



Published in final edited form as:

*Popul Dev Rev.* 2014 June 1; 40(2): 331–360. doi:10.1111/j.1728-4457.2014.00676.x.

## Beyond Self-Reports: Changes in Biomarkers as Predictors of Mortality

Dana A. Gleia<sup>a,\*</sup>, Noreen Goldman<sup>b</sup>, Germán Rodríguez<sup>c</sup>, and Maxine Weinstein<sup>d</sup>

<sup>a</sup>Senior Research Investigator, Center for Population and Health, Georgetown University

<sup>b</sup>Hughes-Rogers Professor of Demography and Public Affairs, Office of Population Research, Princeton University

<sup>c</sup>Senior Research Demographer, Office of Population Research, Princeton University

<sup>d</sup>Distinguished Professor, Center for Population and Health, Georgetown University

### Abstract

The proliferation of biosocial surveys has increased the importance of weighing the costs and benefits of adding biomarker collection to population-based surveys. A crucial question is whether biomarkers offer incremental value beyond self-reported measures, which are easier to collect and impose less respondent burden. We use longitudinal data from a nationally representative sample of older Taiwanese ( $n = 639$ , aged 54+ in 2000, examined in 2000 and 2006 with mortality follow-up through 2011) to address that question with respect to predicting all-cause mortality. A summary measure of biomarkers improves mortality prediction (as measured by the area under the receiver operating characteristic curve) compared with self-reports alone, but individual biomarkers perform better than the summary score. We find that incorporating change in biomarkers over a six-year period yields a small improvement in mortality prediction compared with one-time measurement. But, is the incremental value worth the costs?

### INTRODUCTION

Publication of *Between Zeus and the Salmon: The Biodemography of Longevity* (Wachter and Finch 1997) heightened interest in the collection of biological measures of chronic disease within demographic household surveys. Prior to the 1990s, the inclusion of such measures was generally limited to epidemiologic studies, which were often based on geographically limited samples and generally included only limited social data. However, since then, growing numbers of population-based surveys have collected biological measures alongside questionnaires that rely on self-reports of health and extensive household and individual information. As these “biosocial” surveys have proliferated, increasing levels of research funding are being devoted to their collection. Thus, it is important to weigh the benefits of biomarker collection against the costs: financial expense, logistical complications, increased respondent burden, ethical issues, and threats to privacy.

\*Corresponding Author: Dana A. Gleib, 5985 San Aleso Ct., Santa Rosa, CA 95409-3912, USA. Phone (707) 539-5592; dag77@georgetown.edu.

There are many ways to quantify the value of the biomarkers being collected. Here, we focus on only one of the potential uses: their value in predicting mortality. We use mortality because it is a well defined and salient health outcome with little measurement error, assuming that vital status is determined from virtually complete death registration records. Friedman and Kern (2014) argue that all-cause mortality is the single best measure of health, noting that it is used worldwide as a key indicator of public health. Biomarkers of individual diseases are, of course, of particular clinical interest, however the biomarkers evaluated here relate to a broad range of diseases. Thus, we would obtain an incomplete – perhaps even biased – assessment of their predictive power if we used a specific disease or cause of death as the benchmark. In addition to strong validity, all-cause mortality has the advantage of avoiding the problem that studies of a single disease may miss some deaths that were attributed to a different cause, but resulted indirectly from the disease of interest (Friedman and Kern 2014).

As biosocial surveys have evolved, some studies have begun longitudinal collection of biomarkers with the expectation that measurement at multiple survey waves will improve our understanding of the link between biomarkers and mortality (Crimmins and Vasunilashorn 2011). As yet, however, little data have been available to test this assumption. Prior studies have demonstrated that biomarkers are useful predictors of mortality, but few have assessed the prognostic value of change in biomarkers to determine whether they offer greater value than measurement of biomarker levels at a single wave. Collecting a wide range of biomarkers is costly, so it is also important to determine whether some biomarkers are more valuable than others or whether most of the power derives from markers representing a particular biological system.

In this analysis, we use longitudinal data from the 2000 and 2006 waves of the Social Environment and Biomarkers of Aging Study (SEBAS) in Taiwan, one of the first biosocial surveys to collect two rounds of biomarkers, to investigate four hypotheses motivated by the existing literature. First, we expect that biomarkers will have greater value for predicting mortality than self-reported measures of health when compared side by side. Second, we hypothesize that biomarkers will retain incremental prognostic value above and beyond self-reports. Third, we expect that measurement of change in biomarkers will yield better discrimination (i.e., ability to distinguish between those who died and those who did not) than a single measurement of biomarkers. Fourth, based on prior studies that have included a wide range of biomarkers (e.g., Crimmins and Vasunilashorn 2011, Newman et al. 2009, Sidorenkov, Nilssen, and Grjibovski 2012, Swindell et al. 2010), we anticipate that markers drawn from multiple biological systems will contribute prognostic value.

## **PREDICTING MORTALITY**

### **Self-reported indicators of health status**

Health indicators obtained from self-reports are typically important mortality predictors. Indeed, many (but not all) of the indicators incorporated into prognostic indices designed for clinical use (Yourman et al. 2012) could be obtained via self-report, although their validity would need to be assessed. For the purposes of predicting longer-term mortality (over at least four years) in community-based populations, two such indices have demonstrated good

(Schonberg et al. 2009) or very good (Lee et al. 2006b) discrimination, although none of the existing prognostic indices has attained excellent discrimination (Yourman et al. 2012). Among the health indicators in these indices are: mobility/functional limitations, history of cancer, current diabetes, smoking status, self-assessed health status, and number of hospitalizations in the past year (Schonberg et al. 2009, Lee et al. 2006b); see also [www.epronosis.org](http://www.epronosis.org). Other indicators include current chronic obstructive pulmonary disease (COPD), congestive heart failure, and body mass index, but self-reports may not reliably capture these variables.

As suggested above, self-reported measures have a downside: potential inaccuracy. Misreporting may result from recall bias or from unwillingness to acknowledge health problems. Respondents may also be unaware of some underlying conditions that confer health risk. Then, too, there is a subjective component: for example, two people with similar levels of physical ability may differ in their perceived—and consequently, reported—difficulty performing a specified task (Melzer et al. 2004). All of these factors can contribute to measurement error. One appeal of biological markers is that they may be less subject to such errors.

## Biomarkers

Some studies that examine the relationship between biomarkers and mortality use individual biomarkers as predictors, whereas others develop a composite measure based on multiple markers. Studies also differ in the number and range of markers they use: some rely on a few closely related markers that reflect a common biological system, while others include a larger number of markers that represent multiple systems (e.g., cardiovascular, metabolic, immune/inflammatory, neuroendocrine, renal, hepatic). Given the vast literature linking biomarkers and mortality, here we focus on the smaller set of studies that include markers drawn from multiple systems. We divide our review into two parts: 1) studies that use individual markers; and 2) those that use multi-system summary scores.

### **Studies Using Individual Markers Representing Multiple Biological Systems—**

Many studies have investigated the association between survival and a collection of individual biomarkers drawn from multiple biological systems (e.g., Crimmins, Kim, and Vasunilashorn 2010, Crimmins and Vasunilashorn 2011, Fried et al. 1998, Kistorp et al. 2005, Lan et al. 2007, Lee et al. 2006a, Newman et al. 2009, Sidorenkov, Nilssen, and Grijbovski 2012, Swindell et al. 2010, Turra et al. 2005). These studies demonstrate that cardiovascular, metabolic, inflammatory, neuroendocrine, and other markers that reflect disease (e.g., kidney, liver) are useful predictors of mortality. An important question is: how do we quantify the predictive ability of a particular set of biomarkers?

The most common approach, by far, for quantifying the predictive ability of a model is the C-statistic or Area Under the Receiver Operating Characteristic Curve (AUC). Receiver operating characteristic (ROC) analysis quantifies the model's accuracy in discriminating between two states (in this case, alive or dead) as the discrimination threshold varies. Variation in the “cut-off” threshold used to classify subjects into the two groups typically involves a trade-off between sensitivity (those who died are classified as “dead”) and specificity (those who survived are classified as “alive”). A lower cut-off yields greater

sensitivity but lower specificity. The ROC curve plots sensitivity against 1-specificity for all possible cut-off values of the classifier. The area under that curve is used to summarize the global performance of the classifier. The AUC can be interpreted as the probability that those who actually died are assigned a higher predicted probability of death than those who survived (Pencina and D'Agostino 2004). So, an AUC of 0.5 would indicate that the model performs no better than chance alone and an AUC of 1.0 would represent perfect accuracy (100% sensitivity and specificity). Statistical significance alone is not sufficient to demonstrate that a marker substantially improves model performance, and even a large effect size does not necessarily translate into a meaningful improvement in discriminatory ability (Pencina et al. 2008). The improvement in the C-statistic/AUC associated with an additional indicator is often small in magnitude even when the odds ratio associated with that marker is large; Pencina et al. (2008) suggest that an increase in AUC of 0.01 may be considered a meaningful improvement.

Here, we limit our review to studies that: 1) examined individual biomarkers drawn from multiple systems as predictors of all-cause mortality, where the markers include standard cardiovascular/metabolic markers (e.g., blood pressure, lipids, measures of glucose metabolism, and obesity) as well as at least one inflammatory marker (e.g., interleukin-6, IL-6; C-reactive protein, CRP); and 2) assessed the discriminatory ability of the set of biomarkers. We identified six such studies (see Table 1) based on five unique datasets in four countries: Costa Rica (Rosero-Bixby and Dow 2012), Germany (Haring et al. 2011), the Netherlands (Newson et al. 2010, Stork et al. 2006), and Taiwan (Goldman et al. 2006b, Goldman et al. 2009). Some of these studies also included neuroendocrine markers such as cortisol, dehydroepiandrosterone sulfate (DHEAS), epinephrine and norepinephrine. The six studies also incorporated “other” biomarkers that cannot be easily classified into one of the above-mentioned groups. Half included a measure of renal function (e.g., creatinine clearance, urinary albumin). Two studies included serum albumin (Goldman et al. 2009, Newson et al. 2010), which is a non-specific measure of underlying disease. Two studies (Goldman et al. 2006b, Goldman et al. 2009) incorporated insulin-like growth factor 1 (IGF-1), which affects muscle growth and is an indirect marker of inflammation. Haring et al. (2011) added a marker of liver function (glutamyltransferase). Two studies (Stork et al. 2006, Newson et al. 2010) incorporated carotid plaque burden, which is a marker of atherosclerosis. Rosero-Bixby and Dow (2012) included performance assessments (i.e., grip strength, gait speed, peak expiratory flow). Another study included a measure of bone density (Newson et al. 2010).

Most of these studies found that the addition of biomarkers led to a large gain in predictive ability. One exception was Haring et al. (2011) who reported that the biomarkers provided only a minuscule gain in AUC (0.004), but their model included only the four “most informative” biomarkers (out of 10). In contrast, the Costa Rican study (Rosero-Bixby and Dow 2012) included 22 biomarkers and the Taiwan studies (Goldman et al. 2009, Goldman et al. 2006b) included 13–16 biomarkers; these three studies found that the set of biomarkers accounted for a 0.06–0.08 increase in AUC. One of the Dutch studies included only three markers (IL-6, CRP and the number of carotid plaques), but those markers contributed to a 0.15 improvement in AUC (Stork et al. 2006). The other Dutch study (Newson et al. 2010) did not estimate the increase in AUC associated with the seven biomarkers (i.e., diastolic

blood pressure, CRP, number of carotid plaques, and several other markers) compared with a model that included only the control variables.

As noted by Pencina et al. (2012), the gain in AUC depends heavily on the strength of the baseline model used for comparison. All of the studies mentioned above controlled for sex, age, smoking, and some self-reported measures of baseline health status. Haring et al. (2011) adjusted for the most extensive list of covariates including multiple measures of socioeconomic status, health behaviors, disease history, self-assessed health status, perceived quality of life, and waist circumference. They had the strongest baseline model (AUC=0.88) among the six studies considered here. The limited set of selected biomarkers combined with an extensive set of control variables may explain why the incremental value attributable to biomarkers was much smaller in the Haring et al. (2011) study than the others. Not surprisingly, the largest gain in AUC (0.15) was generated by the study with the weakest baseline model (AUC=0.62) (Stork et al. 2006).

**Studies Using Multi-System Biomarker Summary Scores**—The allostatic load framework (McEwen and Stellar 1993) helped promote the use of composite measures of biological function across multiple systems (although not all multisystem scores are predicated on the allostatic load framework). A number of prior studies in the U.S. and Taiwan have found that multi-system measures consistently predict all-cause mortality (Goldman et al. 2006a, Seeman et al. 2001, Seeman et al. 2004, Gruenewald et al. 2006, Karlamangla, Singer, and Seeman 2006, Borrell, Dallo, and Nguyen 2010, Crimmins, Kim, and Seeman 2009, Wang et al. 2006, Levine 2013)—see Table 1 for more details. Early formulations of such scores incorporated standard cardiovascular/metabolic markers as well as neuroendocrine markers. More recent versions typically include at least one inflammatory marker and often include some “other” marker(s) (e.g., serum albumin, creatinine clearance, homocysteine).

In most cases the magnitude of the association between the multi-system score and mortality was substantial even among the studies that controlled for self-reported indicators of baseline health status (Goldman et al. 2006a, Seeman et al. 2004, Seeman et al. 2001, Wang et al. 2006). For example, Seeman et al. (2001) showed that the odds of dying within seven years were more than six times as high for someone with high-risk levels for at least 7 of the 10 biomarkers compared with someone with no high-risk markers. Wang et al. (2006) reported that the mortality rate was four times as high for individuals with high multi-system scores (i.e., top quintile) relative to those with low scores (i.e., bottom two quintiles) adjusted for age, sex, and conventional risk factors, although the multi-system score yielded only a small improvement in predictive ability (C-statistic increased from 0.80 to 0.82). However, standard clinical biomarkers were included as control variables rather than being incorporated into the multisystem score. None of the other studies quantified the predictive ability of the multisystem score.

With one exception, prior studies investigating the link between a multi-system score and mortality relied on markers measured at only one survey wave. Karlamangla et al. (2006) were the only researchers that evaluated the contribution of changes in markers across multiple systems: they found that both starting levels and changes in biomarkers over a 2.5

year period were associated with subsequent mortality. However, this study did not control for any self-reported health indicators, nor did it evaluate whether changes in biomarkers provided significantly better discriminatory ability than a one-time measurement.

### What this study adds

In sum, existing prognostic indexes for predicting mortality among community-based populations of older adults have achieved C-statistic/AUC values ranging between 0.66 and 0.84 (Yourman et al. 2012). Our study extends prior studies by quantifying the incremental improvement in discrimination associated with measurement of change in biomarkers compared with one-time measurement.

## METHODS

### Data

The data come from a cohort study in Taiwan (SEBAS), augmented by the 2003 wave of its parent study, the Taiwan Longitudinal Study of Aging (TLSA). Taiwan provides a good context for study because the level of life expectancy ( $e_0 = 79.2$  in 2010) is similar to that of other industrialized countries such as the U.S. ( $e_0 = 78.9$  in 2010) (Human Mortality Database 2013), and the cause of death structure is also similar to that of high income Western countries. For example, 9 out of the top 10 causes of death in 2011 were the same in Taiwan (Department of Health, Executive Yuan, R.O.C. (Taiwan) 2012) and the U.S. (Hoyert and Xu 2012). Many of the factors (e.g., age, sex, education, marital status, self-reported measures of health) that predict mortality in Taiwan are consistent with those observed in Western countries (Zimmer, Martin, and Lin 2005).

The SEBAS cohort was based on a nationally representative sample of Taiwanese aged 54 and older in 2000, selected randomly using a multi-stage sampling design with oversampling of older persons (71+) and urban residents (Chang et al. 2012). In 2000, in-home interviews were completed with 1497 respondents, 1023 of whom also completed the physical examination. Exam participants did not differ significantly from nonparticipants in ways likely to introduce serious bias (Goldman et al. 2003). Six years later, a follow-up was conducted with those who completed the 2000 exam and survived to 2006: 757 completed the in-home interview and 639 participated in the physical examination. Details regarding response rates, sample attrition, and exam participation are provided elsewhere (Chang et al. 2012).

The physical examination followed a similar protocol in both waves. Several weeks after the household interview, participants collected a 12-hour overnight urine sample (7pm to 7am), fasted overnight, and visited a nearby hospital the following morning for a physical examination that included collection of a blood specimen and measurements of blood pressure, height, weight, waist and hip circumference. Compliance was high: in 2000, 96 percent fasted overnight and provided a urine specimen deemed suitable for analysis; the comparable figure was 88 percent in 2006.

Union Clinical Laboratories (UCL) in Taipei analyzed the blood and urine specimens. In addition to the routine standardization and calibration tests, triplicate sets of specimens were

contributed by individuals outside the target sample ( $n=9$  in 2000;  $n=10$  in 2006): two sets were sent to UCL and the third set was sent to Quest Diagnostics in the US (San Juan Capistrano, CA). The results indicated high intra-lab reliability for duplicates sent to UCL (2000:  $> 0.71$ ; 2006:  $> 0.85$ ). With a few exceptions (glycosylated hemoglobin, HbA1c; Cortisol; Urinary Creatinine), inter-lab correlations between results from UCL and Quest Diagnostics exceeded 0.90 in both waves.

Survival status as of January 1, 2012, was ascertained by linkage to the death certificate file maintained by the Taiwan Department of Health and to the household registration database maintained by the Ministry of the Interior. The analysis sample was based on the longitudinal cohort that completed the exam in both 2000 and 2006 ( $n=639$ , 104 of whom died by December 31, 2011). The mean length of mortality follow-up was 5.1 years (max, 5.3 years). Some ( $n=89$ ) of the respondents had missing data for at least one covariate, typically a biomarker. To maximize the use of the data, we followed standard practices of multiple imputation (Schafer 1999, Rubin 1996) to handle missing data. We created five imputed datasets using regression techniques to fill in missing values using as predictors all of the variables in this analysis plus several auxiliary variables that were correlated with non-response (e.g., interviewed by proxy, cognitive function, limitations in activities of daily living, walking speed). Then, we estimated the model for each imputed dataset and combined the five sets of estimates using Rubin's rules (Royston, Carlin, and White 2009). All measures of fit and predictive ability were calculated for each dataset and then averaged following the same rules. In order to test the robustness of the results to treatment of missing data, we also re-estimated the final models for the sample with complete data ( $n=550$ ).

## Measures

**Biomarkers**—We included 19 biomarkers that have been shown by prior studies to be associated with all-cause mortality. They comprised four groups of markers: 1) eight standard cardiovascular/metabolic risk factors—systolic blood pressure, diastolic blood pressure, total cholesterol, high-density lipoprotein (HDL) cholesterol, triglycerides, glycosylated hemoglobin, body mass index, and waist circumference; 2) four inflammatory markers—IL-6, CRP, soluble intercellular adhesion molecule 1 (sICAM-1), soluble E-selectin; 3) four neuroendocrine markers—dehydroepiandrosterone sulfate (DHEAS), cortisol, epinephrine, norepinephrine; and 4) three other markers that do not represent a common biological subsystem—creatinine clearance, albumin, homocysteine. See Table 2 for more detailed information about each marker.

We calculated an overall summary score based on these 19 markers and subscores for each of the four groups of markers. Although the last group comprises three unrelated markers, we compute the subscore for completeness (i.e., sum of the four subscores equals the overall score). Each score was calculated by counting the number of markers for which the respondent exhibited a high-risk level. High-risk was defined by established cutoffs for the standard cardiovascular/metabolic factors and CRP (see Appendix Table A-1 for a list of markers and cutoffs). For all other markers—which have no generally accepted clinical cutoffs—we defined high-risk based on the weighted distribution of the 2000 sample: bottom quartile for DHEAS, creatinine clearance, and albumin; top quartile for other

markers. Although other types of summary scores have been proposed (Juster, McEwen, and Lupien 2010, Seeman et al. 2010), we use a formulation similar to that employed by many other studies to represent a composite of biomarker values.

**Self-Reported Health Indicators**—Based on self-reports from 2006, we included six health indicators: 1) global self-assessed health (“Regarding your current state of health, do you feel it is excellent, good, average, not so good, or poor?”); 2) an index of mobility limitations; 3) whether the respondent was ever diagnosed with diabetes; 4) history of cancer; 5) number of hospitalizations in the past 12 months; and 6) smoking status (never, former, current). Mobility was based on self-reported difficulty performing eight physical tasks without assistance, each of which was coded on a four-point scale (0=no difficulty, 1=some difficulty, 2=great difficulty; 3=unable): stand for 15 minutes, squat, raise both hands overhead, grasp or turn objects with his or her fingers, lift or carry an object weighing 11–12kg, walk 200–300m, run 20–30m, and climb two or three flights of stairs. Based on the recommendations of Long & Pavalko (2004), we constructed the scale by summing the eight items (potential range 0–24), adding a constant (0.5), and taking the logarithm of the result, which allows for relative rather than absolute effects.

**Social and Demographic Characteristics**—We controlled for key demographic and social characteristics that are known to be important predictors of mortality. Demographic variables comprised age, sex, urban residence, and ethnicity (Mainlander vs. Taiwanese). We also included educational attainment because of the well-established link between socioeconomic status and longevity both within and across countries. Finally, in light of the large literature documenting an association between social relationships—including marital status—and mortality risk (see meta-analysis by Holt-Lunstad, Smith, and Layton 2010), we adjusted for social integration and perceived social support.

An index of social integration was constructed following the strategy used by Cornwell & Waite (2009) to develop a social disconnectedness scale. Using 10 indicators from the 2003 TLSA (e.g., network size, network range, marital status, participation in social organizations; see Appendix Table A-2 for details), we standardized each item and calculated the mean across valid items if at least 8 items were valid ( $\alpha=0.72$ ).

An index of perceived social support was based on four questions (coded 0–4) from the 2003 interview: family/friends willing to listen; family/friends make you feel cared for; satisfaction with emotional support received from family; can count on family to take care of you when you are ill. We calculated the mean across valid items if at least 3 items were valid ( $\alpha=0.84$ ).

### Analytical Strategy

Our modeling strategy begins by assessing the incremental contribution of self-reported health indicators versus biomarkers for predicting mortality relative to a baseline model that includes only the social and demographic control variables. Thus, we test our first hypothesis: that biomarkers predict mortality better than self-reported health indicators. In the second part of the analysis (pertaining to our second hypothesis), we evaluate the added value of biomarkers above and beyond self-reports. That is, do biomarkers yield prognostic



value net of self-reported measures? In the next stage (related to our third hypothesis), we determine whether changes in biomarkers yield more predictive power than one-time measurements. Finally, with regard to our fourth hypothesis, we explore the prognostic value of biomarkers representing different systems.

Descriptive statistics are weighted to account for oversampling and for differential response rates by age, sex, and other covariates. Using unweighted data, we estimate age-specific mortality using a Gompertz hazards model with time measured in terms of age. Initial tests (not shown) show no significant evidence that the age slope of mortality varies by sex. Nonetheless, there is evidence of non-proportional hazards (i.e., the effect varies by age) for a few covariates: perceived social support, current smoker, and the change (2000-06) in DHEAS. Therefore, we include interactions between these variables and age.

In order to compare effect size across predictors, we standardize—to have a mean of zero and standard deviation (SD) of one—each of the continuous measures prior to fitting the hazards models. Thus, the coefficient represents the effect per SD. Many of the biomarkers have a skewed distribution, which influences the mean and the SD. Therefore, for models that employ the markers in continuous form, we transform markers with a skewed distribution taking logs or using a power transformation (see Table A-3) in order to better approximate normality. These transformations substantially improve the fit of the model.

To compare various models, we use measures of relative goodness of fit and discrimination. We use Akaike's information criteria (AIC) to measure *relative goodness of fit*. The AIC compares the fit of one model relative to another (whether nested or not) based on the maximized log-likelihood with a penalty for the number of parameters estimated (StataCorp 2011, Akaike 1974).

To measure *discrimination*, we use the AUC (described earlier). We calculate the AUC by comparing the model-based predicted probability of dying by the end of follow-up (see Appendix for details) with the observed binary outcome (death vs. survival). One drawback of the AUC measure is that it can be misleading if the ROC curves cross (i.e., the model with higher AUC is not superior for all values of the classification threshold; that is, it does not have better sensitivity at all levels of specificity) (Hand 2010). In such cases, one must be cautious about concluding that the model with higher AUC is “better” because this assessment depends on the tradeoff between specificity and sensitivity. In light of this potential problem, we inspect the ROC curves for all the comparisons presented here. Among nested models in which the change in AUC is significant, we find no appreciable crossover of the ROC curves; thus, we can conclude that the model with the higher AUC is better. However, in cases where the change in AUC is not significant and for some comparisons of non-nested models, we find a slight crossover of some of the ROC curves. In these cases, we have qualified our discussion of the results.

## RESULTS

The mean within-individual change (2006 – 2000) in overall biological risk (i.e., in the biomarker summary score) is 0.6 (Table 3), but there is considerable variation across

individuals. Nearly half the respondents exhibit an increase, while one-third display a decrease. Sizeable percentages demonstrate an increase or decrease of at least two points in overall biological risk.

Table 4 shows comparisons among models in terms of relative model fit (AIC) and discrimination (AUC). We begin with a baseline model that includes only social and demographic characteristics. Panel 1 shows a model that assesses the contribution of self-reported health indicators, while Panel 2 shows a series of models that evaluates the gain attributable to biomarkers. Panel 3 presents a set of models that tests the incremental value of biomarkers net of self-reported health indicators. The coefficients from selected models are presented in Table 5; results from the remaining models are given in the Appendix (Tables A-4, A-5, and A-6). The key comparisons of interest are summarized in Figure 1 in terms of AUC improvements over the baseline model.

### Do biomarkers predict mortality better than self-reported health indicators?

In the first stage of the analysis, we evaluate the incremental prognostic value of self-reported health indicators versus biomarkers by comparing the respective models with the baseline model, which yields good discrimination (AUC = 0.75, Table 4, Model 0).<sup>i</sup> Net of the social and demographic characteristics, self-reported health indicators greatly improve the discriminatory ability of the model (Figure 1). Among the self-reported health indicators, mobility limitations and number of hospitalizations in the past year significantly predict higher mortality (Table 5).<sup>ii</sup> In auxiliary analyses (not shown), we test the association between changes (2000-06) in self-reported health indicators and mortality. Net of the 2006 levels, changes in self-reports do very little to improve model fit or discrimination.

The models presented in Panel 2 of Table 4 test various specifications for the biomarkers. Model 2 includes the biomarker summary score from 2006. Model 3 adds the change (2006 – 2000) in biomarker risk. Models 4a and 5a substitute individual biomarkers for the biomarker summary scores. As shown graphically in Figure 1, levels of biomarkers in 2006 (as measured by the summary score or individual markers) account for a substantial improvement in the AUC, but changes in biomarkers also contribute incremental gains. Individual biomarkers (Models 4a and 5a) generate larger gains in discriminatory ability than the biomarker summary scores (Model 2 and 3), although there appears to be slight crossover of the ROC curves in some cases. Compared with the gain in the AUC attributable to the self-reported health indicators (0.07, Model 1), individual biomarkers (including both the levels in 2006 and the changes between 2000 and 2006) yield a higher gain in AUC (0.10, Model 5a), whereas the biomarker summary score (including level in 2006 and change 2000-06) demonstrates a smaller improvement (0.05, Model 3). Although there is slight crossover among these three ROC curves in some regions, the model with higher AUC is usually better (i.e., for most tradeoffs between sensitivity and specificity), and it is never appreciably worse. In sum, our findings suggest that the two types of data – health

<sup>i</sup>Although there is some disagreement about benchmarks for the C-statistic/AUC, Yourman et al. (2012) use the following categorization: values below 0.60 = poor, 0.60–0.69 = moderate, 0.70–0.79 = good, 0.80–0.89 = very good, and 0.90 or higher = excellent.

<sup>ii</sup>The effect of current smoking varies by age: at age 60, current smokers appear to have *lower* mortality than those who never smoked (HR =  $e^{-1.97} = 0.14$ ,  $p < 0.05$ ), but by age 90, smoking is associated with higher mortality (HR =  $e^{-1.97+30 \times 0.11} = 3.78$ ,  $p \sim 0.02$ ).

information reported by the respondent and biological measures based on physical examination and analyses of blood and urine specimens – are both valuable for estimating whether or not an individual is likely to die in the next five years. In most cases, individual biomarkers perform better than the biomarker summary score.

### **Do biomarkers yield prognostic value net of self-reported measures?**

Given the comparative ease of collecting self-reports, a more important question is whether biomarkers retain discriminatory ability after we take into account the self-reported health indicators. In this second stage of the analysis, we evaluate the added value net of all the self-reported measures (see Panel 3 of Table 4); thus, all comparisons are relative to Model 1.

In the presence of controls for self-reported indicators, the biomarker summary score (Model 7) yields a 0.03 gain in the AUC.<sup>iii</sup> Individual biomarkers (Model 9a) generate a 0.06 increase in the AUC, but at the cost of many additional parameters. Indeed, among the 19 markers included, very few exhibit a significant net effect for either the 2006 level or the change, and the magnitude of the effects are generally small (see Appendix Table A-5, Model 9a).

Thus, we conclude that adding the collection of biomarkers to a population-based health survey that relies on self-report will enhance predictions of 5-year survival. However, the measured effect of any particular biomarker is likely to be small, and may not be statistically discernible without a huge sample.

### **Do changes in biomarkers provide better discrimination than one-time measures?**

The analyses to this point suggest that changes in biomarkers provide more prognostic value than one-time measurement of those same biomarkers. In this third stage of the analysis, we formally test whether this is in fact the case. All of these models include self-reported health indicators.

Despite a significant coefficient, change in the biomarker summary score yields negligible improvement in the AUC (0.004, Model 7 versus Model 6). In contrast, changes in individual biomarkers (Model 9a) generate a larger improvement in AUC (0.02) compared with individual biomarkers measured only in 2006 (Model 8a). Although a joint Wald test of all parameters related to changes in individual biomarkers is significant, the gain in AUC is not significant and a comparison of the two ROC curves shows some overlap. Thus, the incremental value owing to changes in biomarkers is equivocal. Figure 1 shows graphically the larger contribution attributable to changes in individual markers compared with changes in the summary score.

---

<sup>iii</sup>The 2006 level ( $HR = e^{0.70}$  = per SD) was significantly associated with a higher mortality rate, while the change (2006 minus 2000) in the biomarker summary score had a negative effect ( $HR = e^{-0.29}$  = per SD) (Table 5). This result implied that respondents with high levels of biomarker risk throughout the six-year period experienced higher mortality than those who had lower risk in 2000 but reached high levels of risk more recently.

Overall, we conclude that collecting a second round of biological measures *may* improve mortality prediction compared with one-time measurement. Nonetheless, the incremental gain is likely to be modest.

### Are some biomarkers more valuable than others?

Our set of biomarkers covers multiple physiological systems, so the final stage of the analysis considers whether some biomarkers provide more prognostic value than others after controlling for self-reports. First, we examine the contribution of subscores that represent different biological systems (Appendix Table A-6). The inflammation subscore offers the most prognostic value (Model 7b); higher inflammation in 2006 is significantly associated with higher mortality, although the gain in the AUC is not significant. Subscores based on standard cardiovascular/metabolic markers (Model 7a), neuroendocrine markers (Model 7c), and other unrelated markers (Model 7d) make negligible contributions to the AUC. None of the subscores provides as much prognostic value as the overall summary score (Model 7, Table 4). When all four subscores are included in the same model (Model 7e, Table A-6), we find that the 2006 levels for the subscores comprising cardiovascular/metabolic, inflammation, and other unrelated markers are positively associated with mortality, while the change (2006 – 2000) is not significant for any of the subscores. Neither the 2006 level nor change in the neuroendocrine subscore has a significant effect.

Next, we consider the value of individual biomarkers (including the level in 2006 and the change, 2006 minus 2000). As noted earlier, very few of the 19 markers demonstrate a significant effect net of self-reports (Model 9a, Table A-5). Given that many of the individual markers are correlated with each other and that the net effect of any one marker is likely to be small, it is reasonable to ask whether a smaller set of markers may yield similar prognostic value. Thus, we fit a model that includes only the best 10 biomarkers based on the improvement in model fit (AIC) attributable to each marker. The “best 10” biomarkers comprises all four of the inflammatory markers, all three of the other unrelated markers, two metabolic markers (HDL & BMI), and one neuroendocrine marker (DHEAS). Nonetheless, the only individual parameters that reach statistical significance net of self-reports are 2006 levels of IL-6, sICAM-1, and homocysteine as well as changes in DHEAS (Model 9b, Table 5).<sup>iv</sup> Overall, the best 10 biomarkers achieve an AUC (0.86, Model 9b) that is higher than the biomarker summary score (0.85, Model 7), but not as good as the model that includes all 19 biomarkers (0.88, Model 9a). We can conclude that the individual biomarkers as a group perform better than the summary score (i.e., Model 9a dominates Model 7 at all levels of specificity), but we cannot say the same for the reduced set (10 best markers) because of slight crossover of the ROC curves (Model 7 vs. 9b). Nonetheless, the 10 best markers outperform the summary score at most levels of sensitivity and specificity.

In short, inflammatory markers provide the strongest prediction of 5-year survival, whereas neuroendocrine markers offer the least value. Nevertheless, it is preferable to measure a

<sup>iv</sup>Levels of IL-6 and homocysteine in 2006 have the biggest effect on mortality ( $HR = e^{0.55} = 1.73$  and  $e^{0.57}$  per SD, respectively). The effect of DHEAS declines with age: at age 60, a decline in DHEAS is associated with higher mortality, but by age 80, the effect is negligible.

combination of markers drawn from multiple physiological systems than to rely on a single cluster of biologically-related markers.

### Robustness of the results

To determine the robustness of the results to the treatment of missing data, we re-estimate the models for the sample with complete data ( $n=550$ ). Most of the coefficients are very similar to those presented here. One exception is history of diabetes, which appears to have a stronger effect when we use listwise deletion than under multiple imputation. This result may be attributable to a relationship between completeness of information, diabetes and survival that is better handled in the multiple imputation model.

We also investigate the sensitivity of the results to using a Cox (non-parametric) model instead of the Gompertz (parametric) model. The coefficients and AUC values from the corresponding Cox models—which include the same set of interactions with age to allow for non-proportional hazards—are very similar to the results shown here.

## DISCUSSION

The past 15 years have witnessed the marriage of biomarker collection with large-scale, population-based social surveys. Dozens of demographic and health surveys that include standard measures of self-reported health, social status, social relationships, health behaviors, exposure to stress, and other aspects of the social environment have also incorporated biomarkers. Currently, there are more than 30 biosocial surveys in use (Table 5 in Love et al. 2010, Table 9-1 in National Research Council and Institute of Medicine 2013, The Biomarker Network 2013).

Wachter (2001, p. 330) hinted at the unspoken promise offered by biomarkers when he noted, “We know so much, but only so much, about ourselves. Self-reports as the mainstay of social surveys have taken us a long way, but now survey research is stretching itself to move beyond them.” Now that the field of biodemography has accumulated decades of research, it is fair to ask whether biomarkers have lived up to expectation. Several biosocial surveys are embarking on a second round of biomarker collection, so it is especially timely to consider the value of longitudinal biomarker collection.

In this paper, we have attempted to address these issues with respect to mortality prediction. In terms of this particular benchmark, our results suggest that when biomarkers and self-reported indicators of health status are compared side by side, both make a substantial contribution. Yet, when comparing the value of biomarkers with self-reports, we should not ignore differences in cost, both monetary and non-monetary. Self-reports are much easier and cheaper to collect and impose less burden on the respondent. Thus, a more crucial question is whether biomarkers offer incremental value beyond the contribution of self-reported measures.

Based on our results, the answer is a qualified “yes.” A summary measure of biomarkers yields an improvement in discrimination compared with self-reports alone, but individual biomarkers perform better than the summary score. We find that incorporating change in

biomarkers over a six-year period yields a small improvement in mortality prediction compared with a one-time snapshot of biomarker levels. One might legitimately argue that the value of repeated biomarker collection is greater than this comparison suggests. If the objective is prediction of longer-term mortality (e.g., for a period distant from the biomarker collection), then an updated set of biomarker values substantially enhances prediction. Specifically, biomarkers from 2006 predict mortality for the follow-up period 2006–2011 considerably better (AUC = 0.856, Model 8a) than the baseline biomarkers obtained in 2000 (AUC = 0.839, not shown). Thus, longitudinal collection of biomarkers may well be warranted in such instances.

In light of the expense and respondent burden involved in collecting a wide range of biomarkers, it is important to identify the biomarkers that are most valuable. Since markers representing the same system may be correlated with one another, we might ask whether a reduced set of markers performs almost as well. Our results indicate that, with the possible exception of DHEAS, neuroendocrine markers (i.e., stress hormones) make little contribution to predicting mortality, whereas inflammatory markers yield the most prognostic value. Among the 19 individual biomarkers considered here, three appear to be especially good predictors of mortality: two inflammatory markers (IL-6, sICAM-1) and homocysteine, which has been linked with cardiovascular disease and may promote atherosclerosis (American Heart Association 2012). Compared with a model using the biological risk summary score (based on all 19 biomarkers), a model that includes the 10 best individual biomarkers performs only slightly better (as measured by the AUC), whereas the model with all 19 individual markers yields more prognostic value. Although the summary score formulation used here is similar to that used in many other studies, our findings underscore the need for a better summary measure, one that captures more of the information embodied in continuous measures of myriad individual biomarkers.

There are several limitations to this study. First, our statistical power is limited by the relatively small sample and short-term mortality followup (104 deaths over five years). In many cases, the improvement in discrimination is not significant even though the magnitude of the gain in the AUC is well above the level ( $p < 0.01$ ) considered “meaningful” and the coefficient(s) associated with the biomarkers are jointly significant (Pencina et al. 2008). Second, although exam participation was high for studies of this type, the subset of respondents who survived to the second wave (six years later) and completed both the interview and examination was a select group. The effects of selection and attrition may explain the counter-intuitive effect of smoking among younger individuals. Among those who completed the 2000 exam, smokers were more likely to die by 2006 than non-smokers. Among survivors, those who were still smoking in 2006 had slightly lower mortality than those who never smoked, but the death rate was much higher for those who had quit smoking. These analyses suggest that respondents who suffered ill health as a result of smoking either quit or died before they could be interviewed. Third, measurement of biomarkers is not always straightforward, especially for markers that vary widely from day to day or even hour to hour. Some of the observed changes in levels may reflect measurement error and random variation that is not substantively meaningful. For example, the negligible effect of neuroendocrine markers on mortality may stem from the way in which they are measured: overnight urine collection captures only resting levels of markers

that fluctuate widely throughout the day and in response to stressful conditions. Some researchers have begun experimenting with using hair samples to obtain retrospective measures of cortisol levels and sympathetic nervous system activity over the past three months. This new approach seems promising and may prove to be more powerful and reliable than current methods based on urine and saliva samples. Fourth, showing that an indicator has prognostic value does not imply a causal connection. For example, some markers (e.g., IL-6, homocysteine) may reflect undiagnosed health conditions. To reduce risk of mortality, we may need to look beyond these “symptoms” to identify the root causes that could be amenable to intervention.

The cost, complex logistics, and additional burden placed on the respondent posed by adding biomarkers to large-scale, community-based surveys mean that we must evaluate carefully the “bio” in medical biodemography. This study has provided one small piece of that evaluation. Our scope here has—necessarily—been limited; we have focused only on mortality, and have not accounted for—or compared the results with—other “objective” measures such as functional assessments. Such measures might be equally powerful in predicting downstream mortality and health outcomes, potentially less invasive, and possibly less expensive. We cannot pass judgment on the value of biomarkers based on a single criterion: the ability to predict mortality. There are other health outcomes, such as depression or cognitive function, for which biomarkers may hold prognostic value. Another important use for biomarkers is to explore the biological mediators that account for the links between social factors and health outcomes. Of course, one prerequisite for demonstrating that a biomarker acts as a mediator is establishing that the biomarker does in fact predict health outcomes. Friedman and Kern (2014) caution against using limited-time measurements of biomarkers as an endpoint because such markers do not necessarily predict future health and survival, especially since they naturally fluctuate as part of homeostasis. Thus, biomarkers must pass that initial test, and we would argue that the ability to predict all-cause mortality is a crucial part of that test. We have provided some evidence to place in the balance, but the underlying question remains open: Is the blood worth the toil, tears, and sweat?

## Acknowledgments

This work was supported by the National Institute on Aging (grant numbers R01AG16790, R01AG16661) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant number R24HD047879).

Funding for the TLSA came from the Taiwan Department of Health, the Taiwan National Health Research Institute [grant number DD01-86IX-GR601S] and the Taiwan Provincial Government. SEBAS was funded by the Demography and Epidemiology Unit of the Behavioral and Social Research Program of the National Institute on Aging [grant numbers R01 AG16790, R01 AG16661]. The Bureau of Health Promotion (BHP, Department of Health, Taiwan) provided additional financial support for SEBAS 2000. We acknowledge the hard work and dedication of the staff at the Center for Population and Health Survey Research (BHP), who were instrumental in the design and implementation of the SEBAS and supervised all aspects of the fieldwork and data processing.

We are also grateful to Chioun Lee for her helpful comments and suggestions on an early version of this manuscript.

## REFERENCES

- Akaike, Hirotugu. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 1974; 19:716–723.
- American Heart Association. Homocysteine, Folic Acid and Cardiovascular Disease. 2012 <<[www.heart.org/HEARTORG/GettingHealthy/NutritionCenter/Homocysteine-Folic-Acid-and-Cardiovascular-Disease\\_UCM\\_305997\\_Article.jsp](http://www.heart.org/HEARTORG/GettingHealthy/NutritionCenter/Homocysteine-Folic-Acid-and-Cardiovascular-Disease_UCM_305997_Article.jsp)>>.
- Borrell, Luisa N.; Dallo, Florence J.; Nguyen, Norma. Racial/ethnic disparities in all-cause mortality in U.S. adults: the effect of allostatic load. *Public Health Reports (Washington, D.C.: 1974)*. 2010; 125(6):810–816.
- Chang, Ming-Cheng; Lin, Hui-Sheng; Chuang, Yi-Li; Goldman, N.; Peterson, Christine E., et al. Social Environment and Biomarkers of Aging Study (SEBAS) in Taiwan, 2000 and 2006: main documentation for SEBAS longitudinal public use data (released 2012). Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]; 2012.
- Chobanian, Aram V.; Bakris, George L.; Black, Henry R.; Cushman, William C.; Green, Lee A., et al. The seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA : The Journal of the American Medical Association*. 2003; 289(19):2560–2572. [PubMed: 12748199]
- Cockcroft, Donald W.; Gault, MH. Prediction of creatinine clearance from serum creatinine. *Nephron*. 1976; 16(1):31–41. [PubMed: 1244564]
- Cornwell, Erin Y.; Waite, Linda J. Measuring social isolation among older adults using multiple indicators from the NSHAP study. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*. 2009; 64(Suppl 1):38–46.
- Crimmins, Eileen M.; Vasunilashorn, Sarinnapha. Links between biomarkers and mortality. In: Rogers, Richard G.; Crimmins, Eileen M., editors. *International Handbook of Adult Mortality*. New York: Springer; 2011. p. 381–398.
- Crimmins, Eileen; Kim, Jung K.; Vasunilashorn, Sarinnapha. Biodemography: new approaches to understanding trends and differences in population health and mortality. *Demography*. 2010; 47(Suppl):S41–S64. [PubMed: 21302421]
- Crimmins, Eileen M.; Kim, Jung K.; Seeman, Teresa E. Poverty and biological risk: the earlier “aging” of the poor. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*. 2009; 64(2):286–292.
- Department of Health (Taiwan). Definition of Adult Obesity [English translation]. Taipei: Department of Health, Taiwan; 2002. <<<http://food.doh.gov.tw/foodnew/Files/Health/AdultDefinition.jpg>>>. [Accessed December 17 2008]
- Department of Health, Executive Yuan, R.O.C. (Taiwan). Deaths in Taiwan, 2011. Department of Health, Taiwan; 2012. <<<http://www.doh.gov.tw/EN2006/DisplayFile.aspx?url=http%3a%2f%2fwww.doh.gov.tw%2fufile%2fdoc%2fDeaths+in+Taiwan%2c+2011-20121017.docx&name=Deaths+in+Taiwan%2c+2011-20121017.docx>>>. [Accessed February 27 2013]
- Fried, Linda P.; Kronmal, Richard A.; Newman, Anne B.; Bild, Diane E.; Mittelmark, Maurice B., et al. Risk factors for 5-year mortality in older adults: the Cardiovascular Health Study. *JAMA : The Journal of the American Medical Association*. 1998; 279(8):585–592. [PubMed: 9486752]
- Friedman, Howard S.; Kern, Margaret L. Personality, well-being and health. *Annual Review of Psychology*. 2014; 65:719–742.
- Glei, Dana A.; Goldman, Noreen; Weinstein, Maxine. *Demographic Studies on Survival and Health*. Springer-Max Planck Series. Heidelberg: Springer Dordrecht; Health as Reflected by Biomarkers in Older Populations. Forthcoming
- Goldman, Noreen; Glei, Dana A.; Lin, Yu-Hsuan; Weinstein, Maxine. Improving mortality prediction using biosocial surveys. *American Journal of Epidemiology*. 2009; 169(6):769–779. [PubMed: 19147742]
- Goldman, Noreen; Lin, I-Fen; Weinstein, Maxine; Lin, Yu-Hsuan. Evaluating the quality of self-reports of hypertension and diabetes. *Journal of Clinical Epidemiology*. 2003; 56(2):148–154. [PubMed: 12654409]



- Goldman, Noreen; Turra, Cassio M.; Glei, Dana A.; Lin, Yu-Hsuan; Weinstein, Maxine. Physiological dysregulation and changes in health in an older population. *Experimental Gerontology*. 2006a; 41(9):862–870. [PubMed: 16908111]
- Goldman, Noreen; Turra, Cassio M.; Glei, Dana A.; Seplaki, Christopher L.; Lin, Yu-Hsuan, et al. Predicting mortality from clinical and nonclinical biomarkers. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*. 2006b; 61(10):1070–1074.
- Gruenewald, Tara L.; Seeman, Teresa E.; Ryff, Carol D.; Karlamangla, Arun S.; Singer, Burton H. Combinations of biomarkers predictive of later life mortality. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(38):14158–14163. [PubMed: 16983099]
- Hand, David J. Evaluating diagnostic tests: The area under the ROC curve and the balance of errors. *Statistics in Medicine*. 2010; 29(14):1502–1510. [PubMed: 20087877]
- Haring, Robin; Feng, You-Shan; Moock, Jorn; Volzke, Henry; Dorr, Marcus, et al. Self-perceived quality of life predicts mortality risk better than a multi-biomarker panel, but the combination of both does best. *BMC Medical Research Methodology*. 2011; 11:103. [PubMed: 21749697]
- Holt-Lunstad, Julianne; Smith, Timothy B.; Layton, JB. Social relationships and mortality risk: a meta-analytic review. *PLoS Medicine*. 2010; 7(7):e1000316. [PubMed: 20668659]
- Hoyert, Donna L.; Xu, Jiaquan. Deaths: preliminary data for 2011. *National vital statistics reports*. Vol. 61. Hyattsville, MD: National Center for Health Statistics; 2012.
- Human Mortality Database. University of California, Berkeley (USA);Max Planck Institute for Demographic Research (Germany);". 2013 <<[www.mortality.org](http://www.mortality.org)>>.
- Juster, Robert-Paul; McEwen, Bruce S.; Lupien, Sonia J. Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience and Biobehavioral Reviews*. 2010; 35(1):2–16. [PubMed: 19822172]
- Karlamangla, Arun S.; Singer, Burton H.; Seeman, Teresa E. Reduction in allostatic load in older adults is associated with lower all-cause mortality risk: MacArthur studies of successful aging. *Psychosomatic Medicine*. 2006; 68(3):500–507. [PubMed: 16738085]
- Kistorp, Caroline; Raymond, Ilan; Pedersen, Frants; Gustafsson, Finn; Faber, Jens, et al. N-terminal pro-brain natriuretic peptide, C-reactive protein, and urinary albumin levels as predictors of mortality and cardiovascular events in older adults. *JAMA : The Journal of the American Medical Association*. 2005; 293(13):1609–1616. [PubMed: 15811980]
- Lan, Tzuo-Yun; Chiu, Heng-Chia; Chang, Hsing-Yi; Chang, Wen-Chiung; Chen, Hui-Yang, et al. Clinical and laboratory predictors of all-cause mortality in older population. *Archives of Gerontology and Geriatrics*. 2007; 45(3):327–334. [PubMed: 17383026]
- Lee, Kenny WJ.; Hill, John S.; Walley, Keith R.; Frohlich, Jiri J. Relative value of multiple plasma biomarkers as risk factors for coronary artery disease and death in an angiography cohort. *CMAJ : Canadian Medical Association Journal = Journal De L'Association Medicale Canadienne*. 2006a; 174(4):461–466.
- Lee, Sei J.; Lindquist, Karla; Segal, Mark R.; Covinsky, Kenneth E. Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA : The Journal of the American Medical Association*. 2006b; 295(7):801–808. [PubMed: 16478903]
- Levey AS, Coresh J, Balk E, Kausz AT, Levin A, et al. National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Annals of Internal Medicine*. 2003; 139(2):137–147. [PubMed: 12859163]
- Levine, Morgan E. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*. 2013; 68(6):667–674.
- Libby, Peter; Ridker, Paul M. Novel inflammatory markers of coronary risk: theory versus practice. *Circulation*. 1999; 100(11):1148–1150. [PubMed: 10484532]
- Long JS, Pavalko Eliza. Comparing alternative measures of functional limitation. *Medical Care*. 2004; 42(1):19–27. [PubMed: 14713736]
- Love, Gayle D.; Seeman, Teresa E.; Weinstein, Maxine; Ryff, Carol D. Bioindicators in the MIDUS National Study: protocol, measures, sample, and comparative context. *Journal of Aging and Health*. 2010; 22(8):1059–1080. [PubMed: 20876364]

- McEwen, Bruce S.; Stellar, Eliot. Stress and the individual. Mechanisms leading to disease. *Archives of Internal Medicine*. 1993; 153(18):2093–2101. [PubMed: 8379800]
- Melzer, David, Tzuo-Yun Lan, Brian DM, Tom, Dorly J.H. Deeg, Jack M Guralnik. Variation in thresholds for reporting mobility disability between national population subgroups and studies. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*. 2004; 59(12): 1295–1303.
- National Cholesterol Education Program, National Heart, Lung and Blood Institute. Third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III): executive summary. Bethesda, MD: National Institutes of Health; 2001.
- National Heart, Lung, and Blood Institute (NHLBI). Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: executive summary. Bethesda, MD: National Institutes of Health; 1998.
- National Research Council and Institute of Medicine. U.S. Health in International Perspective: Shorter Lives, Poorer Health. Panel on Understanding Cross-National Health Differences among High-Income Countries. In: Woolf, SH.; Aron, L., editors. Committee on Population, Division of Behavioral and Social Sciences and Education, and Board on Population Health and Public Health Practice, Institute of Medicine. Washington, DC: The National Academies Press; 2013.
- Newman, Anne B.; Sachs, Michael C.; Arnold, Alice M.; Fried, Linda P.; Kronmal, Richard, et al. Total and cause-specific mortality in the cardiovascular health study. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*. 2009; 64(12):1251–1261.
- Newson, Rachel S.; Witteman, Jacqueline C.; Franco, Oscar H.; Stricker, Bruno H.; Breteler, Monique M., et al. Predicting survival and morbidity-free survival to very old age. *Age (Dordrecht, Netherlands)*. 2010; 32(4):521–534.
- Pearson, Thomas A.; Mensah, George A.; Alexander, RW.; Anderson, Jeffrey L.; Cannon, Richard O., 3rd, et al. Markers of inflammation and cardiovascular disease: application to clinical and public health practice: A statement for healthcare professionals from the Centers for Disease Control and Prevention and the American Heart Association. *Circulation*. 2003; 107(3):499–511. [PubMed: 12551878]
- Pencina, Michael J.; Ralph B, D'Agostino; Ralph B, D'Agostino, Jr; Ramachandran S, Vasan. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in Medicine*. 2008; 27(2):157–72. discussion 207–12. [PubMed: 17569110]
- Pencina, Michael J.; Ralph B, D'Agostino. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statistics in Medicine*. 2004; 23(13):2109–2123. [PubMed: 15211606]
- Pencina, Michael J.; Ralph B, D'Agostino; Pencina, Karol M.; Janssens, ACW.; Greenland, Philip. Interpreting incremental value of markers added to risk prediction models. *American Journal of Epidemiology*. 2012; 176(6):473–481. [PubMed: 22875755]
- Rodbard, Helena W.; Blonde, Lawrence; Braithwaite, Susan S.; Brett, Elise M.; Cobin, Rhoda H., et al. American Association of Clinical Endocrinologists medical guidelines for clinical practice for the management of diabetes mellitus. *Endocrine Practice : Official Journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists*. 2007; 13(Suppl 1):1–68.
- Rosero-Bixby, Luis; Dow, William H. Predicting mortality with biomarkers: a population-based prospective cohort study for elderly Costa Ricans. *Population Health Metrics*. 2012; 10(1):11. [PubMed: 22694922]
- Royston, Patrick; Carlin, John B.; White, Ian R. Multiple imputation of missing values: new features for *mim*. *Stata Journal*. 2009; 9(2):252–264.
- Rubin, Donald B. Multiple imputation after 18+ years (with discussion). *Journal of the American Statistical Association*. 1996; 91:473–489.
- Schafer, Joseph L. Multiple imputation: a primer. *Statistical Methods in Medical Research*. 1999; 8(1): 3–15. [PubMed: 10347857]

- Schonberg, Mara A.; Davis, Roger B.; McCarthy, Ellen P.; Marcantonio, Edward R. Index to predict 5-year mortality of community-dwelling adults aged 65 and older using data from the National Health Interview Survey. *Journal of General Internal Medicine*. 2009; 24(10):1115–1122. [PubMed: 19649678]
- Seeman, Teresa E.; Crimmins, Eileen; Huang, Mei-Hua; Singer, Burton; Bucur, Alexander, et al. Cumulative biological risk and socio-economic differences in mortality: MacArthur Studies of Successful Aging. *Social Science & Medicine*. 2004; 58(10):1985–1997. [PubMed: 15020014]
- Seeman, Teresa E.; McEwen, Bruce S.; Rowe, John W.; Singer, Burton H. Allostatic load as a marker of cumulative biological risk: MacArthur Studies of Successful Aging. *Proceedings of the National Academy of Sciences of the United States of America*. 2001; 98(8):4770–4775.
- Seeman, Teresa; Gruenewald, Tara; Karlamangla, Arun; Sidney, Steve; Liu, Kiang, et al. Modeling multisystem biological risk in young adults: The Coronary Artery Risk Development in Young Adults Study. *American Journal of Human Biology: The Official Journal of the Human Biology Council*. 2010; 22(4):463–472. [PubMed: 20039257]
- Sidorenkov, Oleg; Nilssen, Odd; Grjibovski, Andrej M. Determinants of cardiovascular and all-cause mortality in northwest Russia: a 10-year follow-up study. *Annals of Epidemiology*. 2012; 22(1): 57–65. [PubMed: 21982128]
- StataCorp. Stata: Release 12. Statistical Software. College Station, TX: StataCorp LP; 2011.
- Stork; Stefan; Feelders, Richard A.; van den Beld, Annemieke W.; Steyerberg, Ewout W.; Savelkoul, Huub FJ., et al. Prediction of mortality risk in the elderly. *The American Journal of Medicine*. 2006; 119(6):519–525. [PubMed: 16750966]
- Swindell, William R.; Ensrud, Kristine E.; Cawthon, Peggy M.; Cauley, Jane A.; Cummings, Steve R., et al. Indicators of “healthy aging” in older women (65–69 years of age). A data-mining approach based on prediction of long-term survival. *BMC Geriatrics*. 2010; 10 55-2318-10-55.
- The Biomarker Network. Descriptions of Selected Population Studies with Biomarkers. USC/UCLA Center of Biodemography and Population Health. 2013 <<<http://gero.usc.edu/CBPH/network/resources/studies/>>>.
- Turra, Cassio M.; Goldman, Noreen; Seplaki, Christopher L.; Glei, Dana A.; Lin, YH., et al. Determinants of mortality at older ages: The role of biological markers of chronic disease. *Population and Development Review*. 2005; 31(4):677–701.
- Wachter, Kenneth W.; Finch, Caleb E., editors. *Between Zeus and the Salmon: The Biodemography of Longevity*. Washington, DC: National Academy Press, Committee on Population, Commission on Behavioral and Social Sciences and Education, National Research Council; 1997.
- Wachter, Kenneth W. Biosocial Opportunities for Surveys. In: Finch, CE.; Vaupel, JW.; Kinsella, K., editors. *Cells and Surveys: Should Biological Measures be Included in Social Science Research?*. Washington, D.C: National Academy Press; 2001. p. 329-336.
- Walzog, Barbara; Gaetgens, Peter. *Adhesion Molecules: The Path to a New Understanding of Acute Inflammation*. *News in Physiological Sciences : An International Journal of Physiology Produced Jointly by the International Union of Physiological Sciences and the American Physiological Society*. 2000; 15:107–113.
- Wang, Thomas J.; Gona, Philimon; Larson, Martin G.; Tofler, Geoffrey H.; Levy, Daniel, et al. Multiple biomarkers for the prediction of first major cardiovascular events and death. *The New England Journal of Medicine*. 2006; 355(25):2631–2639. [PubMed: 17182988]
- World Health Organization (WHO). *Waist circumference and waist-hip ratio: report of a WHO expert consultation*. Geneva: WHO; 2008.
- World Health Organization, International Society of Hypertension Writing Group. 2003 World Health Organization (WHO)/International Society of Hypertension (ISH) statement on management of hypertension. *Journal of Hypertension*. 2003; 21(11):1983–1992. [PubMed: 14597836]
- Yourman, Lindsey C.; Lee, Sei J.; Schonberg, Mara A.; Widera, Eric W.; Smith, Alexander K. Prognostic indices for older adults: a systematic review. *JAMA : The Journal of the American Medical Association*. 2012; 307(2):182–192. [PubMed: 22235089]
- Zimmer, Zachary; Martin, Linda G.; Lin, Hui-Sheng. Determinants of old-age mortality in Taiwan. *Social Science & Medicine*. 2005; 60(3):457–470. [PubMed: 15550295]

## APPENDIX

### PREDICTED PROBABILITY OF DYING BY THE END OF FOLLOW-UP

To calculate the AUC using ROC analysis, we use the model coefficients to compute the predicted probability of dying by June 30, 2011 for each respondent, which is then compared with the observed outcome (i.e., whether or not the respondent actually died). The Gompertz proportional hazards model takes the following form:

$$\log\lambda(t)=x\beta+\gamma t, \quad (1)$$

where  $t$  represents time measured in age,  $\lambda(t)$  is the hazard rate at time  $t$  (age),  $\gamma$  denotes the age slope,  $\chi$  represents a covariate, and  $\beta$  is the corresponding regression coefficient. In our case, we fit a model that allows for non-proportional hazards. That is,  $\gamma$  is a function of  $\chi$  (i.e., the covariate  $\chi$  is interacted with age)

$$\gamma=\gamma_0+\gamma_1x \quad (2)$$

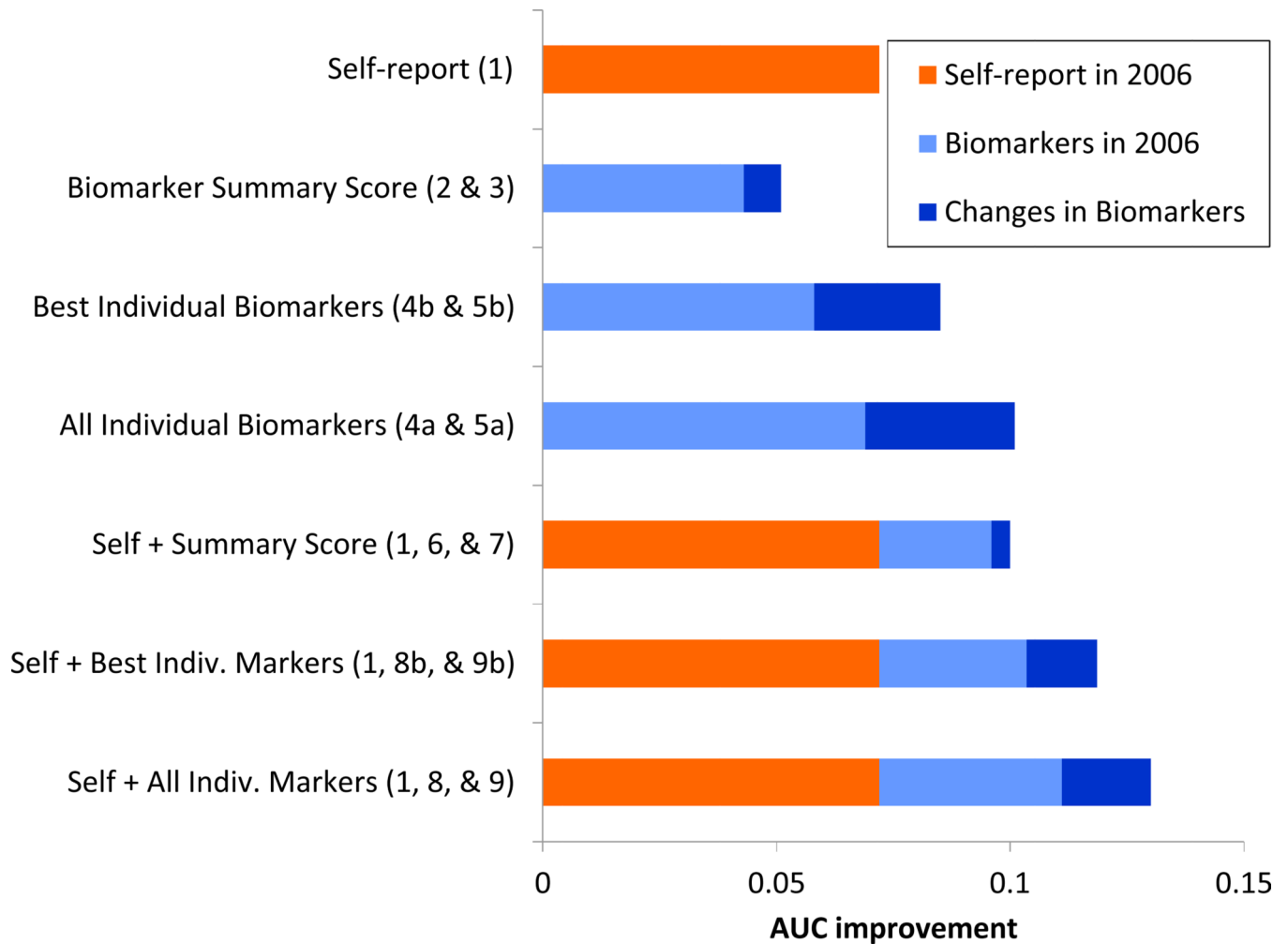
For this model, the conditional probability of surviving from the date of the survey ( $t_0$ ) to the end of follow-up ( $t_1$ ) can be computed as:

$$S(t_1|t_0)=exp\left\{-e^{x\beta}(e^{\gamma t_1}-e^{\gamma t_0})/\gamma\right\}. \quad (3)$$

Thus, for each respondent, we: a) calculate the linear prediction ( $\chi\beta$ ) based on the observed value(s) for the covariate(s) and the model coefficient(s); b) compute  $\gamma$  given the observed value(s) of any covariates that were interacted with time  $t$  (age); and c) estimate conditional survival for using equation (3).

The probability of dying between  $t_0$  and  $t_1$  is simply the complement:

$$\hat{q}(t_1|t_0)=1-\hat{S}(t_1|t_0). \quad (4)$$



**Figure 1.**

Improvement in predictive ability for models based on self-reported indicators of health status, biomarkers, and changes in biomarkers

Note: The improvement in the AUC is measured relative to the baseline model that controls for sex, Mainlander ethnicity, urban residence, education, social integration, and perceived social support (Table A-4, Model 0). The numbers shown in parentheses indicate the model numbers upon which the AUC improvement is based.

Table 1

Selected studies using biomarkers from multiple systems to predict all-cause mortality<sup>a</sup>

Study; Dataset; Country	Sample Size/Ages; Number Deaths; Mean follow up/End of observation	Biomarkers Included				AUC (Controls Only)	AUC gain (Add Biomarkers)
		Cardiovascular/Metabolic	Inflammatory	Neuro-endocrine	Other		
<i>Studies Using Individual Biomarkers Representing Multiple Systems</i>							
Rosero-Bixby & Dow (2012); CRELES; Costa Rica	n=2313 aged 60+; n=564 deaths; 5 years ending 2010	SBP DBP TC HDL Ratio TC/HDL LDL TG HbA1c GLU Waist BMI WHR Knee height	CRP	DHEAS U. CORT U. EPI U. NE	CrCl Grip strength Walking speed PEF	0.72	0.06
Haring et al. (2011); Study of Health in Pomerania; Germany	n=4261 aged 20-79; n=456 deaths; 9.7 years ending 12/15/2009	HbA1c	Fbg		U. Albumin GGT	0.88	0.004
Goldman et al. (2009); SEBAS 2000; Taiwan	n=933 aged 54+; n=162 deaths; 6 years ending 12/31/2006	Hypertension TC HDL BMI Waist HbA1c	WBC % NEUT IL-6	DHEAS U. CORT U. EPI U. NE	CrCl Albumin IGF-1	0.79	0.06

Study; Dataset; Country	Sample Size/Ages; Number Deaths; Mean follow up/ End of observation	Biomarkers Included				Control Variables	AUC (Controls Only)	AUC gain (Add Biomarkers)
		Cardiovascular/ Metabolic	Inflammatory	Neuro-endocrine	Other			
Goldman et al. (2006b); SEBAS 2000; Taiwan	n=927 aged 54+; n=61 deaths; 3 years ending in 2003	SBP DBP TC Ratio TC/HDL HbA1c BMI	IL-6	DHEAS U. CORT U. EPI U. NE U. DOP	IGF-1	Use of: BP meds Diabetic meds  Age Sex Urban Smoking # Current chronic conditions Mobility limitations Cognitive function CES-D SAH Pain level	0.79	0.08
Stork et al. (2006); Zoetermeer, Netherlands	n=403 men aged 70+; n=75 deaths; 4 years ending in 2000	Included in "Controls"	CRP IL-6		# Carotid plaques	Age SBP Pulse Smoking Use of: BP meds Diabetic meds Hx of MI BMI	0.62	0.15
Newson et al. (2010); Rotterdam Study; Netherlands	n=2008 aged 55-85; n=639 deaths before 85 <sup>th</sup> birthday; 7.9 years ending 1/1/2007	DBP	CRP		Albumin Uric acid Femoral neck bone mineral density Ankle brachial index # carotid plaques	Age Sex Smoking Left ventricular function SAH Cognitive status Any IADL limitations Body weight stable past year	0.75	N/A
<i>Studies Using Multi-System Biomarker Summary Scores</i>								
Borrell et al. (2010); NHANES III; U.S.	n=13,715 aged 25+; n=2491 deaths; 8.7 years ending 12/31/2000	SBP DBP Pulse TC HDL HbA1c WHR	CRP		Albumin	Age Sex Race/Ethnicity Education Income	N/A	N/A

Study; Dataset; Country	Sample Size/Ages; Number Deaths; Mean follow up/ End of observation	Biomarkers Included				Control Variables	AUC (Controls Only)	AUC gain (Add Biomarkers)
		Cardiovascular/ Metabolic	Inflammatory	Neuro-endocrine	Other			
Crimmins et al. (2009); NHANES III; U.S.	n=14,443 aged 20+; n=2134 deaths; 8.3 years ending 12/31/2000	SBP DBP Pulse TC HDL HbA1c BMI	CRP		Albumin	Age Sex Race Poor Smoking Heavy Drinking Exercise BMI	N/A	N/A
Goldman et al. (2006a); SEBAS 2000; Taiwan	n=935 aged 54+; n=72 deaths; 3 years ending in 2004	SBP DBP TC Ratio TC/HDL TG GLU HbA1c BMI WHR	IL-6		IGF-1	Age Sex Urban # Chronic conditions Mobility CES-D Cognitive function Smoking	N/A	N/A
Gruenewald et al. (2006); MacArthur Study of Successful Aging; U.S.	n=1189 aged 70-79; n=492 deaths; 12 years ending in 2000	SBP DBP HDL Ratio TC/HDL HbA1c WHR	IL-6 CRP Fbg	DHEA U. CORT U. EPI U. NE	Albumin	Modeled separately by sex	N/A	N/A
Karlamangla et al. (2006); MacArthur Study of Successful Aging; U.S.	n=171 aged 70-79; n=19 deaths; 4.5 years ending in 1995	SBP DBP TC HDL HbA1c WHR		DHEAS U. CORT U. EPI U. NE		Age	N/A	N/A
Wang et al. (2006); Framingham Heart Study; U.S.	n=3209 mean age 59; n=207 deaths; 7.4 years (date follow up ended is not stated)		CRP		B-type natriuretic peptide Renin Hcy UACR	Age Sex Smoking Hypertension TC HDL Diabetes BMI SCR Presence of CVD	0.80	0.02
Seeman et al. (2004); MacArthur Study of Successful Aging; U.S.	n=657 aged 70-79; n=141 deaths; 7.5 years ending 12/31/1995	SBP DBP HDL Ratio TC/HDL HbA1c WHR	IL-6 CRP Fbg	DHEAS U. CORT U. EPI U. NE	Albumin CrCl PEF	Age Sex Race Education # Chronic conditions	N/A	N/A



Study; Dataset; Country	Sample Size/Ages; Number Deaths; Mean follow up/ End of observation	Biomarkers Included				AUC (Controls Only)	AUC gain (Add Biomarkers)
		Cardiovascular/ Metabolic	Inflammatory	Neuro-endocrine	Other		
Seeman et al. (2001) MacArthur Study of Successful Aging; U.S.	n=720 aged 70–79; n=153 deaths; 7 years ending in 1997	SBP DBP HDL Ratio TC/HDL HbA1c WHR		DHEAS U. CORT U. EPI U. NE		N/A	N/A
Levine et al. (2013); NHANES III; U.S.	n=9389 aged 30–75; n=1843 deaths; 18 years ending 12/31/2006	SBP TC HbA1c	CRP CMV optical density		SCr Albumin S. ALP FEV SUN	0.83	N/A

Abbreviations: BMI = body mass index; BP = blood pressure; CES-D = Center for Epidemiologic Studies Depression scale; CMV = cytomegalovirus; CrCl = creatinine clearance; CRP = c-reactive protein; DBP = diastolic blood pressure; DHEA = dehydroepiandrosterone; DHEAS = dehydroepiandrosterone sulfate; Fbg = fibrinogen; FEV = forced expiratory volume; GGT = gamma glutamyltransferase; GLU = fasting glucose; HbA1c = glycosylated hemoglobin; Hcy = homocysteine; HDL = high-density lipoprotein cholesterol; Hx = history; IADL = instrumental activities of daily living; IGF-1 = insulin-like growth factor 1; MI = myocardial infarction; NEUT = neutrophils; PEF = peak expiratory flow; QoL = quality of life; SAH = self-assessed health status; S. ALP = serum alkaline phosphatase; SBP = systolic blood pressure; SCr = serum creatinine; SUN = serum urea nitrogen; TC = total cholesterol; UACR = urinary albumin-to-creatinine ratio U. CORT = urinary cortisol; U. EPI = urinary epinephrine; U. NE = urinary norepinephrine; WHR = waist-hip ratio.

<sup>a</sup> For studies of individual biomarkers, we list in this table only those that: 1) included at least one standard cardiovascular/metabolic markers (e.g., blood pressure, lipids, measures of glucose metabolism, and obesity) as well as at least one inflammatory marker (e.g., interleukin-6, IL-6; C-reactive protein, CRP); and 2) assessed the discriminatory ability of the set of biomarkers. For studies using a multi-marker summary score, we included all studies that used a summary score based on markers from multiple systems to predict all-cause mortality.

**Table 2**

## Description of biomarkers included in the analysis

<b>Biomarker</b>	<b>Description</b>
<u>Cardiovascular/Metabolic</u>	
Systolic blood pressure (SBP)	Maximum arterial blood pressure after contraction (systole) of the heart's left ventricle.
Diastolic blood pressure (DBP)	Minimum arterial blood pressure when the heart fills with blood (diastole).
Total cholesterol	Waxy, fat-like substance that regulates the permeability of cell membranes. Excess cholesterol in the blood can combine with other substances and stick to the walls of the arteries forming a plaque. Includes both LDL (bad) and HDL (good) cholesterol.
High-density lipoprotein (HDL) cholesterol	Lipoprotein in the blood characterized by a high ratio of protein relative to triglyceride and cholesterol; helps remove cholesterol from the arteries.
Triglycerides	Most common type of fat in the human body; widespread in adipose tissue and typically circulate in the blood in the form of lipoproteins.
Glycosylated hemoglobin (HbA1c)	Measures the level of hemoglobin A1c (glycoprotein formed when glucose binds to hemoglobin A) in the blood; represents average blood sugar concentrations over the previous 2–3 months.
Body mass index (BMI)	Measure of body fat computed as the ratio of the body weight (in kg) divided by height (in m) squared. BMI is typically classified as underweight (<18.5), normal (18.5–24.9), overweight (25–29.9), or obese (30+), but some evidence suggests that Asians have increased health risk at a lower cutoff for obesity.
Waist circumference	Marker of abdominal fat content.
<u>Inflammation</u>	
Interleukin-6 (IL-6)	Messenger cytokine that stimulates the synthesis of acute phase proteins (e.g., CRP, fibrinogen) in the liver as part of the inflammatory process.
C-reactive protein (CRP)	Protein produced in the liver that increases markedly with acute inflammation.
Soluble intercellular adhesion molecule 1 (sICAM-1) and Soluble E-selectin (sE-selectin)	Cell adhesion molecules such as ICAM-1 and E-selectin are proteins that help cells bind to one another. Upon activation by inflammatory cytokines, ICAM-1 and E-selectin are expressed by endothelial cells that line blood vessels and facilitate the transfer of leukocytes from the blood to inflamed tissue. sICAM-1 and sE-selectin are soluble forms shed by activated cells and are measurable in blood.
<u>Neuroendocrine</u>	
Dehydroepiandrosterone sulfate (DHEAS)	Steroid hormone (androgen) produced by the adrenal gland in both sexes.
Cortisol	Steroid hormone produced by the adrenal gland that has anti-inflammatory and immunosuppressive properties. It is a metabolite of the primary stress hormone cortisone and serves as the main glucocorticoid in humans. Levels may be elevated in response to physical or psychological stress.
Epinephrine	Catecholamine secreted by the adrenal medulla as part of the acute stress response (fight-or-flight). It is the principal hormone that causes increased blood pressure and also increases the heart rate.
Norepinephrine	Catecholamine produced by the adrenal medulla. It is a precursor to epinephrine that narrows blood vessels and raises blood pressure.
<u>Other markers</u>	
Serum Creatinine	Breakdown product of creatine, which is an important part of muscle. Creatinine is produced from the metabolism of protein (e.g., when muscles burn energy). Most is filtered out of the blood by the kidneys and excreted in urine. Used to evaluate kidney function.
Creatinine Clearance (CrCl)	Measures the rate at which a waste (creatinine) is cleared from the blood by the kidneys. Used to estimate the glomerular filtration rate (GFR), which reflects kidney function. Can be measured by comparing the level of creatinine in urine (based on 24h collection) with the level in the blood, but

Biomarker	Description
	is typically estimated using a prediction equation (e.g., Cockcroft-Gault) based on serum creatinine, age, sex, and body weight.
Serum albumin	The main protein in human blood; helps maintain the osmotic pressure of the blood. Low levels may be a sign of liver or kidney disease or reflect insufficient nutrition.
Homocysteine	Amino acid produced by the body, usually as a byproduct of meat consumption. Linked with increased risk for cardiovascular disease; elevated levels may promote atherosclerosis.

*Note:* Adapted from Table 1 in Glei et al. (Forthcoming).

*Sources:* American Association for Clinical Chemistry ([labtestsonline.org](http://labtestsonline.org), accessed 12 August 2013); American Heart Association ([http://www.heart.org/idc/groups/heart-public/@wcm/@hcm/documents/downloadable/ucm\\_300308.pdf](http://www.heart.org/idc/groups/heart-public/@wcm/@hcm/documents/downloadable/ucm_300308.pdf), accessed 12 August 2013); Libby & Ridker (1999); Levey et al. (2003); MedicineNet.com MedTerms dictionary (<http://www.medterms.com/script/main/hp.asp>, accessed 12 August 2013); MedlinePlus (<http://www.nlm.nih.gov/medlineplus/mplusdictionary.html>, accessed 8 August 2013); National Heart, Lung, and Blood Institute (1998); National Kidney Disease Education Program (<http://nkdep.nih.gov/resources/quick-reference-uacr-gfr-508.pdf>, accessed 12 August 2013); Walzog & Gaehtgens (2000); World Health Organization ([http://apps.who.int/bmi/index.jsp?introPage=intro\\_3.html](http://apps.who.int/bmi/index.jsp?introPage=intro_3.html), accessed 22 January 2014).

**Table 3**

Descriptive statistics for social and demographic characteristics, self-reported indicators of health status, biomarker summary scores, and survival status

	Analysis sample (N=639)
<b><u>Social and demographic characteristics</u></b>	
Age at the 2006 exam (60–97), mean (SD)	72.0 (7.4)
Female, %	44.4
Mainlander, %	12.9
Urban resident, %	42.3
Years of completed education (0–17), mean (SD)	5.3 (4.5)
Social integration (–1.5 to 1.6), mean (SD)	0.1 (0.5)
Perceived availability of social support (0.5–4.0), mean (SD)	3.1 (0.7)
<b><u>Self-reported health indicators</u></b>	
Self-assessed health status (1–5, 5=excellent), mean (SD)	3.0 (1.0)
Index of mobility limitations (–0.7 to 3.2), mean (SD)	0.7 (1.3)
History of diabetes, %	19.9
History of cancer, %	4.8
Number of hospitalizations in the past 12 months (0–11), mean (SD)	0.3 (0.8)
Smoking status in 2006	
Never, %	59.1
Former, %	22.2
Current, %	18.7
<b><u>Biomarker summary scores</u></b>	
Overall biomarker risk in 2000 (0–13), mean (SD)	4.5 (2.5)
Overall biomarker risk in 2006 (0–14), mean (SD)	5.1 (2.6)
Change (2006 – 2000) in biomarker risk (–6 to +9), mean (SD)	0.6 (2.3)
Declined by two or more points, %	17.3
Declined by one point, %	15.2
Not change, %	18.1
Increased by one point, %	15.9
Increased by two points, %	13.5
Increased by three or more points, %	20.0
<b>Died between the 2006 exam and December 31, 2011, %</b>	<b>16.2</b>

Measures of fit and discrimination for models predicting mortality as a function of social and demographic characteristics, self-reported indicators of health status, and biomarkers

**Table 4**

Model	Description	Average log L <sup>a</sup>	Number of parameters	Average AIC <sup>b</sup>	AUC <sup>c</sup>	Change in AUC
0	Baseline: Social and demographic characteristics only	-112.4	9	242.9	0.745	
<b>Panel 1: Self-reported health indicators</b>						
1	Model 0 + Self-reported health indicators	-83.7	17	201.3	0.817	vs Model 0 0.072***
<b>Panel 2: Biomarkers</b>						
2	Model 0 + Biomarker summary score 2006	-94.9	10	209.8	0.788	0.043**
3	Model 0 + Biomarker summary score (2006 & change 2000-06)	-89.4	11	200.8	0.796	0.051**
4a	Model 0 + Individual biomarkers 2006	-76.8	28	209.5	0.814	0.069***
4b	Model 0 + Best individual biomarkers 2006 <sup>d</sup>	-82.8	19	203.6	0.803	0.058**
5a	Model 0 + Individual biomarkers (2006 & changes 2000-06)	-57.4	48	210.9	0.846	0.101***
5b	Model 0 + Best individual biomarkers (2006 & changes 2000-06) <sup>d</sup>	-68.9	30	197.7	0.830	0.084***
<b>Panel 3: Biomarkers net of self-reported health indicators</b>						
6	Model 1 + Biomarker summary score 2006	-70.8	18	177.6	0.841	vs Model 1 0.024
7	Model 1 + Biomarker summary score (2006 & change 2000-06)	-67.8	19	173.5	0.845	0.028*
8a	Model 1 + Individual biomarkers in 2006	-53.6	36	179.2	0.856	0.039*
8b	Model 1 + Best individual biomarkers in 2006 <sup>d</sup>	-59.0	27	172.1	0.849	0.031*
9a	Model 1 + Individual biomarkers (2000 & changes 2000-06)	-37.6	56	187.3	0.875	0.058***
9b	Model 1 + Best individual biomarkers (2000 & changes 2000-06) <sup>d</sup>	-48.1	38	172.2	0.864	0.047*

\*\*\*  
 $p < 0.001$ ,

\*\*  
 $p < 0.01$ ,

\*  
 $p < 0.05$

<sup>a</sup>The model is fitted on each of five multiply imputed datasets, and then we compute the mean of the five corresponding log L values.

<sup>b</sup>Based on the average log L across five multiply imputed datasets. Lower values of AIC indicate better model fit.

<sup>c</sup>Higher values of AUC indicate better discrimination. An AUC of 0.5 would indicate that the model was no better than random chance.

<sup>d</sup>Model includes the best 10 (out of 19) biomarkers as determined by comparing model fit (AIC) for 19 separate models in which the two parameters (2006 level & 2000-06 change) for each marker were added to the baseline model. The best 10 markers comprise HDL, BMI, IL-6, CRP, sICAM-1, sE-selectin, DHEAS, creatinine clearance, albumin, and homocysteine.

**Table 5**

Coefficients from selected models predicting mortality using social and demographic characteristics, self-reported indicators of health status, and biomarkers

	<b>Model 1</b>	<b>Model 7</b>	<b>Model 9b</b>
Age slope <sup>a</sup>			
Age	0.08***	0.06**	0.08**
Age × Perceived social support	0.05***	0.04**	0.04***
Age × Current smoker	0.11*	0.13**	0.15**
Age × Change in DHEAS (square-root)			0.04***
Female	-0.61*	-0.85**	-0.11
Mainlander	-0.62*	-0.65*	-0.70*
Urban resident	0.03	-0.09	-0.15
Education <sup>b</sup>	0.12	0.19	0.32*
Social integration <sup>b</sup>	-0.10	-0.07	-0.07
Perceived social support <sup>b</sup>	-1.05***	-0.94***	-1.02***
Self-assessed health status <sup>b</sup>	-0.20	-0.19	-0.19
Index of mobility limitations <sup>b</sup>	0.38**	0.30*	0.37*
History of diabetes	0.33	-0.04	-0.18
History of cancer	0.34	0.01	0.55
Number of hospitalizations <sup>b</sup>	0.28***	0.27***	0.23**
Former smoker	0.12	-0.01	0.19
Current smoker	-1.97*	-2.64**	-2.90**
Biomarker risk score in 2006 <sup>b</sup>		0.70***	
Change (2006 – 2000) in biomarker risk <sup>b</sup>		-0.29*	
HDL (log) in 2006 <sup>b</sup>			-0.18
Change (2006 – 2000) in HDL (log) <sup>b</sup>			0.13
BMI (log) in 2006 <sup>b</sup>			-0.16
Change (2006 – 2000) in BMI (log) <sup>b</sup>			-0.07
IL-6 (log) in 2006 <sup>b</sup>			0.55***
Change (2006 – 2000) in IL-6 (log) <sup>b</sup>			-0.29
CRP (log) in 2006 <sup>b</sup>			0.02
Change (2006 – 2000) in CRP (log) <sup>b</sup>			0.00
sICAM-1 (square root) in 2006 <sup>b</sup>			0.29*
Change (2006 – 2000) in sICAM-1 (square root) <sup>b</sup>			-0.06
sE-selectin (log) in 2006 <sup>b</sup>			0.07
Change (2006 – 2000) in sE-selectin (log) <sup>b</sup>			-0.02
DHEAS (square-root) in 2006 <sup>b</sup>			-0.14
Change (2006 – 2000) in DHEAS (square-root) <sup>b</sup>			-0.86***

	<b>Model 1</b>	<b>Model 7</b>	<b>Model 9b</b>
Creatinine clearance in 2006 <sup>b</sup>			0.27
Change (2006 – 2000) in Creatinine clearance <sup>b</sup>			-0.18
Albumin (cubed) in 2006 <sup>b</sup>			-0.02
Change (2006 – 2000) in Albumin (cubed) <sup>b</sup>			-0.05
Hcy (log) in 2006 <sup>b</sup>			0.57**
Change (2006 – 2000) in Hcy (log) <sup>b</sup>			-0.06
Intercept <sup>c</sup>	-4.86***	-4.32***	-5.21***

<sup>a</sup>The age slope represents the exponential increase in the mortality rate per year of age.

<sup>b</sup>This variable was standardized (Mean=0, SD=1) prior to fitting the model; so, the coefficient represents the effect per SD of the specified variable.

<sup>c</sup>Time was measured in terms of years after age 60. Thus, the intercept represents the mortality rate at age 60.



**Table A-1**

Biomarker summary score: cutoffs used to define high-risk levels for each biomarker

	<b>Definition of high risk:</b> Clinical cutoffs (where available); Otherwise, high risk quartile <sup>a</sup> (bottom 25% for DHEAS, CrCl, & Albumin; Top 25% for all others)
<b>Cardiovascular/Metabolic</b>	
SBP (mmHg)	>140 (Chobanian et al. 2003, World Health Organization, International Society of Hypertension Writing Group 2003)
DBP (mmHg)	>90 (Chobanian et al. 2003, World Health Organization, International Society of Hypertension Writing Group 2003)
Ratio total/HDL cholesterol	>5 (AHA website)
HDL (mg/dL)	<40 (National Cholesterol Education Program, National Heart, Lung and Blood Institute 2001)
Triglycerides (mg/dL)	>200 (National Cholesterol Education Program, National Heart, Lung and Blood Institute 2001)
HbA1c (%)	>6.5 (Rodbard et al. 2007)
BMI	<18.5 or >27 (Department of Health (Taiwan) 2002)
Waist circumference (cm)	> 88cm(F)/102cm(M) (World Health Organization (WHO) 2008)
<b>Inflammation</b>	
IL-6 (pg/mL)	>3.69
CRP (mg/L)	>3 (Pearson et al. 2003)
sICAM-1 (ng/mL)	>290.9
sE-selectin (ng/mL)	>56.5
<b>Neuroendocrine</b>	
DHEAS (µg/dL)	<42.2
Cortisol (µg/g creatinine)	>28.2
Epinephrine (µg/g creatinine)	>4.44
Norepinephrine (µg/g creatinine)	>26.4
<b>Othermarkers</b>	
CrCl (ml/min) <sup>a</sup>	<52.6
Albumin (g/dL)	<4.4
Homocysteine (µmol/L)	>16.7

Abbreviations: BMI = body mass index; CrCl = creatinine clearance; CRP = C-reactive protein; DBP = diastolic blood pressure; DHEAS = dehydroepiandrosterone sulfate; HbA1c = glycosylated hemoglobin; HDL = high-density lipoprotein cholesterol; IGF-1 = insulin-like growth factor 1; IL-6 = interleukin-6; SBP = systolic blood pressure; sE-selectin = soluble E-selectin; sICAM-1 = soluble intercellular adhesion molecule 1

<sup>a</sup>Quartile cutoffs are based on the weighted distribution in 2000 among the longitudinal cohort (n=639).

<sup>b</sup>Estimated using the Cockcroft-Gault formula (Cockcroft and Gault 1976).

**Table A-2**

Index of social integration: description and coding of each component

Indicator	Definition	Coding
Network size	Number of friends and relatives with whom the respondent lives or has regular contact	Recoded <5, 5–7, 8–10, 11–14, 15–19, 20–29, 30+.
Network range	Number of types of relationships in social network	One point each for spouse/partner, kids, other relatives, non-relatives; range=0–4.
Married/partner	Dummy indicating that the respondent is married or lives with a companion.	
Household size		
Does not live alone	Dummy indicating that the respondent does not live alone.	
Number of friends	Number of close friends and neighbors with whom the respondent has weekly contact	Recoded 0, 1–2, 3–4, 5–9, 10–19, 20+.
Religious attendance	How often the respondent attends church or temple	Response categories: never, rarely, sometimes, often.
Socializing	How often the respondent socializes with friends, neighbors, or relatives.	Response categories: never, less than once a month, two to three times a month, once or twice a week, nearly daily.
Participation in social organizations	Whether respondent participates in the following activities/organizations: <ol style="list-style-type: none"> <li>1 Group activities (e.g., singing, dancing, tai chi, or karaoke)</li> <li>2 Neighborhood association (e.g., women’s association or arts &amp; crafts classes)</li> <li>3 Religious organization (e.g., church or temple committee)</li> <li>4 Occupational associations for farmers, fishermen, or other professional group, civic group, Lion’s Club, etc.</li> <li>5 Political association (e.g., political party)</li> <li>6 Social service groups (e.g., Lifeline, relief association, benevolent societies, charities, etc.)</li> <li>7 Village or lineage association</li> <li>8 Elderly club (e.g., Elderly Association, Evergreen Recreation Club, etc.)</li> </ol>	One point for each type of organization in which the respondent participates; range = 0–7.

**Table A-3**

Summary statistics for individual biomarkers and changes in biomarkers

	Units	Transformation	Mean (SD) for the Transformed Markers:	
			Level in 2006	Change (2006 – 2000)
SBP	mmHg	log	4.90 (0.15)	-0.01 (0.16)
DBP	mmHg	log	4.27 (0.15)	-0.13 (0.15)
Ratio TC/HDL	ratio	log	1.44 (0.27)	0.00 (0.23)
HDL	mg/dL	log	3.83 (0.27)	-0.02 (0.22)
Triglycerides	mg/dL	log	4.57 (0.51)	-0.08 (0.42)
HbA1c	%	-1/(HbA1C) <sup>2</sup>	-0.03 (0.01)	0.01 (0.01)
BMI	$\frac{\text{weight(kg)}}{(\text{height(m)})^2}$	log	3.19 (0.15)	0.00 (0.08)
Waist circumference	cm	none	84.90 (9.94)	-0.54 (5.96)
IL-6	pg/mL	log	1.06 (0.79)	0.26 (0.89)
CRP	mg/L	log	-2.01 (1.12)	0.53 (1.50)
sICAM-1	ng/mL	square root	16.54 (2.87)	1.12 (2.28)
sE-selectin	ng/mL	log	3.57 (0.58)	-0.17 (0.43)
DHEAS	µg/dL	square root	8.88 (3.20)	0.25 (2.06)
Cortisol	µg/g	log	2.68 (0.87)	-0.28 (0.96)
Epinephrine	µg/g	log	1.25 (0.58)	0.12 (0.62)
Norepinephrine	µg/g	log	3.17 (0.53)	0.21 (0.53)
Creatinine Clearance	ml/min	none	58.11 (19.98)	-5.25 (11.24)
Albumin	g/dL	cubed	83.56 (17.26)	-9.06 (15.21)
Homocysteine	µmol/L	log	2.48 (0.39)	-0.20 (0.31)

**Table A-4**

Biomarker summary scores: coefficients from models predicting mortality using social and demographic characteristics, self-reported indicators of health status, and biomarker summary scores

	Model 0	Model 2	Model 3	Model 6
Age slope <sup>a</sup>				
Age	0.11***	0.09***	0.09***	0.06**
Age × Perceived social support	0.04***	0.03**	0.03**	0.04***
Age × Current smoker			0.14**	0.14**
Female	-0.46*	-0.67**	-0.77**	-0.82**
Mainlander	-0.55*	-0.58*	-0.57*	-0.64*
Urban resident	-0.06	-0.18	-0.21	-0.08
Education <sup>b</sup>	0.03	0.12	0.16	0.16
Social integration <sup>b</sup>	-0.14	-0.12	-0.09	-0.09
Perceived social support <sup>b</sup>	-0.96***	-0.82***	-0.81**	-0.98***
Self-assessed health status <sup>b</sup>				-0.18
Index of mobility limitations <sup>b</sup>				0.34*
History of diabetes				0.07
History of cancer				0.11
Number of hospitalizations <sup>b</sup>				0.28***
Former smoker				-0.02
Current smoker				-2.73**
Biomarker risk score in 2006 <sup>b</sup>		0.60***	0.80***	0.53***
Change (2006 – 2000) in biomarker risk <sup>b</sup>			-0.35**	
Intercept <sup>c</sup>	-5.13***	-4.83***	-4.77***	-4.35***

<sup>a</sup>The age slope represents the exponential increase in the mortality rate per year of age.

<sup>b</sup>This variable was standardized (Mean=0, SD=1) prior to fitting the model; so, the coefficient represents the effect per SD of the specified variable.

<sup>c</sup>Time was measured in terms of years after age 60. Thus, the intercept represents the mortality rate at age 60.

Individual biomarkers: coefficients from models predicting mortality using social and demographic characteristics, social factors, self-reported indicators of health status, and individual biomarkers

Table A-5

	Model 4a	Model 4b	Model 5a	Model 5b	Model 8a	Model 8b	Model 9a
Age slope ( $\gamma^a$ )							
Age	0.09***	0.08***	0.10***	0.11***	0.07**	0.06**	0.09**
Age $\times$ Perceived social support	0.04**	0.04***	0.03*	0.03**	0.05***	0.05***	0.05**
Age $\times$ Current smoker				0.14**	0.15**	0.14**	0.14**
Age $\times$ Change in $\sqrt{DHEAS}$			0.05***	0.04***			0.04***
Female	-0.22	-0.25	-0.37	-0.25	-0.28	-0.21	-0.45
Mainlander	-0.53	-0.54	-0.48	-0.60*	-0.70*	-0.66*	-0.68*
Urban resident	-0.02	-0.12	-0.03	-0.12	-0.10	-0.19	-0.06
Education <sup>b</sup>	0.16	0.19	0.23	0.22	0.23	0.26*	0.33*
Social integration <sup>b</sup>	-0.14	-0.16	-0.08	-0.13	-0.15	-0.14	-0.06
Perceived social support <sup>b</sup>	-1.00***	-1.01***	-0.82**	-0.81**	-1.13***	-1.16***	-1.06***
Self-assessed health status <sup>b</sup>					-0.25	-0.23	-0.19
Index of mobility limitations <sup>b</sup>					0.33*	0.34*	0.35*
History of diabetes					0.19	0.10	0.26
History of cancer					0.49	0.40	0.65
Number of hospitalizations <sup>b</sup>					0.26***	0.25***	0.26***
Former smoker					0.11	0.13	0.06
Current smoker					-2.67*	-2.82**	-2.75*
SBP (log) in 2006 <sup>b</sup>	0.18		0.28		0.16		0.26
Change (2006 – 2000) in SBP (log) <sup>b</sup>			-0.11				-0.08
DBP (log) in 2006 <sup>b</sup>	0.07		0.04		0.07		0.02
Change (2006 – 2000) in DBP (log) <sup>b</sup>			-0.01				0.03
TC/HDL (log) in 2006 <sup>b</sup>	0.19		0.44*		0.32		0.56*
Change (2006 – 2000) in TC/HDL (log) <sup>b</sup>			-0.15				-0.09

	Model 4a	Model 4b	Model 5a	Model 5b	Model 8a	Model 8b	Model 9a
HDL (log) in 2006 <sup>b</sup>	-0.01	-0.11	-0.01	-0.19	0.07	-0.07	0.05
Change (2006 – 2000) in HDL (log) <sup>b</sup>			-0.07	0.12			-0.01
TG (log) in 2006 <sup>b</sup>	-0.15		-0.31		-0.21		-0.39*
Change (2006 – 2000) in TG (log) <sup>b</sup>			-0.08				-0.08
-1/(HbA1c) <sup>2</sup> in 2006 <sup>b</sup>	-0.06		-0.16		-0.09		-0.25
Change (2006 – 2000) in -1/(HbA1c) <sup>2</sup> <sup>b</sup>			0.01				0.09
BMI (log) in 2006 <sup>b</sup>	-0.03	-0.04	0.03	-0.14	0.07	-0.08	0.20
Change (2006 – 2000) in BMI (log) <sup>b</sup>			-0.09	-0.13			-0.09
Waist in 2006 <sup>b</sup>	-0.03		-0.19		-0.18		-0.40
Change (2006 – 2000) in Waist <sup>b</sup>			-0.10				0.01
IL-6 (log) in 2006 <sup>b</sup>	0.32**	0.32**	0.48**	0.42**	0.40**	0.37**	0.58***
Change (2006 – 2000) in IL-6 (log) <sup>b</sup>			-0.15	-0.22			-0.19
CRP (log) in 2006 <sup>b</sup>	0.09	0.06	0.08	0.12	-0.00	-0.00	-0.01
Change (2006 – 2000) in CRP (log) <sup>b</sup>			-0.09	-0.06			-0.08
$\sqrt{\text{sICAM-1}}$ in 2006 <sup>b</sup>	0.18	0.18	0.23	0.27*	0.19	0.18	0.26
Change (2006 – 2000) in $\sqrt{\text{sICAM-1}}$ <sup>b</sup>			-0.05	-0.04			-0.06
sE-selectin (log) in 2006 <sup>b</sup>	0.13	0.08	0.17	0.10	0.09	0.04	0.21
Change (2006 – 2000) in sE-selectin (log) <sup>b</sup>			-0.03	-0.03			-0.09
$\sqrt{\text{DHEAS}}$ in 2006 <sup>b</sup>	-0.18	-0.20	-0.26	-0.20	-0.10	-0.14	-0.21
Change (2006 – 2000) in $\sqrt{\text{DHEAS}}$ <sup>b</sup>			-0.95***	-0.87***			-0.87**
Cortisol (log) in 2006 <sup>b</sup>	0.03		0.10		0.04		-0.01
Change (2006 – 2000) in Cortisol (log) <sup>b</sup>			-0.07				0.11
EPI (log) in 2006 <sup>b</sup>	0.19		0.11		0.13		0.13
Change (2006 – 2000) in EPI (log) <sup>b</sup>			-0.02				-0.08
NE (log) in 2006 <sup>b</sup>	-0.17		0.12		-0.08		0.14
Change (2006 – 2000) in NE (log) <sup>b</sup>			-0.33*				-0.29
CrCl in 2006 <sup>b</sup>	-0.01	-0.14	0.13	0.11	0.18	0.11	0.27

	Model 4a	Model 4b	Model 5a	Model 5b	Model 8a	Model 8b	Model 9a
Change (2006 – 2000) in CrCl <sup>b</sup>			-0.27	-0.27			-0.24
Albumin (cubed) in 2006	-0.24*	-0.23*	-0.13	-0.11	-0.09	-0.06	-0.04
Change (2006 – 2000) in Albumin (cubed)			-0.04	-0.11			0.01
Hcy (log) in 2006 <sup>b</sup>	0.32*	0.32**	0.50**	0.36*	0.46**	0.52***	0.60**
Change (2006 – 2000) in Hcy (log) <sup>b</sup>			-0.24	-0.05			-0.24
Intercept <sup>c</sup>	-5.25***	-5.04***	-5.66***	-5.53***	-4.98***	-4.77***	-5.48***

\*  $p < 0.05$ ,

\*\*  $p < 0.01$ ,

\*\*\*  $p < 0.001$

<sup>a</sup>The age slope ( $\gamma$ ) represents the exponential increase in the mortality rate per year of age.

<sup>b</sup>This variable was standardized (Mean=0, SD=1) prior to fitting the model; so, the coefficient represents the effect per SD of the specified variable.

<sup>c</sup>Time was measured in terms of years after age 60. Thus, the intercept represents the mortality rate at age 60.

Subscores for biomarker clusters: coefficients from models predicting mortality using social and demographic characteristics, self-reported indicators of health status, and biomarker subscores

Table A-6

	Model 7a	Model 7b	Model 7c	Model 7d	Model 7e
Age slope ( $\gamma^a$ )					
Age	0.08***	0.07***	0.08***	0.05*	0.05*
Age $\times$ Perceived social support	0.04***	0.05***	0.05***	0.05***	0.04***
Age $\times$ Current smoker	0.12**	0.14**	0.11*	0.11*	0.14**
Female	-0.69*	-0.54	-0.78*	-0.46	-0.71*
Mainlander	-0.60*	-0.64*	-0.60*	-0.67*	-0.65*
Urban resident	-0.03	-0.07	0.02	0.03	-0.07
Education <sup>b</sup>	0.12	0.19	0.11	0.17	0.19
Social integration <sup>b</sup>	-0.10	-0.14	-0.09	-0.10	-0.10
Perceived social support <sup>b</sup>	-0.20	-0.16	-0.20	-0.21	-0.19
Self-assessed health status <sup>b</sup>	0.34*	0.37**	0.38**	0.38**	0.30*
Index of mobility limitations <sup>b</sup>	-0.00	0.23	0.39	0.30	0.02
History of diabetes	0.06	0.17	0.29	0.41	0.06
History of cancer	0.30***	0.27***	0.27***	0.25***	0.24***
Number of hospitalizations <sup>b</sup>	0.06	0.05	0.11	0.14	0.05
Former smoker	-2.08*	-2.71**	-1.99*	-2.10*	-2.67**
Current smoker	-0.20	-0.16	-0.20	-0.21	-0.19
Cardiovascular/metabolic subscore in 2006 <sup>b</sup>	0.36**				0.30*
Change (2006 – 2000) in cardiovascular/metabolic subscore <sup>b</sup>	-0.26*				-0.21
Inflammation subscore in 2006 <sup>b</sup>		0.49***			0.34**
Change (2006 – 2000) in inflammation subscore <sup>b</sup>		-0.14			-0.06
Neuroendocrine subscore in 2006 <sup>b</sup>			0.19		0.24
Change (2006 – 2000) in neuroendocrine subscore <sup>b</sup>			-0.06		-0.19
Other markers subscore in 2006 <sup>b</sup>				0.50***	0.41**



	Model 7a	Model 7b	Model 7c	Model 7d	Model 7e
Change (2006 – 2000) in other markers subscore <sup>b</sup>					
Intercept <sup>c</sup>	-4.71***	-4.76***	-4.76***	-4.53***	-4.26***
AUC	0.824	0.833	0.823	0.824	0.846
Change in AUC (vs. Model 1)	0.007	0.016	0.006	0.007	0.029*

\*  $p < 0.05$ ,

\*\*  $p < 0.01$ ,

\*\*\*  $p < 0.001$

<sup>a</sup>The age slope ( $\gamma$ ) represents the exponential increase in the mortality rate per year of age.

<sup>b</sup>This variable was standardized (Mean=0, SD=1) prior to fitting the model; so, the coefficient represents the effect per SD of the specified variable.

<sup>c</sup>Time was measured in terms of years after age 60. Thus, the intercept represents the mortality rate at age 60.