Highlights

A novel measurement approach and evidence for multi-system physiological dysregulation during aging

Cohen et al.

- We show a novel approach to measuring physiological dysregulation.
- This approach is based on multivariate distance (MD) of a biomarker profile.
- MD is associated with age and mortality in the elderly.
- Multiple physiological systems are implicated in dysregulation.
- New aging-related biomarkers were detected.

A novel statistical approach shows evidence for multi-system physiological dysregulation during aging

Alan A. Cohen[a1], Emmanuel Milot[a], Jian Yong[a], Christopher L. Seplaki[b], Tamas Fulop[c], Karen Bandeen-Roche[d], and Linda P. Fried[e]

[a]Groupe de recherche PRIMUS, Dept. of Family Medicine, University of Sherbrooke, CHUS-Fleurimont, 3001 12[e] Ave N, Sherbrooke, QC J1H 5N4, Canada

[b]Department of Community & Preventive Medicine, University of Rochester School of Medicine and Dentistry, 265 Crittenden Blvd, CU 420644, Rochester, NY 14642, USA

[c]Centre de recherche sur le viellissement, Dept. of Medicine, University of Sherbrooke, CSSS-IUGS, pavillon Argyll, 375 rue Argyll, Sherbrooke, QC J1J 3H5, Canada

[d]Dept. of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Center on Aging and Health, 2024 E. Monument Street Suite 2-700, Baltimore, MD 21287, USA

[e]Mailman School of Public Health, Columbia University, 722 W. 168th Street, R1408, New York, NY 10032, USA

[1] Corresponding author: Alan.Cohen@USherbrooke.ca, 819-346-1110 x12589 (office), 819-580-2655 (cell), 819-864-5424 (fax).

Abstract

Previous studies have identified many biomarkers that are associated with aging and related outcomes, but the relevance of these markers for underlying processes and their relationship to hypothesized systemic dysregulation is not clear. We address this gap by presenting a novel method for measuring dysregulation via the joint distribution of multiple biomarkers and assessing associations of dysregulation with age and mortality. Using longitudinal data from the Women's Health and Aging Study, we selected a 14-marker subset from 63 blood measures: those that diverged from the baseline population mean with age. For the 14 markers and all combinatorial sub-subsets we calculated a multivariate distance called the Mahalanobis distance (MHBD)[2] for all observations, indicating how "strange" each individual's biomarker profile was relative to the baseline population mean. In most models, MHBD correlated positively with age, MHBD increased within individuals over time, and higher MHBD predicted higher risk of subsequent mortality. Predictive power increased as more variables were incorporated into the calculation of MHBD. Biomarkers from multiple systems were implicated. These results support hypotheses of simultaneous dysregulation in multiple systems and confirm the need for longitudinal, multivariate approaches to understanding biomarkers in aging.

Keywords: Dysregulation, biomarker, multivariate, aging, physiology

[2] Abbreviations: Mahalanobis distance (MHBD), Multivariate Distance (MD), Women's Health and Aging Study (WHAS), blood urea nitrogen (BUN), dehydroepiandrosterone-sulfate (DHEA-S), high-density lipoprotein (HDL).

1. Introduction

1.1 General introduction

Epidemiological studies of aging have long used biomarkers both as a potential way to understand underlying processes in aging and as a way to predict patient outcomes (Crimmins et al. 2008; Glei et al. 2007; Martin-Ruiz et al. 2011; Walston 2005; Walston et al. 2006). Most such studies have used single markers, or have used composite indices based on relatively simple accumulation criteria (e.g., number of criteria met). However, the individual markers measured are almost always integrated in complex physiological regulatory networks involved in maintaining organismal homeostasis: this is why they are chosen. In such complex dynamic systems, the interpretation of the level of any one marker generally depends on the levels of many other markers (Cohen et al. 2012).

Many studies now invoke physiological dysregulation (or related concepts such as allostatic load or homeostenosis) as a primary factor in aging (Crimmins et al. 2003; Gruenewald et al. 2009; Karlamangla et al. 2002; McEwen and Wingfield 2003), a hypothesis that appears capable of explaining the multiple mechanistic theories of aging (Medvedev 1990) and the diversity of lifespans and aging patterns observed across species (Finch 1990). Under this scenario, aging is largely a system-level property caused by regulatory breakdown, not any lone biological mechanism such as an up-regulated gene or oxidative stress (Cohen et al. 2012). Multiple studies have shown associations between summary indices of allostatic load and aging outcomes (Crimmins et al. 2003; Glei et al. 2007; Seeman et al. 2001; Szanton et al. 2009). A small number of studies have applied sophisticated statistical approaches to measurement of the relationships among biomarkers in the context of dysregulation, generally with confirmatory but complex results (Arbeev et al. 2011; Gruenewald et al. 2009; Karlamangla et al. 2002; Seplaki et al. 2006; Yashin et al. 2007). However, one of the major

challenges for the field remains the operationalization of multi-system dysregulation and allostatic load (Singer et al. 2004).

Here, we present a novel approach for measuring multi-system physiological dysregulation. This method uses the joint distribution of multiple biomarkers to assign individuals a score indicating how normal or abnormal their overall profile is relative to a reference population. The joint distribution permits incorporation of the likelihood of different combinations of biomarker levels and allows for abnormal biomarker levels to be either high or low. Such a score has many potential uses in studies of aging: 1) Under a hypothesis of dysregulation, scores are expected to increase with age. Moreover, simultaneous dysregulation in multiple systems should result in a stronger association between multivariate distance and age or health outcomes as the number of variables used to compute the distance increases. This can be tested. 2) The contrast between the performance of scores incorporating different number of biomarkers can indicate the complexity of the system and whether the effects of biomarkers are independent from each other. 3) The performance of models incorporating different sets of biomarkers can be used to test hypotheses about which physiological systems are implicated in dysregulation and the degree to which their dysregulation is independent. 4) Long-term, such scores are promising as clinical tools for measuring degree and type of dysregulation in patients. We demonstrate the first three applications, showing apparent dysregulation in biomarkers from multiple physiological systems. Longitudinal data allow us to show both changes within individuals over time and higher risk of mortality among those with abnormal biomarker profiles.

1.2 Specific approach

Multivariate probability distributions can be used to calculate the probability of observing various combinations of variables. Conceptually this is different from using several univariate

distributions together – for example, although it is not particularly rare for a person to have a height of 190 cm or a weight of 45 kg, the combination is quite rare. Statistical distance, when applied to multivariate probability distributions, is a measure of how rare a certain combination of variables is relative to a reference population (the multivariate mean – the mean of all variables – has a statistical distance of 0). Additionally, most physiological systems are regulated to maintain parameters within a certain range, and there is often potential for abnormal values to occur in both directions (e.g. Seplaki et al. 2005). Multivariate statistical distance makes no assumptions about direction of change in different markers, and can integrate changes of all markers in both directions into a single analysis. We refer generally to analyses using multivariate statistical distance as Multivariate Distance (MD) analysis.

Specifically, we use 63 biomarkers from a longitudinal cohort of elderly women, the Women's Health and Aging Study. We use a measure of multivariate statistical distance, the Mahalanobis distance (MHBD) (De Maesschalck et al. 2000; Mahalanobis 1936), to assign each observation of each individual a score for distance from the baseline population mean (assumed to indicate "normal" physiological state) as measured simultaneously for many variables. We hypothesized that a "strange" overall biomarker profile, i.e. one with high MHBD, would be indicative of physiological dysregulation, and that this would be apparent as increases in MHBD with age and associations between MHBD and mortality risk. Further, we predicted that models including more variables in MHBD would yield better predictions. We demonstrate the approach on a subset of biomarkers chosen using statistical criteria; future studies could use biological or alternative statistical criteria.

2. Methods

2.1 Data

We used data from 429 women aged 70-93 in the Women's Health and Aging Study (WHAS) II. This study has been described in detail elsewhere (Fried et al. 1995; Guralnik et al. 1995). Briefly, WHAS is a population-based, prospective study of community-dwelling women with no more than minimal physical disability. Participants were drawn from eastern Baltimore City and Baltimore County, Maryland between March 1, 1994 and May 1, 1995 (Fried et al. 2000). All women gave their informed consent prior to their inclusion in the study. Of 1630 women screened, 880 were eligible, and 430 agreed to participate in WHAS II and had a baseline examination with biomarkers measured. Eligible non-participants were less educated, had lower incomes, and had lower self-rated health compared to WHAS II participants. Follow-ups were conducted roughly 1.5, 3, 6, 7.5, and 9 years later. On average, the time between the first and the last recorded visits for women with at least two visits (390/429) was 8.4 years. Each examination consisted of a comprehensive medical history, medication inventory, physical and neurological examination, neuropsychological battery, and blood draw (Fried et al. 2000). For mortality, 172 participants (40%) died during the course of the study, with data available through July 23, 2010 and a mean follow-up of 10.7 years.

Blood samples were collected at seven time points, but not all collected samples have been analyzed for all parameters, and most have not been analyzed beyond the fourth time point. Sixty-three biomarkers were selected for analysis based on sufficient longitudinal sampling and co-occurrence in samples with other biomarkers of interest. Depending on the specific markers in any given analysis, there are 2-4 effective data collection points – always the minimum number for any included variable. Markers were mostly blood metabolites and included hormones, inflammatory markers, basic blood count measures, micronutrient levels, lipid levels, and ion levels. A full list is provided in Table 1.

2.2 Analysis

2.2.1 Variable groupings

Simultaneous MD analysis of all 63 biomarker variables is not feasible because: (a) the different patterns of age-related changes across markers may cancel each other out in aggregate; and (b) the assumption of multivariate normality, necessary for our methods, becomes untenably constraining at higher dimensions. There are many potential ways to choose subsets of variables for analysis, including methods that use both biological and statistical criteria. Our goal was not to identify the best or only groups of interest, but rather to choose relevant and interesting groups with which to demonstrate MD analysis. In particular, because we were interested in dysregulation, we used statistical criteria to choose variables that either increased or decreased in variability with age.

Change in variability with age was measured using the deviance of each observation for each variable – the absolute value of its difference from the variable's population mean. This approach is similar to that taken previously by Seplaki et al. (2005). In order to approximate as much as possible a younger, healthier population, population mean was calculated based solely on the first visits of patients. Deviances were calculated after appropriate log- or square-root-transformations for normality; this preserves the sensitivity of the deviances to aberrant values in both directions, regardless of the original scale of the variable.

For each variable, we then pooled deviances for all patients at all time points and assessed the Pearson correlation with age. Of the 63 variables included in the analyses, 14 showed significant (p<0.05) correlations with age and 5 showed significant negative correlations. We call these the positive and negative suites, respectively. These suites serve as the basis for all subsequent analyses. Thus, the statistical criterion selected biomarkers showing an increasing or decreasing average deviation from the baseline population with age. This could be due to a simple linear increase or decrease in average biomarker values with age, without a concomitant increase in total variance, or to an increase in total variance (with more extreme

positive and/or negative values), with or without a linear trend in average values (see Figure S1 for examples). The positive suite may represent variables experiencing a loss of regulatory control with age, and the negative suite may represent variables experiencing a loss of capacity to respond to changing conditions with age. We emphasize that this variable screening procedure is neither meant to be exhaustive (it does not detect non-monotonic associations with age, for example) nor highly specific (there may be false positives for either statistical or biological reasons), and is but one example for how variables might be chosen. However, these variable groupings are sufficiently coherent and small to allow a meaningful demonstration of MD analysis.

2.2.2 Mahalanobis distance

Mahalanobis distance (MHBD) is a measure of multivariate statistical distance for a multivariate normal distribution, given by the formula:

$$D_M(x) = \sqrt{(x-\mu)^T S^{-1}(x-\mu)} \quad (1)$$

where x is a multivariate observation (a vector of simultaneously observed values for the variables in question, such as all the biomarker values for a given patient at a given time point), μ is the equivalent-length vector of population means for each variable, and S is the population variance-covariance matrix for the variables. If all variables are uncorrelated then this is equivalent to scaling each biomarker by its variance and then summing the squared deviances for an observation:

$$D_M(x) = \sqrt{\sum_{i=1}^{B} \frac{(x_i - \mu_i)^2}{\sigma^2(x_i)}} \quad (2)$$

where B is the number of biomarkers and $\sigma^2(x_i)$ the variance in the i^{th} biomarker. The behavior of MHBD thus depends not only on the identity of the variables included in the

calculation but also on their covariance. For instance, BUN/creatinine ratio and total cholesterol show opposite trends with age (Figure S1) and are uncorrelated to each other (r=0.024, p=0.45, n=971). MHBD (scaled by its standard deviation, see below) calculated from these two variables has an average of 1.78. When we permuted the values among individuals to obtain the maximal correlation possible given the observed data (r=0.99), MHBD decreased to 1.36, namely 76% of the previous value and closer to MHBD calculated with either variable taken separately (respectively 1.32 and 1.20 for cholesterol and BUN/creatinine ratio).

In this study, we calculated μ and S based on the baseline population (all individuals at their first visit) rather than the full population of all measurement points. This allowed us, as much as possible, to compare current physiological state to a healthy reference population. We used standard normal transformations of the raw biomarkers (log or square-root as necessary, then minus the mean and divided by the standard deviation) in order to give equal weight to all variables in the analysis.

Multivariate normality is generally a strong assumption, and it is particularly so for the case of a complex dynamic system, where the relationships between the variables are expected to follow particular patterns that may not be captured by the assumptions related to standard distributions. Nonetheless, it is a conservative assumption in that, by making it, we are likely to miss many patterns that would be detected if we knew the true distribution. To the extent that the assumption is false, we are likely to decrease the probability of generating significant results, so it is a good starting point. We calculated MHBD for each individual at each time point. This was done separately for the positive suite, the negative suite, and each possible subset of variables within each suite (16,383 and 31 combinations, respectively).

Statistical properties of MHBD depend on the number of variables used to calculate it. The scale depends on the scales and number of the variables included. The lower bound is at zero,

and the distribution is usually roughly log-normal, with a peak density a bit higher than zero. Proportional to the scale of a given MHBD, the peak tends to shift away from zero as more variables are included in the calculation. To account for this distribution, MHBDs were log-transformed when included in correlations and regressions with age, though results were not sensitive to using the raw MHBD (data not shown). MHBDs were not log-transformed in analyses of mortality because we suspected that the risks increased exponentially with MHBD. Because the scale of MHBD changes depending on the variables included, we standardized MHBD by its standard deviation, or when appropriate the log of MHBD by the standard deviation of log-transformed value, for use in comparisons across analyses.

2.2.3 Relationship to age and mortality

For each MHBD calculated, we assessed its correlation with age (Pearson correlation coefficient). Significant correlations could result from either individual or population changes. To measure individual changes, we calculated the slope of MHBD with age for each individual having at least two values of all variables used to calculate the MHBD. We then averaged this slope across individuals, and performed a t-test to see if it was significantly positive or negative.

To analyze the relationship between MHBD and mortality, we used Cox proportional hazards models to assess the association between MHBD and mortality risk. Specifically, we controlled for the age at the first visit and used the absolute time of the follow-up to model a time-to-event process (i.e. until either death or censoring because visits ended).

2.2.4 Effect size and variables included

In order to assess whether inclusion of more variables tended to augment or decrease our ability to detect effects, we compared the effect sizes and $-\log_{10}$ (*p*-values) to the number of

variables included in each MHBD model. Intuitively, we might suspect that inclusion of more variables should augment the predictive value of a model. However, this is not necessarily true for MHBD, where inclusion of a variable less related to age can increase the noise and thus decrease the signal relative to a model with fewer variables (see also Seplaki et al. 2005). Moreover, in our dataset, sample size decreases as more variables are included due to missing data, diminishing the power of the higher-order analyses.

Lastly, in addition to asking how the predictive value of MHBD changes as more variables are included, we assessed the importance of each variable for the predictive power. We used the results of the 16,383 and 31 positive and negative subgroup analyses in univariate regression models to predict how the models' effect sizes (the dependent variable) depended on whether a given biomarker was included in the calculation of MHBD (dichotomous; the independent variable). This was repeated for the correlation, slope, and mortality analyses for all 21 biomarkers in the positive and negative suites.

3. Results

3.1 Variable correlations with age

Many of the 63 variables correlated with age (Table 1). We found 14 that correlated positively with age, 19 that correlated negatively with age, 14 with deviances that correlated positively (the positive suite, Table 1), and 5 with deviances that correlated negatively (the negative suite, Table 1). As expected given the relatively narrow range of ages and the absence of young individuals in our cohort, most of the correlations were relatively weak ($0.04 < |r| < 0.36$ for significant correlations).

3.2 Relationship of Mahalanobis distance to age

We calculated MHBD for each observation for all 16,383 combinations of the 14 variables in the positive suite, as well as for the 31 combinations of the 5 variables in the negative suite. The majority of variable combinations in the positive suite produced MHBDs positively correlated with age (99.4% significant at $p<0.05$, Fig. 1a) and having positive average slopes with age (99.5% signif., Fig. 1b). In both cases, the inclusion of more variables in the calculation of MHBD resulted in larger standardized effect sizes and more significant *p*-values (Figs 1a-b). The majority of variable combinations in the negative suite produced MHBDs negatively correlated with age as expected (80.6% signif., Fig. 2a), but paradoxically having positive average intra-individual slopes with age (64.5% signif., Fig. 2b).

3.3 Relationship of Mahalanobis distance to mortality

Many variable combinations in the positive suite also produced MHBDs predictive of mortality (survival analyses, 67.4% signif., 99.5% of relative risks in the expected direction (>1), Fig. 1c). On average, model performance increased as more variables were included, but the best models had intermediate numbers of variables (Fig. 1c). Effect size in both cases shows a significant negative quadratic relationship with variable number ($p<0.0001$). Although MHBD is log-normally distributed, we did not transform it for the mortality analyses because the risk was hypothesized to increase exponentially with MHBD score. For the negative suite, only one combination of variables significantly predicted mortality.

3.4 Effects of individual variables

We used linear regression to analyze how inclusion or exclusion of each variable from the model affected the strength of the predictions in the models above (Table 2). As expected, some variables contributed more information than others. For example, inclusion of osteocalcin in calculating MHBD tended to strongly increase the correlation, the intra-

individual slope, and mortality prediction. The general trend for higher-order models to perform better depends on which variables are included. For example, sensitivity analyses excluding the models with osteocalcin show a much stronger association between variable number and slope (Figure S2; slope trend, r=0.44 compared to r=0.26; p-value trend, r=-0.09 compared to r=-0.25; p<0.0001 for all).

4. Discussion

4.1 General discussion

Here we have presented a novel analytical approach for studying longitudinal changes in suites of biomarkers with age based on the concept of statistical distance. This distance is a measure of the strangeness or abnormality of an individual's biomarker profile at a moment in time, relative to a reference population mean. Our analyses show that, when calculated based on appropriate sets of variables, this statistical distance increases with age within individuals and predicts subsequent mortality, providing support for the role of dysregulation in aging. Furthermore, predictive power for both age and mortality increase as more variables are included in the calculation of statistical distance.

Concordant with studies on allostatic load and aging (Crimmins et al. 2003; Glei et al. 2007; Gruenewald et al. 2009; Karlamangla et al. 2002; Seeman et al. 2001; Seplaki et al. 2006; Szanton et al. 2009), several lines of evidence in this study support the hypothesis of dysregulation in aging and MHBD as a measure of it. First, the biomarkers selected here are largely not those traditionally considered biomarkers of aging and functional decline (with the exception of cholesterol, see Table 2). This suggests global, multi-systemic dysregulation. Second, MHBD in the positive suite both increases with age and predicts mortality, as predicted for physiological dysregulation.

Third, predictive power tends to increase as more variables are included in the model. Unlike a regression model, this would not be predicted if all variables had independent effects; in that case, the effect of the multivariate distance would approximate the average effect of all the variables used to calculate it. For example, imagine we calculate MHBD based on just one variable that is strongly associated with age and mortality. It performs very well. If we add in a second variable, uncorrelated with the first, that has only a very weak association with age and mortality, this performance should go down because half of the information on statistical distance (that provided by the second variable) is largely unrelated to age and mortality. Thus, an individual could have a high MHBD because of "good" values for several variables, "poor" values for several variables, or any mix of both. On the other hand, under the dysregulation hypothesis it is the joint distribution of variables (not their independent distributions) that signals aging; in this case, the more variables included in the calculation the stronger the expected aging signal. Thus, the increasing predictive power of models with more biomarkers provides substantial support for the hypothesis that it is the interactions between the variables that are critical.

Preliminary analyses (not shown) did not detect associations of statistical distance with age or subsequent mortality when the distance was calculated from all available biomarkers. This highlights the need to choose relevant suites of variables for this type of analysis. In this study, we used statistical selection criteria: namely, a significant correlation between age and the absolute deviation of the variable from the population mean. Even after this, it was necessary to separate the positively and negatively correlated biomarkers. We believe that the positive suite represents biomarkers that increasingly escape regulatory control with age, whereas the negative suite represents markers that increasingly fail to respond properly to changing conditions (either internal or external). Future studies should examine the relevance

of biologically rather than statistically identified suites of variables, such as inflammatory markers, electrolytes, and oxygen transport markers.

Our results support hypotheses of dysregulation in aging but do not provide definitive proof, as we do not show direction of causality. It is possible that other underlying processes cause both aging and dysregulation, or that dysregulation is a result of disease, which increases in frequency with age and preceding mortality. In order to further test for a causal role of dysregulation in aging, future studies should follow patterns of dysregulation over longer periods of time, with various combinations of biomarkers, and in relationship to specific disease processes as they progress. For example, studies should address links between dysregulation and frailty (Fried et al. 2001; Rockwood et al. 2005), with an eye to understanding if frailty is a global physiological process or a result of dysregulation in one or several specific systems. As part of such studies, the multivariate statistical distance approach presented here can be applied to both functional suites of biomarkers (inflammatory markers, hormones, etc.) and to statistically selected suites. In addition, future analyses should combine multivariate distance with other approaches such as principal components analysis, grade-of-membership models, and structural equations modeling. For example, Arbeev et al. (2011) presented detailed dynamic models of biomarkers, allostatic load, and health outcomes during aging; statistical distance could be incorporated directly into such models as a way to summarize multiple markers.

4.2 Limitations

It is tempting to interpret the biological significance of the individual biomarkers identified in this study, as well as their interactions. Indeed, it is noteworthy that the markers come from many different systems. Additionally, most of the markers that emerge here as the best predictors of aging and mortality are not traditionally considered aging biomarkers. Blood

urea nitrogen, basophil counts, ion levels, and hematocrit are (to our knowledge) all novel as aging biomarkers and deserve specific follow-up study. In contrast, creatinine, hemoglobin, and bilirubin have been examined individually in other studies (Ble et al. 2005; Bulpitt et al. 2009; Den Elzen et al. 2011). However, we caution that the effects of any one biomarker may be due to particularities of the WHAS data set or to random chance (given the large number of statistical tests performed), and we prefer to refrain from over-interpreting the specific biomarkers that emerged in this study.

Our analyses were designed largely without consideration of the biological functions of the biomarkers we measured. This was intentional: so little is known about the long-term structure and function of physiological regulatory networks that incorporation of specific *a priori* biological hypotheses would likely have biased our findings and prevented us from detecting novel patterns. Indeed, the biomarkers we found to be the most important are not the ones that we would have predicted *a priori*, such as high-density lipoprotein (HDL) cholesterol, C-reactive protein, and hemoglobin A1C. Our approach is justified as an exploratory method, but the specific effects of individual markers will need to be confirmed in other studies and datasets. It will also be critical to conduct complementary analyses that reflect clear hypotheses based on the literature, and such analyses are currently underway.

The analyses presented here assume a multivariate normal distribution. Multivariate normality, rare in general, is particularly unlikely to apply to these biomarker suites, which are parts of physiological regulatory networks and are thus expected to follow complex dynamics not readily explained by standard statistical distributions. Despite this conservative assumption, we were able to detect strong signals in the data. Future research will explore non-parametric approaches to estimating statistical distance, including kernel density estimation and data depth (Bouezmarni and Rombouts 2010; Liu et al. 1999; Scott 2008). Although we performed a large number of statistical tests, the fact that 67% of all models

were significant at α=0.05 (much higher than the 5% expected by chance) suggests that concerns about false positives and multiple testing are unwarranted.

Lastly, our study is currently limited to a single population of elderly women in one American city, and thus may not be generalizable to other populations. We were not able to follow age changes throughout the life course, and it is possible, even likely, that different patterns would emerge on a similar analysis of women aged 45-70, or of men, for example.

5. Conclusions

Using a novel method for measuring multi-systemic dysregulation during aging, we have shown that abnormal physiological states are associated with increasing age and risk of mortality. The increasing performance of measures including more biomarkers shows that abnormal state results from the interactions among variables, not just their independent effects. Our results agree with previous studies on physiological dysregulation but used a different set of markers. This suggests that dysregulation is a generalized process that can be measured through many different sets of biomarkers. This is supported by the diverse physiological systems represented by the biomarkers in this study. Our results are consistent with, though not proof of, dysregulation itself as a fundamental cause of aging. Our method can be widely applied to similar data sets or other suites of biomarkers, and in the long run such comparisons will help establish the biological pathways involved in aging-associated dysregulation.

6. Acknowledgments

AAC is a member of the FRQ-S-supported *Centre de recherché sur le vieillissement* and *Centre de recherche Étienne Le-Bel*, and is a funded Research Scholar of the FRQ-S. This research was supported by CIHR grant #s 110789,120305, 119485 and by NSERC Discovery

Grant #402079-2011. Dr. Seplaki is supported by Mentored Research Scientist Development Award number K01AG031332 from the National Institute on Aging. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

7. Literature Cited

Arbeev, K. G., S. V. Ukraintseva, I. Akushevich, A. M. Kulminski, L. S. Arbeeva, L. Akushevich, I. V. Culminskaya et al. 2011. Age trajectories of physiological indices in relation to healthy life course. Mechanisms of Ageing and Development 132:93-102.

Ble, A., J. C. Fink, R. C. Woodman, M. A. Klausner, B. G. Windham, J. M. Guralnik, and L. Ferrucci. 2005. Renal Function, Erythropoietin, and Anemia of Older Persons: The InCHIANTI Study. Arch Intern Med 165:2222-2227.

Bouezmarni, T., and J. V. K. Rombouts. 2010. Nonparametric density estimation for multivariate bounded data. Journal of Statistical Planning and Inference 140:139-152.

Bulpitt, C. J., R. L. Antikainen, H. L. Markowe, and M. J. Shipley. 2009. Mortality according to a prior assessment of biological age. Current Aging Science 2:193-199.

Cohen, A. A., L. B. Martin, J. C. Wingfield, S. R. McWilliams, and J. A. Dunne. 2012. Physiological regulatory networks: ecological roles and evolutionary constraints. Trends in Ecology & Evolution 27:428-435.

Crimmins, E., S. Vasunilashorn, J. K. Kim, and D. Alley. 2008. Chapter 5 Biomarkers Related To Aging In Human Populations, Pages 161-216 *in* S. M. Gregory, ed. Advances in Clinical Chemistry, Elsevier.

Crimmins, E. M., M. Johnston, M. Hayward, and T. Seeman. 2003. Age differences in allostatic load: an index of physiological dysregulation. Experimental Gerontology 38:731-734.

De Maesschalck, R., D. Jouan-Rimbaud, and D. L. Massart. 2000. The Mahalanobis distance. Chemometrics and Intelligent Laboratory Systems 50:1-18.

Den Elzen, W. P. J., C. Martin-Ruiz, T. von Zglinicki, R. G. J. Westendorp, T. B. L. Kirkwood, and J. Gussekloo. 2011. Telomere length and anaemia in old age: results from the Newcastle 85-plus Study* and the Leiden 85-plus Study. Age and Ageing 40:494-500.

Finch, C. E. 1990, Longevity, Senescence, and the Genome. Chicago, University of Chicago Press.

Fried, L. P., K. Bandeen-Roche, P. H. Chaves, and B. A. Johnson. 2000. Preclinical mobility disability predicts incident mobility disabilityin older women. Journal of gerontology. Series A, Biological sciences and medical sciences 55:M43-M52.

Fried, L. P., K. D. Kasper, J. M. Guralnik, and E. M. Simonsick. 1995. The Women's Health and Aging Study: an introduction, Pages 1-8 *in* J. M. Guralnik, L. P. Fried, E. M. Simonsick, K. D. Kasper, and M. E. Lafferty, eds. The Women's Health and Aging Study: health and social characteristics of old women with disability. Bethesda, MD, National Institute on Aging.

Fried, L. P., C. M. Tangen, J. Walston, A. B. Newman, C. Hirsch, J. Gottdiener, T. E. Seeman et al. 2001. Frailty in Older Adults: Evidence for a Phenotype Journal of gerontology. Series A, Biological sciences and medical sciences 56:M146-M157.

Glei, D. A., N. Goldman, Y.-L. Chuang, and M. Weinstein. 2007. Do Chronic Stressors Lead to Physiological Dysregulation? Testing the Theory of Allostatic Load. Psychosom Med 69:769-776.

Gruenewald, T. L., T. E. Seeman, A. S. Karlamangla, and C. A. Sarkisian. 2009. Allostatic Load and Frailty in Older Adults. Journal of the American Geriatrics Society 57:1525-1531.

Guralnik, J. M., L. P. Fried, J. D. Kasper, E. M. Simonsick, and M. E. Lafferty. 1995. The Women's Health and Aging Study: health and social characteristics of older women with disability. Washington DC, National Institute on Aging.

Karlamangla, A. S., B. H. Singer, B. S. McEwen, J. W. Rowe, and T. E. Seeman. 2002. Allostatic load as a predictor of functional decline: MacArthur studies of successful aging. Journal of Clinical Epidemiology 55:696-710.

Liu, R. Y., J. M. Parelius, and K. Singh. 1999. Multivariate Analysis by Data Depth: Descriptive Statistics, Graphics and Inference. The Annals of Statistics 27:783-840.

Mahalanobis, P. C. 1936. Mahalanobis distance. Proceedings National Institute of Science of India 49:234-256.

Martin-Ruiz, C., C. Jagger, A. Kingston, J. Collerton, M. Catt, K. Davies, M. Dunn et al. 2011. Assessment of a large panel of candidate biomarkers of ageing in the Newcastle 85+ study. Mechanisms of Ageing and Development 132:496-502.

McEwen, B. S., and J. C. Wingfield. 2003. The concept of allostasis in biology and biomedicine. Hormones and Behavior 43:2-15.

Medvedev, Z. A. 1990. An attempt at a rational classification of theories of ageing. Biological Reviews 65:375-398.

Rockwood, K., X. Song, C. MacKnight, H. Bergman, D. B. Hogan, I. McDowell, and A. Mitnitski. 2005. A global clinical measure of fitness and frailty in elderly people. CMAJ 173:489-495.

Scott, D. W. 2008, Multivariate Density Estimation. New York, John Wiley & Sons, Inc.

Seeman, T. E., B. S. McEwen, J. W. Rowe, and B. H. Singer. 2001. Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. Proceedings of the National Academy of Sciences of the United States of America 98:4770-4775.

Seplaki, C. L., N. Goldman, D. Glei, and M. Weinstein. 2005. A comparative analysis of measurement approaches for physiological dysregulation in an older population. Experimental Gerontology 40:438-449.

Seplaki, C. L., N. Goldman, M. Weinstein, and Y.-H. Lin. 2006. Measurement of Cumulative Physiological Dysregulation in an Older Population. Demography 43:165-183.

Singer, B. H., C. D. Ryff, and T. Seeman. 2004. Operationalizing allostatic load, Pages 113-149 *in* J. Schuli, ed. Allostasis, Homeostasis, and the Costs of Physiological Adaptation. Cambridge, UK, Cambridge University Press.

Szanton, S. L., J. K. Allen, C. L. Seplaki, K. Bandeen-Roche, and L. P. Fried. 2009. Allostatic Load and Frailty in the Women's Health and Aging Studies. Biological Research For Nursing 10:248-256.

Walston, J. 2005. Biological Markers and the Molecular Biology of Frailty, Pages 83-90 *in* J. R. Carey, J.-M. Robine, J. Pierre Michel, and Y. Christen, eds. Longevity and Frailty. Research and Perspectives in Longevity, Springer Berlin Heidelberg.

Walston, J., E. C. Hadley, L. Ferrucci, J. M. Guralnik, A. B. Newman, S. A. Studenski, W. B. Ershler et al. 2006. Research Agenda for Frailty in Older Adults: Toward a Better Understanding of Physiology and Etiology: Summary from the American Geriatrics Society/National Institute on Aging Research Conference on Frailty in Older Adults. Journal of the American Geriatrics Society 54:991-1001.

Yashin, A. I., K. G. Arbeev, I. Akushevich, A. Kulminski, L. Akushevich, and S. V. Ukraintseva. 2007. Stochastic model for analysis of longitudinal data on aging and mortality. Mathematical Biosciences 208:538-551.

Table 1: Biomarkers analyzed and their correlations with age

	n	Raw variable		Deviance from baseline mean		Category/system
		r	*p*	*r*	*p*	
The positive suite						
Red blood cell count (r)	1967	-0.03	0.24	**0.09**	**<0.0001**	Blood, circulation
Hemoglobin (r)	1967	**-0.09**	**0.0001**	**0.07**	**0.002**	Blood, circulation
Hematocrit (r)	1967	-0.02	0.28	**0.06**	**0.006**	Blood, circulation
Osteocalcin (l)	999	**0.34**	**<0.0001**	**0.29**	**<0.0001**	Bone metabolism
Calcium (r)	1990	**0.05**	**0.02**	**0.09**	**<0.0001**	Ions
Sodium (r)	1990	-0.02	0.32	**0.14**	**<0.0001**	Ions
Potassium (r)	1990	**0.07**	**0.004**	**0.07**	**0.003**	Ions
Chloride (r)	1990	**-0.12**	**<0.0001**	**0.04**	**0.05**	Ions
Total cholesterol (r)	1287	**-0.17**	**<0.0001**	**0.10**	**0.0005**	Lipids
BUN-creatinine ratio (r)	1991	**0.36**	**<0.0001**	**0.25**	**<0.0001**	Liver, kidney, proteins, excretion
Creatinine (l)	1991	**-0.13**	**<0.0001**	**0.22**	**<0.0001**	Liver, kidney, proteins, excretion
Albumin (r)	1990	**-0.16**	**<0.0001**	**0.06**	**0.005**	Liver, kidney, proteins, excretion
Bilrubin (direct) (l)	1083	**0.08**	**0.007**	**0.07**	**0.03**	Liver, kidney, proteins, excretion
Basophils (r)	972	-0.03	0.34	**0.09**	**0.005**	White blood cell counts
The negative suite						
Alanine transaminase (l)	1990	**0.08**	**0.0003**	**-0.07**	**0.001**	Liver, kidney, proteins, excretion
Alkaline 22aemoglobin (l)	1989	-0.02	0.45	**-0.06**	**0.01**	Liver, kidney, proteins, excretion
Vitamin D hydroxyl (r)	805	-0.04	0.25	**-0.12**	**0.0009**	Micronutrients
Vitamin D dihydroxyl (r)	1070	-0.02	0.49	**-0.07**	**0.02**	Micronutrients
Ferritin (l)	1085	-0.01	0.75	**-0.07**	**0.02**	Oxygen transport
Other biomarkers						
Mean corpuscular volume (r)	1967	0.01	0.68	-0.03	0.17	Blood, circulation
Red cell distribution width (l)	1966	0.00	0.86	-0.03	0.17	Blood, circulation
Platelet count (r)	1957	0.03	0.16	0.02	0.40	Blood, circulation

Mean corpuscular 23aemoglobin conc. (r)	1967	**-0.16**	**<0.0001**	-0.02	0.44	Blood, circulation
Mean corpuscular haemoglobin (r)	1967	**-0.08**	**0.0004**	0.01	0.64	Blood, circulation
Mean platelet volume (r)	875	**-0.07**	**0.05**	-0.02	0.64	Blood, circulation
Estradiol (l)	1286	**-0.19**	**<0.0001**	0.05	0.09	Endocrine
Thyroxine (r)	1093	-0.04	0.18	-0.05	0.11	Endocrine
IGF-1 (r)	1075	0.01	0.84	-0.05	0.13	Endocrine
Thyroid stimulating hormone (s)	1281	**0.16**	**<0.0001**	0.03	0.23	Endocrine
Parathyroid hormone (intact) (l)	984	0.02	0.54	0.02	0.47	Endocrine
DHEA-S (l)	1450	**-0.09**	**0.0005**	0.01	0.57	Endocrine
Parathyroid hormone (serum mid-region) (l)	994	0.02	0.50	0.02	0.59	Endocrine
Interleukin-6 (l)	1094	-0.02	0.57	0.03	0.29	Inflammation
C-reactive protein (l)	1078	**-0.08**	**0.01**	-0.02	0.42	Inflammation
Magnesium (r)	1050	0.05	0.12	0.05	0.11	Ions
Phosphate (r)	1084	0.01	0.75	0.00	0.87	Ions
Triglycerides (l)	1287	**-0.10**	**0.0006**	0.02	0.51	Lipids
Cholesterol-HDL ratio (l)	1084	**-0.09**	**0.004**	0.01	0.73	Lipids
HDL (r)	1291	0.03	0.33	0.00	0.95	Lipids
Aspartate aminotransferase (l)	1989	**0.20**	**<0.0001**	0.04	0.08	Liver, kidney, proteins, excretion
G-glutamyl transferase (l)	1084	0.01	0.71	-0.03	0.28	Liver, kidney, proteins, excretion
Lactate dehydrogenase (r)	1082	-0.03	0.31	0.03	0.34	Liver, kidney, proteins, excretion
Uric acid (r)	1084	**-0.07**	**0.02**	0.03	0.40	Liver, kidney, proteins, excretion
Total Protein (r)	1990	**-0.09**	**0.0001**	0.02	0.47	Liver, kidney, proteins, excretion
Globulin (r)	1084	**-0.07**	**0.02**	-0.02	0.48	Liver, kidney, proteins, excretion
Albumin-globulin ratio (r)	1084	0.01	0.66	-0.02	0.53	Liver, kidney, proteins, excretion
Bilirubin (total) (l)	1990	-0.03	0.19	0.01	0.64	Liver, kidney, proteins, excretion
Folate (l)	1085	**0.18**	**<0.0001**	-0.05	0.11	Micronutrients
Vitamin B12 (l)	1083	0.01	0.76	-0.04	0.20	Micronutrients
Vitamin A (r)	767	-0.04	0.24	0.03	0.36	Micronutrients

Vitamin B6 (l)	594	**0.11**	**0.005**	-0.03	0.53	Micronutrients
Total iron-binding capacity (r)	1083	0.01	0.63	-0.03	0.40	Oxygen transport
Saturated transferring (r)	1086	0.02	0.61	0.01	0.70	Oxygen transport
Iron (r)	1083	0.02	0.45	0.01	0.87	Oxygen transport
Glycohemoglobin (l)	912	0.01	0.69	-0.06	0.08	Sugar metabolism
Protein-bound glucose (l)	1083	**0.07**	**0.02**	-0.04	0.17	Sugar metabolism
Glucose (l)	1989	**-0.15**	**<0.0001**	0.03	0.20	Sugar metabolism
Hemoglobin a1c (l)	1059	**-0.08**	**0.01**	-0.004	0.91	Sugar metabolism
White blood cell count (r)	1967	**0.10**	**<0.0001**	-0.04	0.06	White blood cell counts
Monocytes (r)	973	0.00	0.93	0.05	0.14	White blood cell counts
Eosinophils (l)	973	**0.07**	**0.03**	-0.02	0.49	White blood cell counts
Differential poly (r)	972	**0.12**	**0.0001**	0.00	0.91	White blood cell counts
Lymphocytes (r)	973	**-0.15**	**<0.0001**	0.00	0.96	White blood cell counts

Results where $p<0.05$ are in **bold**; results where $p<0.05$ and the correlation is negative are in **red**. All variables were transformed as necessary for normality, coded (r) = raw, (l) = log, and (s) = square-root. All correlations were performed on data pooled across individuals and time points. DHEA-S indicates dehydroepiandrosterone-sulfate; HDL indicates high-density lipoprotein.

Table 2: Effect of inclusion of each raw variable on the predictivity of Mahalanobis distance for different models

	Raw correlation		Intra-individual slope		Mortality relative risk	
	β	*p*	β	*p*	β	*p*
The positive suite						
Red blood cell count	0.006	<0.0001	-0.0065	0.002	-0.0468	<0.0001
Hemoglobin	-0.0076	<0.0001	-0.0167	<0.0001	-0.0059	<0.0001
Hematocrit	-0.0009	0.5065	-0.0177	<0.0001	-0.022	<0.0001
Osteocalcin	0.1447	<0.0001	0.2638	<0.0001	0.0919	<0.0001
Calcium	0.0178	<0.0001	-0.0019	0.3578	0.0315	<0.0001
Sodium	0.005	0.0001	0.0091	<0.0001	-0.0014	0.2149
Potassium	-0.0048	0.0003	0.0046	0.0275	0.0255	<0.0001
Chloride	0.003	0.0231	-0.0017	0.4274	-0.0002	0.8727
Total cholesterol	-0.0112	<0.0001	-0.0021	0.3119	0.0205	<0.0001
BUN-creatinine ratio	0.0068	<0.0001	0.0103	<0.0001	-0.0239	<0.0001
Creatinine	-0.0017	0.2066	0.01	<0.0001	0.0363	<0.0001
Albumin	-0.0011	0.3905	-0.0023	0.2672	-0.0022	0.052
Bilrubin (direct)	-0.0272	<0.0001	0.0127	<0.0001	0.0364	<0.0001
Basophils	-0.0062	<0.0001	0.0042	0.0482	0.024	<0.0001
The negative suite						
Vitamin D hydroxyl	-0.013	0.098	-0.004	0.804	0.056	0.009
Alkaline phosphatase	0.001	0.889	0.048	0.001	0.048	0.026
Vitamin D dihydroxyl	-0.016	0.041	0.062	<0.0001	-0.070	0.001
Alanine transaminase	-0.007	0.389	0.003	0.865	-0.008	0.724
Ferritin	-0.026	<0.0001	-0.001	0.932	0.051	0.018

Results of regression models predicting the effect sizes of other models based on the inclusion or exclusion of each biomarker in the calculation of MHBD. For example, the β-coefficient of 0.006 for the raw correlation of MHBD with red blood cell count indicates that the *r* for the correlation between MHBD and age was 0.006 higher on average if red blood cell count was one of the variables used to calculate MHBD. The second set of columns refers to the within-individual changes in MHBD over time, and the third to the relative risk of mortality based on Cox regression. The positive suite analyses are based on the 16,383 combinations of the 14 variables includes, and the negative suite analyses on the 31 combinations of the 5 variables included.

Fig. 1. Changes in predictive power of MHBD with increasing numbers of variables from the positive suite used in its calculation. Each circle represents an analysis based on one of the 16,383 combinations of the 14 variables in the positive suite. Color indicates p-value: black: $p \geq 0.1$; blue: $0.05 \leq p < 0.1$; cyan: $0.01 \leq p < 0.05$; yellow- green: $0.001 \leq p < 0.01$; orange: $0.0001 \leq p < 0.001$; red: $p < 0.0001$. The line represents a linear regression of number of variables on relevant effect size. Effect size trend shows the results of a Pearson correlation analysis of variable number with relevant effect size, and P-value trend shows the results of a Pearson correlation analysis of variable number with $\log_{10}$(p-value). (a) Correlation of MHBD with age; (b) mean intra-individual slope of MHBD with age; (c) relative risk based on Cox proportional hazards of mortality before next visit. In (b) – (c), effect sizes are standardized as indicated in text.

Fig. 2. Changes in predictive power of MHBD with increasing numbers of variables from the negative suite used in its calculation. Each circle represents an analysis based on one of the 31 combinations of the 5 variables in the negative suite. Color indicates p-value: black: $p \geq 0.1$; blue: $0.05 \leq p < 0.1$; cyan: $0.01 \leq p < 0.05$; yellow- green: $0.001 \leq p < 0.01$; orange: $0.0001 \leq p < 0.001$; red: $p < 0.0001$. The line represents a linear regression of number of variables on relevant effect size. Effect size trend shows the results of a Pearson correlation analysis of variable number with relevant effect size, and P-value trend shows the results of a Pearson correlation analysis of variable number with $\log_{10}$(p-value). (a) Correlation of MHBD with age; (b) mean intra-individual slope of MHBD with age; (c) relative risk based on Cox proportional hazards of mortality before next visit. In (b) – (c), effect sizes are standardized as indicated in text.

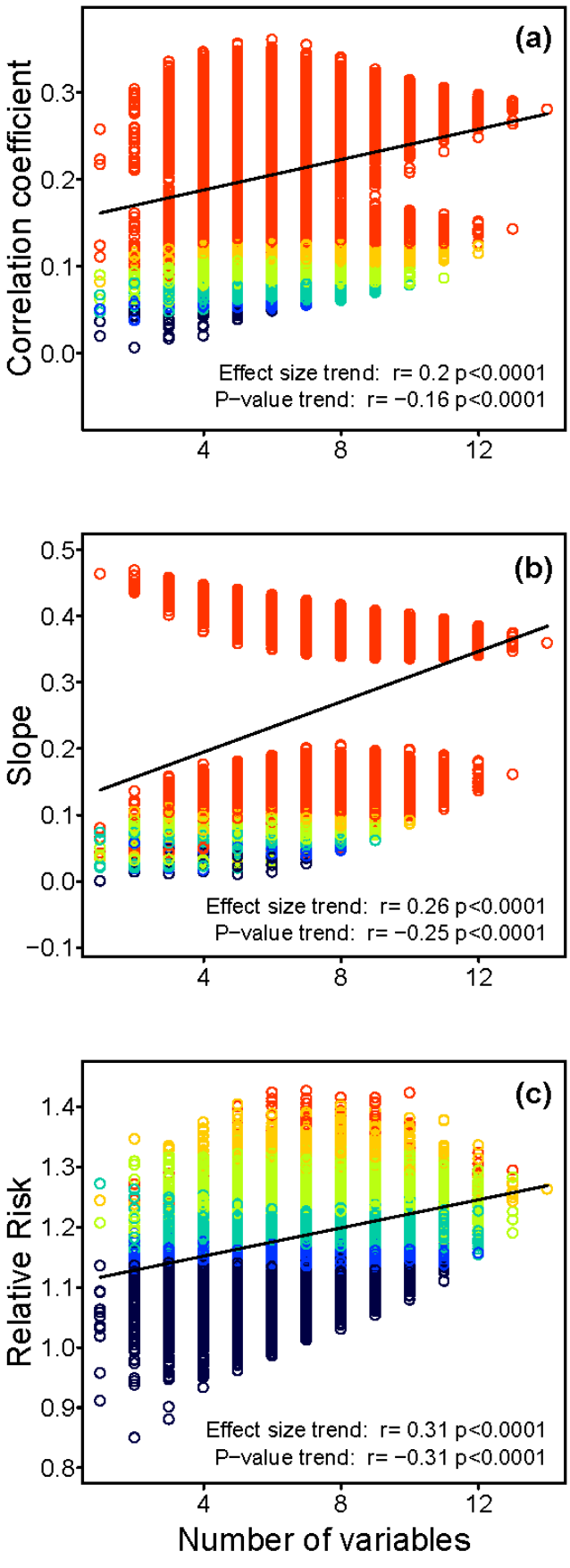
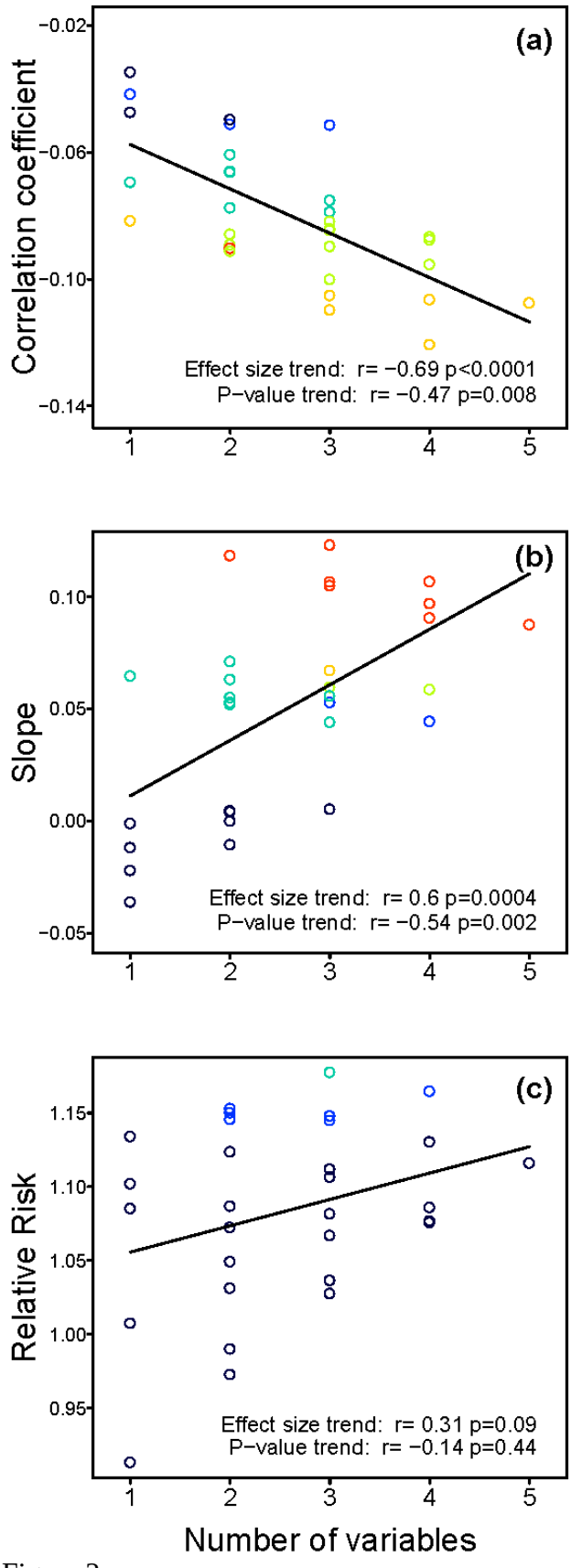

(a)
Correlation coefficient
0.0
0.1
0.2
0.3
Effect size trend: r= 0.2 p<0.0001
P−value trend: r= −0.16 p<0.0001
4
8
12
(b)
Slope
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
Effect size trend: r= 0.26 p<0.0001
P−value trend: r= −0.25 p<0.0001
4
8
12
(c)
Relative Risk
0.8
0.9
1.0
1.1
1.2
1.3
1.4
Effect size trend: r= 0.31 p<0.0001
P−value trend: r= −0.31 p<0.0001
4
8
12
Number of variables

Figure 1


(a)
Correlation coefficient
−0.02
−0.06
−0.10
−0.14
Effect size trend: r= −0.69 p<0.0001
P−value trend: r= −0.47 p=0.008
1
2
3
4
5
(b)
Slope
0.10
0.05
0.00
−0.05
Effect size trend: r= 0.6 p=0.0004
P−value trend: r= −0.54 p=0.002
(c)
Relative Risk
1.15
1.10
1.05
1.00
0.95
Effect size trend: r= 0.31 p=0.09
P−value trend: r= −0.14 p=0.44
Number of variables

40 Figure 2

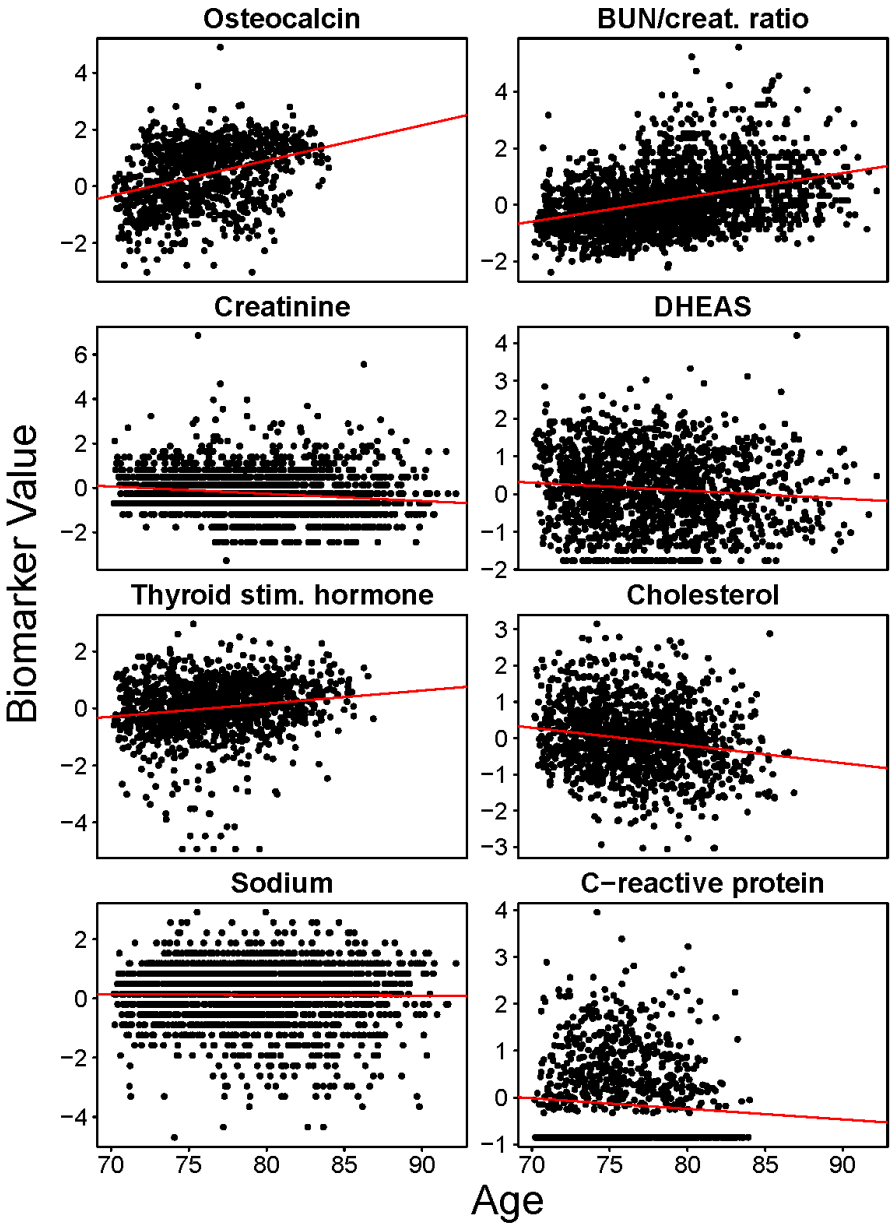


Figure S1 Trends in eight selected biomarkers with age. To be retained by our statistical selection criterion (i.e. a significant correlation between deviation from the basal "healthy" population and age) a marker could show different types of patterns (see Methods). For instance, osteocalcin increases with age but not its variance. As a result, both the raw measurement and its deviance from the basal population are positively correlated with age and there is no increase in extreme negative values. A mirror situation is observed for creatinine where the measurement is correlated negatively with age and its deviance positively. The positive correlation for the thyroid stimulating hormone is not mainly due to a steady linear increase with age but rather to the occurrence of more extreme negative values at intermediate ages. Thus, the deviance is not itself correlated with age. Sodium shows a correlation only in the deviance, with older women tending to show a greater variance (hence more frequent extreme values). The BUN-creatinine ratio shows an increase with age of both the raw measurement and the deviance. DHEAS shows a slight albeit significant decrease with age but not its deviance (no more or less extreme values at higher age, whatever positive, negative or both). Cholesterol is the mirror of osteocalcin, with the raw measurement correlated negatively and its deviance positively with age. Finally, for CRP both the measurement and the deviance show a negative correlation with age although in the latter case it is very weak and does not cross the significance level. Since deviance is always is relative to the variation observed at younger ages, i.e. at the first visit, all these markers except TSH, DHEAS and CRP were retained by our statistical criterion.

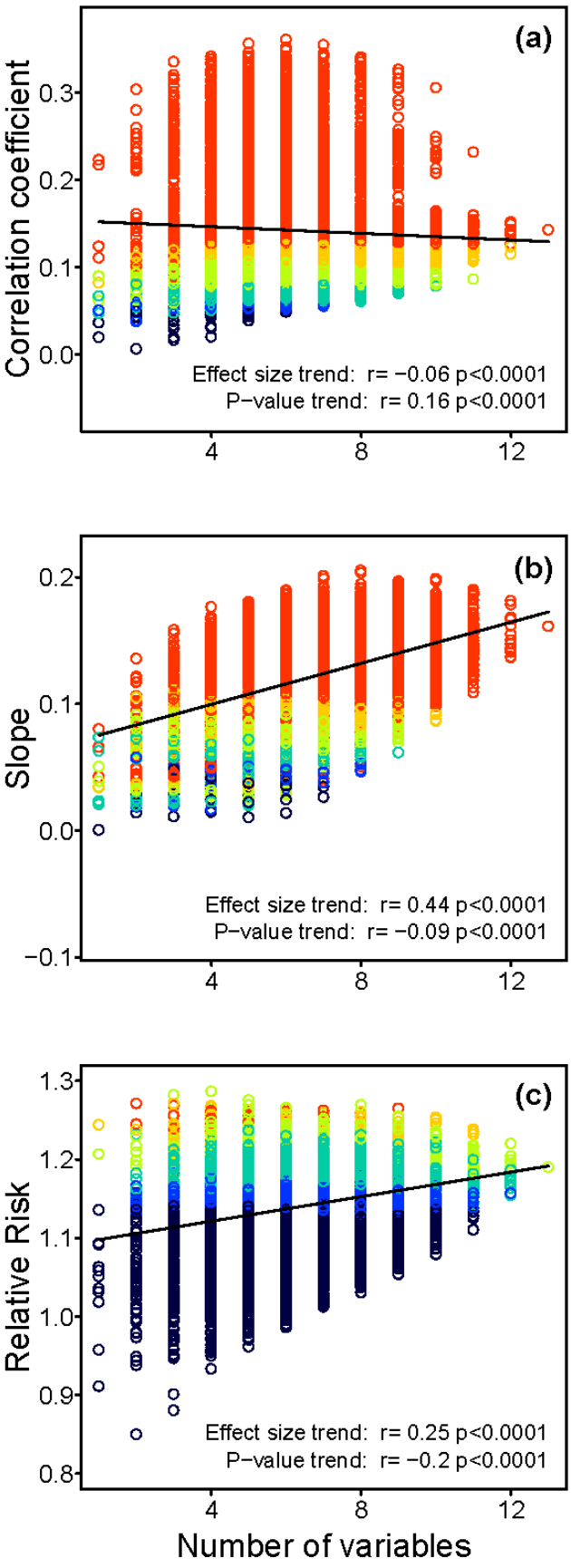


Fig. S2. Figure 1 omitting Osteocalcin. Changes in predictive power of MHBD with increasing numbers of variables from the positive suite used in its calculation. Each circle represents an analysis based on one of the 16,383 combinations of the 14 variables in the positive suite. Color indicates p-value: black: $p \geq 0.1$; blue: $0.05 \leq p < 0.1$; cyan: $0.01 \leq p < 0.05$; yellow-green: $0.001 \leq p < 0.01$; orange: $0.0001 \leq p < 0.001$; red: $p < 0.0001$. The line represents a linear regression of number of variables on relevant effect size. Effect size trend shows the results of a Pearson correlation analysis of variable number with relevant effect size, and P-value trend shows the results of a Pearson correlation analysis of variable number with $\log_{10}$(p-value). (a) Correlation of MHBD with age; (b) mean intra-individual slope of MHBD with age; (c) relative risk based on Cox proportional hazards of mortality before next visit. In (b) – (c), effect sizes are standardized as indicated in text.

Figure 1

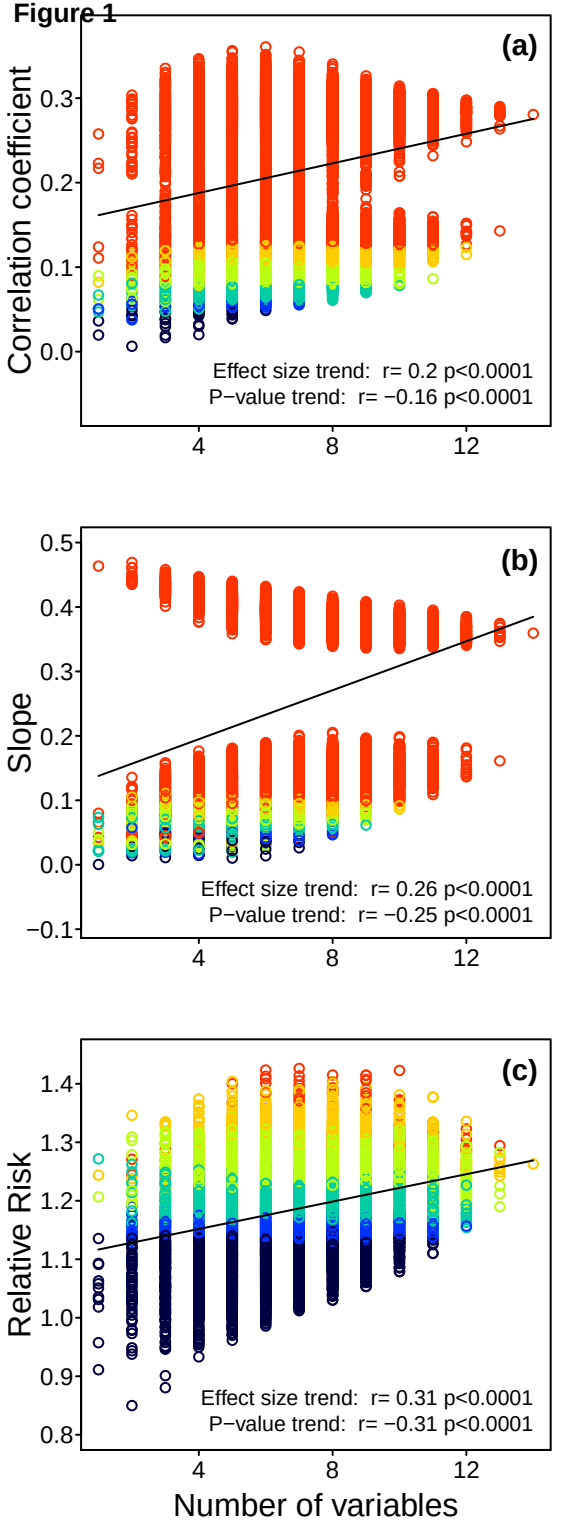
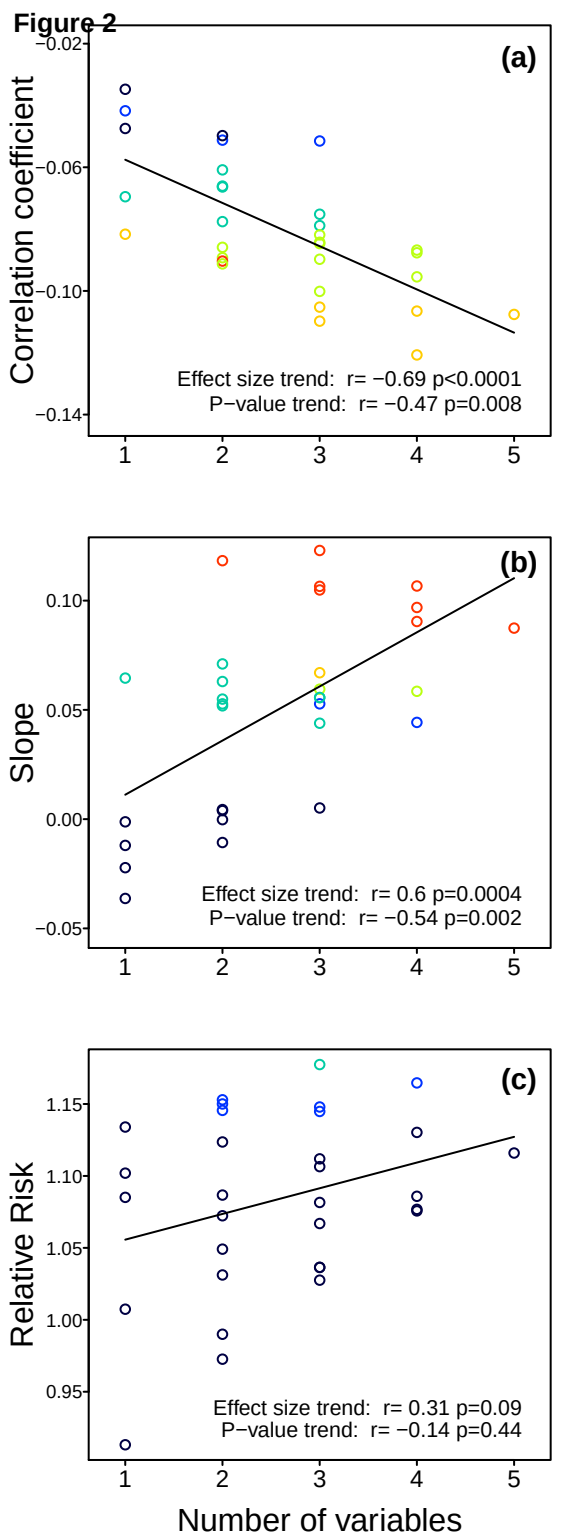

(a)
Correlation coefficient
0.0
0.1
0.2
0.3
Effect size trend: r= 0.2 p<0.0001
P−value trend: r= −0.16 p<0.0001
4
8
12
(b)
Slope
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
Effect size trend: r= 0.26 p<0.0001
P−value trend: r= −0.25 p<0.0001
4
8
12
(c)
Relative Risk
0.8
0.9
1.0
1.1
1.2
1.3
1.4
Effect size trend: r= 0.31 p<0.0001
P−value trend: r= −0.31 p<0.0001
4
8
12
Number of variables

Figure 2


(a)
Correlation coefficient
−0.02
−0.06
−0.10
−0.14
Effect size trend: r= −0.69 p<0.0001
P−value trend: r= −0.47 p=0.008
1
2
3
4
5
(b)
Slope
0.10
0.05
0.00
−0.05
Effect size trend: r= 0.6 p=0.0004
P−value trend: r= −0.54 p=0.002
1
2
3
4
5
(c)
Relative Risk
1.15
1.10
1.05
1.00
0.95
Effect size trend: r= 0.31 p=0.09
P−value trend: r= −0.14 p=0.44
1
2
3
4
5
Number of variables