

A Clinical Trial using Methylation Age to Evaluate Current Anti-Aging Practices

(preprinted from Rejuvenation Research, 2019)

Josh Mitteldorf*†

* School of Medicine, Washington University of St Louis 63130 USA

† National Institute of Biological Sciences, Beijing 102206 China

Abstract

Recent advances in the technology of “aging clocks” based on DNA methylation suggest that it may be possible to measure changes in the rate of human aging over periods as short as a year or two. To the extent that methylation (and other biomarkers) are valid surrogates for biological age, the testing of anti-aging interventions has thus become radically cheaper, faster, and more practical. Together with colleagues at UCLA, I am organizing a clinical trial to evaluate some of the most popular anti-aging strategies currently deployed by “early adopters” in the lay community of personal health activists. We are recruiting 5,000 subjects, age 45-65, and interviewing them in detail about their diet, drugs and supplements, exercise, social, and other practices that plausibly contribute to modulate the rate of aging. They agree to submit blood samples for analysis of methylation age at the beginning, middle, and end of a two-year test period, together with 23-and-me or equivalent genetic data. Primary endpoint is the difference in methylation age over the course of two years. We are in the process of developing a specialized clock, optimized for individual differences over time. Results will be viewed as an exploratory study to identify synergistic combinations of age-retarding treatments. It is our expectation that there is a great deal of redundancy in the strategies that have been researched and promoted to the aware public; thus, most combinations can retard the rate of aging by only a few percent, consistent with the best-known single measures. However, it is our hope that among the many strategies that our subjects have adopted, there will be some combinations that synergize and achieve age retardation by 25 percent or more. A mock-up analysis of computer-generated data has been used to fix parameters of the study and confirm that we will be able to detect such combinations with good probability, should they exist. All data (redacted for privacy) will be open-sourced, available to the scientific community and to the public.

Introduction

In the body of this article, I will argue

1. That testing, more than development, is the bottleneck in the flow of knowledge about human anti-aging;
2. That single interventions are unlikely to afford major increments in life expectancy in humans;
3. That interactions among combinations of known anti-aging interventions are the most important unknown in the field;
4. That the daunting number of combinations may be addressed by enrolling large numbers of early adopters who are already using diverse combinations of strategies;
5. That the newest methylation clock, called “DNAm PhenoAge” has the potential to tell us which of these people are best succeeding in their quest to slow the aging clock;
6. That further optimization of this clock, specialized to the proposed application, is feasible; and
7. That exploratory data analysis can be used to identify the best combinations of known interventions that are already being deployed by members of the community which actively seeks to enhance their long-term health.

These ideas have been integrated into the design of a clinical trial based on collection of information about what well-informed and highly-motivated individuals are already doing to slow their own aging. We do not ask participants to do anything differently from what they are already doing; this is a clinical trial based on questionnaires and methylation testing only. The project has benefited from generous contributions of time and an in-kind donation from Zymo Research, which has agreed to provide methylation testing.

Background

What strategy or combination of strategies has the most potential for extending human lifespan? In the past, answering this question with epidemiological studies has been enormously slow and expensive; hence we have relied on a patchwork of reasonable guesses, based on animal studies and biochemistry, rather than human epidemiology. To evaluate an aging intervention in humans would require treating a test group and identifying a comparable untreated control group. Then there is a long latency period before a small percentage of the control group dies of age-related disease and (if the treatment is effective), a smaller percentage of the test group dies during the same period. Statistical significance depends on the difference between the small mortality rates in these two groups. Only the subjects who die offer statistical information, while the much larger number who do not die are a cost of doing business.

A numerical example will help you see why epidemiological studies of age-related

mortality are so slow and expensive. Suppose you start with 10,000 60-year-olds. You assign 5,000 to your test group and another 5,000 to your control group. If you wait 10 years, you expect 320 of the controls to die in that time. Suppose that your anti-aging treatment effectively slows the rate of aging decline by 20%. (You would certainly like to be able to detect effects of such magnitude.) The result would be that in the treated group, the expected number of people that die might be 255 instead of 320. The difference between 255 and 320 is marginally statistically significant by the standard criterion ($p < 0.05$). The numbers 320 and 255 are averages, and in a real test the actual results would be larger or smaller than these averages. There is a 20% probability that the numbers in the test and control groups would be too close to detect a significant difference, which is to say that this study with this number of subjects would have an 80% chance of detecting the efficacy of your intervention.

Following 10,000 people for 10 years costs ~ \$100 million at an American medical school. One might reduce the numbers by starting with 75-year-olds rather than 60-year-olds, because the mortality rate of the control group is much higher, offering more statistical certainty. On the other hand, there are treatments that work for 60-year-olds that do not work for 75-year-olds, since it is easier to maintain a healthy condition than to restore health once it is lost.

Given the prohibitive cost and time delays, epidemiological testing of any substantial number of treatment combinations has been out of the question. Hence anti-aging science has relied on a hierarchy of test animals, from yeast at the bottom of the pyramid to mice at the top. Only a few studies have been done with dogs, and just two replicates with Rhesus monkeys, the latter at enormous expense with marginal results [1, 2].

Recent improvements in the accuracy of age estimation based on methylation offer the possibility of disruptive change in the way anti-aging treatments are evaluated. A much smaller test population is needed, because every subject contributes data, not just the ones who have died. The time required for the study is reduced from decades to years, and the cost is greatly reduced. It has been estimated [3] that the time necessary for a study is reduced by a factor of ten and the cost by a factor of 100.

Methylation testing has not been used previously to evaluate any health intervention; there is evidence but yet no proof that treatments extending lifespan will tend to retard methylation age, or, conversely, that treatments that retard methylation age will tend to extend lifespan. Nevertheless, we have chosen to take a chance on an unproven test paradigm because the implications of a successful trial justify the risk. We regard the test as both a proof of principle that methylation testing can be used to evaluate anti-aging measures in humans, and as an exploratory study to identify promising combinations of measures already in use.

Our test population

A loose community already exists of thousands of people who identify as early adopters and are currently experimenting on themselves in the hope of extending their lifespans and healthspans. These people are generally well-informed regarding the latest research findings, but they have different belief systems, different information and different resources; hence there is a good deal of variation in their regimens of diet, exercise, supplements and prescription medications. We will be recruiting subjects from sources including the following:

- Longecity
- Gerontology Research Group
- Customer base of Life Extension Foundation
- Aging Matters, FoundMyFitness, Self-hacked, Rogue Health, and other blogs
- A4M, Rejuvenation Biotech, and RAADfest annual meetings

Optionally, individual profiles may be linked to genomic information. There is reason to believe that different subjects respond to medications, exercise and diet in different ways, and to the extent that this depends on genetic differences, the information might lay the foundation for a future science of individualized medicine [4].

Testing, not development, is the bottleneck

A great number of techniques have been discovered for extending lifespan in invertebrate animal models. A smaller but substantial number have been validated in rodents. Many of these have been tried in humans and are proven safe. A substantial database of these has been established at Geroprotectors.org [5] With few exceptions, nothing is known about their long-term effects in humans. In addition, there are substances that have shown promise in small human trials that remain unexplored and unreproduced [6, 7].

There exists a common perception that none of these will have dramatic effects on human lifespan, and that greater breakthroughs remain to be discovered. This may be true or not; we cannot know at present. If it is true, then any future discoveries will have to be validated in human testing before they are accepted into (preventive) medical practice. If the discovery is rather conventional, it is unlikely to produce paradigm-breaking results; on the other hand, if the idea were profoundly new and innovative, they will not be accepted without extensive testing, and conventional epidemiological testing would undoubtedly regard as prohibitively expensive for high-risk innovations. In either case, the ability to test for anti-aging benefits cheaply and quickly will be essential

The following is a partial list of treatments we have included in our questionnaire:

- Calorie restriction
- Various fasting schedules

- Various diet regimens including low-carb, low-protein, low-fat, vegetarian, high-fiber, Atkins-type, Mediterranean, paleo, and Pritikin
- Various exercise regimens, including aerobic, resistance training, interval training, and stretching
- Yoga, tai qi, meditation
- Family and work environments
- Prescription drugs, including metformin, rapamycin, statins, angiotensin inhibitors, and deprenyl
- NSAIDs
- Various vitamins
- Minerals, including Magnesium, Chromium, Zinc, Selenium, and Iodide
- Hormones and pre-hormones, including DHEA, melatonin, SAME, CoQ10, pregnenolone, and NAC
- Herbs and supplements, including fish oil, ashwagandha, rhodiola, bacopa, mushroom derivatives, fo-ti, and ginseng

Combining treatments

Many individual treatments are known that modestly extend lifespan in rodents and, presumably, in humans, but the interactions among these treatments have barely been explored. Most of the people in our sample group use many of these treatments, based on vague notions of bet-hedging and an un-articulated intuition that perhaps the benefits will be additive. But it is unlikely the benefits to add linearly; if they did, we would have seen by now some dramatic and obvious examples of age reversal. There is much redundancy among the mechanisms of action of the known interventions, so the underlying metabolic pathways tend to become saturated. Hence, our default expectation is that most combinations of treatments will offer little more benefit than any one of the treatments separately.

But more rarely, we may find pairs of interactions that synergize; in other words, in a few cases we might expect that the mean life extension from two treatments A and B is equal to or even greater than the sum $A + B$ of the benefits from the treatments separately. For example, rapamycin and metformin are reported to synergize [8], and angiotensin inhibitors work via a pathway distinct from either of these [9].

It is unlikely that any single remedy will completely neutralize or reverse aging. Testing interactions among combinations of treatments may be even more important than testing treatments separately. But for even a handful of treatments, the number of possible interactions between them presents a daunting challenge. Separate trials for each possible combination are prohibitively costly and complicated.

Based on this reasoning, we have decided to plan our study around a great number of independent variables. We hope to use principles of multiplexing and multivariate analysis to make sense of the data that is generated.

Our first priority in analysis will be to focus on outliers, the subset of the test population who have scored the greatest decrease in methylation age over the two-year test. (Even if none of the treatments actually lead to age reversal, the standard error of the test is such that over a two-year timespan there will be substantial numbers of people who *appear to* have reduced their biological age.) This subset will be enriched in subjects who have had the greatest success in reducing the rate of aging, and we can look for commonalities among the programs of people who have been most successful. As demonstrated below (**How many subjects are needed?**), we expect that a “fat tail” in the distribution of rates of aging retardation is will be signature of anti-aging combinations that are particularly effective. We will consider our exploratory program to be successful to the extent that we identify a small number of combinations of treatments that are candidates for more detailed study in the future.

History of Methylation Clocks

Cytosine is one of the four nucleic acids that combine to form DNA, and in most higher organisms, some of the cytosine within DNA has an extra methyl group, making 5-methyl cytosine. Methylation is modulated not within genes but at adjacent regions with high concentrations of cytosine, called CpG islands. High levels of methylation tend to suppress expression of adjacent genes, and can also have effects (by mechanisms unknown on expression of genes that are not nearby.

DNA Methylation is the best-studied and easiest to measure among many known mechanisms of epigenetic control. It has been known since the 1990s that human methylation patterns vary systematically, both by tissue type and by age. There is ongoing discussion whether changing methylation patterns are a driver of aging or a passive marker [10].

The first practical methylation clock were developed by Hannum [11] and independently by Horvath [12], building on an idea of Teschendorff [13]. Horvath identified patient records for methylation measurements of tissue samples from 8,000 individuals, with associated ages. Methylation is recorded as a number between 0 and 1 for each Cytosine, indicating the proportion of that site that is methylated. He scanned the available genome for sites that changed most with age and varied least among different tissue types. His algorithm identified 353 sites, and derived a set of 353 multipliers, such that multiplying levels of methylation at each site and adding the products produced a number that could be mapped onto chronological age. About 55% of the sites gain methylation with age and 45% lose methylation.

The original Horvath clock correlates 0.95 with chronological age. The standard error in predicting any one individual's age is ± 4 years. Averages of N individuals increase the accuracy of the clock by \sqrt{N} , so that the average of 100 individuals is accurate to 0.4 years. For our purposes, the relevant question is: measuring the same individual at two different times, how accurate is the difference in Horvath age compared to the elapsed time? There is little data on this thus far, but we might safely assume that it is well under 4 years, since standard error of 4 years represents mostly individual departures from the average, and the PhenoAge clock described below is somewhat more accurate. For calculations in this document concerning

accuracy and numbers of subjects needed, I have assumed that an optimized methylation clock will measure differences over time for the same individual with an accuracy of 1.5 years.

A methylation map of the genome represents 28 million numbers between 0 and 1, representing the proportion (in a sample of cells) of each of 28 million CpG sites that are methylated. 850,000 of these are covered in the EPIC test from Illumina, Inc, the largest commercial kit for methylation profiling. This represents less than 3% of all CpGs, but the selection is concentrated in areas that control expression of local (27%) and distal (7%) genes [14]. A “clock” is merely an algorithm for calculating an age from these 850,000 numbers. In practice, only a few hundred out of the 850,000 are used in the clock function, a sample of a sample.

Concurrently with Horvath, Hannum [11] independently developed a clock that was more accurate and more parsimonious when specialized to blood samples. Five years on, there are several other clocks based on methylation. The smallest is based on a single site, but several others sample hundreds of methylation sites each. All clocks developed so far are linear in the sense that they are simple weighted averages of the methylation levels of (several hundred) sampled sites. The weighted average is not yet an age, because methylation levels do not change linearly over a lifespan; after the weighted average is computed, another function is derived (empirically) that maps the readout as an age. In other words, a clock consists of several hundred multiplier numbers, together with a function that translates the weighted average into “years”.

The limitation to weighted averages is unnecessarily restrictive. It is probable that accuracy can be achieved beyond any of the extant clocks using separate nonlinear functions for each methylation site. (See **Improvements and Adaptations** below.)

As of this writing, the best predictor of age-related decline is the [Levine/Horvath clock \[15\]](#), also called DNAm PhenoAge. It is based on 513 methylation sites and it is calibrated not to chronological age, but to a tighter measure of age-based health, derived from blood lipid profiles, inflammatory markers, insulin resistance, and other markers of age-related health, a composite which Horvath calls “phenotypic age”. The PhenoAge clock is (by design) less well correlated with chronological age than the original, but it is better able to predict mortality than either the classic Horvath clock or chronological age itself. By this measure, the scatter has been greatly reduced. The PhenoAge clock is not yet optimized for the purpose of assessing year-to-year changes within an individual, nor is it specialized to the age range 45-65 of our study. We plan to use PhenoAge as a starting point to develop a clock specialized for our study.

Statistical evidence that the PhenoAge Clock reliably measures biological age

The most direct evidence comes from a meta-analysis based on historic data. Starting with frozen blood samples from the past, Levine and Horvath calculate PhenoAge as it was when the sample was drawn, and compare their results to the health and mortality histories after the blood draw. The Horvath Clock “predicts” the remaining life expectancy of the subjects more accurately than their chronological age [16]. There is less data available for the new PhenoAge

clock, but strong indications it performs much better than the Horvath clock for this purpose [17, 18]

Also, many of the life styles that promote long life have been confirmed to slow the PhenoAge clock, while, conversely, obesity and high blood pressure and insulin resistance have been found to accelerate aging as measured by the PhenoAge clock. Better statistics are available for the original Horvath clock, which has been studied for 5 years, but indications are that the PhenoAge clock performs better. Fig 1. Is copied from [18].

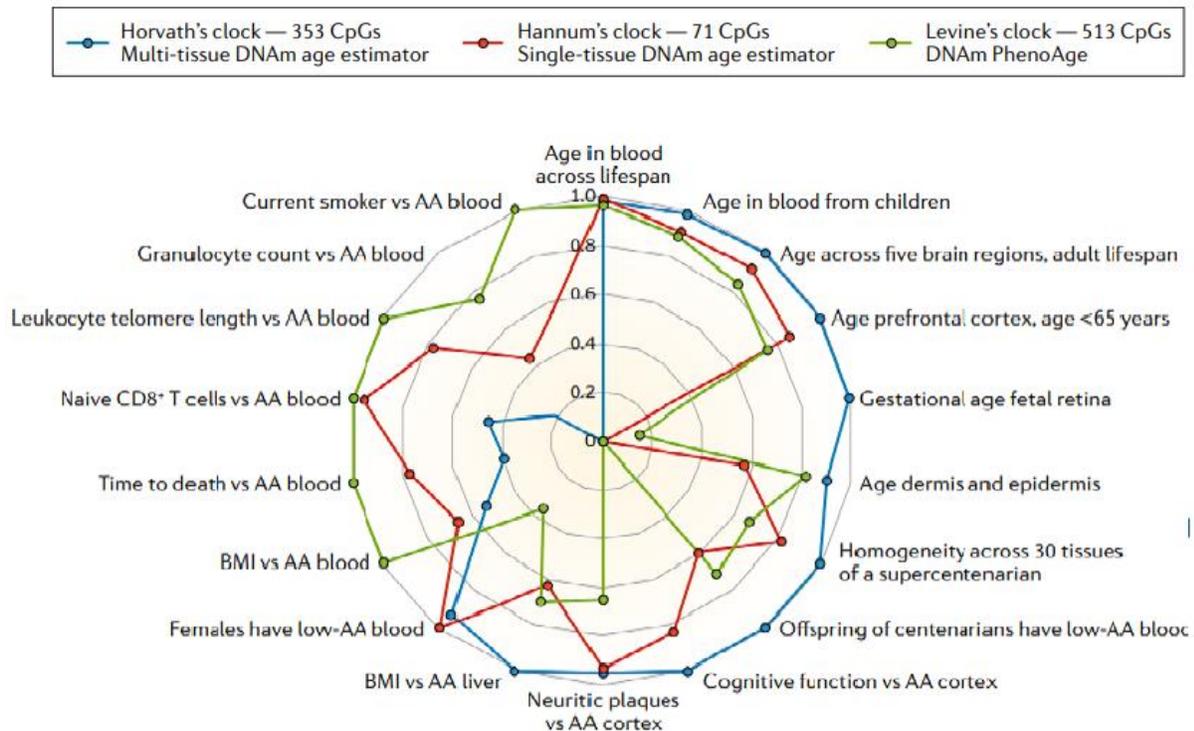


Figure 1 Comparison of three DNA methylation-based biomarkers of aging. The multi-tissue DNA methylation-based (DNAm) age estimator (blue line), also known as Horvath's clock, stands out in terms of its correlation with chronological age across multiple tissue types, its high accuracy in children, its strong correlation with gestational age (differentiation day) in neuronal cell culture models and the homogeneity of its age estimates across tissues, for example 30 tissue samples collected from a supercentenarian (>110 years). The phenotypic age estimator (green line) or DNAmPhenoAge stands out in terms of its predictive accuracy for time to death, its association with smoking status and its association with various markers of immunosenescence. In general, DNAm PhenoAge and DNAm Age as calculated by the single-tissue age estimator known as Hannum's clock (red line) outperform other blood-based biomarkers in regard to lifespan prediction.

- Epigenetic age correlates with progression of Alzheimer's and Parkinson's Disease [16]
- Same for Arthritis [19]
- Menopause moves the methylation clock forward. Early menopause is associated with accelerated methylation aging, and late menopause with younger methylation age.
- Epigenetic age is accelerated by obesity, blood sugar, insulin, and inflammation.S

- Epigenetic age is retarded by carotene supplementation, exercise, education, and by diets high in vegetables, fruits and nuts.
- Stem cell transplants lower epigenetic age dramatically (from a study of leukemia patients [20]). Epigenetic age is set back ~8 years for a short period, but then accelerates to a set-forward a few years after treatment.

(All the above claims drawn from [18].)

Theoretical foundation of the Horvath Clock

The original Horvath clock was developed by a statistical process that took into account only chronological age. It was later validated as a better predictor than chronological age for risk of all-cause mortality and several diseases of old age. This fact is powerful evidence that methylation is measuring something fundamental about the aging process. If an individual's methylation age is higher or lower than his chronological age, the difference is a predictor of his disease risk and how long he will live. This can only be true if methylation is associated with a fundamental cause of age-related decline and is not merely a “marker” of aging.

Development of an embryo to a fetus to a child to an adult takes place under epigenetic control [21, 22]. An emerging theory the last 7 years is that the epigenetic program continues after growth ends, and that aging is the result. De Magalhaes [23], Rando [24], [Blagosklonny](#) [25] [Johnson](#) [26] and [Mitteldorf](#) [27] all have independently proposed an epigenetic basis for aging. This fits with the observed fact that much of the change in methylation is continuous, from development through aging [28]. According to the epigenetic theory, the root cause of aging is that different sets of genes are expressed at different times of life. Whether the epigenetic changes responsible for aging constitute an evolved adaptation [27, 29] or whether it is a kind of inertia that continues past its usefulness to the individual [23, 30] is an interesting theoretical question, but irrelevant to the question of the clock's usefulness in testing interventions. In either case, the fact that epigenetics drives aging lends credence to the idea that methylation should be closely associated with biological age.

Parabiosis experiments [31-33] support the idea that factors circulating in the blood can rejuvenate remote tissues. This is indirect support for the epigenetic foundations of aging, because these blood factors come from gene expression in cells—especially but not exclusively endocrine cells. The experiments suggest that aging at the cellular level is reversible via a young signaling environment. It is to be hoped that epigenetic reprogramming of endocrine cells will trigger release of youthful signals. More modestly, slowing of the epigenetic clock ought to slow changes in the age-dependent circulating signals, retarding aging system-wide.

Skepticism, Doubts, and Counter-arguments

- Methylation clock untethered to gene expression** — Methylation sites queried in the aging clocks have been derived empirically without reference to theory or to biochemistry. Attempts to correlate these sites with expression of known genes have

achieved quite limited success thus far [14]. Expression of some inflammatory cytokines have been associated with increased methylation age, and, paradoxically, telomerase expression has been linked to **accelerated** methylation aging [18]. But most sites of the Horvath, Hannum, and Levine clocks have yet to be linked to gene expression. This constitutes a gap in our understanding of how the clocks work, and cause for questioning whether their basis is fundamental to aging. The most parsimonious explanation is that methylation controls expression of genes far from the methylation site in ways that have yet to be catalogued.

- B. **'Epigenetic drift'** — Many authors still write about changes in methylation during aging as “[epigenetic drift](#)” [34], as though these changes were wholly stochastic. This represents an attempt to reconcile the self-destructive potential of epigenetic changes over a lifespan with a perspective dismissive of the idea that programmed aging might be an evolutionary adaptation. While it is true that gene expression does become more random with age, it is also true that there are specific gene expression changes associated with aging—the methylation clock is based on such programmed changes. Many of the known changes appear to have wholly detrimental effects on the metabolism, and indeed some of these changes correspond to the human aging phenotype [17].
- C. **Perhaps gene expression changes are a response to damage, the body's attempt to mitigate aging.** This is the suspicion that haunts the aging clock. If this is the case, interventions that thwart the mitigation would come out looking like age reversal, but in fact they would have the opposite effect, increasing risk of disease and mortality. Support for this idea comes from the prejudice that says “the body would never purposefully destroy itself.” The strongest evidence for this idea is that genetic variations associated with longer telomeres are correlated with older methylation age [10]. But many others among the programmed changes have been shown to be detrimental. For example, “DNA PhenoAge acceleration was found to be associated with increased activation of pro-inflammatory and interferon pathways and decreased activation of the transcriptional and translational machineries, the DNA damage response and nuclear mitochondrial signatures” [quote from [18], referenced therein to [17]]
- D. **Not all anti-aging interventions affect the PhenoAge or Horvath clocks.** This could be a substantial problem if it turns out that there are effective anti-aging strategies that work in ways undetectable by the PhenoAge clock or, conversely, if some strategies are able to set back the PhenoAge clock without reducing risk of age-related disease. “For example, within a 9-month follow-up period, the substantial weight loss resulting from bariatric surgery was not associated with a reduction in epigenetic age of human liver tissue samples” [quote from [18]; referenced therein to [35]] To the extent we think that bariatric surgery is a legitimate anti-aging strategy, this hints at a problem that could be more general.

Improvements and adaptations of the Horvath clock

For present purposes, the PhenoAge clock is a substantial improvement on the original Horvath clock, attained by calibrating it against health indicators and not just chronological age. There are at least three more ways in which the methylation age test can be improved

- A. The original clock and all its successors thus far have been based on simple linear combinations of different methylation sites. But we know that different methylation sites change with age on different schedules. With mathematics that is just a bit more complicated, it is possible to adapt the clock's sensitivity to each site in accordance with its unique age trajectory. This change to the algorithm ought to improve the clock's accuracy for **any** application.
- B. The clock might be specialized to the application of testing anti-aging effects on individual humans, i.e., comparing biological age for the same individual at two different times. Some of the scatter in the plot of DNAmAge is due to variation from one individual to another, and some is due to other random factors that don't depend on the individual. In the past, there was little data available for the same individual at two different times, but this is changing, and now it is feasible to separate the two kinds of scatter. The clock can then be specialized to report age differences even more accurately.
- C. Again, for the particular application proposed, there is no need for a clock that works generally on any age, from pre-birth to centenarian. If all the people in the study are between the ages of 45 and 65, then the clock might be specialized to be more accurate in this age range, at the expense of losing accuracy for younger and older subjects—who aren't part of the study. It may be worthwhile to take this idea even further and have specialized sub-clocks calibrated for narrow age ranges. Derivation of the algorithms of these specialized clocks will demand more training data than is now available.

Tightening the accuracy of the clock affects the amount of information that can be derived from a given number of experimental subjects. The DNAm PhenoAge clock correlates 96% with phenotypic age. If this can be raised to 98%, the number of human subjects required to detect a given anti-aging benefit is reduced by half. Conversely, for the same effort and expense, we will be able to derive more information about more treatment strategies.

If we can indeed construct a clock with 98% correlation accuracy, a new benefit will be available: Accuracy at that point will be approaching the point where a single individual might detect benefits over the course of two years. This will be a gateway to individualized medicine, as individuals will be able to use the methylation clock for feedback on whatever strategies they use consistently. Self-experimentation is elevated from guesswork based on subjective response in real time to a science based on objective measurement.

How many subjects are needed?

This is an exploratory study, and we do not know what we will find. That said, it is worthwhile to target the size of the study based on a hypothesized benefit and sufficient statistical power to detect that benefit should it exist in our data.

To this end, I created a computer simulation based on the following assumptions:

- There are 20 plausible anti-aging measures, each of which has a prevalence between 0.2 and 0.8 in the population we are studying. They are distributed through the population independently.
- The 20 measures have an average benefit of 0.05, interpreted to mean that they slow the rate of aging by 5%. The benefits are normally distributed with a standard deviation of 0.05, so that the best of them is likely to slow aging by 15%, while some of the worst actually accelerate aging.
- Almost all combinations of these benefits are redundant. The combined effectiveness of a set of anti-aging measures is equal to the best single measure pursued separately.
- But (by assumption) one combination of three (out of the $C(20,3)=1140$ possible triples) has an extraordinary synergy, and slows aging by X%. If such a magic triplet exists, it is the most compelling purpose of our study to assure that it is identified.

As we saw above, the best available methylation clocks have uncertainty for each individual test of about 1.5 years. The difference between two measured ages will have an uncertainty $\sqrt{2}$ times higher than this, under the worst case assumption that the two errors are independent random variables. But aging markers are now recognized as an important research tool, and several groups are making rapid progress toward development of new markers [Horvath, personal communication; Kennedy, personal communication]. For purposes of this calculation, I have assumed that a specialized clock can be developed that measures the difference in age for a given individual with a normal distribution that has width 1.5 year.

Without going through the full simulation, the following calculation can provide a rough guide to what can be achieved with 5,000 subjects. Every given triple of aging measures will be adopted on average by roughly 1/8 of the subjects ($\frac{1}{2} * \frac{1}{2} * \frac{1}{2}$). We assumed that the distribution of the errors in measured ages has a standard deviation of 1.5 year, and averaging 625 subjects, the standard deviation of the mean is about $1.5/\sqrt{625} = 0.06$ year. We need three standard deviations for an effect to stand out above the noise, and that is 0.18 year. Over a two-year time period, under ideal assumptions, we would be able to detect (0.18 years / 2 years), which is a 9% deceleration of the rate of aging. The calculation below shows that the simulated result is somewhat less sensitive than this, and in real life with dropouts, inaccurate reporting, individual variation, and other departures from ideal assumptions, it may be reasonable to detect a 20 or 30% effect on the rate of aging at the smallest. If a combination of available treatments exists that offers benefits of this magnitude, we would certainly be interested to know about it.

In the simulated analysis, I compared the average benefit for each of the possible triples, and counted a success if the one presumed synergistic combination was ranked first or second (out

of 1140) in analysis of the computer-generated data. Since both generation and analysis of the data is quick once the program is written, I was able to evaluate many thousands of scenarios to estimate the average probability that the single synergistic triple would be detected.

Based on these assumptions, I calculate the number of experimental subjects needed to assure an 80% chance of detecting a synergistic benefit of a given size. The result is plotted in Fig. 2.

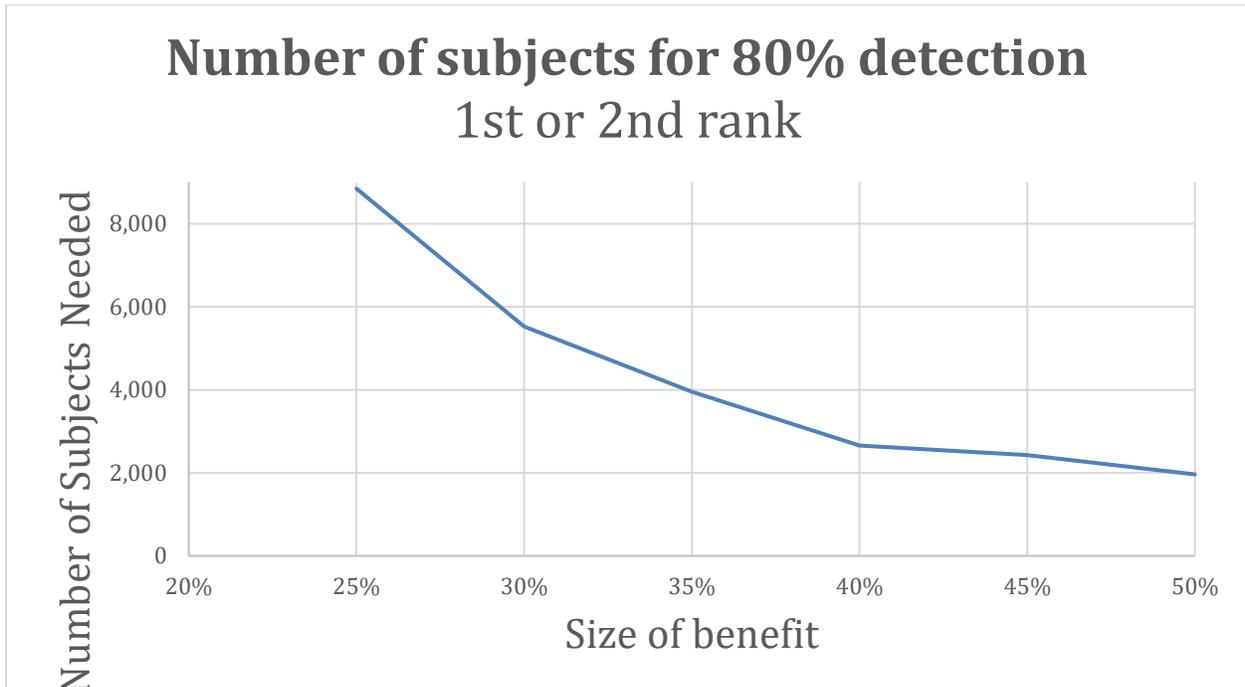


Figure 2 The number of subjects needed in order to have an 80% assurance of detecting a synergy, plotted against the size of the synergy on the x axis.

Given the planned size of our study (5,000 subjects), what is the probability of detecting a synergy of a given size? This result is plotted in Fig 3.

Probability of Detecting a Synergistic Combination

based on 5,000 subjects

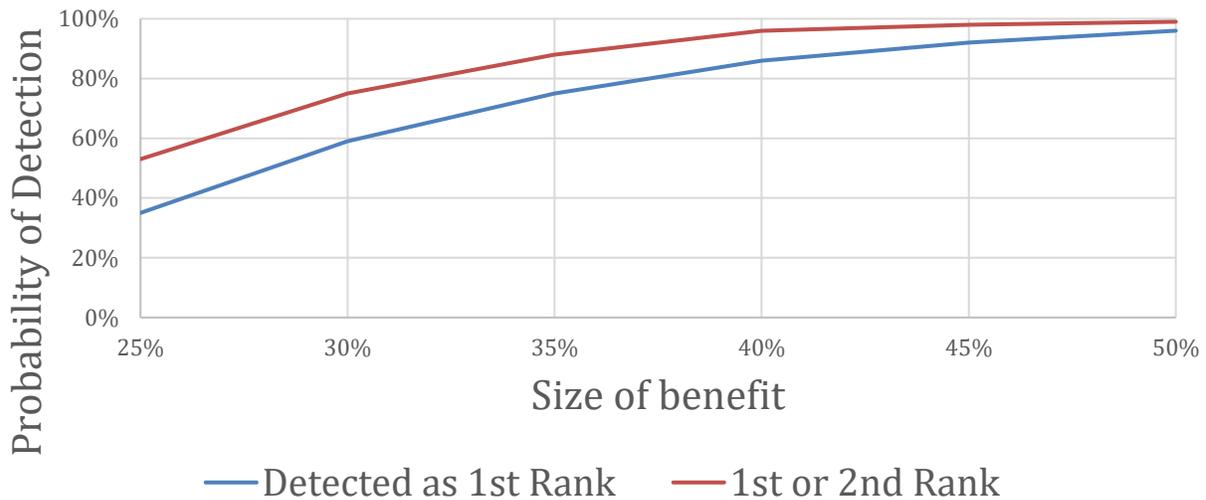


Figure 3 For a study including 5,000 subjects, what is the probability of detecting a three-way synergy as described above, as a function of the size of the synergy.

How will we know if we have identified a synergistic combination?

This is an exploratory study, and we will look at the data on its face for suggestions of interesting patterns of any kind. One disappointing but informative possibility is that among this collection of well-informed early adopters no coherent subset has achieved more than a 15% deceleration of the aging rate (+20 months biological age in the course of 24 months chronological age). The most promising result will be the one outlined just above, namely, the identification of a combination of two or three or four anti-aging measures that synergize to offer age retardation of 25% or more. One sign that this is the case would be a long, asymmetric tail on the distribution of age retardation calculated for all possible three-way combinations. The reason for this is as follows:

Imagine, as in the simulation above, a universe of 20 anti-aging measures, of which each subject in the study has adopted some and not adopted others. There are $C(20,3)=1140$ triples among these 20 measures. For each of these triples, calculate the average age retardation for just that subset of individuals who have adopted all three measures (no matter what other measures they may also have adopted). Under the null assumption of no synergy, the distribution of the 1141 age retardations will look like fig 4. (Of course, the averages are calculated on results as measured, including a random error of order 1.5 years, as assumed above.)

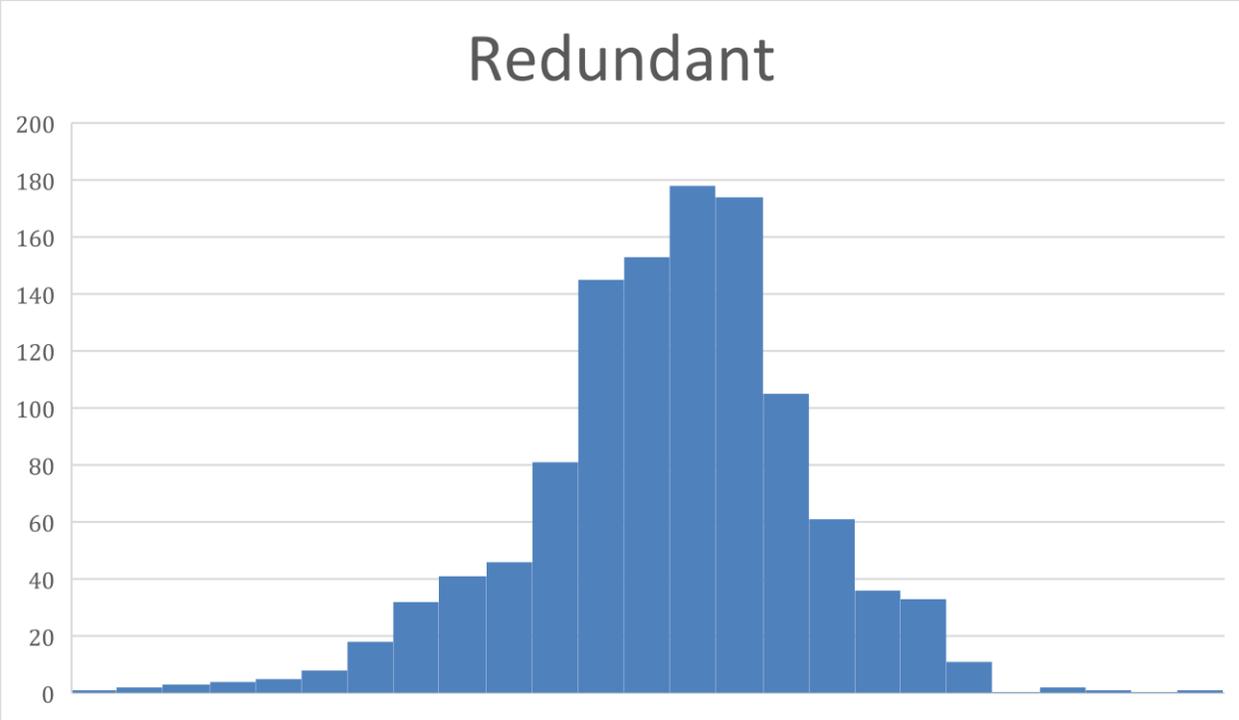


Figure 4 Distribution of average age retardation for subjects that include each of 1141 combinations of 3 measures out of a universe of 20. Calculated under the assumption that there is no synergy, and all measures are mutually redundant.

But if there is one combination of three that offers age retardation of 25%, the same plot looks like fig. 5.

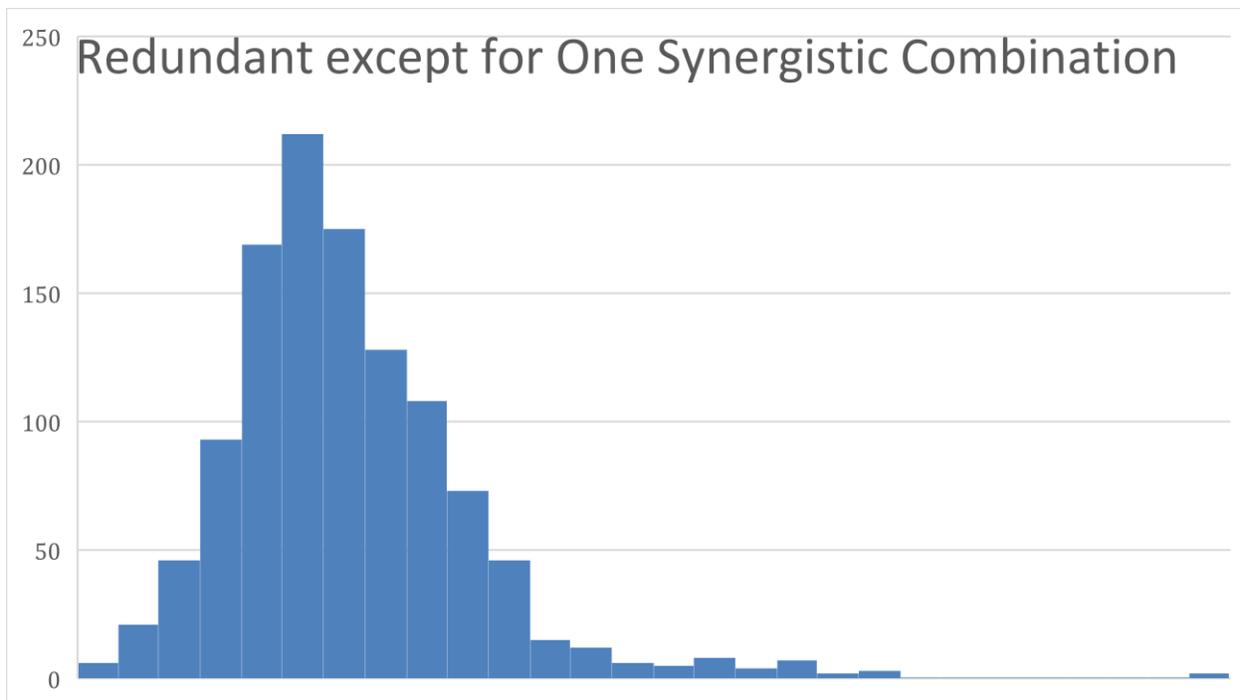


Figure 5 Calculated as in Fig 4 for the case where there is a single synergistic triple. Note the long tail that connects the outlier on the right, consisting of all those triples that overlap with the synergistic triple, having 2 out of 3 measures in common.

The synergistic triple is alone at the far right. The reason for the long tail in the distribution connecting the synergistic triple to the rest of the distribution is that there are 51 other triples that overlap with the synergistic triple, sharing 2 out of 3 measures in common. For each of these, about half of the subjects counted in the average actually have deployed the synergistic combination, pulling their average up.

Figures 4 and 5 are included merely as examples. Other patterns might indicate other patterns of synergy, and we plan to approach the data without preconceptions about what it might reveal.

Summary

In the future, there will undoubtedly be exceptional strategies that have substantial effect on lifespan but no effect on methylation clocks. Conversely, there will be Type I errors—strategies that set the methylation clock back and yet have no effect on life expectancy. But we expect that these will be exceptions. There is extant evidence for a substantial overlap between interventions that extend lifespan and interventions that decelerate the methylation clocks. Given that exploration with this measure is projected to be 10 times faster and 100 times cheaper than traditional epidemiology, it is worthwhile to pursue methylation assessments now for all known and suspected anti-aging strategies.

If these projected savings are realized, human trials will be less expensive (and, of course, more relevant) than rodent testing for a large class of interventions. At present, there is no

methylation clock for mice with comparable accuracy to the human clock, but clocks for rodents and other mammals are under development in the Horvath lab. Such a clock could slash the cost of rodent testing, so that a great number of interventions could be tested with an unprecedented level of efficiency. (Calibration of a mouse clock may not require the tens of thousands of subjects that have been used for humans because we have inbred mouse populations that are genetically homogeneous.)

High-throughput testing of the dozens of proposed anti-aging interventions that are presently backlogged should soon be a practical and affordable reality.

REFERENCES

1. Mattison, J.A., et al., *Impact of caloric restriction on health and survival in rhesus monkeys from the NIA study*. Nature, 2012.
2. Colman, R.J., et al., *Caloric restriction reduces age-related and all-cause mortality in rhesus monkeys*. Nature communications, 2014. **5**.
3. Mitteldorf, J., *An Incipient Revolution in the Testing of Anti-aging Strategies* Biochemistry (Mosc.), 2018. **83**(forthcoming).
4. Schork, N.J., *Personalized medicine: time for one-person trials*. Nature, 2015. **520**(7549): p. 609-611.
5. Moskalev, A., et al., *Geroprotectors.org: a new, structured and curated database of current therapeutic interventions in aging and age-related disease*. Aging (Albany NY), 2015. **7**(9): p. 616-628.
6. Anisimov, V.N., *[Pineal gland, biorhythms and aging of an organism]*. Usp Fiziol Nauk, 2008. **39**(4): p. 40-65.
7. Anisimov, V.N. and V.K. Khavinson, *Peptide bioregulation of aging: results and prospects*. Biogerontology, 2010. **11**(2): p. 139-149.
8. Singh, A.K., et al., *Synergistic effect of rapamycin and metformin against age-dependent oxidative stress in rat erythrocyte*. Rejuvenation Research, 2017(ja).
9. Blagosklonny, M.V., *From rapalogs to anti-aging formula*. Oncotarget, 2017. **8**(22): p. 35492-35507.
10. Horvath, S. and K.J.N.R.G. Raj, *DNA methylation-based biomarkers and the epigenetic clock theory of ageing*. 2018: p. 1.
11. Hannum, G., et al., *Genome-wide methylation profiles reveal quantitative views of human aging rates*. Mol Cell, 2013. **49**(2): p. 359-367.
12. Horvath, S., *DNA methylation age of human tissues and cell types*. Genome biology, 2013. **14**(10): p. R115.
13. Teschendorff, A.E., et al., *Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer*. Genome research, 2010. **20**(4): p. 440-446.
14. Pidsley, R., et al., *Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling*. Genome biology, 2016. **17**(1): p. 208.
15. Levine, M.E., et al., *An epigenetic biomarker of aging for lifespan and healthspan*. Aging (Albany NY), 2018. **10**(4): p. 573-591.
16. Chen, B.H., et al., *DNA methylation-based measures of biological age: meta-analysis predicting time to death*. Aging (Albany NY), 2016. **8**(9): p. 1844-1859.
17. Levine, M.E., et al., *An epigenetic biomarker of aging for lifespan and healthspan*. bioRxiv, 2018: p. 276162.

18. Horvath, S. and K. Raj, *DNA methylation-based biomarkers and the epigenetic clock theory of ageing*. Nature Reviews Genetics, 2018: p. 1.
19. Horvath, S. and B.R. Ritz, *Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients*. Aging (Albany NY), 2015. **7**(12): p. 1130-1142.
20. Stölzel, F., et al., *Dynamics of epigenetic age following hematopoietic stem cell transplantation*. Haematologica, 2017. **102**(8): p. e321-e323.
21. Goldberg, A.D., C.D. Allis, and E. Bernstein, *Epigenetics: a landscape takes shape*. Cell, 2007. **128**(4): p. 635-638.
22. Murgatroyd, C. and D. Spengler, *Epigenetics of Early Child Development*. Frontiers in Psychiatry, 2011. **2**(16).
23. de Magalhães, J.P., *Programmatic features of aging originating in development: aging mechanisms beyond molecular damage? The FASEB Journal*, 2012. **26**(12): p. 4821-4826.
24. Rando, T.A. and H.Y. Chang, *Aging, rejuvenation, and epigenetic reprogramming: resetting the aging clock*. Cell, 2012. **148**(1): p. 46-57.
25. Blagosklonny, M.V., *Aging and Immortality: Quasi-Programmed Senescence and Its Pharmacologic Inhibition*. Cell Cycle, 2006. **5**(18): p. 2087-2102.
26. Johnson, A.A., et al., *The Role of DNA Methylation in Aging, Rejuvenation, and Age-Related Disease*. Rejuvenation Research, 2012. **15**(5): p. 483-494.
27. Mitteldorf, J., *How does the body know how old it is? Introducing the epigenetic clock hypothesis*. Biochemistry (Moscow), 2013. **78**(9): p. 1048-1053.
28. Smith, Z.D. and A. Meissner, *DNA methylation: roles in mammalian development*. Nature Reviews Genetics, 2013. **14**(3): p. 204.
29. Mitteldorf, J., *An epigenetic clock controls aging*. Biogerontology, 2016. **17**(1): p. 257-265.
30. Blagosklonny, M.V. and M.N. Hall, *Growth and aging: a common molecular mechanism*. Aging (Albany NY), 2009. **1**(4): p. 357-62.
31. Katcher, H., *Studies that shed new light on aging*. Biochemistry (Moscow), 2013. **in this issue**.
32. Conboy, M.J., I.M. Conboy, and T.A. Rando, *Heterochronic parabiosis: historical perspective and methodological considerations for studies of aging and longevity*. Aging cell, 2013. **12**(3): p. 525-530.
33. Rebo, J., et al., *A single heterochronic blood exchange reveals rapid inhibition of multiple tissues by old blood*. Nature Communications, 2016. **7**: p. 13363.
34. Mendelsohn, A.R. and J.W. Larrick, *Epigenetic drift is a determinant of mammalian lifespan*. Rejuvenation research, 2017. **20**(5): p. 430-436.
35. Horvath, S., et al., *Obesity accelerates epigenetic aging of human liver*. Proceedings of the National Academy of Sciences, 2014. **111**(43): p. 15538-15543.