HRS Documentation Report

HRS Epigenetic Clocks

Report prepared by
Eileen Crimmins, University of Southern California
Jung Ki Kim, University of Southern California
Jonah Fisher, University of Michigan
Jessica Faul, University of Michigan

Survey Research Center
University of Michigan
Ann Arbor, Michigan
September 2020
I. Introduction

This document describes a data set consisting of values for 13 epigenetic clocks developed from DNA methylation data derived from the 2016 Health and Retirement Study Venous Blood Study. These constructed measures are released by HRS as sensitive health data.

The full QC’d HRS 2016 DNA methylation data from which these were constructed are being released as a HRS restricted health data release as well as by NIAGADS (https://www.niagads.org/). This file is approximately 19 GB. The full iDAT files will be released only via NIAGADS (500 GB).

A. Rationale

DNA Methylation (DNAm) is one mechanism by which exposure to adverse life circumstances and environments are linked to health outcomes related to aging. DNAm occurs with the addition of a methyl group to a CpG site in DNA. A number of researchers have identified portions of the genome where methylation changes are related to either age or, more recently, to health outcomes linked to age. The clocks combine information for a small number of of CpGs (typically 100-500) to produce indicators of epigenetic aging. Methylation clocks representing epigenetic age are estimated in epigenetic years and the idea is that ticks of the clock represent aging. Whether someone is aging faster or slower than his/her chronological age can be indicated by the difference between epigenetic age and chronological age.

Thirteen epigenetic clocks have been constructed using the HRS data. Eleven of these were constructed by both Morgan Levine (Yale) as well as HRS staff in order to ensure reliability. GrimAge, the twelfth clock, was constructed by HRS staff member Jonah Fisher with assistance of Steve Horvath. The DundedinPoAm38 clock was estimated by Thalida Arpawong (USC) with assistance of Karen Sugden (Duke). Morgan Levine and colleagues (Liu, Leung and Levine 2019) have provided detailed comparisons of 11 of these clocks in terms of CpG characteristics, tissue specificity, and co-regulatory gene network signatures.

B. The HRS Methylation Sample of Respondents

DNA methylation assays were done on a non-random subsample (n=4,104) people who participated in the 2016 Venous Blood Study. The sample includes all the participants of the 2016 Healthy Cognitive Aging Project (HCAP) who have provided blood samples, plus younger participants designated for future HCAP assessments, and a subsample of HCAP non-participants. This subsample fully represents the entire HRS sample. A total of 4,018 samples passed QC. The sample is 58% Female and has a median age of 68.7 years. It is racially diverse: Non Hispanic White (n=2,669, 66.4%), Non Hispanic Black (n=658, 16.4%), Hispanic (n=567, 14.11%), Non Hispanic Other (n=124, 3%). The sample is also socioeconomically diverse. The educational distribution is less than High School (16.8%), High School / GED (52.12%), Some College (5.97%), College + (24.1%), Other (1%).

C. Collection

The 2016 VBS blood collection was managed by Hooper Holmes Health & Wellness. The phlebotomy service was provided with the names, addresses, and phone numbers of consenting respondents and contacted respondents to set appointments. Collection materials were mailed to the phlebotomists’ homes in advance of the scheduled visit. Every attempt was made to schedule the blood draw within 4 weeks of the HRS core interview. Fasting was recommended and preferred but not required. Phlebotomists noted the fasting status of the samples. We collected 50.5 mL of blood in 6 tubes – 1 x 8
mL CPT tube, 3 x 10 mL double gel serum separator tubes (SST), 1 x 10 mL EDTA whole blood tube, and a 2.5 mL PAXgene RNA tube. The SST tubes are centrifuged in the field before being shipped overnight to the CLIA-certified Advanced Research and Diagnostic Laboratory at the University of Minnesota. Tube processing is done within 24 hours of arrival at the lab (within 48 hours of collection). DNA for methylation analysis was done using DNA extracted from the EDTA tube.

More information on the 2016 Venous Blood Study, including details on sampling, consent, and administration, is provided VBS 2016 Data Description.

D. Protocols for DNA Methylation Data
DNA methylation data are based on assays done using the Infinium Methylation EPIC BeadChip at the University of Minnesota. Samples were randomized across plates by key demographic variables (i.e. age, cohort, sex, education, race/ethnicity) with 40 pairs of blinded duplicates. Analysis of duplicate samples showed a correlation >0.97 for all CpG sites.

The minfi package in R software was used for data preprocessing, and quality control. 3.4% of the methylation probes (n=29,431 out of 866,091) were removed from the final dataset due to suboptimal performance (using a detection P-value threshold of 0.01). Analysis for detection P-value failed samples was done after removal of detection P-value failed probes. Using a 5% cut-off (minfi) we remove 58 samples. We also removed sex mismatched samples and any controls (cell lines, blinded duplicates). High quality methylation data is available for 97.9% samples (n=4,018).

Prior to the estimation of the 13 clocks missing beta methylation values were imputed with the mean beta methylation value of the given probe across all samples.

E. Subsample Weights
Respondents with at least one valid venous blood result (VBS16VALID) were assigned a VBS weight. The weights were adjusted for the differential probabilities of participation by dividing the HRS 2016 sample weight by the predicted probability of having a valid venous blood result among community-dwelling 2016 HRS respondents born prior to 1960, excluding all members of the LBB cohort. The resulting interim weight was trimmed at the 1st and 99th percentiles and was then post stratified back to the entire 2016 HRS sample born prior to 1960 by age, sex, and race/ethnicity. Two separate respondent-level weights were created for the VBS 2016 Innovative Sub Sample and should be used for analyses of data from that sample. VBS16WGTRA should be used for analyses including DNA methylation and epigenetic clocks. Sample weights can be found in the HRS Tracker data file.
II. Individual Clocks

A. Horvath 1

Horvath 1, the first multi-tissue epigenetic clock, was developed using 8,000 samples from 82 Illumina DNA methylation array datasets, incorporating 51 healthy tissues and cell types in order to estimate the DNA methylation age of most tissues and cell types. The clock is defined based on DNA methylation at 353 CpGs that form an aging clock, and shows strong correlation with age (r=0.96-0.97). Horvath et al. (2013) found DNAm age acceleration was related to multiple types of cancer.

HRS Horvath 1 details are presented below.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horvath 1</td>
<td>4018</td>
<td>23.31</td>
<td>114.52</td>
<td>91.21</td>
<td>65.28</td>
<td>65.72</td>
</tr>
</tbody>
</table>

Reference

B. Hannum

Hannum’s epigenetic clock is a blood-based age estimator, based on DNA methylation at 71 CpGs selected from the Illumina 450,000 array (Hannum 2013). Hannum et al. developed this clock based on the whole blood of 656 humans at ages 19 to 101. They reported a strong correlation with age for this clock (r=0.96) and that the rate of DNAm ageing is influenced by gender and genetic variants.

HRS Hannum details are presented below.

<table>
<thead>
<tr>
<th></th>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hannum</td>
<td>4018</td>
<td>25.06</td>
<td>107.79</td>
<td>82.73</td>
<td>53.91</td>
<td>54.59</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Reference

DNAm PhenoAge was developed using composite clinical biomarkers combined into a multi-system measure of biological age, called phenotypic age, which was developed to estimate an individual’s mortality risk using 9 markers of tissue and immune function (albumin, creatinine, serum glucose, CRP, lymphocyte percent, mean (red) cell volume, red cell distribution width, alkaline phosphatase, white blood cell count) and age. Phenotypic age was predicted by DNAm PhenoAge based on 513 CpGs in whole blood from the same sample. Levine et al. (2018) found that while this clock was developed using whole blood data, values from all tested tissues and cells are correlated with age and predict mortality better than chronological age-based clocks. DNAm PhenoAge has been shown to predict multiple aging outcomes such as mortality, cancer, healthspan, physical function and Alzheimer’s disease; the rate of DNAmPhenoAge acceleration was related to biomarkers such as high CRP, glucose, triglycerides waist-to-hip ratio and low HDL cholesterol (Levine et al. 2018).

HRS Levine details are presented below.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levine</td>
<td>4018</td>
<td>26.72</td>
<td>101.68</td>
<td>74.96</td>
<td>56.85</td>
<td>57.48</td>
</tr>
</tbody>
</table>

Reference

D. Horvath 2

This epigenetic clock, based on 391 CpGs, was developed to better measure the age of human fibroblasts and other skin cells such as keratinocytes, buccal cells, endothelial cells, lymphoblastoid cells, skin, blood, and saliva samples (Horvath et al. 2018). This clock has high age correlations in sorted neurons, glia, brain, liver and bone samples, to predict lifespan and to relate to many age-related conditions.

This skin & blood clock shares 45 CpGs with the blood-based clock from Hannum (2013) and 60 CpGs with the pan tissue clock from Horvath (2013). However, epigenetic age acceleration in the skin & blood clock shows only moderate correlations with that of Hannum's and Horvath's 2013 clock.

HRS values for the Horvath 2018 clock are presented below.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horvath 2018</td>
<td>4018</td>
<td>36.97</td>
<td>101.29</td>
<td>64.32</td>
<td>69.04</td>
<td>69.61</td>
</tr>
</tbody>
</table>

Reference

E. Lin

This 99 CpG model was originally derived from the HumanMethylation27K BeadChip data and subsequently modified for the 450,000 BeadChip. It was developed on DNAm profiles of normal blood samples and trained on life expectancy.

HRS Lin details are presented below.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin</td>
<td>4018</td>
<td>1.91</td>
<td>133.28</td>
<td>131.37</td>
<td>57.73</td>
<td>58.41</td>
</tr>
</tbody>
</table>

References


Weidner et al. (2014) developed a simple DNA methylation age based on 3 age-related CpGs (cg02228185 in ASPA, cg25809905 in ITGA2B, and cg17861230 in PDE4C), to estimate epigenetic aging in blood. They selected these three CpGs based on recursive feature elimination and pyrosequencing analysis. This clock produced age predictions with average accuracy of 5.4 years.

HRS Weidner details are presented below.

<table>
<thead>
<tr>
<th></th>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weidner</td>
<td>4018</td>
<td>25.22</td>
<td>148.87</td>
<td>123.65</td>
<td>65.56</td>
<td>67.29</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Reference

Vidal-Bralo et al. (2016) developed a DNAm age predictor based on 8 CpGs, which were selected as the most informative CpGs in a training set of 390 healthy persons. This clock was developed specifically targeting adults who show slower rates of change compared to pre-adolescents in order to more accurately calibrate DNAm age for adults.

HRS Vidal-Bralo details are presented below.

<table>
<thead>
<tr>
<th>Vidal-Bralo</th>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4018</td>
<td>36.47</td>
<td>109.95</td>
<td>73.48</td>
<td>63.37</td>
<td>63.76</td>
<td>0.10</td>
<td></td>
</tr>
</tbody>
</table>

Reference

H. **GrimAge**

GrimAge was developed based on the 7 DNAm surrogates of plasma proteins and smoking pack years in a two-stage procedure (Lu et al. 2019). First, they defined surrogate DNAm biomarkers of physiological risk and stress factors with plasma proteins (including adrenomedullin, CRP, plasminogen activation inhibitor 1 (PAI-1) and growth differentiation factor 15 (GDF15)) and DNAm-based estimator of smoking pack-years. Then, time-to-death was regressed on these biomarkers and an estimator of smoking years to estimate a composite biomarker of lifespan, GrimAge. They named it “DNAm GrimAge” because high values of this measure means grim news in terms of mortality and morbidity risk. Lu et al. (2019) report that the rate of GrimAge-based aging has predictive ability for time to death, coronary heart disease, cancer and age-related conditions.

HRS GrimAge details are presented below.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GrimAge</td>
<td>4018</td>
<td>42.67</td>
<td>99.61</td>
<td>56.94</td>
<td>67.71</td>
<td>68.13</td>
</tr>
</tbody>
</table>

Reference

Yang et al. (2016) developed a mitotic-like clock using 385 PCGT promoter CpGs. This is based on the DNAm-based age-correlated model called epiTOC (Epigenetic Timer Of Cancer) that features three properties including being constitutively unmethylated across 11 different fetal tissue types, showing age-associated hypermethylation, and targeting the promoters marked by the PRC2 complex in human embryonic stem cells (ESCs). This mitotic-like clock was shown to be universally accelerated in cancer and pre-cancerous lesions.

HRS Yang details are presented below.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yang</td>
<td>4018</td>
<td>0.03</td>
<td>0.23</td>
<td>0.20</td>
<td>0.07</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

Reference

Zhang et al. (2017) developed a DNAm age based on 10 CpGs that showed a strong association with all-cause mortality, which was selected from replicated results (58 out of 11,063 CpGs with FDR<0.05) from an epigenome-wide association study (EWAS) for all-cause mortality. This epigenetic clock is said to predict disease and mortality better than the original chronological DNAm clocks. This clock specifically identifies those with increased risk of death by cancer and cardiovascular disease.

HRS Zhang details are presented below.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang</td>
<td>4018</td>
<td>-2.53</td>
<td>0.60</td>
<td>3.14</td>
<td>-1.10</td>
<td>-1.08</td>
</tr>
</tbody>
</table>

Reference

The Bocklandt clock was developed in 2011 using saliva from twin pairs ages 21 to 55 years. The methylation in three sites, EEDARADD, TOM1LI, and NPTX2 genes, was linear with age, and a predictor including two CpGs in the promoter region of EDARADD and NPTX2 explained 73% of the variance in age and predicted age with an average accuracy of 5.2 years (Blocklandt et al. 2011).

HRS Bocklandt details are presented below.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bocklandt</td>
<td>4018</td>
<td>0.10</td>
<td>0.89</td>
<td>0.78</td>
<td>0.39</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Reference

Garagnani et al. (2012) used the Illumina Infinium Human Methylation450 BeadChip on whole blood DNA to identify methylation levels of 3 regions, the CpG islands of ELOVL2, FHL2 and PENK genes, strongly correlated with age. This was confirmed using whole blood from 501 persons ages 9 to 99 years and they identified one CpG (cg16867657) in ELOVL2 as a promising biomarker of aging \( r=0.92 \).

HRS Garagnani details are presented below.

<table>
<thead>
<tr>
<th></th>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garagnani</td>
<td>4018</td>
<td>0.43</td>
<td>0.99</td>
<td>0.56</td>
<td>0.71</td>
<td>0.72</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Reference

M. DunedinPoAm38

A recent blood DNA methylation measure, DunedinPoAm38, was developed to represent individual variation in the pace of biological aging. Based on data from the Illumina 450k Array run on samples from the Dunedin cohort, estimates were derived by using elastic-net regression models to calculate a methylation Pace of Aging (mPoA) score (Belsky et al. 2020). The pace of aging was calculated with composited slopes across the 18 biomarkers that measure the rate of aging in the cardiovascular, metabolic, renal, hepatic, pulmonary, periodontal, and immune systems. Then, the pace of aging composite was scaled to represent the mean trend in the cohort among Dunedin Study members with methylation data at age 38. The Pace of Aging methylation algorithm was trained on 3 waves of biomarker data from participants, including data collected at ages 26, 32, and 38. DunedinPoAm is estimated in years per chronological year (years/chron year).

HRS DunedinPoAm38 details are presented below.

<table>
<thead>
<tr>
<th></th>
<th>Sample Size</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
<th>Median</th>
<th>Mean</th>
<th>SE (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DunedinPoAm38</td>
<td>4018</td>
<td>0.74</td>
<td>1.46</td>
<td>0.72</td>
<td>1.06</td>
<td>1.07</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Reference

III. If You Need to Know More

This document is intended to serve as a brief overview to the HRS Epigenetic Clock data product. If you have questions or concerns that are not adequately covered here or on our Web site, or if you have any comments, please contact us. We will do our best to provide answers.

A. HRS Internet Site

Health and Retirement Study public release data and additional information about the study are available on the Internet. To access public data or to find out more about restricted data products and procedures, visit the HRS Web site.

B. Contact Information

If you need to contact us, you may do so by one of the methods listed below.

Internet: Help Desk at the HRS Web site (http://hrsonline.isr.umich.edu)

E-mail: hrsquestions@umich.edu

Postal Service:
Health and Retirement Study
The Institute for Social Research
426 Thompson Street
Ann Arbor, Michigan 48104

C. Citing this Document

Please include the following citation in any research reports, papers, or publications based on these data along with the citation for the reference epigenetic clock:

In text: “The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (NIA U01AG009740) and is conducted by the University of Michigan.”

IV. References


