MIDUS
Midlife in the United States
A National Study of Health & Well-Being

MIDUS Genomics Project

DNA Methylation Age

Data Documentation

For File:

M2MR1_GEN_DNAmAge_N2118_20230822

August 2023

University of Wisconsin Institute on Aging and UCLA Social Genomics Core Laboratory. NIA Grant #U19 AG051426

DNA Methylation Age Data Documentation

Summary: DNA methylation (DNAm) profiling was conducted on n=1,310 whole blood DNA samples from the MIDUS 2 (M2) and MIDUS Refresher 1 (MR1) samples in 2019. In 2022, the DNAm data were scored to compute 6 widely used measures of "epigenetic age" including the original measures from Horvath et al. [1] and Hannum et al. [2], as well as the PhenoAge [3] and GrimAge [4] scores and the DunedinPACE epigenetic age acceleration factor [5]. In 2023, a recently developed GrimAge2 score was computed [6] along with a revised version of the original GrimAge [4] (GrimAge1v2, derived from a new online score calculator). Scores and related technical variables are available in the following file:

M2MR1_GEN_DNAmAge_N2118_20230822.sav

This data file is available through the MIDUS Colectica Portal (http://midus.colectica.org/). It contains 6 epigenetic age scores ("epigenetic clocks"), as well as quality control metrics and technical control variables.

Variables are named according to MIDUS conventions (see the Naming and Coding Conventions included with the MIDUS Refresher Survey documentation), thus the variable names for the Methylation data begin with the unique 6 character set BRA6DM (B for M2 sample, RA for Refresher wave 1, 6 for Genomics project, D for data derived from DNA, M for Methylation).

The SPSS data file "M2MR1_GEN_DNAmAge_N2118_20230822.sav" contains 21 variables as follows:

- Administrative Variables:
  - M2MRID – contains the public identifiers for the MIDUS core sample and the MIDUS Refresher
  - SAMPLMAJ – the identifier created by the MIDUS Administrative Core to indicate the participants 'sample of origin' (e.g. MIDUS Refresher, Twin, etc.)
  - M2MRCASE – indicates if sample was obtained as part of the MIDUS 2 or MIDUS Refresher data collection
  - GENCONSENT – flag variable indicating whether the participant consented to genetics or not
  - BRA6DTISSUE – indicates the tissue sample source from which DNA was extracted
  - BRA6DMAVAIL – categorical flag variable indicating if DNA methylation data is available or not
- Technical Variables:
  - BRA6DMQCCORR – quality metric indicating sample DNAm profile correlation with a "gold standard" reference profile
  - BRA6DMQCDIF – quality metric indicating sample DNAm profile difference from a "gold standard" reference profile
  - BRA6DMQCFEMALE – quality metric estimating participant sex as inferred from X chromosome methylation abundance
  - BRA6DMDMXCHR – quantitative measure of X chromosome methylation

    - BRA6DMARRAYID – individual microarray used to assay samples (technical control variable)
    - BRA6DMPLATE – individual 96-well plate storing samples for assay (technical control variable)
    - BRA6DMWELL – sample's individual well within the 96-well plate (technical control variable)
- Age Score Variables:
    - BRA6DMDMAGEHORVATH - DNA methylation age - Horvath
    - BRA6DMDMAGEHORVATH2 - DNA methylation age – Horvath2 (Skin)
    - BRA6DMDMAGEHANNUM - DNA methylation age - Hannum
    - BRA6DMDMAGEPHENOAGE - DNA methylation age - PhenoAge
    - BRA6DMDMAGEGRIMAGE - DNA methylation age – GrimAge
    - BRA6DMDMAGEGRIMAGE2 - DNA methylation age – GrimAge2
    - BRA6DMDMAGEGRIMAGE1V2 - DNA methylation age – GrimAge1v2
    - BRA6DMDMAGEDUNEDINPACE - DNA methylation age – DunedinPACE

Details about the technical and epigenetic age variables are provided below.

The blood samples used for DNAm profiling were obtained, along with other samples, as part of a fasting blood draw completed in the morning of the second day of the Biomarker visit. Whole blood samples were collected using a BD Vacutainer Tube with EDTA anticoagulant, frozen for storage, and subject to DNA extraction and DNAm profiling as described below. Details about collecting and processing this sample are included in the MIDUS Refresher 1 Biomarker Project Blood, Urine, Saliva documentation which is available at ICPSR (https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36901) or via the MIDUS Colectica Portal (http://midus.colectica.org/). The portal houses interactive codebooks for all the publicly available MIDUS projects. The Portal includes search and explore functions, links to documentation, and a custom download function. A link to the portal is also available on the MIDUS website (http://midus.wisc.edu/) under QuickLinks.

Epigenetic Age Scores

Background

DNAm profiles refer to the pattern of variation in DNA methylation levels (0%-100%) at each of the many CpG nucleotide sequences scattered across the human genome. Patterns of DNAm change systematically with age in many biological tissues [7], including blood cells [8, 9]. These age-associated changes in DNAm are one example of "biological aging," which may proceed more rapidly or more slowly than chronological aging. To the extent that a DNAm profile ("epigenetic age") is older than the tissue donor's chronological age, then the tissue is said to show "epigenetic age acceleration" [7]. Epigenetic age acceleration can be assessed by the simple difference between biological age and chronological age, or by the residual from a regression of biological age on chronological age. Much research interest has focused on identifying environmental or psychosocial influences on epigenetic age acceleration (e.g., biological "weathering" by adverse environments or the "youth-preserving" effects of resilience factors). Several different measures of epigenetic age have been developed, using machine learning algorithms trained to predict criteria such as chronological age, risk of age-related disease, mortality, etc. These different measures of DNAm age are sometimes referred to as "epigenetic clocks." "First generation" epigenetic clocks such as the ones from Horvath et al. [1] and Hannum et al. [2] were

"trained" to predict chronological age, but epigenetic age acceleration measures derived from these clocks are only modestly predictive of future disease or mortality, and show only sporadic associations with psychosocial or environmental risk factors. "Second generation" epigenetic aging measures such as GrimAge [4, 6] and DunedinPACE [5] were developed to more effectively predict disease and mortality through a combination of age and DNAm correlates of health risk factors such as diet, smoking, adiposity, inactivity, inflammation, and metabolic dysregulation. As a result, second generation epigenetic age measures correlate with a range of environmental, behavioral, and psychosocial risk factors in addition to health and longevity. Both first and second generation epigenetic age measures are significantly influenced by the abundance of specific immune cell types within the assayed tissue sample (particularly neutrophils and monocytes), which increases systematically with age [10]. The DunedinPACE [5] score is distinct from the epigenetic clocks in measuring the relative pace of recent aging as a multiplicative factor, rather than an estimated biological age in years.

How were the MIDUS epigenetic age scores derived?

Participants in the MIDUS biomarker project provided whole blood samples from which DNA was later extracted, tested for suitable DNA yield and DNA integrity, and subjected to genome-wide methylation profiling using Illumina Methylation EPIC microarrays. The resulting "beta values" (estimated % methylation at each assayed CpG site) were normalized to control for technical sources of variance (using the noob function in the R minfi package), registered onto the list of CpG sites assayed on the Illumina Methylation 450K microarray (which is the basis for most epigenetic age scores), screened using standard quality control metrics for DNAm array data (all samples passed), and scored using previously published algorithms for 4 first-generation measures of epigenetic age (i.e., in years; Hannum et al. [2], Horvath et al. [1], Horvath et al.'s "skin and blood epigenetic clock" [11], and Levine et al.'s PhenoAge [3]), the second-generation GrimAge epigenetic clocks [4, 6], and the DunedinPACE measure of epigenetic age acceleration [5] (which yields a relative age acceleration factor rather than an estimated biological age in years). These epigenetic age measures are accompanied by a set of quality control metrics: measures of each sample's correlation with or difference from a "gold standard" reference blood DNAm profile, and estimates of biological sex based on quantitative variation in X chromosome methylation.

Quality control metrics were good for all samples, with only minor quantitative variations among samples. However, users should be aware that a small number of epigenome-based estimates of biological sex do not match biological sex reported from survey data. The epigenetic clocks were all tested for the expected correlation with chronological age (all $r > .9$) and also show high intercorrelation (generally around $r = .9$). The DunedinPACE age acceleration measure shows moderate correlation with chronological age and with the epigenetic clocks. Epigenetic age acceleration is slightly greater in males relative to females, which is expected given the shorter average lifespan of males.

References

1. Horvath, S., *DNA methylation age of human tissues and cell types.* Genome Biol, 2013. **14**(10): p. R115. PMC4015143
2. Hannum, G., et al., *Genome-wide methylation profiles reveal quantitative views of human aging rates.* Mol Cell, 2013. **49**(2): p. 359-367. PMC3780611
3. Levine, M.E., et al., *An epigenetic biomarker of aging for lifespan and healthspan.* Aging (Albany NY), 2018. **10**(4): p. 573-591. PMC5940111

4. Lu, A.T., et al., *DNA methylation GrimAge strongly predicts lifespan and healthspan.* Aging (Albany NY), 2019. **11**(2): p. 303-327. PMC6366976
5. Belsky, D.W., et al., *DunedinPACE, a DNA methylation biomarker of the pace of aging.* Elife, 2022. **11**. PMC8853656
6. Lu A.T., et al., *DNA methylation GrimAge version 2*. Aging (Albany NY). 2022. **14**(23): p. 9484-9549.
7. Seale, K., et al., *Making sense of the ageing methylome.* Nat Rev Genet, 2022. **23**(10): p. 585-605.
8. Heyn, H., et al., *Distinct DNA methylomes of newborns and centenarians.* Proc Natl Acad Sci U S A, 2012. **109**(26): p. 10522-7. PMC3387108
9. Zhao, M., et al., *Distinct epigenomes in CD4(+) T cells of newborns, middle-ages and centenarians.* Sci Rep, 2016. **6**: p. 38411. PMC5137168
10. Jaffe, A.E. and R.A. Irizarry, *Accounting for cellular heterogeneity is critical in epigenome-wide association studies.* Genome Biol, 2014. **15**(2): p. R31. PMC4053810
11. Horvath, S., et al., *Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies.* Aging (Albany NY), 2018. **10**(7): p. 1758-1775. PMC6075434