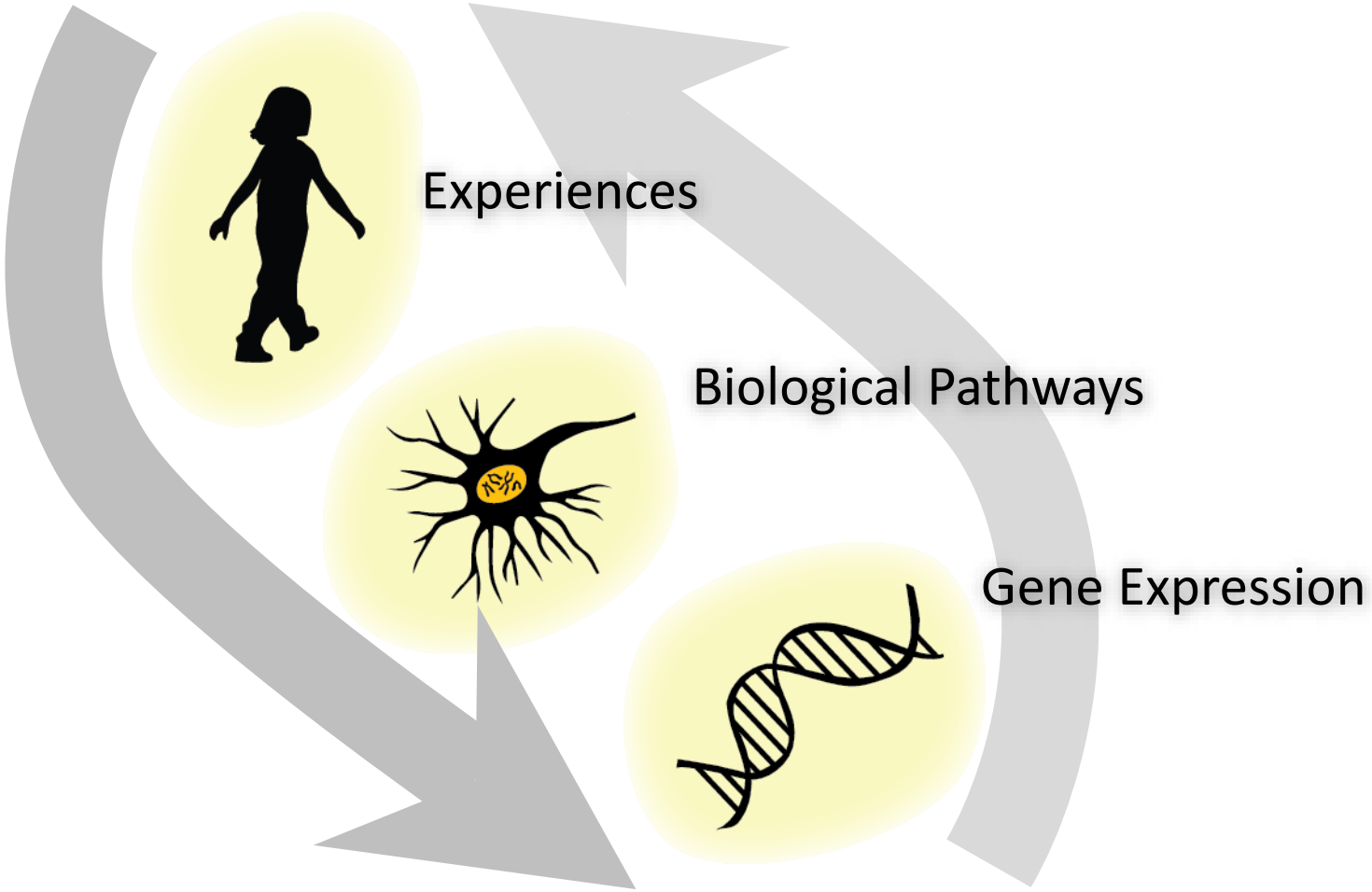
Stanford
MEDICINE

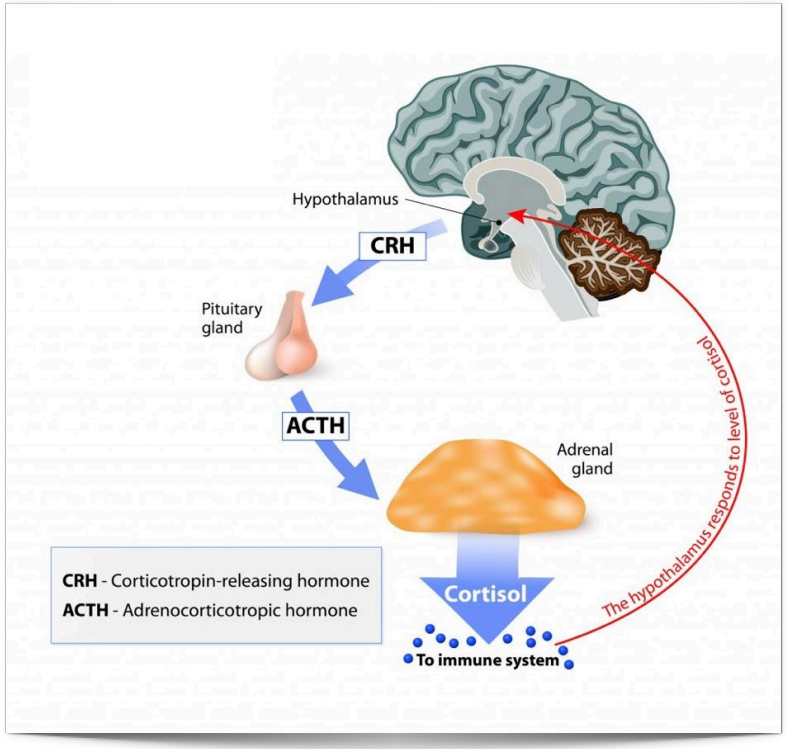
Epigenetic aging as a mediator of racial and ethnic inequalities in mortality: results from the National Health and Nutrition Examination Survey, 1999-2002 mortality follow up study

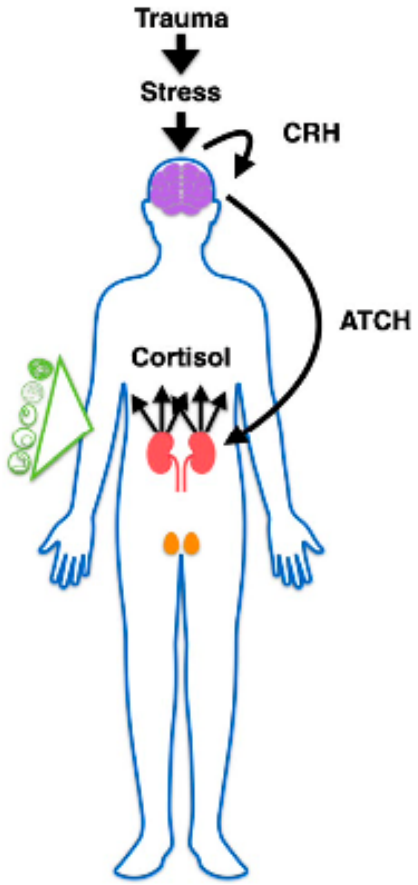
Social Epigenetics

The process by which life experiences influence the ways our genes are expressed.
Persistent and mitotically heritable alterations in genomic information that do NOT involve changes in DNA sequence.



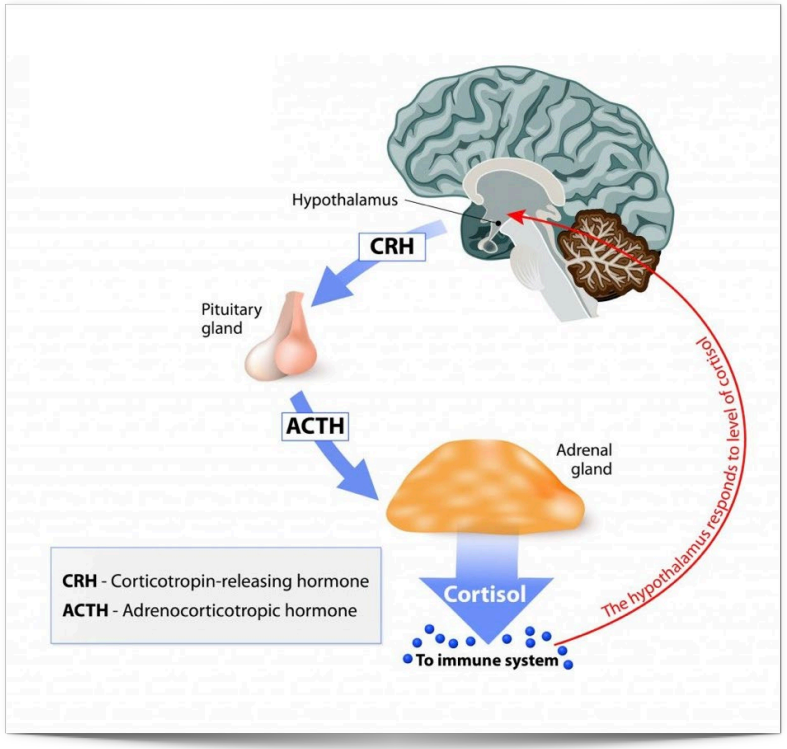
Racial and ethnic inequalities and disease risk

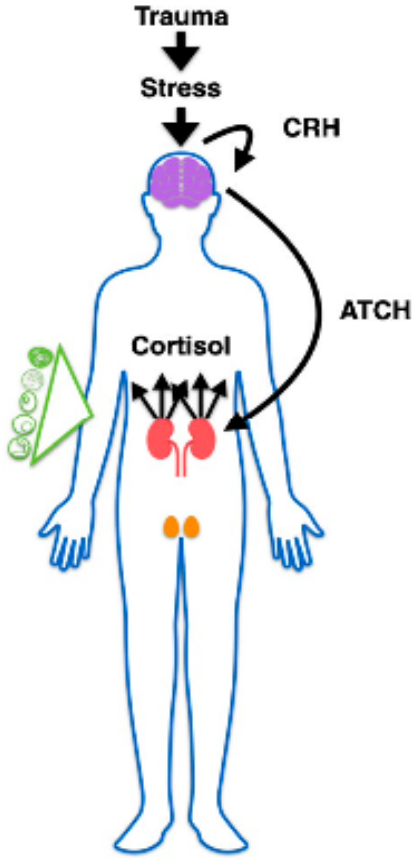




Social Disadvantage → Chronic Stress → Dysregulated Immune Function → Disease Risk

Epigenetic clocks as a mediator between racial/ethnic inequalities and disease



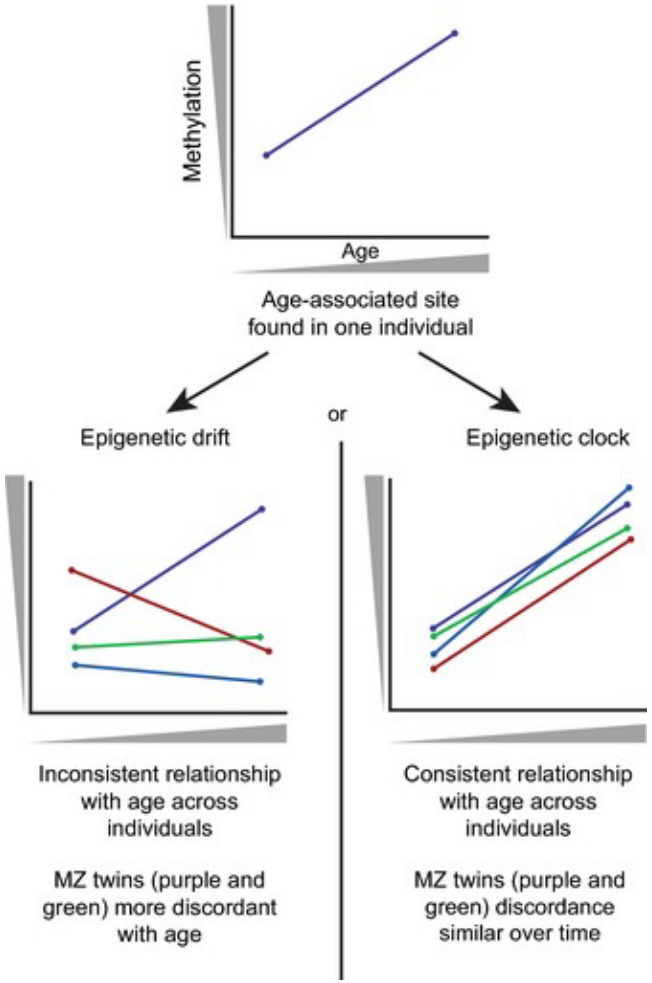


Social Disadvantage → Chronic Stress → Dysregulated Immune Function → Disease Risk

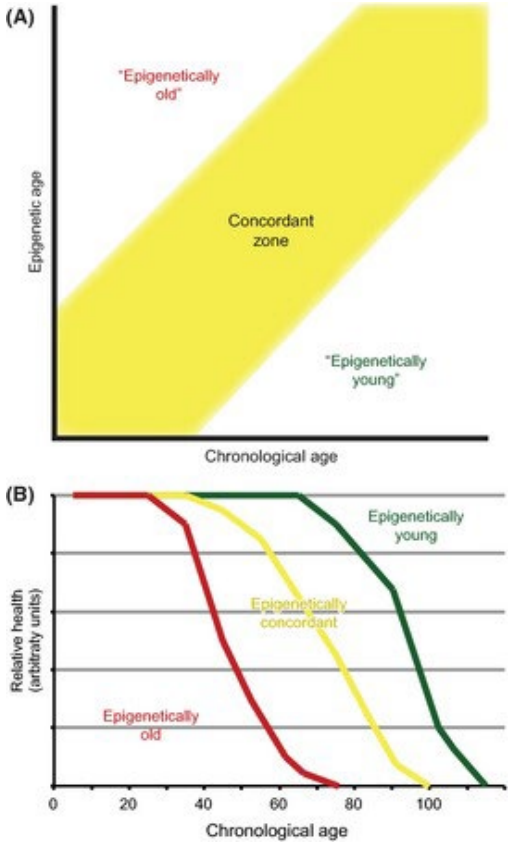
DNAm?

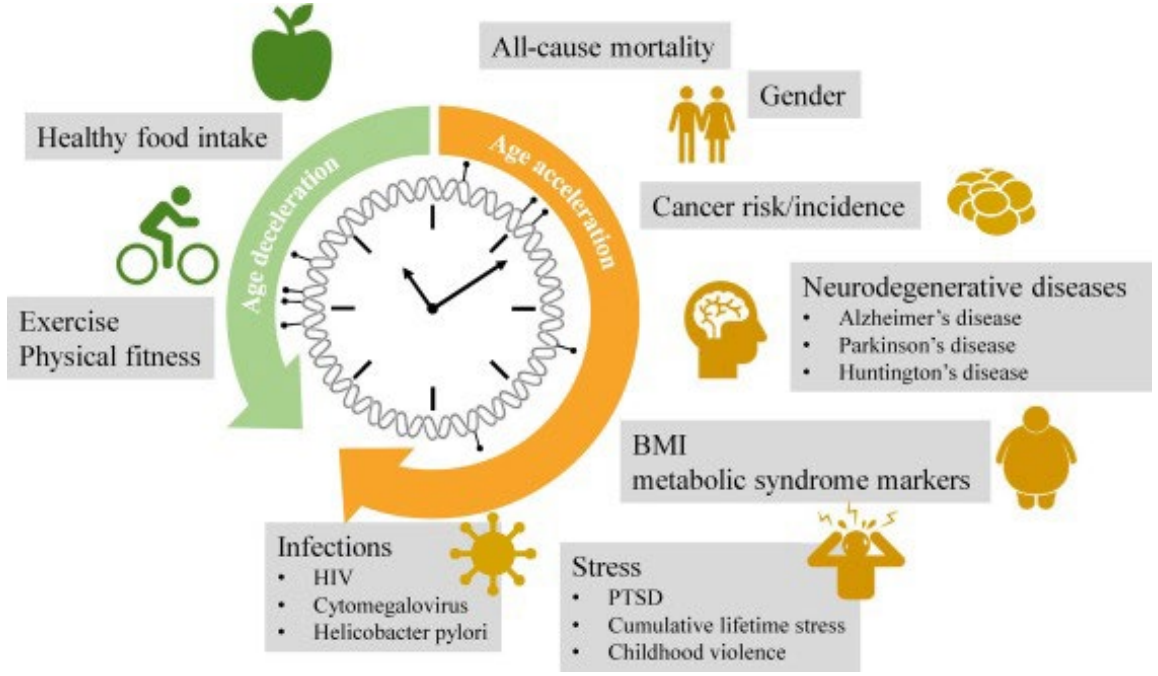
DNAm?

Epigenetic clocks: First generation



Jones, Goodman, and Kobor, 2015





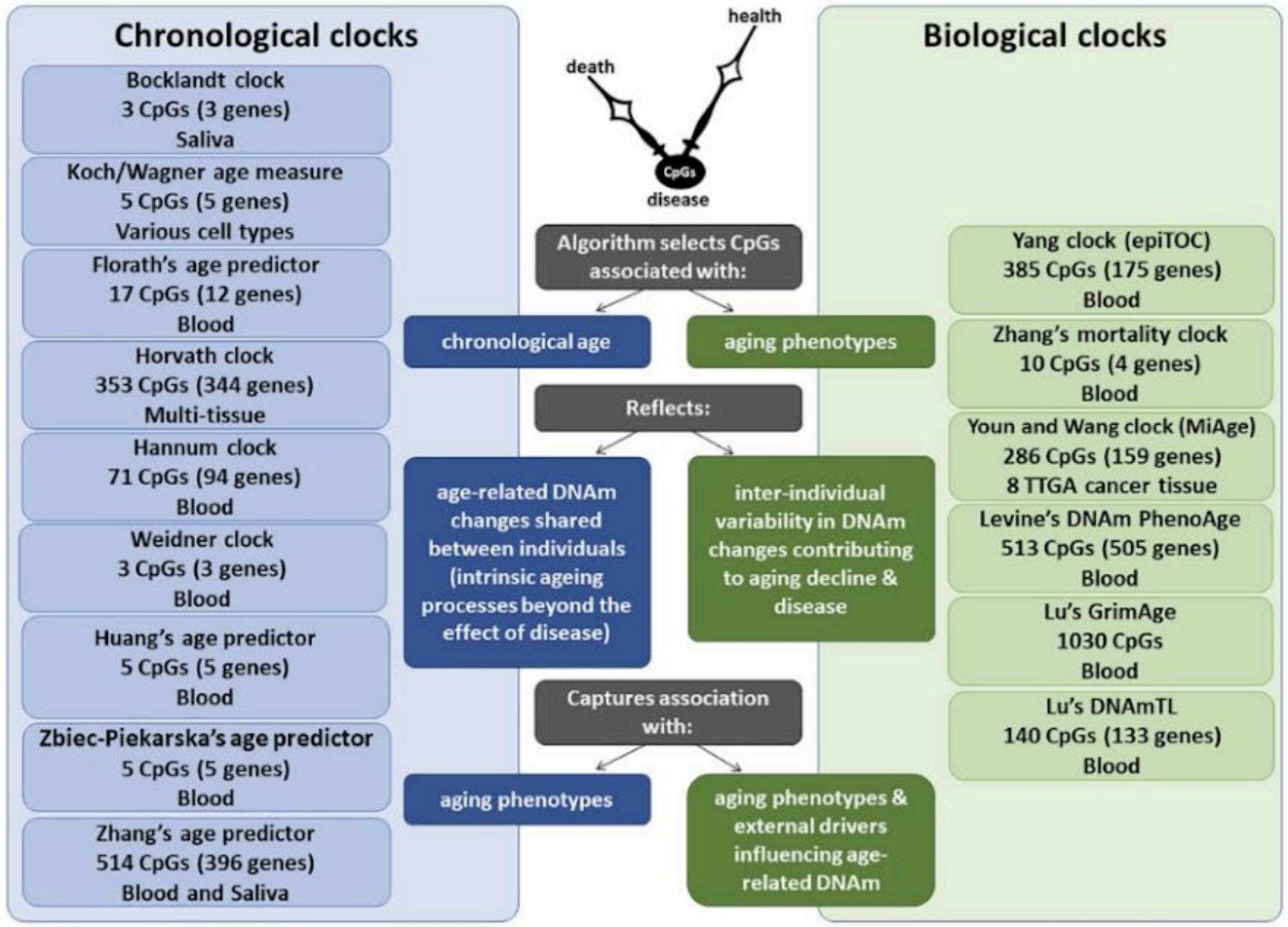
Declerck and Berghe, 2018

Epigenetic clocks: Second and third generation

1st Gen. Clock
Trained using **Chronological Age**
Age ~ CpG Methylation + Age + Sex + …

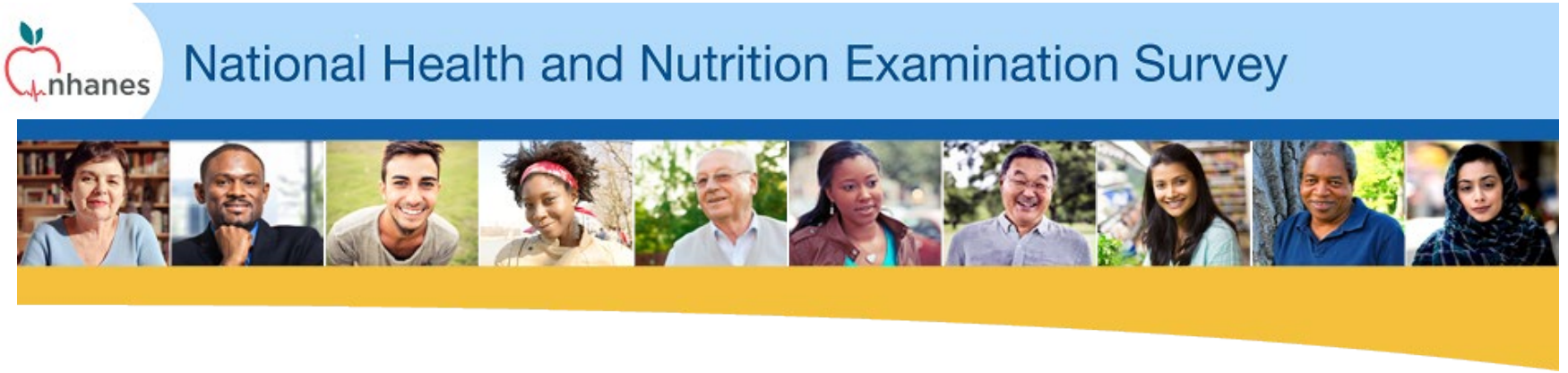
2nd Gen. Clock
Trained using **Aging Phenotypes**
Biomarker ~ CpG Methylation + Age + Sex + …

3rd Gen. Clock
Trained using Aging Phenotype and measurements, **produces a instantaneous rate of aging**



Bergesma and Rogaeva, 2020

NHANES background

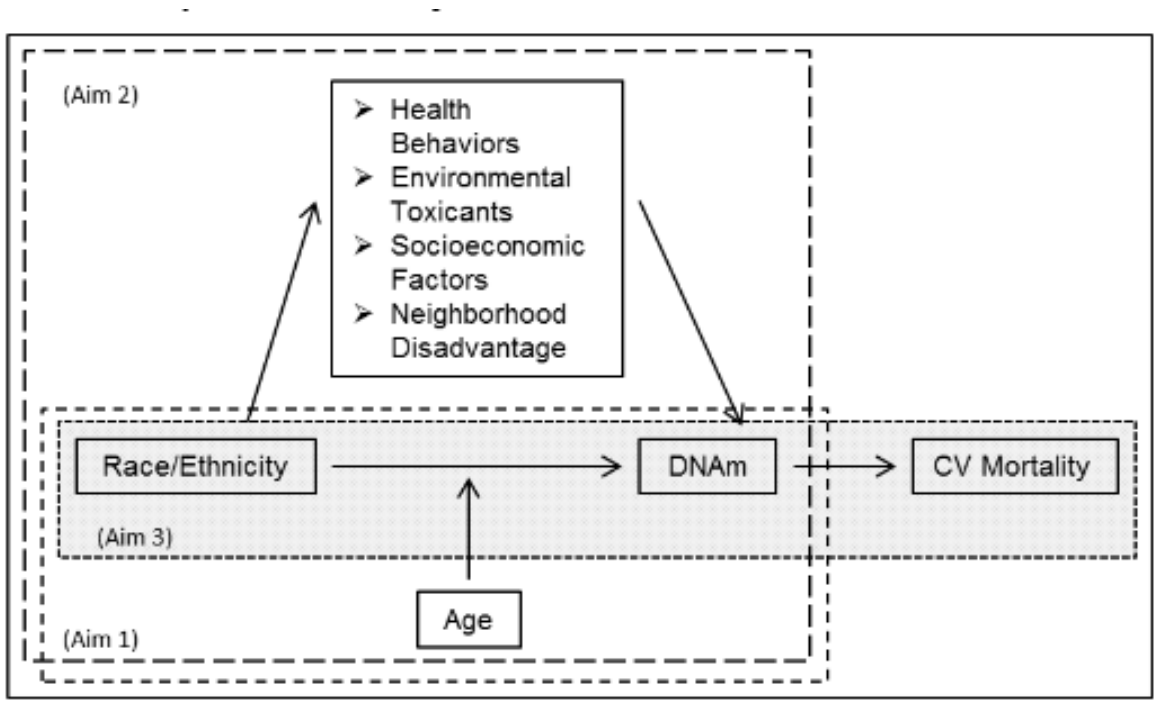


- This study started in the early 1960s and is led by the National Center for Health Statistics as part of the CDC to assess the health and nutritional status of adults and children in the United States.
- The ~5,000-person sample is drawn from the entire US population allowing results to be more generalizable than previously investigated.
- NHANES has detailed measures related to potential underlying causes of racial/ethnic disparities:
  - Health behaviors
  - Environmental toxicant exposures
  - Social and economic disadvantage
  - 20 years mortality follow-up

Study Design

- Sample size: 2,535
  - 42% non-Hispanic White, 22% non-Hispanic Black, 35% Hispanic
  - Age 50-85
  - From NHANES 1999-2002

- DNA methylation measured using the Illumina Infinium MethylationEPIC BeadChip
- 13 DNAm clocks and biomarkers will be generated
- A Health Disparities Epigenetic Mapping Project website will be developed to allow researchers easy access to findings.
- All DNAm generated will be made public through the CDC Research Data Center



Health and Retirement Study comparisons with NHANES

Variable	NHANES	HRS
Unique DNAm Biomarkers	Telomere, DunedinPACE	Garagnani, Bocklandt
Sample Size	2,535	4,018
Age	50-85 years (Median 67 y)	(Median 69 y)
Race/Ethnicity	42% non-Hispanic White, 22% non-Hispanic Black, 35% Hispanic	66% non-Hispanic White, 16% non-Hispanic Black, 14% Hispanic
Tissue Type Measured	Venous blood	Venous blood
Sex	49% Female	56% Female
Education	<High School (47%), High School (22%), Some College (18%), College + (14%)	<High School (17%), High School (52%), Some College (6%), College + (24%)

Health and Retirement Study comparisons with NHANES

Similarities	Differences
Nationally representative	NHANES cross-sectional; HRS longitudinal
Multidisciplinary content	NHANES more CV, lifestyle screens, and family history
Rapid public data release	NHANES reports on medications and mental health
National research data resource	HRS performs cancer screens
Relies on self-reporting	HRS reports on vaccinations
Oversample for minorities	HRS provides more economic details
Similar sampling sizes (6,000-10,000)	NHANES utilizes Medicare/Medicaid claims; HRS Veterans' Affairs
Perform physical examinations	HRS has birthdates

NHANES project to date

Clock	Group/Publication	Platform	Probe Number	Units	Train n	Test n	Age Range	Race	Tissue	Normalization	Error	Purpose
Horvath	Horvath	27K & 450K	353	Years	3,900	3,200	0-100	Unknown - 26%, White = 93%, Black = 3%, Asian = 1%, Hispanic = 3%	Pan Tissue (51 types)	Horvath's BMIQ	MAD = 3.6 years	Predict chronological age across tissues and a large age range
Hannum	Hannum/Zhang	450K	71	Years	482	174	19-101	White = 65%, Hispanic = 35%	Blood	Horvath's BMIQ	4.9 years	Predict chronological age.
PhenoAge	Liu/Levine	450K	513	Years	9,926	6,209	20-100	non-Hispanic African American & White	Blood	Horvath's BMIQ (multiple to test)	-	Predict phenotypic age which is a combination of chronological age and 9 mutli-system clinical chemistry biomarkers. Mortality risk.
BloodSkin	Horvath/Raj	450K & EPIC	391	Years	896	1,326	0-94	White, Indigenous, African American, Hispanic	Fibroblasts, keratinocytes, BECs, endothelial, lymphoblastoid, blood, saliva	Horvath's BMIQ	3.4 years	Predict chronological age in fibroblasts and cell types used in ex vivo studies.
Lin	Lin/Wagner	27K	99	Years	575	2,100	19-101	White	Blood	Background Subtracted	5 years	Predicts mortality risk.
Weidner	Weidner/Wagner	450K	3	Years	446		79	White	Blood	Background Subtracted	9-11 years	Predict chronological age.
Vidal-Bralo	Vidal-Bralo/Gonzalez	27K & 450K	8	Years	390	335, 92 & 557	20-89	White	Whole Blood	None?	MAD = 6.07 years	Predicts chronological age. Limited to sites that can be measured on MS-SNuPE
Yang	Teschendorff/Yang	450K	385	Beta level	650	300	70	White = 65%, Hispanic = 35%	Fetal tissue, 12 tissue types, assessed in blood	BMIQ	-	Mitotic-like clock trained in normal and cancer tissues predicting stem cell division by restricting to promoter CpGs in Polycomb group target genes.
Zhang	Visscher/Zhang	450K & EPIC	514	Years	13,661	-	2-104	White	Blood and Saliva (n = 260)	z-score	-	Building a more predictive epigenetic clock
GrimAge	Lu/Horvath	450K	-	Years	1,731	625	mean age 66	White	Blood	Horvath's BMIQ (but also mixed)		A composite biomarker based on 7 DNAm surrogates and smoking packing years to predict lifespan
DunedinPoAm	Belsky/Moffitt	450K & EPIC	173 & 46	Rate of Aging	810	-	26-38	White	Blood	Any normalization method	NA	Predicts the rate of change of 18 organ-system integrity indicators from ages 26, 32,38 years.
DunedinPACE	Belsky/Moffitt	EPIC	173	Rate of Aging	818	-	26-45	White	Blood	Any normalization method	NA	Predicts the rate of change of 19 blood biomarkers from ages 26, 32,38 and 45 years.
Telomere	Horvath/Lu	450K & EPIC	140	Kilobases	2,256	9,345	22-93	40% European, 60% African.	Blood	Mixed; GS, noob, BMIQ, watermelon.	-	To predict telomere length, trained on LTL (leukocyte telomere length) measured by southern blot terminal restriction fragments (TRFs).

- Epigenetic clocks and biomarkers have been produced – accessing the data for analysis in the secured CDC facilities is the final hurdle.
- Basic pre-processing, cell proportions and sex predictions completed.
- Data to be released soon.

Adjusting for age in the model

- They state that some variables are associated with age and so age should be adjusted for in the model.
- They calculate EAA in the same way:
  - Epigenetic Age Acceleration (EAA) = residuals(DNAmAge ~ Age)
- They recommend one of two models:
  - EAA ~ Variable + Age
  - DNAmAge ~ Variable + Age

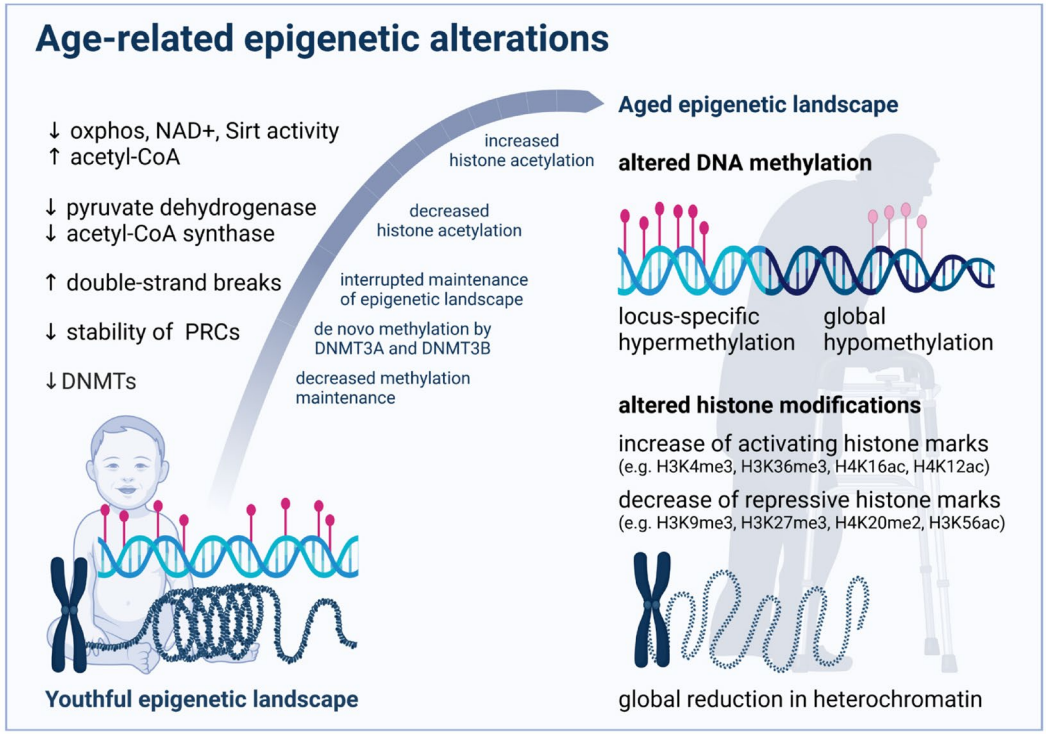
Practice of Epidemiology

Use of Correct and Incorrect Methods of Accounting for Age in Studies of Epigenetic Accelerated Aging: Implications and Recommendations for Best Practices

Nancy Krieger*, Jarvis T. Chen, Christian Testa, Ana Diez Roux, Kate Tilling, Sarah Watkins, Andrew J. Simpkin, Matthew Suderman, George Davey Smith, Immaculata De Vivo, Pamela D. Waterman, and Caroline Relton

* Correspondence to Dr. Nancy Krieger, Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Kresge 717, Boston, MA 02115 (e-mail: nkrieger@hsph.harvard.edu).

Accounting for non-linear effects of age



EPIGENETICS
2019, VOL. 14, NO. 9, 912–926
https://doi.org/10.1080/15592294.2019.1623634
Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Check for updates

Human epigenetic ageing is logarithmic with time across the entire lifespan

Sagi Snir[a], Colin Farrell[b], and Matteo Pellegrini[b]

[a]Department of Evolutionary Biology, University of Haifa, Haifa, Israel; [b]Department of Molecular, Cell and Developmental Biology, University of California, Los Angeles, CA, USA

Calculating epigenetic age acceleration to account for non-linearity:

- EAA2 = residuals(DNAmAge ~ Age + Age^2)

Two model suggestions:

1. EAA2 ~ Variable + Age + Age^2
2. DNAmAge ~ Variable + Age + Age^2

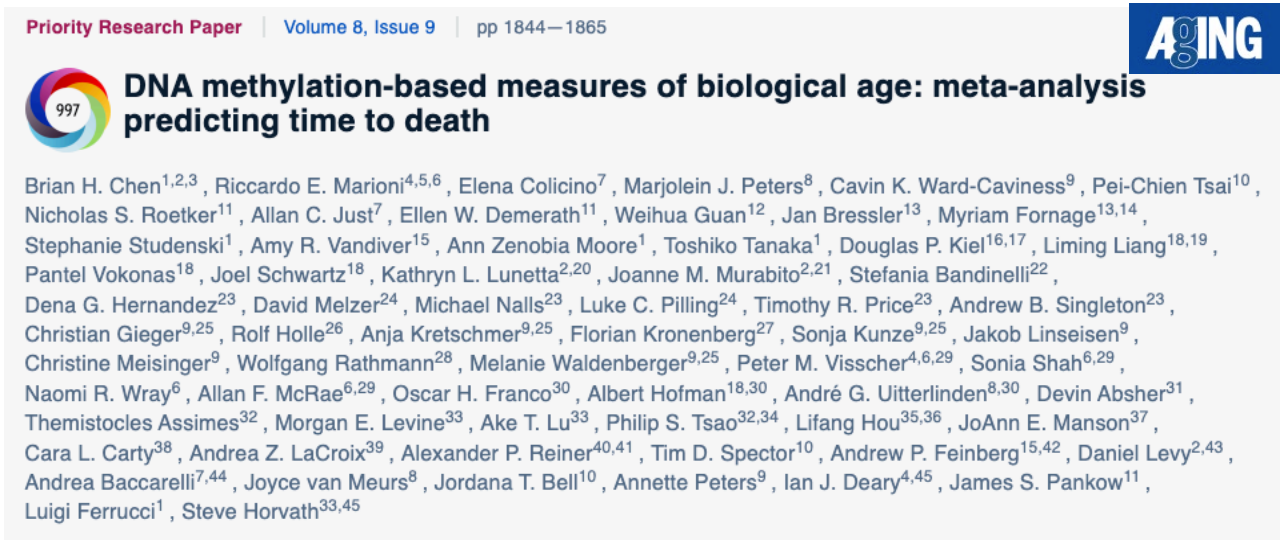
Including cell type proportions (CTPs)

Horvath calculator acknowledges cell types with following measures:

- EEAA = residuals(DNAmAge ~ Age)
- IEAA = residuals(DNAmAge ~ Age + CTPs)

Three model suggestions (all produce same results):

1. IEAA ~ Variable + Age + Age^2 + CTPs
2. EEAA ~ Variable + Age + Age^2 + CTPs
3. DNAmAge ~ Variable + Age + Age^2 + CTPs



Priority Research Paper | Volume 8, Issue 9 | pp 1844—1865

AGING

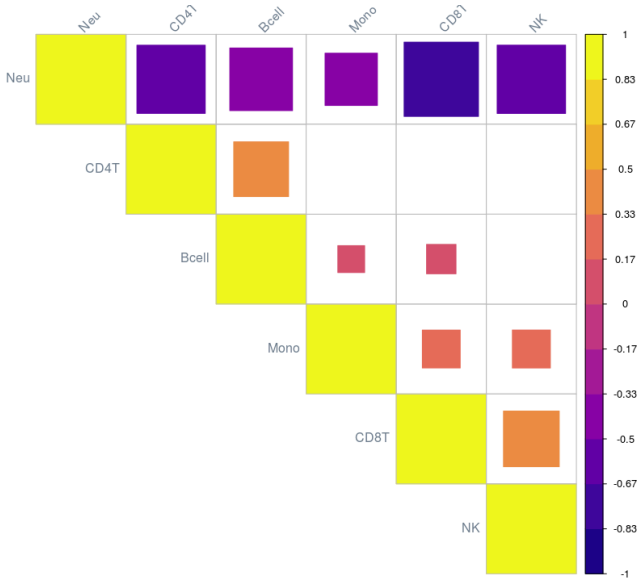
DNA methylation-based measures of biological age: meta-analysis predicting time to death

Brian H. Chen[1,2,3], Riccardo E. Marioni[4,5,6], Elena Colicino[7], Marjolein J. Peters[8], Cavin K. Ward-Caviness[9], Pei-Chien Tsai[10], Nicholas S. Roetker[11], Allan C. Just[7], Ellen W. Demerath[11], Weihua Guan[12], Jan Bressler[13], Myriam Fornage[13,14], Stephanie Studenski[1], Amy R. Vandiver[15], Ann Zenobia Moore[1], Toshiko Tanaka[1], Douglas P. Kiel[16,17], Liming Liang[18,19], Pantel Vokonas[18], Joel Schwartz[18], Kathryn L. Lunetta[2,20], Joanne M. Murabito[2,21], Stefania Bandinelli[22], Dena G. Hernandez[23], David Melzer[24], Michael Nalls[23], Luke C. Pilling[24], Timothy R. Price[23], Andrew B. Singleton[23], Christian Gieger[9,25], Rolf Holle[26], Anja Kretschmer[9,25], Florian Kronenberg[27], Sonja Kunze[9,25], Jakob Linseisen[9], Christine Meisinger[9], Wolfgang Rathmann[28], Melanie Waldenberger[9,25], Peter M. Visscher[4,6,29], Sonia Shah[6,29], Naomi R. Wray[6], Allan F. McRae[6,29], Oscar H. Franco[30], Albert Hofman[18,30], André G. Uitterlinden[8,30], Devin Absher[31], Themistocles Assimes[32], Morgan E. Levine[33], Ake T. Lu[33], Philip S. Tsao[32,34], Lifang Hou[35,36], JoAnn E. Manson[37], Cara L. Carty[38], Andrea Z. LaCroix[39], Alexander P. Reiner[40,41], Tim D. Spector[10], Andrew P. Feinberg[15,42], Daniel Levy[2,43], Andrea Baccarelli[7,44], Joyce van Meurs[8], Jordana T. Bell[10], Annette Peters[9], Ian J. Deary[4,45], James S. Pankow[11], Luigi Ferrucci[1], Steve Horvath[33,45]

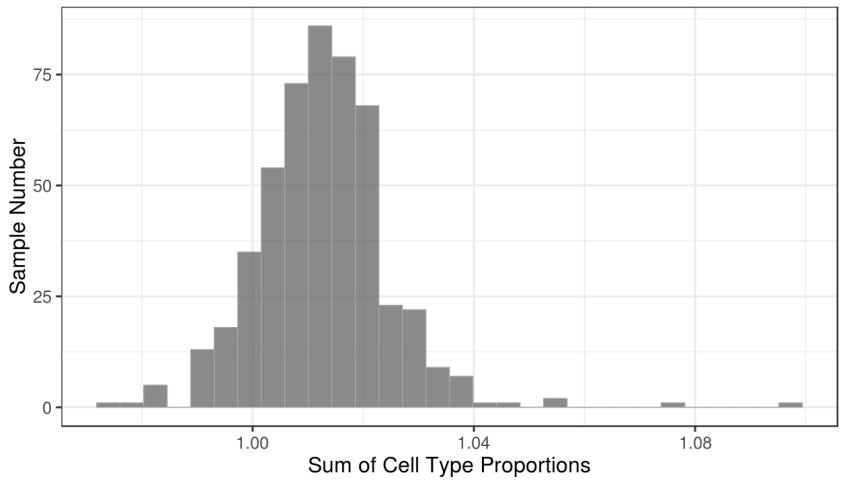
Modelling cell type proportions

Assumptions for linear regression:

- Linear relationship – Likely met
- Multivariate Normality – Likely met
- Homoscedasticity – Likely met
- No or little multicollinearity – Problematic



CTPs are highly correlated



CTPs are compositional data

Modelling cell type proportions

Proposed solutions so far:

1. **Raw** – solves neither issue
2. **Regress out cell type prior to running model** – removes variation instead of accounting for it
3. **Remove cell types until correlated below threshold** – some variation left unadjusted, how to pick
4. **PCs-1 from PCA of CTPs** – solves correlation only as returns compositional data almost unchanged
5. **PCs-1 from PCA on isometric log ratio transformed CTPs** – solves both issues but not as interpretable
6. **Constrained (log-contrast) linear regression with no intercept with raw CTPs** – solves both issues but not as interpretable

Acknowledgments

Stanford MEDICINE | Epidemiology & Population Health

SCHOOL OF PUBLIC HEALTH
UNIVERSITY OF MICHIGAN

Department of Medicine
Duke University School of Medicine

David Rehkopf, ScD

Belinda Needham, PhD

Yongmei Liu, MD, PhD

Hanyang Shen, MPH

Kurt Lohman