

A cross-national comparison of 12 biomarkers finds no universal biomarkers of aging among individuals aged 60 and older

*David H. Rehkopf, Luis Rosero-Bixby and William H. Dow**

Abstract

There is uncertainty about whether biological and anthropometric measures that are clinical risk factors for disease are universally associated with chronological age, or whether these correlations vary depending on the social and economic context. The answer to this question has implications for the malleability of biological aging. To examine this issue, we use population-based data on individuals aged 60 and older from the Costa Rican Study on Longevity and Healthy Aging, and temporally consistent data from the United States National Health and Nutrition Examination Survey and the United States Health and Retirement Study. Our analysis focuses on 12 biomarkers that have been shown in the literature to have an association with age, and that occur prior to the clinical manifestation of disease. We find that there are few consistent patterns of association with age when these biomarkers are stratified by gender, country, and level of education. This result suggests that these measures of biological aging are highly context-dependent, and that none of the 12 biomarkers we examined are universal biomarkers of aging. Future research that investigates composite measures of biological age should test newly proposed measures across gender, social class, and country.

* David H. Rehkopf (corresponding author), Assistant Professor, Stanford University, Division of General Medical Disciplines, 1070 Arastradero Road, Suite 300, Room 305, Palo Alto, CA 94304, U.S.A.

Email: drehkopf@stanford.edu

Luis Rosero-Bixby, Emeritus Professor, University of Costa Rica and Research Associate, University of California, Berkeley, U.S.A.

William H. Dow, Professor, University of California, Berkeley, U.S.A.

1 Introduction

While there are many definitions of biomarkers of aging, the following is widely used: “A biomarker of aging is a biological parameter of an organism that either alone or in some multivariate composite will, in the absence of disease, better predict functional capability at some late age than will chronological age” (Baker and Sprott 1988). This definition makes explicit a fundamental question in aging research: namely, are aging and disease separate entities, or is the former just a consequence of the latter? A primary motivation for studying biomarkers of aging is to determine whether we can characterize the aging process separately from disease. To support the identification of potential biomarkers of aging, Baker and Sprott described six characteristics of a biomarker of aging: (1) the rate of change of a biomarker must reflect some measurable parameter that can be predicted at a later chronological age; (2) the biomarker should reflect some basic biological process of aging, and not a predisposition to a disease state or some inborn error in metabolism; (3) the biomarker should have a high degree of reproducibility in cross-species comparisons of functional or physiological age and chronological age, particularly within the same classes, and certainly within the same families of species; (4) the biomarker should change independently with the passage of time, and reflect physiological (functional) age; (5) the assessment of the biomarker should be nonlethal in animal systems, and should cause minimal trauma in humans; and (6) the biomarker should be reproducible and measureable during a relatively short time interval. Meanwhile, the American Federation for Aging Research (AFAR) proposed four criteria for identifying a biomarker of aging: (1) the biomarker must predict a person’s physiological, cognitive, and physical functioning in an age-related way; (2) the biomarker must monitor the basic process underlying aging, and not the effects of disease; (3) the biomarker must be testable and not harmful to subjects; and (4) the biomarker should work in laboratory animals as well as in humans (Johnson 2006).

While these criteria reflect a wide range of characteristics, in our analysis we will focus on the assessment of three of the most fundamental criteria: (1) the passage of time, (2) the underlying physiological process, and (3) reproducibility. The first and primary focus of our analysis is to assess whether the biomarker has some fundamental relationship with the passage of time in an individual’s life; that is, with the individual’s chronological age. In the definition of biomarkers of aging, this point is explicated in the requirement that the marker reflect the rate of aging. For example, the Baker and Sprott criteria stated that “biomarkers should change independently with the passage of time;” while the AFAR criteria stipulate that a biomarker should function in an “age-related way.” This point was made even more explicitly in the 1987 National Institute of Aging primate study, which stated that the first step in defining a biomarker of aging is establishing that it has a “significant cross-sectional correlation with age” (Ingram, Nakamura, Smucny, Roth and Lane 2001). While chronological age is only one of many criteria, it is fundamental to the definition of a biomarker of aging. Our primary investigations

use conventional age, measured as the time elapsed since birth, as the metric. Most of the biomarker definitions and research on biomarkers have relied on this definition of age. However, a secondary focus of our analysis also uses the metric of thanatological age, or the time before death. As we are examining all individuals in the data in the years prior to their death, we use this metric to explore the question of whether biomarkers meet another criterion put forth by Butler: namely, that a biomarker of aging “should predict remaining longevity at an age when 90% of the population is still alive” (Butler et al. 2004).

The second criterion we consider in our analysis of the biomarkers of aging is the underlying physiological process. This criterion has been broadly described in the criteria outlined by Baker and Sprott as “some basic biological process of aging.” We thus cast a narrower net than some other studies which examined any physiological parameter that had been measured, even if it had no clear connection to biological aging or pathobiology. We include only biomarkers that have a known function and that are related to illness. These criteria are also consistent with other expectations of biomarkers of aging, as defined by Arking: namely, that a biomarker of aging must “be crucial to the maintenance of health” and “be monitoring some basic, important process” (Arking 2006). Moreover, this approach is in line with the recent definition of a biomarker of aging issued by the American Federation for Aging Research, in which one of the criteria is that a biomarker of aging must “monitor the basic process underlying the aging process, not the effects of disease.” The use of biologically well-characterized risk factors for disease gives us a higher degree of confidence that these are indeed risk factors of disease, and not the effects of disease – a certainty we lack when we look at many of the hypothesized biomarkers of aging that, for example, capture aspects of immune function. This is not to argue that these other factors are not potentially important as biomarkers of aging; rather, we are noting that we can be less certain that they are not consequences of the disease process. Based on these criteria, we included the following 12 potential biomarkers of aging in our analyses: systolic blood pressure, diastolic blood pressure, HbA1c, fasting glucose, C-reactive protein, HDL cholesterol, triglycerides, total cholesterol, urinary creatinine, body mass index (BMI), waist circumference, and leukocyte telomere length. All of these biomarkers have been investigated in prior studies of biomarkers of aging (Crimmins, Vasunilashorn, Kim and Alley 2008; Mather, Jorm, Parslow and Christensen 2011). Six of these factors (systolic blood pressure, diastolic blood pressure, total cholesterol, HDL cholesterol, HbA1c, and waist circumference) are components of the calculation of allostatic load, which has been proposed as a biomarker index of aging (Johnson 2006; Karlamangla, Singer, McEwen, Rowe and Seeman 2002; Seeman, McEwen, Rowe and Singer 2001), and has been shown to be associated with all-cause mortality (Karlamangla, Singer and Seeman 2006). A large number of scholars have also speculated about the use of leukocyte telomere length as a biomarker of aging (Boonekamp, Simons, Hemerik and Verhulst 2013; Der et al. 2012; Mather et al. 2011; Sanders and Newman 2013).

Third, we use a population-based approach to assess reproducibility; i.e., to examine whether each biomarker has a consistent relationship with age. In both early and more recent definitions of a biomarker of aging, the criteria of “high reproducibility in cross-species comparisons” (Baker and Sprott 1988) and “something that works in humans and in laboratory animals, such as mice” are mentioned. Rather than examine whether each biomarker is associated across species, we examine whether each biomarker is associated with age in a similar manner across humans who are exposed to different environments, both within a country and between countries. This approach complements much of the previous lab-based research on biomarkers of aging (Zahn et al. 2007) and non-human mammal-based approaches (Nadon 2006). In our examination, we explicitly apply this approach to humans, acknowledging that the wide range of environments in which humans live could affect the utility of biomarkers for measuring aging. Our approach is based on the theoretical concept of embodiment; i.e., that the environment in which humans live has profoundly important effects on their physiology (Krieger 2005). An indicator could still be considered a biomarker of aging even if the rate at which it deteriorates with age varies across environments; as long as the indicator worsens to some degree with age, we can consider it a “weakly universal” biomarker of aging. An indicator with a slope that does not differ across environments could be termed a “strongly universal” biomarker of aging. However, if in some environments the age slope of an indicator reverses to improve with age, we can reject that particular biomarker as a candidate for a universal biomarker of aging.

The primary purpose of this article is thus to describe how 12 potential biomarkers of aging are correlated with age, or with the number of years since birth. A secondary goal is to examine how the 12 potential biomarkers are correlated with thanatological age, or with the number of years until death. The first minimal criterion we will use to test whether each of these factors qualifies as a biomarker of aging is whether the physiological levels of the biomarker worsen with a greater number of years since birth. The second criterion we will use to test whether each of these biomarkers acts as a biomarker of aging is whether the correlation is consistent across sexes, countries, and levels of education. In proposing these criteria, which are based directly on definitions of biomarkers of aging, we wish to emphasize that they are necessary but not sufficient criteria for identifying biomarkers of aging. Our descriptive, population-based approach is intended to contribute to efforts to establish the validity and reliability of potential biomarkers of aging.

We also wish to note that work has been done to combine biomarkers into composite measures of biological aging. A study on a younger birth cohort in New Zealand used a combination of 18 biomarkers to characterize biological aging, including a majority of the clinical measures we examine here (Belsky et al. 2015). While this study showed that most of these biomarkers were associated with age, this research is somewhat limited because it examined the biomarkers at three ages only (26, 32, and 38), and because it is unclear whether findings from this population can be generalized to other contexts. Another study using data from the United

States from a wider age range has shown that many of the biological risk factors we examine are associated with age (Levine 2013). But as with the research for New Zealand, it is not clear whether these relationships hold across different contexts and levels of education, and whether the associations with age are non-linear. In our study, we extend these analyses to examine whether the individual measures have consistent relationships with age across country, gender, and social class.

2 Data and Methods

2.1 Study Samples

2.1.1 Costa Rica

The data on Costa Rica were drawn from the Costa Rican Study on Longevity and Healthy Aging (CRELES), a longitudinal, nationally representative, probabilistic sample of close to 3,000 adults aged 60 and older, with an over-sampling of older ages (Rosero Bixby, Fernandez and Dow 2010). The current study used data from the first wave of data collection, which took place mainly in 2005. The sample size varied depending on the biomarker, as described in the Table 1 note. The dates of each participant's birth and death were taken from the national registry databases using the participant's "*cédula*" number (the Costa Rican identification card number). A computerized death registry follow-up to establish the deaths of participants was accessed up to March 2014, at which point 60% of observations were survivors with between 5.6 and 9.4 years of follow-up. Because foreigners (3% of the sample) cannot be accurately followed in the death registry, they were excluded from the thanatological age analysis. For all analyses of the CRELES sample, we used data from two observations where available, and accounted for the clustered nature of this sample in calculating our standard errors for the regression models. The number of deaths used for the calculation of thanatological age was 1,214. Because we were interested in using the exact rather than the predicted thanatological age, we used only individuals who were deceased in the thanatological age calculations. No upper age limit was placed on the Costa Rica sample.

2.1.2 United States

To maximize our sample sizes and range of biomarkers, the data we use for the United States have been drawn from two separate surveys. The first U.S. dataset is from the Health and Retirement Study (HRS). The HRS began in 1992 with a nationally representative sample of non-institutionalized US residents born in 1931–1941, and their spouses. The individuals in the original panel were surveyed every two years across broad categories of survey content, with special focuses on

income, work, assets, and self-reported physical health and functioning. To maintain the age 50+ representative nature of the sample, supplemental sampling is used to recruit new samples. The Survey Research Center at the University of Michigan provided detailed documentation on the HRS sampling design and the selection and validation of health measures (Juster and Suzman 1995). We implemented our analytic models using the HRS Rand Files Version L (RAND 2011). Our HRS sample was taken from the samples used in 2006 and 2008, the years when the biomarker data were collected. In order to match the age distribution of the CRELES sample, we excluded from our HRS analytic sample individuals aged 59 and younger. The final analytic sample consisted of 5,870 men and 9,008 women. The number of deaths used for the calculation of thanatological age was 1,650. These deaths were drawn from data linked to the U.S. National Death Index (Rogot et al. 1983). There is no upper age limit on the HRS population.

Data were also obtained from a separate survey with many additional biomarkers, the U.S. National Health and Nutrition Examination Survey (NHANES) 1999–2004. This survey covered non-institutionalized adults aged 60 and older ($n = 2411$ men, $n = 3196$ women). This cross-sectional dataset is representative of the non-institutionalized population of the United States. Because individuals aged 85 and older are top-coded to an age 85+ category, we excluded individuals aged 85 and older from this U.S. sample. Except for the measures shown as not collected in the HRS, which are displayed in descriptive Table 1, all of the analyses of the U.S. context included individuals from both the HRS and the NHANES. The number of deaths used for the calculation of thanatological age was 329. These deaths were drawn from data linked to the U.S. National Death Index (Rogot et al. 1983)

2.2 Measures

Systolic and diastolic blood pressure were measured using the average readings over multiple measures per individual. Blood glucose was measured after fasting. Creatinine was measured from urine. BMI was calculated from weight and height measured at the time of each survey. Leukocyte telomere length was measured using the PCR-based method in all three samples (Cawthon 2002; Needham et al. 2013; Rehkopf et al. 2014). Further details on the measurement of each biomarker in each survey are described elsewhere (NCHS 1999; Rosero-Bixby and Dow 2012; Weinstein, Vaupel, Wachter and Weir 2008). Urinary creatinine, triglycerides, and glucose were not measured among the HRS participants. The following exclusions were made for biologically implausible values: HbA1c $\geq 12\%$, fasting glucose ≥ 400 mg/dl, diastolic blood pressure ≤ 20 mm Hg, triglycerides ≥ 700 mg/dl, total cholesterol ≥ 400 mg/dl, urinary creatinine ≥ 400 mg/dl, leukocyte telomere length t:s ratio > 6 , BMI > 70 kg/m². We also excluded individuals with C-reactive protein levels ≥ 5 mg/dl because the intent of using this biomarker for aging is to capture systemic inflammation, rather than acute inflammatory response to infection (Ridker 2003; Steel and Whitehead 1994).

In order to compare individuals across social environments within each country by education, while taking into account the differences in the implications of education across countries, education was divided into three major categories in each country. For Costa Rica, educational attainment was categorized as follows: less than three years of education, between three and six years of education (elementary school comprises six grades), and at least one year of high school. For the United States, we used the following educational categories: less than high school, high school, and more than high school.

2.3 Statistical analysis

In order to account for the potential non-linear relationships between age and biomarkers, we fitted generalized additive models that allow for a multiple slope spline, but penalize for overfit with internal cross validation. Thus, a linear model is used if it offers the best fit for the data penalized by degrees of freedom (Wood 2006). Figures 1–6 show a straight line if a linear model was the best fit for the data, or a curved line if a non-linear model was the best fit for the data. While the extent of the differences in the associations with age can be observed in these figures, we also fitted models with interaction terms that allow us to test statistically whether there are different associations between the biomarker and age by gender, place, and education.

Finally, we also fitted models that include all covariates in a generalized additive model to examine the overall variance in age that can be explained by all 12 biomarkers within each country context. Because there is no single slope that can be used to quantify the magnitude of the impact of each of the parameters in these models, we present the estimated degrees of freedom, the overall F-statistics, and the p-values that can be used to assess the relative strength of the association of the covariates within each model. A direct comparison of the test statistics across the models is not appropriate because the different sample sizes in each country will affect the sizes of the F-statistics and the p-values in the various models for associations of a similar magnitude. The models are run on the samples that had all covariates measured.

In order to avoid the bias caused by complete case analysis, we imputed missing observations for all variables using a nonparametric missing value imputation with Random Forest, implemented using the R package *missForest* (Stekhoven and Bühlmann 2012). This approach has several advantages relative to multiple imputation methods for missing data, including that it can be used with non-linear models. This method of imputing missing data has been shown to compare favorably with other missing data imputation approaches. In a validation test, the *missForest* algorithm was found to have the smallest prediction difference between the imputed and the non-missing datasets (Waljee et al. 2013).

Sample weights are used for all analyses. The weights range from zero to 4.61, with a mean of 1.0 within each sample, and an overall mean of 0.98.

Table 1:
Quantiles of the distribution (15th, 50th, and 85th) of biomarkers by study population

	CRELES			HRS			NHANES		
	15 th	50 th	85 th	15 th	50 th	85 th	15 th	50 th	85 th
Age	62	69	79	65	72	83	62	69	78
Systolic	120	142	168	113	131	154	116	135	160
Diastolic	71	82	95	66	77	89	58	70	82
HbA1c	5.1	5.5	6.4	5.2	5.7	6.5	5.2	5.5	6.3
Glucose	81	94	128	NA			85	95	116
CRP	0.2	0.3	0.8	0.1	0.2	0.7	0.1	0.3	0.8
HDL cholesterol	31	42	56	37	52	70	37	51	70
Triglycerides	87	144	244	NA			85	138	233
Total cholesterol	168	211	263	155	193	244	170	209	251
Creatinine	30	56	115	NA			37	94	170
BMI	22	26	32	22	26	33	23	27	33
Waist	82	93	106	81	99	116	85	100	115
Telomere length	0.8	1.0	1.2	1.0	1.2	1.6	0.7	0.9	1.1

Note: NA indicates that this biomarker was not collected in the HRS sample. Systolic and diastolic blood pressure are mm Hg, HbA1c is percent, glucose is mg/dl, CRP is mg/dl, HDL cholesterol is mg/dl, triglycerides is mg/dl, total cholesterol is mg/dl, creatinine is mg/dl, BMI is kg/m², waist is cm, telomere length is T/S ratio of leukocyte telomere length. The total sample sizes are as follows: 2,880 for CRELES, 13,634 for HRS, and 4,975 for NHANES. The sample sizes for each biomarker are as follows for CRELES, HRS, and NHANES, respectively: systolic ($n = 2309$, $n = 12914$, $n = 4290$), diastolic ($n = 2309$, $n = 12893$, $n = 4199$), HbA1c ($n = 2295$, $n = 13274$, $n = 4277$), glucose ($n = 2318$, NA, $n = 4194$), CRP ($n = 1264$, $n = 6537$, $n = 2569$), HDL cholesterol ($n = 2313$, $n = 10904$, $n = 2734$), triglycerides ($n = 2315$, NA, $n = 2091$), total cholesterol ($n = 2316$, $n = 12662$, $n = 4210$), creatinine ($n = 1131$, NA, $n = 4336$), BMI ($n = 2304$, $n = 13633$, $n = 4106$), Waist ($n = 2178$, $n = 6999$, $n = 4039$), telomere length ($n = 1248$, $n = 5946$, $n = 1196$).

3 Results

Table 1 presents the 15th, 50th, and 85th percentiles of the distribution of age, and for each of the 12 biomarkers we examined stratified by each of the three datasets; namely, the CRELES sample from Costa Rica, and the NHANES and the HRS samples from the United States. The table note describes the overall sample size for each dataset and the available sample for each biomarker. The distribution of age differs slightly across datasets due to the nature of the age sampling in each of the surveys. In particular, we note that because the public-use NHANES data do not identify by age individuals aged 85 and older, these individuals were not included in our analysis. Because all of our analyses are examined relative to age, small differences in the age distribution do not affect our findings. It is also important to note that only around 15% of the samples consist of individuals aged 80 and older, which limits the potential for bias due to mortality selection in our samples.

While the overall biomarker levels differ slightly across the samples, they are generally in the same range. Levels of systolic and diastolic blood pressure were a

bit higher in the CRELES sample than in the HRS and NHANES samples. Urinary creatinine was higher in the NHANES sample, while waist circumference was lower in the CRELES sample.

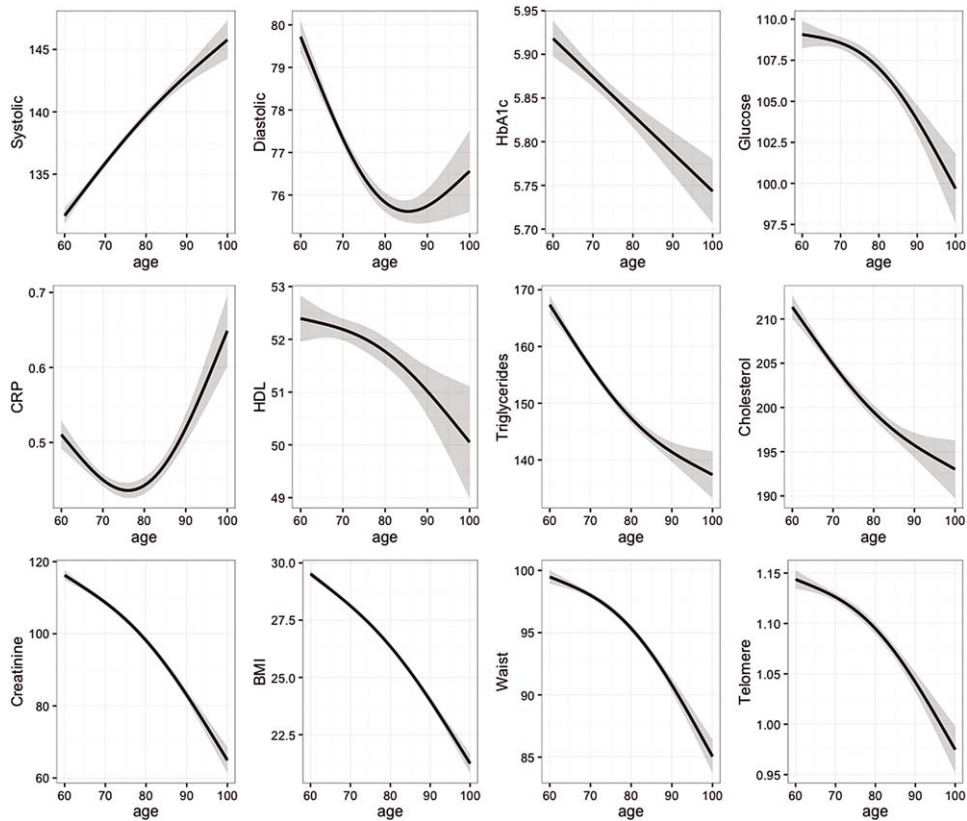
Figure 1 shows the relationship between age and each of the 12 biomarkers in the pooled database that includes all three samples. The plots are based on associations using a generalized additive model that allows for a non-linear dependence between age and the biomarkers. The light gray areas show 95% confidence intervals of the model-based associations. We observe a strong relationship with age across all 12 of these measures, consistent with our *a priori* hypothesis and with the findings of other literature. For biomarkers that are associated with worse health, the direction of the association is not in all cases in the expected direction of higher biomarker levels at older ages. For eight of the 12 biomarkers – diastolic blood pressure, HbA1c, glucose, triglycerides, total cholesterol, BMI, creatinine, and waist circumference – older age is associated with better values. Thus, of the measures that were chosen as biomarkers with important implications for health, the levels worsen with increasing age for only four. Of these four biomarkers that exhibit the expected trajectory with increasing age, two of these relationships generally worsen linearly with age (systolic blood pressure and HDL cholesterol), and two have a non-linear threshold relationship whereby the association with age is stronger after around age 75 (CRP and telomere length).

Figure 2 shows the same associations in the pooled samples, but now the associations are age-stratified by gender (the black line shows the associations among women, while the dark gray line shows the associations among men). When we look at the four biomarkers that have been shown to worsen with age in the overall population, we see that these relationships differ between men and women. While the relationship for systolic blood pressure holds in both genders, it is much stronger among women. The HDL cholesterol biomarker improves with age up to age 80 among women. For CRP, the relationship with age is strong among men only. For telomere length, the shortening of the telomere length with age is found among men only.

When we look at the factors that were not found to worsen with increasing age in the general population, we see that different associations emerge when the results are stratified by gender. For diastolic blood pressure, we find that levels increase with age among women aged 80 and older; for HbA1c, we see that levels rise with age among men aged 75 and older; and for glucose, we find that levels increase among women aged 75 and older.

Figure 3 presents the same analyses as those displayed in Figure 2, but now additionally stratified by country. The black lines are for women, the dark gray lines are for men, the solid lines are for Costa Rica, and the dashed lines are for the United States. About half of the associations between age and biomarkers that we examine are consistent across genders and places. The relationship is most consistent for BMI, with levels of BMI declining with age across genders and countries. The relationships between triglycerides, total cholesterol, and waist circumference are also generally similar across these four groups. For both systolic and diastolic

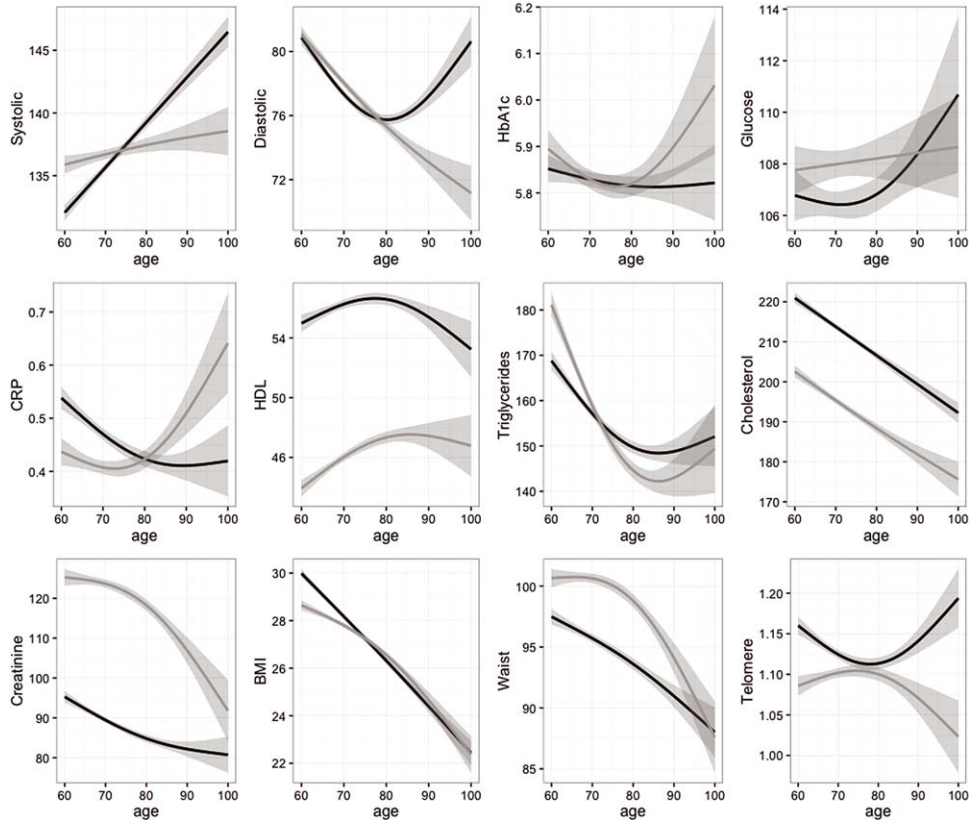
Figure 1:
The association between age and biomarkers



blood pressure, the patterns are generally similar up until around the age of 80. Thus, systolic blood pressure up to the age of 80 is the only biomarker of the 12 we examine that worsens with age across genders and country environments. The most clearly divergent patterns are observed for CRP among women, with the levels increasing with age in Costa Rica and decreasing with age in the U.S.; and for HDL cholesterol, with the levels decreasing with age in the United States and increasing with age in Costa Rica.

Figure 4 presents associations between biomarkers and age stratified by gender and level of education. The black lines are for women, the dark gray lines are for men, the solid lines are for individuals with lower levels of education, and the dashed lines are for individuals with higher levels of education. As in the previous models stratified by gender and education, we find that around half of the biomarkers had similar relationships with age across the four strata. Again, the relationships were

Figure 2:
The association between age and biomarkers by gender

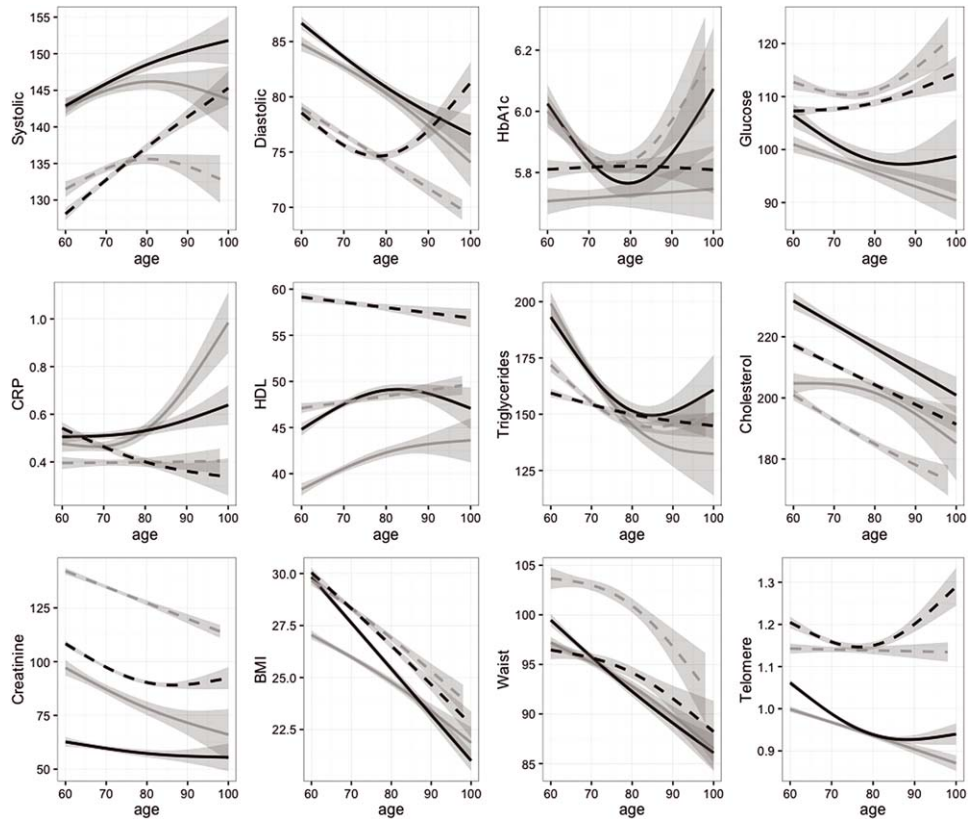


Note: black line is for women, dark gray line is for men.

most similar for BMI, waist, triglycerides, and cholesterol; but in all of these cases the levels declined with age. Moreover, as in the models stratified by country, we see that systolic blood pressure increased with age up to around age 85 – the only portion of the age distribution for any variable that showed worse biomarker levels at higher ages.

The prior analyses were all focused on years of age, which is the standard metric for examining biomarkers of aging. However, other scholars have suggested that thanatological age may be a more relevant metric, because it specifically acknowledges that there is a great deal of life expectancy heterogeneity (Riffe 2015). As years of age from birth compares people with different cumulative insults at a given age, these universal slope comparisons can be difficult to interpret. Thanatological age, which measures years before death, represents a different

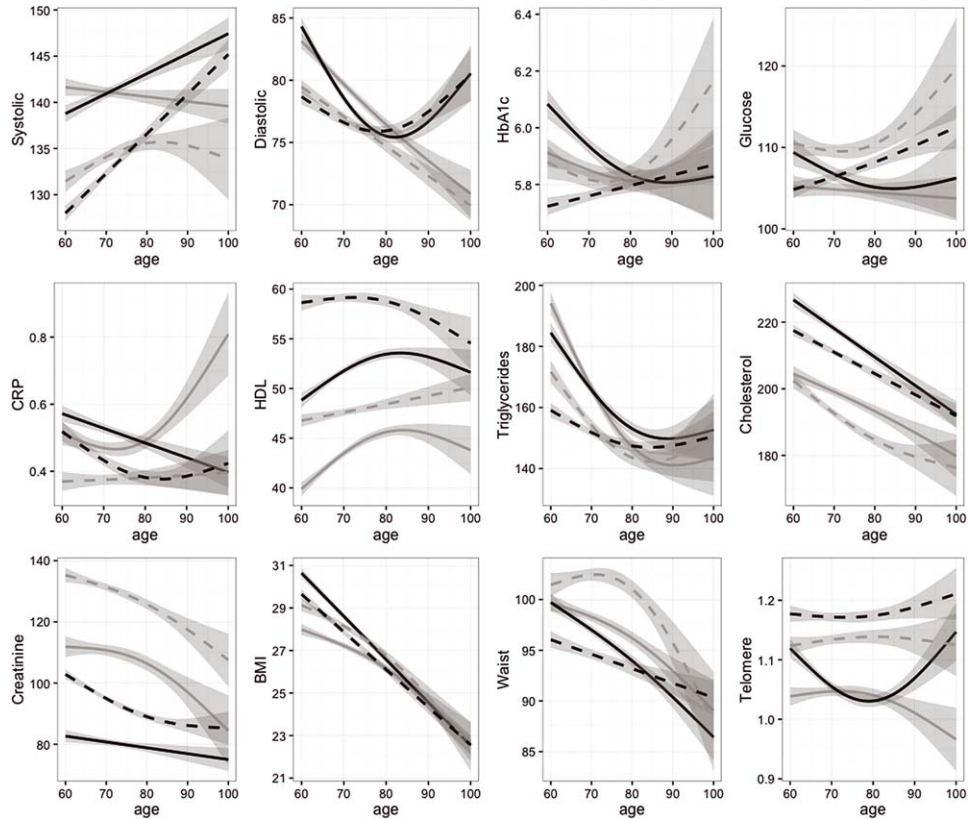
Figure 3:
The association between age and biomarkers by gender and country



Note: black lines are for women, dark gray lines are for men, solid lines are for Costa Rica, dashed lines are for the United States.

metric for comparing biomarker patterns with aging, while alternately standardizing average disease states. For these analyses, we operationalize thanatological age as years before actual death among our participants who died during the follow-up (rather than life expectancy, which can also be done). The results for the pooled samples are shown in Figure 5. When we compare these findings with the results presented in Figure 1, we see both similarities and differences. For eight of the 12 factors we examined, the biomarker associations with age from birth were similar to those with thanatological age (glucose, CRP, HDL cholesterol, triglycerides, cholesterol, creatinine, BMI, and waist circumference). For one other factor, HbA1c, there was a negative association with age, but no association with thanatological age. For three other biomarkers – systolic blood pressure, diastolic blood pressure, and telomere length – the associations between age and thanatological age were in

Figure 4:
The association between age and biomarkers by gender and education

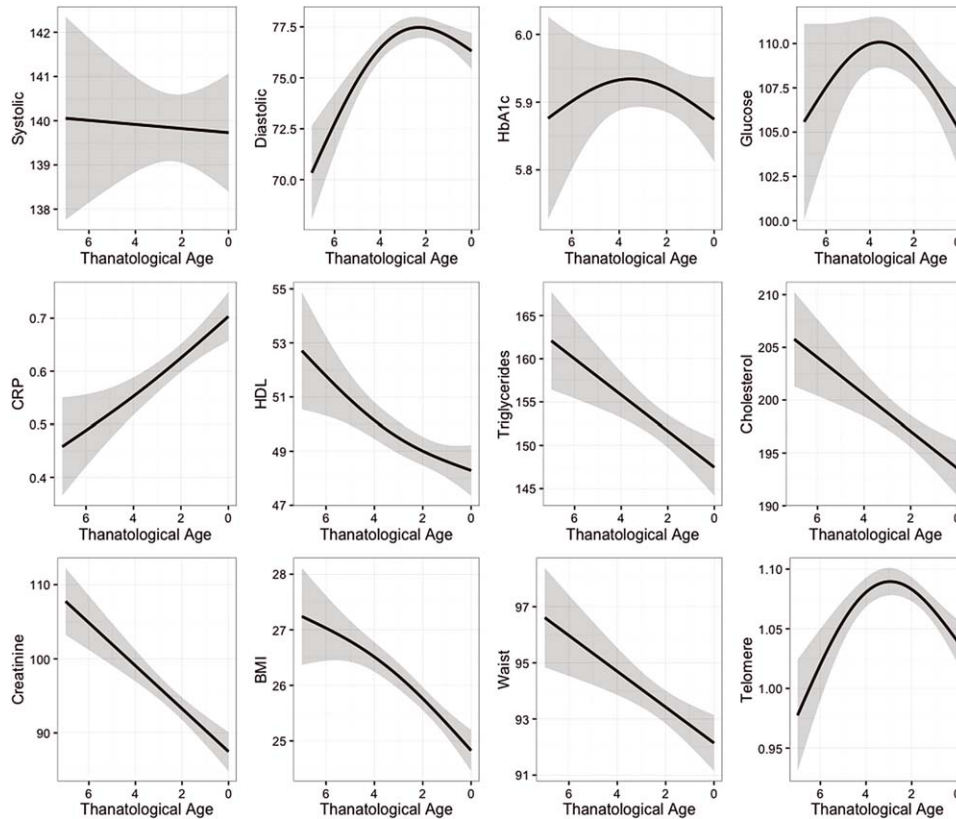


Note: black lines are for women, dark gray lines are for men, solid lines are for lower education, dashed lines are for higher education.

the opposite direction. In the years close to death, levels of systolic blood pressure were lower, but levels of diastolic blood pressure were higher and telomeres were longer. When we look at all 12 biomarkers in the aggregated overall population, we see that diastolic blood pressure, CRP, and HDL cholesterol were the only biomarkers that worsened as the number of years until death declined.

A feature worth noting in Figure 5 is that in half of the markers, the relationship with thanatological age is curvilinear: at about two to four years before death the slope of the association changes for diastolic blood pressure, HbA1c, fasting glucose, BMI, and telomere length. These changes in slope assume two general forms. In one form there is an association with age up until two to four years before death, but this association disappears in the two to four years immediately before death (diastolic blood pressure, HbA1c, and telomere length). In contrast, for

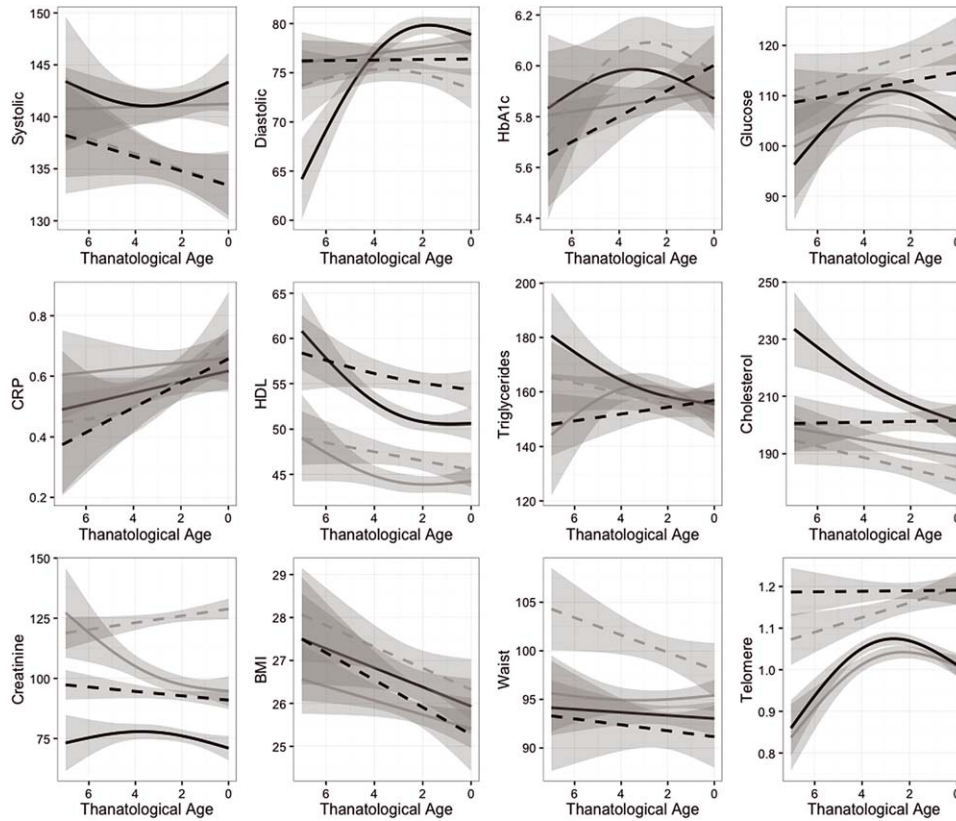
Figure 5:
The association between thanatological age and biomarkers



fasting glucose and BMI, a relationship with age is observed only in the two to four years immediately prior to death, during which levels of both of these biomarkers decrease.

In a final analysis of the relevance of thanatological age, we examine the relationship stratified by gender and level of education. We hypothesized that these thanatological relationships might be more closely aligned with biological aging than relationships with age from birth, because life expectancy differs considerably by both gender and level of education. In contrast to our hypothesis, we find that the heterogeneity of associations between population strata and thanatological age are similar, but not identical, to those with age since birth. For example, while we observe consistent associations for HDL cholesterol, CRP, BMI, and waist circumference with chronological age; we find more deleterious levels closer to death for HDL cholesterol and CRP.

Figure 6:
The association between thanatological age and biomarkers by gender and education



Note: black lines are for women, dark gray lines are for men, solid lines are for lower education, dashed lines are for higher education.

When we examine the curvilinear associations observed in the combined population, we see that most of these non-linear associations are found only within particular subgroups of the population, or not at all. For diastolic blood pressure, the threshold association is found primarily among women with lower levels of education. For HbA1c, fasting glucose, and telomere length, the threshold relationship is observed among those with lower levels of education; whereas an association with age is found only up until two to four years before death. When stratified by these four groups, the lower levels of BMI closer to death are fairly linear across thanatological ages. For urinary creatinine, the curvilinear relationship is found among men. For diastolic blood pressure, the strong relationship with age up until two to four years prior to death is observed among individuals with lower levels of education.

Table 2:
P-values for the interaction terms between age and gender, age and place, and age and education

	Gender p-value	Place p-value	Education p-value
Diastolic	<0.001	<0.001	<0.001
Systolic	<0.001	<0.001	<0.001
Total cholesterol	0.277	0.911	<0.001
BMI	0.158	<0.001	<0.001
Telomere length	<0.001	<0.001	0.002
HbA1c	<0.001	<0.001	<0.001
Creatinine	<0.001	<0.001	<0.001
CRP	0.587	0.661	0.804
Triglycerides	<0.001	<0.001	0.002
Glucose	<0.001	0.035	0.115
HDL cholesterol	0.008	<0.001	<0.001
Waist	0.001	<0.001	0.021

Note: P-values are from the interaction terms with age. All of the models also included controls for the other two factors that were not the interaction.

The previous descriptions are based on a visual assessment of the similarity of model-based estimates of the associations of biomarkers with age from birth and with age to death. Table 2 presents p-values of the interaction terms between age and gender, place, and education in order to test whether there are statistically significant differences in the association between biomarkers and age by gender, place, and education. The only two biomarkers for which no clear interaction with gender are found are for BMI, CRP, and total cholesterol. For all of the other biomarkers, the relationship between age differs markedly by gender. For place, only CRP and total cholesterol does not differ; all of the other interactions between place and age have p-values consistent with an interaction. For education, CRP and glucose do not differ. Overall, these differences are much more numerous than would be expected by chance.

Table 3 shows the strength of the non-linear relationships between all biomarkers and age based on the results of a single regression model for each country. The relative strength of the association, conditional on all of the other biomarkers in the model, could be assessed with the F-statistic within each model. Much more of the variance in age is explained in Costa Rica (59%) than in the United States (28%).

Tables 4 and 5 show the same relationships, but stratified by gender and level of education, respectively. The amount of variance explained is similar for women (34%) and for men (29%) (Table 4). More variance in age is explained for

Table 3:
Full model of biomarkers predicting age in Costa Rica and in the United States

	Costa Rica			United States		
	edf	F-stat	p-value	edf	F-stat	p-value
Diastolic	8.8	40	<0.001	7.4	8.4	<0.001
Systolic	6.9	51	<0.001	5.4	6.6	<0.001
Total Cholesterol	1.6	0.6	0.54	7.2	8.3	<0.001
BMI	8.7	63	<0.001	3.7	4.7	<0.001
Telomere length	8.7	110	<0.001	8.9	9.0	<0.001
HbA1c	5.2	6.2	<0.001	6.8	7.9	<0.001
Creatinine	3.0	9.4	<0.001	7.7	8.5	<0.001
CRP	7.8	15	<0.001	3.2	4.0	<0.001
Triglycerides	8.3	6.2	<0.001	7.5	8.5	<0.001
Glucose	8.0	6.8	<0.001	8.5	8.9	<0.001
HDL cholesterol	7.4	4.0	<0.001	5.9	7.2	<0.001
Waist	8.4	12	<0.001	8.2	8.8	<0.001
	R-sq.(adj) = 0.583			R-sq.(adj) = 0.273		
	Deviance explained = 59%			Deviance explained = 28%		

Note: edf is estimated degrees of freedom. Estimated degrees of freedom are based on penalized degrees of freedom from the spline model.

individuals with lower levels of education (38%) than for individuals with higher levels of education (27%).

4 Discussion

Our analyses examined how 12 potential biomarkers of aging are correlated with age in terms of both the number of years since birth and the number of years until death. The first minimal criterion we used to determine whether each of these biomarkers should be considered a universal biomarker of aging was a test of whether the biomarker was correlated with age; i.e., whether the physiological levels of the biomarker deteriorated with age. Out of the 12 biomarkers tested, only four met this criterion: namely, systolic blood pressure, C-reactive protein, HDL cholesterol, and telomere length. The second minimal criterion that we tested for each of these biomarkers was whether the sign of this correlation was consistent across genders, countries, and levels of education. None of the biomarkers we examined both worsened with age and were consistent across these groups. The results were similar when we examined both age since birth and thanatological age, although two factors met our minimal criteria using thanatological age: namely, CRP and HDL cholesterol. We therefore conclude that of the biological and anthropometric

Table 4:
Full model of biomarkers predicting age by gender

	Women			Men		
	edf	F-stat	p-value	edf	F-stat	p-value
Diastolic	8.8	29	<0.001	8.8	31	<0.001
Systolic	8.3	63	<0.001	6.5	43	<0.001
Total Cholesterol	5.6	8.9	<0.001	5.2	3.6	0.001
BMI	4.9	140	<0.001	7.5	53	<0.001
Telomere length	8.8	54	<0.001	8.9	38	<0.001
HbA1c	8.1	5.5	<0.001	6.0	6.6	<0.001
Creatinine	6.5	28	<0.001	4.3	22	<0.001
CRP	3.3	4.1	0.002	5.0	11	<0.001
Triglycerides	5.7	8.5	<0.001	6.6	12	<0.001
Glucose	8.5	13	<0.001	6.7	2.5	0.009
HDL cholesterol	7.1	9.3	<0.001	7.8	7.0	<0.001
Waist	8.48	21	<0.001	5.2	13	<0.001
	R-sq.(adj) = 0.335			R-sq.(adj) = 0.288		
	Deviance explained = 34%			Deviance explained = 29%		

Note: edf is estimated degrees of freedom. Estimated degrees of freedom are based on penalized degrees of freedom from the spline model.

measures that we examined here, and which have been shown to be associated with mortality, none fit our criteria for being considered a biomarker of aging; i.e., none of these measures were shown to both worsen with age and display consistency across genders, countries, and educational groups. The main contribution of our analysis is that it serves as a population-based demographic complement to research on biomarkers of aging conducted in individual-based physiological studies, animal studies, and *in vitro* studies. Our results suggest that it is critical to examine whether both individual and composite measures of aging are predictive of age, not just in the overall population, but in subsets of the population and in different countries.

Our analysis has several limitations. The most important of these limitations is that we were constrained in the number and in the types of biomarkers we were able to examine by data availability. While two of the three datasets (CRELES and HRS) were specifically focused on aging, the limited availability of other types of biomarkers in the HRS and the NHANES samples precluded their inclusion. This is an important limitation to consider when looking at our results in relation to the findings of other studies, as a number of factors – including IL-6, DHEA, and norepinephrine – that have been used in indices of biomarkers of aging (Gruenewald, Seeman, Ryff, Karlamangla and Singer 2006; Johnson 2006) could not be included in our analysis. Yet despite these drawbacks, an advantage of our selection of biomarkers is that each of these biomarkers clearly meets the criterion of being a

Table 5:
Full model of biomarkers predicting age by education

	Low education			High education		
	edf	F-stat	p-value	edf	F-stat	p-value
Diastolic	8.8	38	<0.001	5.5	51	<0.001
Systolic	8.2	46	<0.001	5.2	82	<0.001
Total Cholesterol	4.9	4.4	<0.001	6.0	6.0	<0.001
BMI	5.2	110	<0.001	6.7	57	<0.001
Telomere	8.8	62	<0.001	8.9	46	<0.001
HbA1c	4.5	4.4	<0.001	2.6	4.1	0.006
Creatinine	5.0	29	<0.001	7.0	34	<0.001
CRP	5.5	6.4	<0.001	1.9	0.4	0.567
Triglycerides	4.5	16	<0.001	8.1	4.0	<0.001
Glucose	7.8	2.9	0.002	7.9	26	<0.001
HDL cholesterol	4.9	5.4	<0.001	5.3	6.9	<0.001
Waist	8.1	11	<0.001	7.8	15	<0.001
	R-sq.(adj) = 0.375			R-sq.(adj) = 0.267		
	Deviance explained = 38%			Deviance explained = 27%		

Note: edf is estimated degrees of freedom. Estimated degrees of freedom are based on penalized degrees of freedom from the spline model.

precursor to chronic disease risk, rather than a consequence of disease. However, the question of whether biomarkers of aging need to be associated with mortality itself remains open. Opinions on this issue differ depending on which definition of biomarker of aging is being used. Another limitation of this analysis is that in our calculation of thanatological age, the range was limited by the amount of follow-up time. Thus, we were able to examine this relationship over a fairly short time span only; i.e., over a period of up to seven years. A further limitation is that we did not account for selection out of the sample due to mortality. The associations with age that we examined may have been affected by the possibility that individuals whose biomarker values worsened with age were more likely to die, and were therefore not represented in our samples. We note, however, that we were able to examine the extent to which this was occurring by fitting non-linear models, and did not observe any general patterns of weakening relationships with age. Nevertheless, it is important to keep in mind that our findings apply only to people aged 60 and older, and that the relationships among people under age 60 may be different, and highly relevant to identifying biomarkers of aging.

Our results appear to support some recent reports in the literature that cast doubt on the possibility of identifying universal biomarkers of aging (Johnson 2006). They are also in line with the findings of other studies indicating that demographic patterns of biomarkers may vary by country context. For example,

while a large number of studies have shown that HDL cholesterol is higher in women than men, more detailed analyses have shown that the size of this gender gap differs by country, with the gap being very large in some countries and almost undetectable in others (Davis et al. 1996). The conclusions of the current analysis are also similar to the findings of prior work showing that the associations between biomarkers and educational levels differ between the U.S. and Costa Rica (Rehkopf, Dow and Rosero-Bixby 2010). In some cases, it has been shown that relationships between biomarkers and factors such as age, gender, and socioeconomic position that appear to be consistent are actually context-dependent, or are not robust to examination when samples from other countries are used. Telomere length represents a particularly striking example of this context dependence. While it is one of the more consistent biomarkers we examined, we still found differences in the relationship with age by gender, and striking non-linear relationships with age, particularly among less educated women in the U.S. The reasons for these and other heterogeneous associations with age should be further investigated in future work.

While we did not find any specific biomarkers that consistently worsened with age across groups, we did find that in our overall models a substantial amount of variation in age was explained: i.e., more than 25%, and up to nearly 60% in Costa Rica. This result is consistent with the observation made by a number of scholars that aggregated indexes of biomarkers may be useful for measuring aging (Der et al. 2012; Gruenewald et al. 2006). However, based on our analysis of changes in individual biomarkers with age, it appears that a relatively large portion of the explained variance is attributable to associations that may be in a direction that is the opposite of the one we expected. Of particular interest is our observation that the associations with thanatological age in the years immediately prior to death differed from the associations with age from birth, as this finding suggests that some of these biomarker relationships may be due to physiological changes that occurred close to death, and may thus have been caused by chronic disease processes. While the scope of our data limits our ability to speculate about this possibility, our findings provide further support for the claim that these factors should not be considered biomarkers of aging. More generally, additional study is needed to answer the question of why these associations exist, and to determine the extent to which these associations are attributable to selection out of the sample due to the deaths of individuals with relatively poor biomarker values, or to more intense medical management of diseases.

Our findings suggest that a number of factors that have been shown to be associated with mortality are not universal biomarkers of aging, especially if we use as our criteria for determining whether they are biomarkers of aging not just whether they are correlated with age, but whether they become more deleterious with age. In light of findings for other measures, such as functional status, grip strength, self-reported health, and walking speed, we may discover that measures that are less specific to a single physiological system are more useful than risk factors for chronic diseases for measuring the aging process.

Acknowledgements

The CRELES data were collected by the Central American Population Center of the University of Costa Rica with support from Wellcome Trust grant 072406.

Dr. Rehkopf was supported by the National Institute on Aging (K01AG047280). We also acknowledge funding from R01AG031716.

References

- Arking, R. 2006. *Biology of aging: observations and principles* (Vol. 7): Oxford University Press New York.
- Baker, G. T., 3rd and R. L. Sprott 1988. Biomarkers of aging. *Experimental gerontology* 23(4-5): 223–239. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3058488>.
- Belsky, D. W., A. Caspi, R. Houts, H. J. Cohen, D. L. Corcoran, A. Danese, T. E. Moffitt 2015. Quantification of biological aging in young adults. *Proc Natl Acad Sci USA* 112(30): E4104–4110. doi:10.1073/pnas.1506264112.
- Boonekamp, J. J., M. J. Simons, L. Hemerik and S. Verhulst 2013. Telomere length behaves as biomarker of somatic redundancy rather than biological age. *Aging cell* 12(2): 330–332. doi:10.1111/accel.12050.
- Butler, R. N., R. Sprott, H. Warner, J. Bland, R. Feuers, M. Forster and M. Hyman 2004. Aging: the reality biomarkers of aging: from primitive organisms to humans. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 59(6): B560–B567.
- Cawthon, R. M. 2002. Telomere measurement by quantitative PCR. *Nucleic Acids Res* 30(10): e47. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12000852>.
- Crimmins, E., S. Vasunilashorn, J. K. Kim and D. Alley 2008. Biomarkers related to aging in human populations. *Adv Clin Chem* 46: 161–216.
- Davis, C., D. Williams, R. Oganov, S.-C. Tao, S. Rywik, Y. Stein and J. Little 1996. Sex difference in high density lipoprotein cholesterol in six countries. *American Journal of Epidemiology* 143(11): 1100–1106.
- Der, G., G. D. Batty, M. Benzeval, I. J. Deary, M. J. Green, L. McGlynn, P. G. Shiels 2012. Is telomere length a biomarker for aging: cross-sectional evidence from the west of Scotland? *PLoS One* 7(9), e45166. doi:10.1371/journal.pone.0045166.
- Gruenewald, T. L., T. E. Seeman, C. D. Ryff, A. S. Karlamangla and B. H. Singer 2006. Combinations of biomarkers predictive of later life mortality. *Proc Natl Acad Sci USA* 103(38): 14158–14163. doi:0606215103 [pii] 10.1073/pnas.0606215103.
- Ingram, D. K., E. Nakamura, D. Smucny, G. S. Roth and M. A. Lane 2001. Strategy for identifying biomarkers of aging in long-lived species. *Experimental Gerontology* 36(7): 1025–1034.
- Johnson, T. E. 2006. Recent results: biomarkers of aging. *Experimental Gerontology* 41(12): 1243–1246.
- Juster, F. T. and R. Suzman 1995. An overview of the health and retirement study. *Journal of Human Resources* 30: S7–S56. doi:10.2307/146277.

- Karlamangla, A. S., B. H. Singer, B. S. McEwen, J. W. Rowe and T. E. Seeman 2002. Allostatic load as a predictor of functional decline: MacArthur studies of successful aging. *Journal of Clinical Epidemiology* 55(7), 696–710.
- Karlamangla, A. S., B. H. Singer and T. E. Seeman 2006. Reduction in allostatic load in older adults is associated with lower all-cause mortality risk: MacArthur studies of successful aging. *Psychosomatic Medicine* 68(3): 500–507.
- Krieger, N. 2005. Embodiment: a conceptual glossary for epidemiology. *Journal of Epidemiology and Community Health* 59(5): 350–355.
- Levine, M. E. 2013. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? *J Gerontol A Biol Sci Med Sci* 68(6): 667–674. doi:10.1093/gerona/gls233.
- Mather, K. A., A. F. Jorm, R. A. Parslow and H. Christensen 2011. Is telomere length a biomarker of aging? A review. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 66(2): 202–213. doi:10.1093/gerona/glq180.
- Nadon, N. L. 2006. Of mice and monkeys: National Institute on Aging resources supporting the use of animal models in biogerontology research. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 61(8): 813–815.
- NCHS 1999. National Health and Nutrition Examination Survey. Retrieved from http://wwwn.cdc.gov/nchs/nhanes/search/nhanes99_00.aspx.
- Needham, B. L., N. Adler, S. Gregorich, D. Rehkopf, J. Lin, E. H. Blackburn and E. S. Epel 2013. Socioeconomic status, health behavior, and leukocyte telomere length in the National Health and Nutrition Examination Survey, 1999–2002. *Soc Sci Med* 85, 1–8. doi:10.1016/j.socscimed.2013.02.023.
- RAND 2011. *HRS Data, Version L, Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration*.
- Rehkopf, D. H., W. H. Dow and L. Rosero-Bixby 2010. Differences in the association of cardiovascular risk factors with education: a comparison of Costa Rica (CRELES) and the USA (NHANES). *Journal of Epidemiology and Community Health* 64, 821–828. doi: 10.1136/jech.2009.086926.
- Rehkopf, D. H., W. H. Dow, L. Rosero-Bixby, J. Lin, E. S. Epel and E. H. Blackburn 2014. Seasonal variation of peripheral blood leukocyte telomere length in Costa Rica: A population-based observational study. *American Journal of Human Biology: the Official Journal of the Human Biology Council* 26(3): 367–375. doi: 10.1002/ajhb.22529.
- Ridker, P. M. 2003. Clinical application of C-reactive protein for cardiovascular disease detection and prevention. *Circulation* 107(3): 363–369.
- Riffe, T. 2015. The force of mortality by life lived is the force of increment by life left in stationary populations. *Demogr Res* 32: 827.
- Rogot, E., M. Feinleib, K. A. Ockay, S. H. Schwartz, R. Bilgrad and J. E. Patterson 1983. On the feasibility of linking census samples to the National Death Index for epidemiologic studies: a progress report. *American Journal of Public Health* 73: 1265–1269.
- Rosero Bixby, L., X. Fernandez and W. H. Dow 2010. *CRELES: Costa Rican Longevity and Healthy Aging Study, 2005 (Costa Rica Estudio de Longevidad y Envejecimiento Saludable: Sampling and Methods No. ICPSR26681-v2)*.

- Rosero-Bixby, L. and W. H. Dow 2012. Predicting mortality with biomarkers: a population-based prospective cohort study for elderly Costa Ricans. *Population Health Metrics* 10(11): 1–15.
- Sanders, J. L. and A. B. Newman 2013. Telomere length in epidemiology: a biomarker of aging, age-related disease, both, or neither? *Epidemiologic Reviews* 35(1): 112–131.
- Seeman, T. E., B. S. McEwen, J. W. Rowe and B. H. Singer 2001. Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. *Proceedings of the National Academy of Sciences* 98(8): 4770–4775.
- Steel, D. M. and A. S. Whitehead 1994. The major acute phase reactants: C-reactive protein, serum amyloid P component and serum amyloid A protein. *Immunology Today* 15(2), 81–88.
- Stekhoven, D. J. and P. Bühlmann 2012. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* 28(1): 112–118.
- Waljee, A. K., A. Mukherjee, A. G. Singal, Y. Zhang, J. Warren, U. Balis and P. D. Higgins 2013. Comparison of imputation methods for missing laboratory data in medicine. *BMJ Open* 3(8): e002847. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3733317/pdf/bmjopen-2013-002847.pdf>.
- Weinstein, M., J. W. Vaupel, K. W. Wachter and D. Weir 2008. Elastic powers: The integration of biomarkers into the Health and Retirement Study.
- Wood, S. 2006. *Generalized Additive Models: An Introduction with R*. Boca Raton: Chapman & Hall/CRC.
- Zahn, J. M., S. Poosala, A. B. Owen, D. K. Ingram, A. Lustig, A. Carter and K. Mazan-Mamczarz 2007. AGEMAP: a gene expression database for aging in mice. *PLoS Genetics* 3(11): e201.