Age acceleration: data description

Marta Vives Usano, Carlos Ruiz, Mariona Bustamante

16/05/2017

The age acceleration variables were estimated using the Horvath and Hannum algorithms[1][2][3][4][5]. These algorithms estimate age from DNA methylation data and comparing it with the actual/chronological age provide a measure of age acceleration.

The Horvath method was developed using a dataset of roughly 4000 samples from 20 healthy tissues and cell types assembled using a combination of Illumina HumanMethylation27 BeadChip (27K) and Illumina HumanMethylation450 BeadChip (450K) technology. The Horvath method uses 353 CpG sited which are common to both platforms. More details in https://labs.genetics.ucla.edu/horvath/dnamage/. [4]

The Hannum method was developed based on a single dataset of 656 samples from adult whole blood assembled using Illumina HumanMethylation450 BeadChip (450K) technology. The Hannum method uses 71 CpG sited to estimate age. [5]

Long-term exposure is associated with age acceleration measures but many of these associations are relatively weak.[6] Also has been found that DNA methylation-based biological indicators for current and past smoking exposure are related to DNA methylation defined age acceleration. But no association is found with self-reported smoking indicators or serum cotinine levels.[7]

Age acceleration has been shown to be associated with obesity, lower physical and cognitive function, Alzheimer's disease, HIV, menopause and all-cause mortality.[8]

Table1. Characteristics of the subcohort sample, N=1192

Variable	Mean/N	SD/%	Min	Max
Age (years)	7.836	1.541	5.437	11.984
Epigenetic age (years)	7.832	2.086	3.557	18.894
Sex				
Male	653	54.78		
Female	539	45.22		
Cohort				
BIB	208	17.45		
EDEN	150	12.58		
KANC	200	16.78		
MOBA	214	17.95		
RHEA	200	16.78		
SAB	220	18.46		
Ethnicity				
African	7	0.59		
Asian	19	1.60		
Caucasian	1065	89.34		
Native_American	2	0.17		
Pakistani	81	6.79		
Other	18	1.51		

Age acceleration measures

The main age acceleration outcomes of the Hovarth algorithm are:

DNAmAge: It is referred to the predicted age based on DNA methylation

AgeAccelerationDiff: Age acceleration measure defined as difference (DNAmAge- Chronological age). A positive value indicates that the predicted age is higher than the chronological age.

AgeAccelerationResidual: It is the recommended and universal measure of age acceleration. It defined as the residual resulting from a linear regression model that regresses the Horvath estimate of epigenetic age on chronological age. A positive value indicates that the observed epigenetic age is higher than that predicted, based on chronological age.

AAHOAdjCellCounts and AAHAAdjCellCounts (known as intrinsic epigenetic age acceleration, IEAA): These measures estimate "pure" epigenetic aging effects that are not influenced by differences in blood cell counts. These are measures of age acceleration that adjust for cell counts. Specifically, these are residuals resulting from multivariate regression models that regress an estimate of DNAmage on age and various blood immune cell counts (naïve CD8+ T cells, exhausted CD8+ T cells, Plasmablasts, CD4+ T cells, Natural killer cells, monocytes and granulocytes). AAHOAdjCellCounts and AAHAAdjCellCounts correspond to age acceleration measures based on Horvath (2013) and Hannum (2013), respectively.

BioAge4HAStaticAdjAge (known as extrinsic age acceleration, EEAA): Extrinsec epigenetic age acceleration measures capture both cell intrinsic methylation changes and extracellular changes in blood cell composition. EEAA is defined as the residual variation resulting from a univariate model regressing the resulting age estimate on chronological age. First, the epigenetic measure from Hannum is calculated and then they increase the contribution of immune blood cell types to the age estimate by

forming a weighted average of Hannum's estimate with 3 cell types that are known to change with age (naïve cytotoxic T cells, exhausted cytotoxic T cells and plasmablasts).[6]

Final databases

They are:
- age_acceleration_subcohort.csv: 1192 samples (1X + 1A)
- age_acceleration_panel.csv: 305 samples (1A + 1B)

They contain:
- arrayName: Methylation array name where the sample was processed
- HelixID: HELIX ID variable. cohort abbreviation + numeric ID
- SampleID: HELIX sample ID variable. cohort abbreviation + numeric ID + period in capital letter
- LabID: laboratory ID. cohort abbreviation + numeric ID + period in lowercase letter
- cohort: name of the cohort (BIB, EDEN, KANC, RHEA, SAB)
- Period: 1X, 1A, 1B
- helix_sp: HELIX samples or extra HELIX samples
- e3_sex: sex
- h_ethnicity_c: ethnicity
- Age: chronological age in years
- DNAmAge: epigenetic age in years
- AgeAccelerationDiff: Difference between epigenetic age and chronological age
- AgeAccelerationResidual: residual from a linear regression modelthat regresses the Horvath estimate of epigenetic age on chronological age
- AAHOAdjCellCounts: residuals resulting from multivariate regression models that regress an estimate of DNAmage on age and various blood immune cell counts based on Horvath
- AAHAAdjCellCounts: residuals resulting from multivariate regression models that regress an estimate of DNAmage on age and various blood immune cell counts based on Hannum.

References

[1] Simpkin, A. J. *et al.* Prenatal and early life influences on epigenetic age in children: a study of mother-offspring pairs from two cohort studies. *Human Molecular Genetics.* 25:191-201 (2016).

[2] Marioni, R, E. *et al.* The epigenetic clock and telomere length are independently associated with chronological age and mortality. *International Journal of Epidemiology.* 0:1-9 (2016).

[3] Chen, B. H. *et al.* DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging.* 8:1844-1865 (2016).

[4] Horvath, S. *et al.* DNA methylation age of human tissues and cell types. *Genome Biology.* 14:1-19 (2013).

[5] Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular Cell.* 49(2): 359-367 (2013).

[6] Ward-Caviness, C. K. *et al.* Long-term exposure to air pollution is associated with biological aging. *Oncotarget.* 7(46): 74510-74525 (2016)

[7] Gao, X. *et al.* Relationship of tobacco smoking and smoking-related DNA methylation with epigenetic age acceleration. *Oncotarget.* 7:46878-46889 (2016).

[8] Simpkin A. J. *et al.* The epigenetic clock and physical development during childhood and adolescence: longitudinal analysis from a UK birth cohort. *International Journal of Epidemiology.* 0:1-10 (2017).