THE GERONTOLOGICAL SOCIETY OF AMERICA®

Journals of Gerontology: Biological Sciences
cite as: *J Gerontol A Biol Sci Med Sci*, 2020, Vol. 75, No. 3, 463–465
doi:10.1093/gerona/glaa021
Advance Access publication January 29, 2020

OXFORD

Guest Editorial

Assessment of Epigenetic Clocks as Biomarkers of Aging in Basic and Population Research

Morgan E. Levine, PhD*

Department of Pathology, Yale University School of Medicine, New Haven, Connecticut.

*Address correspondence to: Morgan E. Levine, PhD, Yale University School of Medicine, New Haven, CT. E-mail: morgan.levine@yale.edu

The Geoscience paradigm suggests that targeting the aging process could delay or prevent the risk of multiple major age-related diseases, such as cancer, cardiovascular disease, Alzheimer's disease, diabetes, and osteoporosis, to name only a few(1,2). These diseases have well-established clinical diagnostic and classification criteria, yet despite the fact that aging may be the key driver across these diverse pathologies, aging remains a latent concept with no agreed upon molecular definition. If our goal is to develop interventions to slow the increase in biological aging as a function of chronological time, and thus facilitate health promotion and disease prevention, we need to first come up with clinically valid measures of the underlying biological process and/or classification criteria for what it means to be biologically, rather than chronologically, "aged." This collection of primary papers and reviews focuses on DNA methylation and its promise as a means to define biological age.

As discussed in the primary research paper by Nelson et al. (3), developing biomarkers of aging would serve three important purposes: First, biomarkers can provide both diagnostic and prognostic insight that may inform medical and/or personal decisions; second, they may uncover mechanistic clues that aid basic research; and third, they have the potential to inform intervention strategies. As a result, the quest to develop both valid and reliable biomarkers of aging has been gaining steam over the past decade. While measures of biological age have been developed using a variety of data sources—from proteomics (4), to microbiological profiles of gut microbiota (5), to facial images (6)—at the forefront of the biomarker quest sit the "epigenetic clocks." These aging measures are based on composite scores that combine information on DNA methylation (DNAm) levels at tens to hundreds of CpG dinucleotide locations across the genome (7). Since the first epigenetic clock was developed by Bocklandt et al. in 2011 (8), numerous age predictors based on DNAm have been developed and or applied across a variety of mammalian species (9–12).

While epigenetic clocks routinely exhibit extremely high precision when it comes to age prediction, many scientists are starting to recognize that perhaps the most important test of these proposed biomarkers is not how well they predict chronological time, but rather how well they predict aging outcomes, above and beyond age itself. With the growing inclusion of DNAm data in many of the existing large human cohort studies, researchers are beginning to test the following—when considering individuals of the same chronological age, do those with higher epigenetic age look phenotypically older on average (eg, have higher mortality rates, greater disease burden, and worse physical and cognitive functioning)? This question is at the heart of this special collection, describing evidence for (or in some cases detracting from) the utility of epigenetic clocks for measuring biological aging in blood.

The papers by Bressler et al. (13) and Ryan et al. (14) test the utility of what we consider the first-generation epigenetic clocks—the blood clock by Hannum et al. (2013) (12), and the pan-tissue clock by Horvath (2013) (11). Early epigenetic clocks have been utilized in multiple epidemiological studies (mostly using blood and saliva) to evaluate morbidity and mortality associations. In their systematic review, Ryan et al. (14) conducted a meta-analysis across 61 studies to test for associations between the epigenetic clocks by Hannum and Horvath when it comes to factors pertaining to environment, lifestyle, and health. One of the most robust findings was for body mass index (BMI), in which results suggested that higher BMI was consistently associated with increased epigenetic age acceleration (defined as the residual when clock scores are regressed on age). Frailty was also reliably associated with higher epigenetic age acceleration across three studies (two using Horvath and one using a derivation of Hannum, called EEAA). Unfortunately, for many of the other variables, results were inconclusive.

Although many of the studies utilized in the meta-analysis by Ryan et al. (14) did not disclose the racial/ethnic makeup of their samples, the authors speculate that the majority of results are likely based on associations in Caucasians and may not generalize to other groups. The tendency towards Caucasian-specific finings is an issue that has plagued recent genetic association studies; however, there is some evidence that epigenetic clock associations may show less racial/ethnic bias, thanks to the inclusion of more diverse, multiethnic samples when developing these measures. To evaluate this assumption, the paper by Bressler et al. (13) tested associations between cognitive functioning and epigenetic age using a sample of just over 2000 middle-aged to older African

Americans from the Atherosclerosis Risk in Communities (ARIC) Study. They found that epigenetic age, using the Hannum clock, was positively associated with a measure of cognitive functioning (the Word Fluency Test), independent of age, sex, and education. They then replicated this finding in a European sample from the Generation Scotland: Scottish Family Health Study. Overall, the results suggest that older epigenetic age, as measured by Hannum but not Horvath, is associated with decreased verbal fluency and that this generalizes to both European and African ancestry samples, and cannot be accounted for by cognitive differences as a function of education, sex, or age.

The epigenetic clock by Horvath, and to some extent, the one by Hannum, have been the most widely applied in both epidemiological and basic research. However, this is perhaps due more to their prominence in the field rather than perhaps their utility. The papers by both McCrory et al. (15) and Maddock et al. (16) go one step further by exploring associations with both the first-generation epigenetic clocks that were trained to predict chronological age, as well as the second-generation clocks that were trained to predict aging correlates, like morbidity and mortality. The paper by McCrory et al. (15) sought to link three clocks—Horvath, Hannum, and Levine (also commonly referred to as the PhenoAge clock) (17)—to another biomarker often used in epidemiolocal and sociological studies to model aging, Allostatic Load (AL)(18,19). AL was found to be associated with the Levine clock, but not either of the chronological age prediction clocks. However, associations for both the Horvath and Levine clocks were found when stratifying by sex and examining the system-specific AL components. For instance, in males, metabolic dysregulation was associated with Horvath and Levine, while in females, only Levine was related to metabolic and cardiovascular dysregulation. The more robust associations with AL found for Levine over the other two clocks are not surprising, as Levine was developed to predict a multisystem measure (phenotypic age) that was composed of many of the same clinical measures found in the AL composite score.

Differences in associations and predictive performance in blood between first- and second-generation epigenetic clocks were also observed in the paper by Maddock et al. (16). In this case, the authors examined four clocks—Horvath, Hannum, Levine, and GrimAge. GrimAge combines DNAm-based plasma protein estimates, DNAm pack-year estimates, and observed age and sex as a function of mortality risk (20). As such, GrimAge has been shown to be a very powerful morbidity and mortality predictor. The paper by Maddock et al. (16) provided additional evidence for this by showing that higher epigenetic age using the second-generation clocks (Levine and GrimAge), but not the first-generation clocks (Hannum or Horvath), was related to worse physical and cognitive functioning, assessed using such as grip strength (Levine), FEV1 (Levine and GrimAge), mental speed (Levine and GrimAge), and episodic memory (GrimAge). The authors also examined longitudinal changes in these measures and found that FEV1 declined at a faster rate for individuals with higher baseline GrimAge and/or Levine DNAmPhenoAge. A similar finding was observed for the decline in grip strength as a function of GrimAge.

The epidemiological evidence for more robust aging associations among second-generation clocks may not simply boil down to the use of aging correlates rather than chronological age to train them. As illustrated in the paper by Nelson et al. (3), the power of these measures as diagnostic and prognostic indicators, may stem from the use of longitudinal data in training them. The majority of aging biomarkers, including the first-generation epigenetic clocks, are developed using cross-sectional data, in which the researchers take a variable that proxies aging (eg, chronological age) and apply supervised machine learning, or deep learning, approaches to predict that variable using tens to hundreds of thousands of input variables. The problem with this approach, that was well-argued by Nelson et al., is that it does not account for mortality selection. As such, this biases the algorithm to select markers that are not causal, but instead correlative with aging. The reason for this is that individuals with accelerated aging rates will on average experience higher mortality burden, and therefore be selected out of the population earlier in the life course. As a result, loci that show consistent trends with chronological age, even at higher ages, are likely not causal—causal loci should exhibit diminishing age prediction in later life because the individuals who harbored them were more likely to die, and thus over time, they should be progressively selected out of the cohort. Nelson et al. (3) further illustrated this in simulations showing that by using a cross-sectional study design for biomarker development there was a propensity away from selecting causal loci, to the point where fewer causal loci were selected than if loci had been chosen at random.

Collectively, these papers raise two very important considerations in light of the ever-increasing quest to develop biomarkers of aging. The first corresponds to the concerns raised by Nelson et al. (3) and pertains to how biomarkers are developed. Given the multifaceted changes that occur with aging, there are conceivably thousands of data types that could be used to develop age predictors, and in fact, this is something we are already seeing. To date, there exist countless measures developed as chronological age predictors, using a variety of input data. However, as suggested by Nelson et al. (3), most of these composite age predictors are likely enriched for variables that are not causally involved in the biological aging process, making them less suitable for the aims they were developed to address. Rather than continuing to train chronological age predictors using diverse data, it may be more advantageous to retrain some of the existing measures by predicting longitudinal outcomes. Doing so will hopefully enable us to develop composite measures and pinpoint specific markers that can serve as prognostic tools and therapeutic endpoints, while also informing our understanding of basic biology of aging.

The second important consideration that surfaced in this collection of papers has to do with the application of these measures in epidemiological and demographic studies. While it is not critical that population researchers have a complete understanding of the underlying biology captured in the tools they are applying, knowing the advantages and drawbacks of different measures for answering their research questions will be critical. Even among the epigenetic clocks, there are clear distinctions in regards to associations with various aging outcomes. As such, population researchers should either select a measure based on a priori assumptions of its applicability, or, as many of the papers in this issue have done, test multiple measures simultaneously. In doing so, population researchers might accelerate evaluation of the utility of biomarkers of aging and thus provide better direction for data scientists and biologists working to develop more robust and valid measures of the complex process that encompasses biological aging.

Conflict of Interest

M.E.L. is a consultant for Elysium Health and Life Epigenetics, which are for profit companies involved in consumer-based testing using Epigenetic Clocks. M.E.L. also holds IP pertaining to Epigenetic Clocks she has helped to develop.

References

1. Kennedy BK, Berger SL, Brunet A, et al. Geroscience: linking aging to chronic disease. *Cell*. 2014;159:709–713. doi: 10.1016/j.cell.2014.10.039
2. Sierra F. The emergence of geroscience as an interdisciplinary approach to the enhancement of health span and life span. *Cold Spring Harb Perspect Med*. 2016;6:a025163. doi: 10.1101/cshperspect.a025163
3. Nelson PG, Promislow DEL, Masel J, et al., Biomarkers for aging identified in cross-sectional studies tend to be non-causative. *J Gerontol A Biol Sci Med Sci*. 2019:glz174. doi:10.1093/gerona/glz174
4. Tanaka T, Biancotto A, Moaddel R, et al.; CHI consortium. Plasma proteomic signature of age in healthy humans. *Aging Cell*. 2018;17:e12799. doi: 10.1111/acel.12799
5. Galkin F, et al., Human microbiome aging clocks based on deep learning and tandem of permutation feature importance and accumulated local effects. bioRxiv, 2018;507780.
6. Chen W, Qian W, Wu G, et al. Three-dimensional human facial morphologies as robust aging markers. *Cell Res*. 2015;25:574–587. doi: 10.1038/cr.2015.36
7. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet*. 2018;19:371–384. doi: 10.1038/s41576-018-0004-3
8. Bocklandt S, Lin W, Sehl ME, et al. Epigenetic predictor of age. *PLoS One*. 2011;6:e14821. doi: 10.1371/journal.pone.0014821
9. Meer MV, et al. A whole lifespan mouse multi-tissue DNA methylation clock. *Elife* 2018;7:e40675.
10. Petkovich DA, Podolskiy DI, Lobanov AV, Lee SG, Miller RA, Gladyshev VN. Using DNA methylation profiling to evaluate biological age and longevity interventions. *Cell Metab*. 2017;25:954–960.e6. doi: 10.1016/j.cmet.2017.03.016
11. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14:R115. doi: 10.1186/gb-2013-14-10-r115
12. Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 2013;49:359–367. doi: 10.1016/j.molcel.2012.10.016
13. Bressler J, Marioni RE, Walker RM, et al. Epigenetic age acceleration and cognitive function in African American adults in midlife: the atherosclerosis risk in communities study. *J Gerontol A Biol Sci Med Sci*. 2020;75:473–480. doi:10.1093/gerona/glz245
14. Ryan J, Wrigglesworth J, Loong J, Fransquet P, Woods RL. A systematic review and meta-analysis of environmental, lifestyle, and health factors associated with DNA methylation age. *J Gerontol A Biol Sci Med Sci*. 2020;75:481–494. doi:10.1093/gerona/glz099
15. McCrory C, Fiorito G, McLoughlin S, et al. Epigenetic clocks and allostatic load reveal potential sex-specific drivers of biological aging. *J Gerontol A Biol Sci Med Sci*. 2020;75:495–503. doi:10.1093/gerona/glz241
16. Maddock J, Castillo-Fernandez J, Wong A, et al. DNA methylation age and physical and cognitive aging. *J Gerontol A Biol Sci Med Sci*. 2020;75:504–511. doi:10.1093/gerona/glz246
17. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*. 2018;10:573–591. doi: 10.18632/aging.101414
18. McEwen BS. Stress, adaptation, and disease. Allostasis and allostatic load. *Ann N Y Acad Sci*. 1998;840:33–44. doi: 10.1111/j.1749-6632.1998.tb09546.x
19. Seeman TE, McEwen BS, Rowe JW, Singer BH. Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. *Proc Natl Acad Sci USA* 2001;98:4770–4775. doi: 10.1073/pnas.081072698
20. Lu AT, Quach A, Wilson JG, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)*. 2019;11:303–327. doi: 10.18632/aging.101684