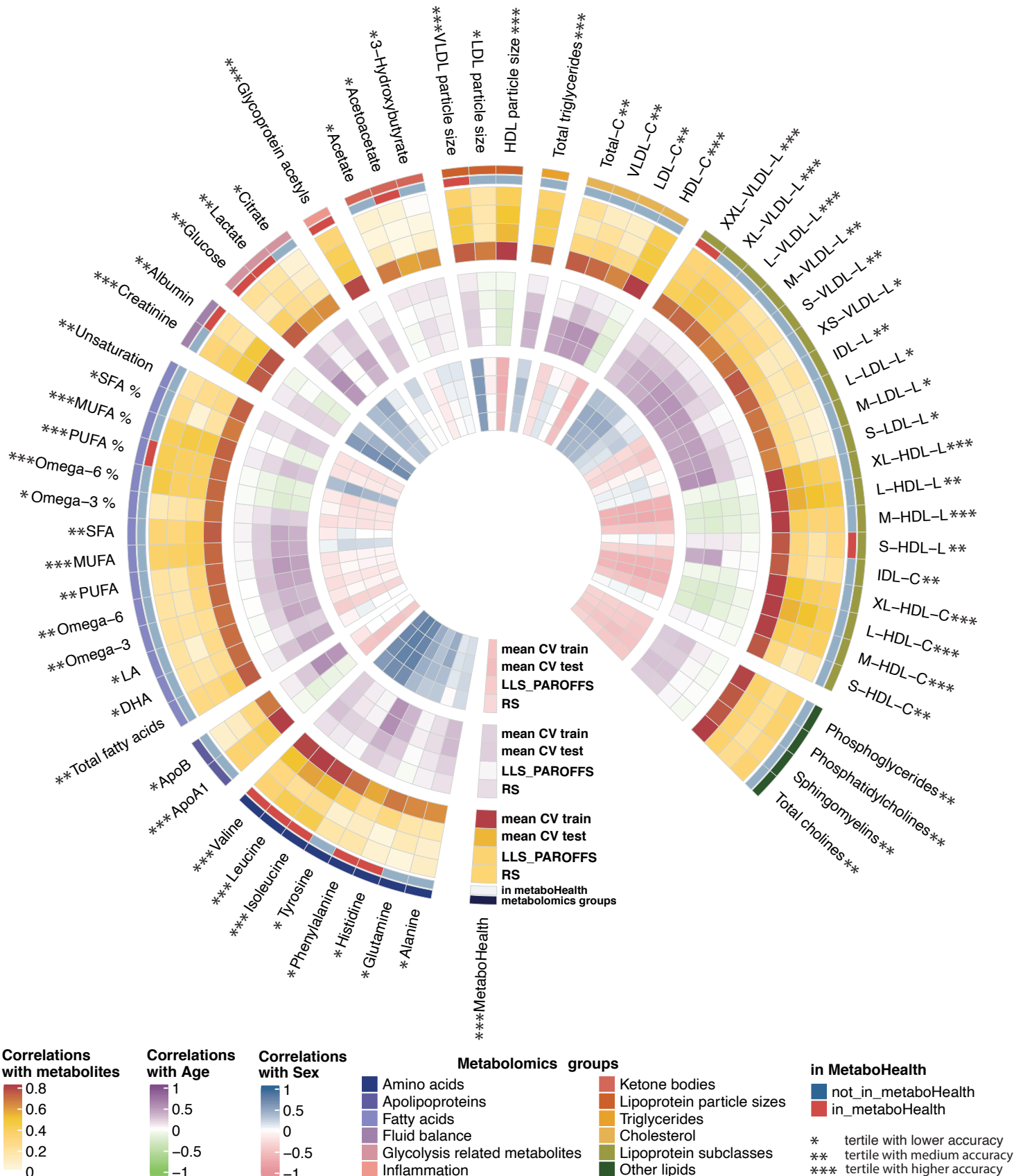


Figure 4: Correlations with pre-trained DNAm scores

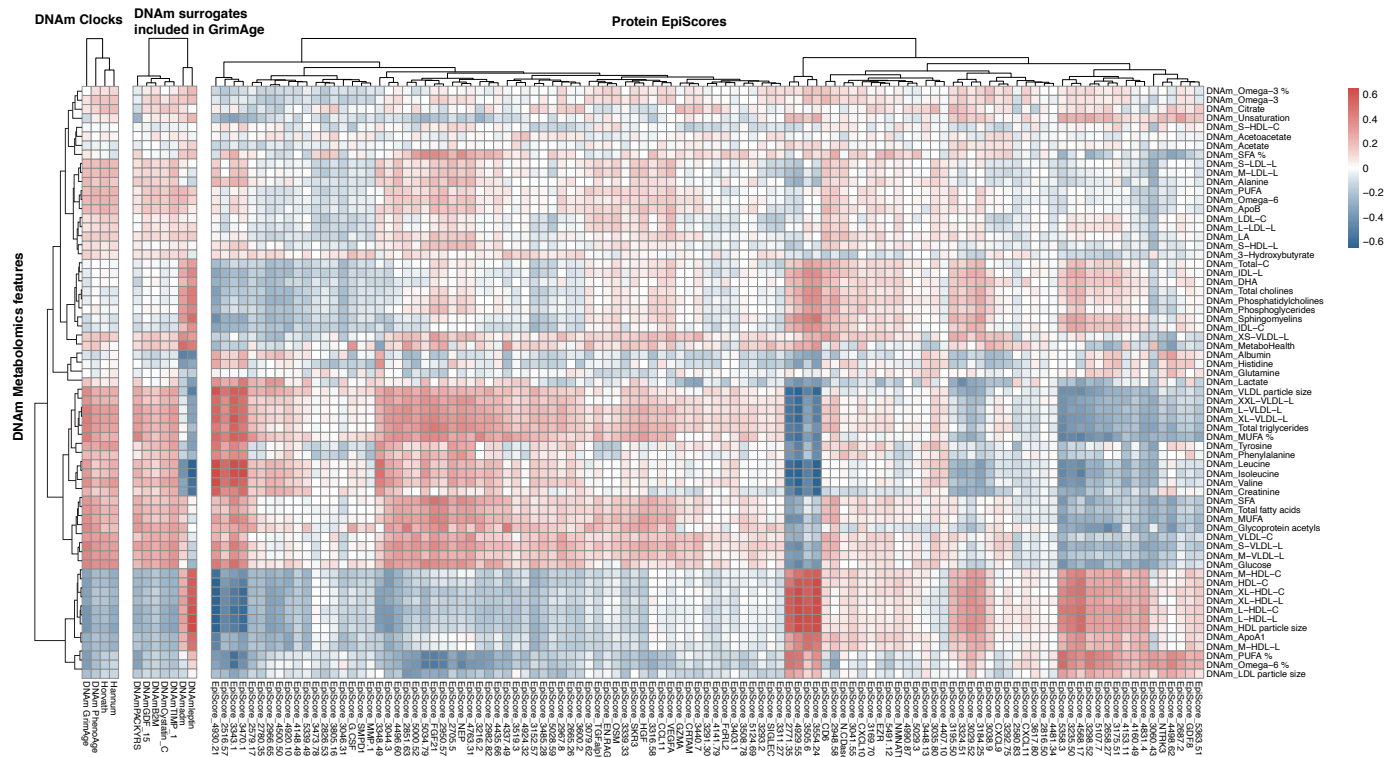


Figure 5: Associations with Time to Death

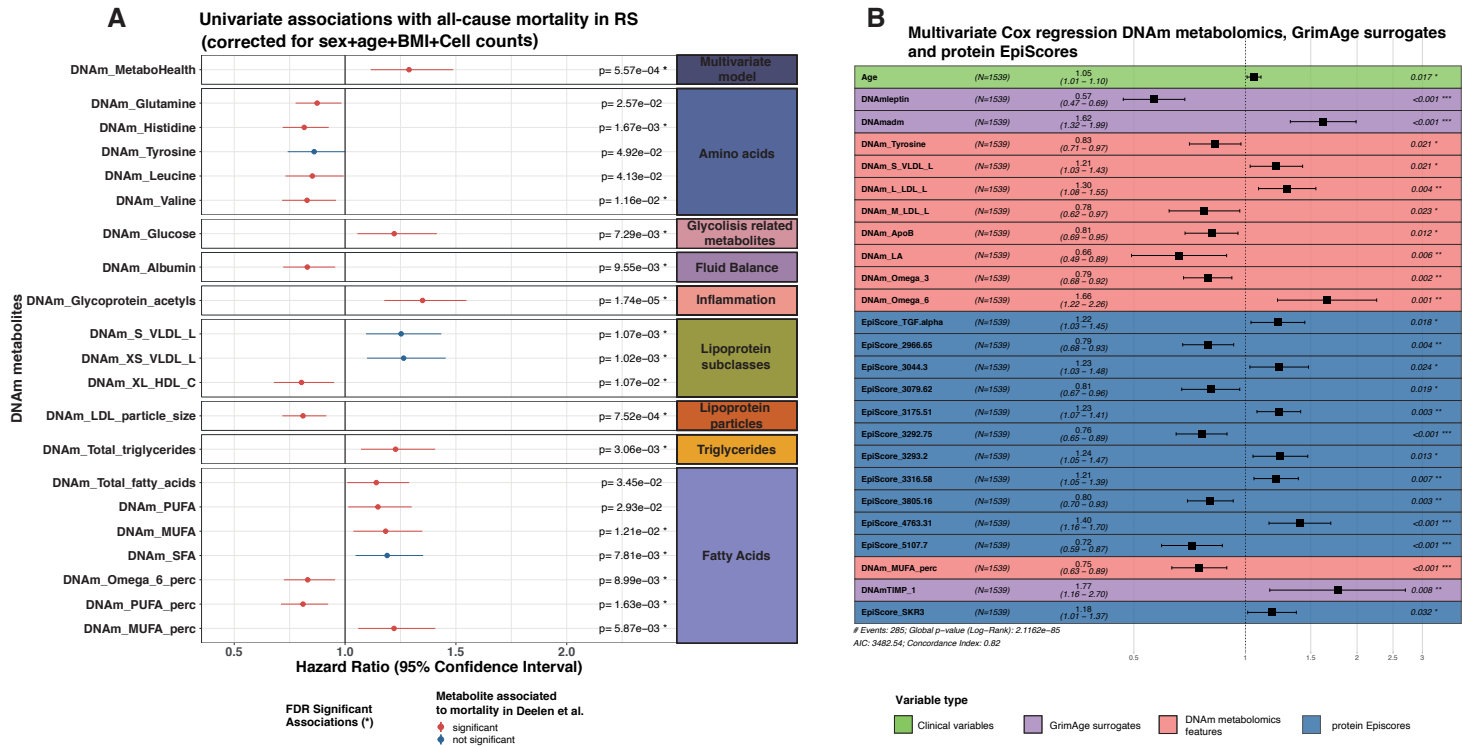


Figure 6: CpG selections of the ElasticNET models

