Published in final edited form as:

Exp Gerontol. 2014 September ; 57: 203–210. doi:10.1016/j.exger.2014.04.016.

Cross-population validation of statistical distance as a measure of physiological dysregulation during aging

Alan A. Cohen^{a,*}, Emmanuel Milot^a, Qing Li^a, Véronique Legault^a, Linda P. Fried^b, and Luigi Ferrucci^c

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

^bMailman School of Public Health, Columbia University, 722 W. 168th Street, R1408, New York, NY 10032, USA

^cTranslational Gerontology Branch, Longitudinal Studies Section, National Institute on Aging, National Institutes of Health, MedStar Harbor Hospital, 3001 S. Hanover Street, Baltimore, MD 21225

Abstract

Measuring physiological dysregulation during aging could be a key tool both to understand underlying aging mechanisms and to predict clinical outcomes in patients. However, most existing indices are either circular or hard to interpret biologically. Recently, we showed that statistical distance of 14 common blood biomarkers (a measure of how strange an individual's biomarker profile is) was associated with age and mortality in the WHAS II data set, validating its use as a measure of physiological dysregulation. Here, we extend the analyses to other data sets (WHAS I and InCHIANTI) to assess the stability of the measure across populations. We found that the statistical criteria used to determine the original 14 biomarkers produced diverging results across populations; in other words, had we started with a different data set, we would have chosen a different set of markers. Nonetheless, the same 14 markers (or the subset of 12 available for InCHIANTI) produced highly similar predictions of age and mortality. We include analyses of all combinatorial subsets of the markers and show that results do not depend much on biomarker choice or data set, but that more markers produces a stronger signal. We conclude that statistical distance as a measure of physiological dysregulation is stable across populations in Europe and North America.

Keywords

Physiological dysregulation; aging; WHAS; InCHIANTI; biomarker; Mahalanobis distance

1. Introduction

Many researchers now believe that physiological dysregulation (or related processes such as allostatic load and homeostenosis) are key players in the aging process, either as causal

*Corresponding author: Alan.Cohen@USherbrooke.ca, (819) 346-1110 x12589, (819) 820-6419 (fax).

drivers of aging or as accompanying processes that nonetheless produce important health consequences (Fried and others 2005; Karlamangla and others 2002; McEwen and Wingfield 2003; Seplaki and others 2005; Taffett 2003). These ideas propose that the complex regulatory networks that maintain homeostasis are not infinitely robust, and that over time the network state might be perturbed in a way that prevents it from returning fully to a baseline state (Cohen and others 2012). For instance, chronic stress results in elevated baseline levels of cortisol, with numerous downstream consequences for health; the failure of cortisol to return to baseline after a stress is an example of dysregulation (Miller and others 2007; Sapolsky and others 2002).

A number of indices have been proposed to measure allostatic load, mostly for sociological or epidemiological studies of population health (Crimmins and others 2003; Karlamangla and others 2002; Seplaki and others 2005; Singer and others 2004; Yashin and others 2007). Most of these indices are highly predictive of a variety of poor health outcomes (Goymann and Wingfield 2004; Gruenewald and others 2009; Schnorpfeil and others 2003). However, the indices are generally composed of a number of biomarkers or criteria already known to indicate poor health; it is thus unsurprising that individuals doing poorly on multiple such measures have poor health outcomes. Accordingly the measures may be useful as a summary of health state, but they do not validate the underlying hypothesis that dysregulation is an important part of aging (Singer and others 2004).

Recently, we proposed a novel way to measure physiological dysregulation based on clinical biomarkers (Cohen and others 2013). Under the hypothesis that a well-functioning, homeostatic physiology should be relatively similar across individuals, but that there are many ways in which physiology might become dysregulated, we proposed statistical distance (specifically Mahalanobis distance, D_M (De Maesschalck and others 2000; Mahalanobis 1936)) as a measure of physiological dysregulation. D_M (applied to biomarkers) is a measure of how strange an individual's profile is relative to everyone else in the population, and greater distance should thus measure greater dysregulation. We applied this measure to a set of 14 biomarkers chosen from the Women's Health and Aging Study (WHAS) II data set based on their increase in deviance from the mean with age (but not necessarily changes in the mean with age). We showed that D_M increased with age within individuals and predicted mortality controlling for age. Additionally, using the combinatorial subsets of the markers, we showed that results were relatively insensitive to marker choice, but that predictive power increased with inclusion of more markers in the calculation of D_M . It thus appeared that D_M is a measure of physiological dysregulation.

However, a number of further validation steps are necessary before D_M can be used widely. We need to establish sensitivity to marker choice – including a wider array of markers – and to the choice of the reference population used to define a “normal” biomarker profile. Also, predictive power for relevant health outcomes needs to be assessed. Here, we tackle the question of reproducibility across populations/data sets, asking whether similar results for the same 14 markers can be obtained in other data sets. We use WHAS I (the complement study to WHAS II, including a less health segment of the population in Baltimore, Maryland, USA) and *Invecchiare in Chianti* (InCHIANTI), a population-based cohort study conducted in Tuscany, Italy.

2. Material and methods

2.1 Data

The Women's Health and Aging Study (WHAS) is a population-based prospective study of community-dwelling women. Originally, WHAS was two separate studies, WHAS I including 1002 women aged 65+ among the 1/3 most disabled in the population (Fried and others 1995), and WHAS II including 436 women aged 70–79 among the 2/3 least disabled (Fried and others 2000). The participants were drawn from eastern Baltimore City and Baltimore County, Maryland. Baseline assessment occurred from November 1992 to February 1995 in WHAS I and from August 1994 to February 1996 in WHAS II. Eligible non-participants were less educated, had lower incomes, and had lower self-rated health compared to WHAS participants. Follow-ups were conducted roughly 1.5, 3, 6, 7.5, and 9 years later. Each examination consisted of a comprehensive medical history, medication inventory, physical and neurological examination, neuropsychological battery, and blood draw (Fried and others 2000). Here, we merge participants from WHAS I and WHAS II into a single data set, WHAS, for comparison with InCHIANTI.

Invecchiare in Chianti (InCHIANTI) is a prospective population-based study of 1156 adults aged 65–102 and 299 aged 20–64 randomly selected from two towns in Tuscany, Italy using multistage stratified sampling in 1998 (Ferrucci and others 2000). Follow-up blood and urine samples were taken in 2001–03, 2005–06, and 2007–08. Because InCHIANTI contains both men and younger individuals, we replicate InCHIANTI analyses on the subset of women aged 70+ in order to have a population comparable to WHAS II in our previous study. Characteristics of both study populations are available in Table 1.

2.2 Biomarker choice

In our previous study (Cohen and others 2013), 14 biomarkers were chosen from among 63 candidate markers based on a positive correlation of their deviances with age (the deviance is the absolute value of the marker level minus the population mean). We intentionally did not impose any *a priori* biological hypotheses on biomarker selection because we felt there was insufficient knowledge of underlying network dynamics to do this accurately, and a very different set of markers could have been chosen. Here, we use the same markers: red blood cell count, hemoglobin, hematocrit, sodium, chloride, potassium, calcium, cholesterol, albumin, creatinine, BUN:creatinine ration, basophil count, osteocalcin, and direct bilirubin, although the last two were not measured in InCHIANTI and are thus excluded from those analyses. We also compared whether we would have chosen the same set of markers had we applied the same criteria to the data sets used here, calculating the correlation of each biomarker with age, and of its deviance with age, in each population.

2.3 Statistical analyses

All statistical analyses were conducted in R v3.0.1 and code is available upon request. All variables were log- or square-root-transformed as necessary to approach normality, and then standardized by subtracting their mean and dividing by their standard deviation. D_M was calculated using the following 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. The parameters μ and S were estimated for each data set from the first visit of each individual, both to assure independence of observations and to use a slightly younger, healthier reference population.

D_M is approximately log-normally distributed, and was thus transformed before analysis of changes with age (Fig. 1). Individual changes in log- D_M with age were modeled using linear regression models for each individual to estimate a slope for each individual; weighted t-tests were then used to assess whether the slope was significantly different from zero, weighted by the number of observations per individual. The mean slope per population was thus a measure of rate of change of log- D_M with age. The relationship between D_M and subsequent mortality was modeled using Cox proportional hazards models (coxph function, survival package) using a counting process with age as time scale. Here, we used untransformed D_M because we hypothesized that exponentially increasing D_M would lead to exponentially increasing risk; results are not qualitatively different using log-transformed D_M (data not shown).

All analyses were repeated for each combinatorial subset of the variables (16,383 combinations for the 14 variables in WHAS and 4095 combinations for the 12 variables in InCHIANTI). Meta-regression models were used to assess the impact of the number of variables and which variables were included on model results. Results were also merged across data sets by biomarker combination to assess correlations of results across data sets for the same biomarker combinations. All analyses were repeated for WHAS I, WHAS II, InCHIANTI, WHAS I and II combined, and the subset of InCHIANTI that is women aged 70+. The latter was chosen to have a population comparable with the original WHAS II study.

3. Results

Correlations of deviances with age were markedly heterogeneous across data sets (Fig. 2). There was very little correspondence across data sets as to which variables would have been retained for use in D_M (i.e., those with significant positive deviance correlations with age, shaded blue in Fig. 2), with only one of the original 12 shared across WHAS I, WHAS II, and InCHIANTI. Restricting InCHIANTI to women aged 70+ so that its composition resembled WHAS II did not improve the correspondence. Raw correlations with age were more often significant than deviance correlations, and showed somewhat greater (but not much) consistency across data sets.

For all data sets, most combinations of biomarkers produced D_M s that increased with age (positive individual slope) and that positively predicted mortality (Figs 3–4). For change with age, 82% of analyses were significant at $\alpha=0.05$ in WHAS and 99% in InCHIANTI; for

mortality, 65% were significant in WHAS and 97% in InCHIANTI. Mortality hazard ratios were consistently positive even when not significant (99.9% of models in WHAS and 100% in InCHIANTI) and had generally large effect sizes (median hazard ratio per unit D_M of 1.26 for WHAS and 1.24 for InCHIANTI). Hence the lower levels of significant results for mortality are likely due to less statistical power as a result of the relatively limited number of deaths in the data sets (up to 122 in WHAS and 193 in InCHIANTI, depending on the biomarker combination and missingness). The average 95% ranges of D_M (i.e., the ranges of values between the 2.5th and 97.5th percentiles; Fig. 1) were 3.18 for InCHIANTI and 3.38 for WHAS; accordingly, we can interpret these hazard ratios as $1.26^{3.38}=2.18$ and $1.24^{3.18}=1.98$ for relative extremes of the population. These effects were slightly but not significantly larger without control for age (Table 2), indicating that most of the predictive power for mortality is age-independent.

In all cases but one there was a significant tendency to have stronger predictions with more variables included in the calculation of D_M (Figs 3–4). In general this effect was quite large, with age slopes generally about twice as large when D_M was calculated with the maximum 14 or 12 variables, compared to just one variable. Hazard ratios were often 50% larger with maximum number of variables, except in WHAS II, where the effect was negligible (Fig. 3f).

We also tested whether results were correlated between WHAS and InCHIANTI for the same variable combination. For slope with age, the correlation was quite strong ($r=0.74$, $p<0.0001$), but this was mostly due to the inclusion or exclusion of one variable, basophil count. Stratifying by basophil count, the correlation was much weaker ($r=0.22$ with basophils and $r=0.26$ without, $p<0.0001$ for both). This was similar to the correlation for hazard ratios ($r=0.28$, $p<0.0001$). These correlations are surprisingly weak: the performance of models in one data set explains only 5–8% of the variance in performance in the other (calculated as the squares of the pairwise correlation coefficients).

Results were also quite heterogeneous for the effects of including or excluding each biomarker in the calculation of D_M (Table 3). Most of the effects (87%) were significant at $\alpha=0.05$, but in all but eight of the 28 cases (14 variables $\times$ 2 outcomes) these significant effects went in opposing directions depending on the data set or subset. The eight cases were as follows: including creatinine, BUN:creatinine ratio, and osteocalcin significantly increased the slope with age, and including bilirubin, sodium, cholesterol, and albumin significantly increased hazard ratios whereas including calcium significantly decreased hazard ratios. The effect for osteocalcin was quite large, explaining the nearly separate point clouds in Fig 3c. Despite one result to the contrary for WHAS II, including basophil count also appears to have a generally strong, positive effect on slope, explaining the separate point clouds in Fig. 4a–b for InCHIANTI. Note, however, that the variables that improve model performance for change in D_M with age are not the same as those that improve mortality prediction.

4. Discussion

The results presented here confirm and expand on our previous study showing that statistical distance is a promising measure for physiological dysregulation during aging (Cohen and others 2013). By using additional data sets (WHAS I and InCHIANTI), we are able to provide independent confirmation of those results, results which are nearly identical in terms of the big picture (Figs 3–4), but which are also surprisingly different in the details. In all data sets, D_M significantly increases with age within individuals, and significantly predicts mortality. Likewise, in nearly all cases predictions improve as more variables are included in the calculation of D_M . These three findings confirm key predictions about how D_M should behave if it is truly a measure of physiological dysregulation (Cohen and others 2013).

One of the key challenges in developing a measure of physiological dysregulation is to avoid circularity (Singer and others 2004). It is unsurprising that by combining multiple measures of poor health one arrives at a measure that predicts poor health or age. Statistical distance, as measured by D_M , circumvents this problem by asking not if each patient is badly off on each marker, but by asking whether the overall profile of markers is far from average, regardless of how we define a healthy state for each. Additionally, we specifically did not restrict marker choice to those known to change with age or with health status; electrolyte levels, for example, are generally quite stable outside of specific pathologies. However, in our original study, we did impose the criterion that the deviance of markers from their mean be positively correlated with age (i.e., more aberrant values at older ages) in order to try to choose markers that would provide a stronger signal (Cohen and others 2013). Here, we showed that even this criterion is largely irrelevant – the correlations of individual markers with age are often quite heterogeneous across data sets and even subsets, and the correlations of the deviances with age are even more so. We would have chosen a completely different suite of markers had our original data set been a different one; nonetheless, using the markers chosen based on WHAS II, we get nearly identical results in all the data sets. We are currently investigating the stability of other combinations of biomarkers in different populations and have found that a similar signal of D_M can be obtained with very different combinations of markers, even mutually exclusive ones (unpublished data). Accordingly, although each biomarker is measured with error and may change from day to day, a relatively consistent signal can be obtained by combining a large number of biomarkers.

Likewise, we expected that a given combination of markers would provide a similar signal in different data sets, but the correlations were surprisingly weak. In each data set, the inclusion or exclusion of each marker was almost always significantly associated with model performance, but the directions of these associations varied across data sets, and few markers showed consistent effects across data sets. No markers showed consistent effects both (a) predicting both age and mortality, and (b) across data sets. In other words, if we find that including a given marker in the calculation of D_M improves the performance of D_M in one data set, we cannot necessarily make any inferences from this toward other data sets, at least among the markers used here.

How can we explain these relatively inconsistent results model-by-model, despite consistent results at a higher level? We believe that the discrepancies across data sets are, counterintuitively, a confirmation of the generality of D_M as a measure of dysregulation. If the performance of D_M depended too heavily on the choice of marker, it would suggest that D_M is not a measure of generalized dysregulatory state, but rather of what is happening with several key markers. Heterogeneity of results suggests that the effects of each marker of D_M performance depend on small differences across data sets in terms of population composition, diet, lifestyle, underlying physiology, and so forth; nonetheless, by combining a sufficient number of markers (and without much regard for which) we are able to circumvent these details and arrive at a fairly robust, generalized signal of dysregulatory state. Given what is known about the complexity of physiological regulation, this is in fact exactly the prediction we would make if D_M truly represents physiological dysregulation.

At a practical level, this study supports the utility of D_M as a measure of dysregulation or generalized health state, whether it be in studies of aging epidemiology, sociological or economic studies of population health, or in clinic. Clinical frailty measures such as Fried's frailty criteria (Fried and others 2001) and the Frailty Index (Rockwood and others 2005) provide useful insight into functional decline during aging (Clegg and others 2013); D_M promises to be a complementary measure of the underlying physiology. While frailty measures are most powerful late in life, D_M appears to pick up a signal much younger, suggesting clinical applications in prevention and non-geriatric populations, as well as coordinated use with frailty measures.

However, several further validation steps are necessary before implementing D_M widely. First, robustness/sensitivity to marker choice and number needs to be established across a wider array of markers. Recommended optimal marker combinations should also be established. Second, we need to understand the sensitivity to reference population (the population used to establish the definition of a "normal" or "average" profile). Given that most of the individuals used here were already elderly and thus in poorer health, there may be a potential to achieve better performance using younger and/or healthier populations to compute μ and S in equation (1). We are currently performing additional validation analyses on the sensitivity of D_M to the choice of the reference population, as well as on the supposition that the population mean is the optimal way to measure the state of least dysregulation. This supposition is almost certainly false in a strict sense, but it remains to be seen whether it is sufficiently close to optimal that more complex methods are not needed. Additionally, there is a possibility that state of least dysregulation differs depending on demographic or genetic factors; we are assessing whether a single reference population can serve well, or whether we need demographic-specific reference populations. Third, predictive value of D_M for specific health outcomes such as frailty and cardiovascular disease needs to be assessed. We would predict that dysregulation is linked with both aging and chronic disease, and that there should be feedback effects with chronic diseases. Extensive analyses on this question are ongoing. Additionally, we predict that dysregulation is the physiological underpinning of frailty, and that it should thus predict incidence of frailty (Clegg and others 2013; Fried and others 2001). Fourth, we will need to analyze whether there is a single global dysregulatory process, or if dysregulation can be usefully

subdivided by physiological or biological system, and, in the latter case, if these dysregulations are correlated. Under one scenario, there is a single global dysregulation; under another, different physiological or biological systems become dysregulated semi-independently, but with the possibility for feedback effects on each other. We are currently attempting to answer these questions.

While such validation is essential before systematic implementation, current results are strong enough to suggest that D_M could be useful immediately in smaller-scale studies. For example, we were recently able to successfully predict two measures of health state in a population of wild birds based on D_M calculated from 11 biomarkers available in an existing data set (Milot and others 2013). Additionally, at a theoretical level, this study confirms the interpretation of D_M as a measure of physiological dysregulation, as well as a role for physiological dysregulation in the aging process. It suggests strongly that too much emphasis on any single molecule may be misleading (as seen from our differing results for each molecule across data sets), and that measures of system-level properties of regulatory state will be necessary to better understand the aging process.

Acknowledgments

AAC is a member of the FRQ-S-supported *Centre de recherche 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. Two anonymous reviewers provided comments that greatly improved the manuscript. This research was supported by the Intramural Research Program of the National Institute on Aging.

References

- Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. *The Lancet*. 2013; 381:752–762.
- Cohen AA, Martin LB, Wingfield JC, McWilliams SR, Dunne JA. Physiological regulatory networks: ecological roles and evolutionary constraints. *Trends in Ecology & Evolution*. 2012; 27:428–435. [PubMed: 22613457]
- Cohen AA, Milot E, Yong J, Seplaki CL, Fülöp T, Bandeen-Roche K, Fried LP. A novel statistical approach shows evidence for multi-system physiological dysregulation during aging. *Mechanisms of Ageing and Development*. 2013; 134:110–117. [PubMed: 23376244]
- Crimmins EM, Johnston M, Hayward M, Seeman T. Age differences in allostatic load: an index of physiological dysregulation. *Experimental Gerontology*. 2003; 38:731–734. [PubMed: 12855278]
- De Maesschalck R, Jouan-Rimbaud D, Massart DL. The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*. 2000; 50:1–18.
- Ferrucci L, Bandinelli S, Benvenuti E, Di Iorio A, Macchi C, Harris TB, Guralnik JM. Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. *Journal of the American Geriatrics Society*. 2000; 48:1618–1625. [PubMed: 11129752]
- Fried LP, Bandeen-Roche K, Chaves PH, Johnson BA. Preclinical mobility disability predicts incident mobility disability in older women. *Journal of gerontology Series A, Biological sciences and medical sciences*. 2000; 55:M43–M52.
- Fried LP, Hadley EC, Walston JD, Newman AB, Guralnik JM, Studenski S, Harris TB, Ershler WB, Ferrucci L. From Bedside to Bench: Research Agenda for Frailty. *Sci Aging Knowl Environ*. 2005; 2005:pe24.
- Fried, LP.; Kasper, KD.; Guralnik, JM.; Simonsick, EM. The Women's Health and Aging Study: an introduction. In: Guralnik, JM.; Fried, LP.; Simonsick, EM.; Kasper, KD.; Lafferty, ME., editors.

- The Women's Health and Aging Study: health and social characteristics of old women with disability. Bethesda, MD: National Institute on Aging; 1995.
- Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman TE, Tracy R, Kop WJ, Burke G, McBurnie MA. Frailty in Older Adults: Evidence for a Phenotype. *Journal of gerontology Series A, Biological sciences and medical sciences*. 2001; 56:M146–M157.
- Goymann W, Wingfield JC. Allostatic load, social status and stress hormones: the costs of social status matter. *Animal Behaviour*. 2004; 67:591–602.
- Gruenewald TL, Seeman TE, Karlamangla AS, Sarkisian CA. Allostatic Load and Frailty in Older Adults. *Journal of the American Geriatrics Society*. 2009; 57:1525–1531. [PubMed: 19682116]
- Karlamangla AS, Singer BH, McEwen BS, Rowe JW, Seeman TE. Allostatic load as a predictor of functional decline: MacArthur studies of successful aging. *Journal of Clinical Epidemiology*. 2002; 55:696–710. [PubMed: 12160918]
- Mahalanobis PC. Mahalanobis distance. *Proceedings National Institute of Science of India*. 1936; 49:234–256.
- McEwen BS, Wingfield JC. The concept of allostasis in biology and biomedicine. *Hormones and Behavior*. 2003; 43:2–15. [PubMed: 12614627]
- Miller GE, Chen E, Zhou ES. If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. *Psychological Bulletin*. 2007; 133:25–45. [PubMed: 17201569]
- Milot E, Cohen AA, Vézina F, Buehler DM, Matson KD, Piersma T. A novel integrative method for measuring body condition in ecological studies based on physiological dysregulation. *Methods in Ecology and Evolution*. 2013:n/a–n/a.
- Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, Mitnitski A. A global clinical measure of fitness and frailty in elderly people. *CMAJ*. 2005; 173:489–495. [PubMed: 16129869]
- Sapolsky RM, Krey LC, McEwen BS. The Neuroendocrinology of Stress and Aging: The Glucocorticoid Cascade Hypothesis. *Sci Aging Knowl Environ*. 2002; 2002:cp21.
- Schnorpfeil P, Noll A, Schulze R, Ehlert U, Frey K, Fischer JE. Allostatic load and work conditions. *Social Science & Medicine*. 2003; 57:647–656. [PubMed: 12821013]
- Seplaki CL, Goldman N, Gleit D, Weinstein M. A comparative analysis of measurement approaches for physiological dysregulation in an older population. *Experimental Gerontology*. 2005; 40:438–449. [PubMed: 15919596]
- Singer, BH.; Ryff, CD.; Seeman, T. Operationalizing allostatic load. In: Schuli, J., editor. *Allostasis, Homeostasis, and the Costs of Physiological Adaptation*. Cambridge, UK: Cambridge University Press; 2004.
- Taffett, GE. Physiology of Aging. In: Cassel, CK.; Leipzig, RM.; Cohen, HJ.; Larson, EB.; Meier, DE.; Capello, CF., editors. *Geriatric Medicine*. Springer; New York: 2003.
- Yashin AI, Arbeev KG, Akushevich I, Kulminski A, Akushevich L, Ukraintseva SV. Stochastic model for analysis of longitudinal data on aging and mortality. *Mathematical Biosciences*. 2007; 208:538–551. [PubMed: 17300818]

Figure 1.

Means and 95% quantiles of D_M with and without transformation and standardization. a) and b): unstandardized; c) and d) standardized (mean=0, sd=1); a) and c): untransformed; b) and d): log-transformed. Blue indicates WHAS and red InCHIANTI; triangles the 2.5th percentile value, circles the mean value, and pluses the 97.5th percentile value. There is a separate set of points for D_M calculated from each possible biomarker combination in each data set.

Figure 2.

Correlations of raw variables and their deviances with age. Variables are sorted from lowest to highest deviance correlation coefficient in WHAS II (third column). Colored boxes indicate significant correlations (blue=positive, red=negative) with darker shading indicating lower p -values. Note that of the 12 original variables retained for WHAS II (those shaded blue at the bottom), only one would have been retained for WHAS I and four for InCHIANTI. On the other hand, eight additional variables not retained for WHAS II would have been retained for WHAS I, and 18 for InCHIANTI.

Figure 3.

Changes in predictive power of D_M in WHAS with increasing numbers of variables used in its calculation. Each circle represents an analysis based on one of the 16383 combinatorial subsets of the 14 variables in WHAS. 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. (a)–(c): average individual slope of D_M with age (units of increase in D_M per year). (d)–(f): hazard ratio of mortality per unit D_M , controlling for age. (a), (d): The full WHAS data set. (b), (e): WHAS I. (c), (f): WHAS II.

Figure 4.

Changes in predictive power of D_M in InCHIANTI with increasing numbers of variables used in its calculation. Each circle represents an analysis based on one of the 4095 combinatorial subsets of the 12 variables in InCHIANTI. Color indicates p-value: black: $p > 0.1$; blue: $0.05 > p > 0.1$; cyan: $0.01 > p > 0.05$; yellow-green: $0.001 > p > 0.01$; orange: $0.0001 > p > 0.0001$; 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. (a) and (b): average individual slope of D_M with age (units of increase in D_M per year). (c) and (d): hazard ratio of mortality per unit D_M , controlling for age. (a) and (c): The full InCHIANTI data set. (b) and (d): The subset of women aged 70+ (for comparison with the original WHAS II data set).

Table 1

Characteristics of study populations

Characteristic	WHAS (N = 1226)	InChianti (N = 1310)
<i>Age at first visit</i>		
Mean $\pm$ SD	77.1 $\pm$ 7	68.3 $\pm$ 15
Range (min – max)	65.8 – 100.3	21.3 – 98.4
Female (%)	1226 (100.0)	727 (55.5)
Race (white, %)	918 (74.9)	1310 (100.0)
Number of deaths (%)	411 (33.5)	366 (27.9)
Person-years	7,780.7	10,412.6
Maximum number of visits (mean)	7 (3.1)	4 (2.8)

Table 2
Effects of control for age on results of survival analyses, using the models based on all biomarkers.

Data set	Controlling for age			Without controlling for age		
	HR (95% CI)	SE	p	HR (95% CI)	SE	p
<i>InChianti</i>						
All	1.25 (1.14–1.37)	0.047	<0.0001	1.32 (1.23–1.41)	0.036	<0.0001
Women age 70+	1.24 (1.06–1.44)	0.079	0.008	1.30 (1.14–1.49)	0.070	0.0002
<i>WHAS</i>						
All	1.36 (1.14–1.62)	0.090	0.0007	1.39 (1.17–1.66)	0.090	0.0002
WHAS I	1.36 (1.05–1.75)	0.130	0.02	1.41 (1.09–1.82)	0.131	0.009
WHAS II	1.24 (1.02–1.51)	0.099	0.03	1.24 (1.03–1.50)	0.094	0.02

HR=hazard ratio; CI=confidence intervals; SE=standard error.

Table 3

Results of metaregression analyses to assess the impact of including or excluding each variable in the calculation of D_M

	WHAS (all)			WHAS I			WHAS II			InCHIANTI			InCHIANTI F70+		
	β	p		β	p		β	p		β	p		β	p	
Individual slope with age															
Albumin	-0,018	<0.0001		-0,009	<0.0001		-0,005	0,00		0,010	<0.0001		-0,002	0,23	
Hemoglobin	-0,013	<0.0001		0,006	0,0001		-0,009	<0.0001		0,004	0,008		0,001	0,35	
Hematocrit	-0,011	<0.0001		0,021	<0.0001		-0,013	<0.0001		0,012	<0.0001		0,012	<0.0001	
Chloride	-0,008	<0.0001		-0,028	<0.0001		0,007	<0.0001		0,008	<0.0001		0,014	<0.0001	
RBCs	-0,003	0,003		0,004	0,003		0,003	0,06		-0,005	0,0001		0,001	0,50	
Cholesterol	0,000	0,85		0,015	<0.0001		0,001	0,37		-0,006	<0.0001		0,009	<0.0001	
Calcium	0,001	0,44		0,010	<0.0001		-0,016	<0.0001		0,003	0,01		-0,007	<0.0001	
Bilirubin (direct)	0,003	0,003		-0,003	0,07		0,016	<0.0001							
Sodium	0,006	<0.0001		-0,006	<0.0001		0,020	<0.0001		0,006	<0.0001		0,001	0,52	
Potassium	0,012	<0.0001		0,021	<0.0001		-0,002	0,18		-0,004	0,001		-0,003	0,06	
BUN:Creat ratio	0,019	<0.0001		0,004	0,003		0,029	<0.0001		0,012	<0.0001		0,014	<0.0001	
Creatinine	0,028	<0.0001		0,028	<0.0001		0,009	<0.0001		0,018	<0.0001		0,023	<0.0001	
Basophils	0,038	<0.0001		0,13	<0.0001		-0,011	<0.0001		0,080	<0.0001		0,089	<0.0001	
Osteocalcin	0,12	<0.0001		0,11	<0.0001		0,19	<0.0001							
Mortality (HR)															
Osteocalcin	-0,042	<0.0001		-0,056	<0.0001		0,023	<0.0001							
Creatinine	-0,031	<0.0001		0,009	<0.0001		-0,155	<0.0001		0,028	<0.0001		0,004	0,002	
Calcium	-0,027	<0.0001		-0,030	<0.0001		-0,018	<0.0001		-0,002	0,05		-0,009	<0.0001	
RBCs	0,002	0,13		-0,006	<0.0001		0,005	0,002		-0,006	<0.0001		-0,011	<0.0001	
Albumin	0,004	0,0003		0,019	<0.0001		0,001	0,38		0,027	<0.0001		0,023	<0.0001	
Basophils	0,007	<0.0001		-0,018	<0.0001		0,037	<0.0001		0,000	0,84		0,004	0,002	
Potassium	0,007	<0.0001		-0,005	<0.0001		0,014	<0.0001		-0,001	0,57		0,005	0,001	
Hemoglobin	0,009	<0.0001		0,015	<0.0001		-0,001	0,41		-0,002	0,05		-0,003	0,04	
Hematocrit	0,011	<0.0001		0,016	<0.0001		-0,007	<0.0001		0,002	0,11		-0,001	0,41	
Cholesterol	0,021	<0.0001		0,003	0,03		0,015	<0.0001		0,030	<0.0001		0,029	<0.0001	
Sodium	0,022	<0.0001		0,025	<0.0001		0,008	<0.0001		0,040	<0.0001		0,029	<0.0001	

	WHAS (all)		WHAS I		WHAS II		lnCHIANTI		lnCHIANTI F70+	
	β	<i>p</i>	β	<i>p</i>	β	<i>p</i>	β	<i>p</i>	β	<i>p</i>
Chloride	0,031	<0.0001	0,041	<0.0001	0,006	<0.0001	0,005	<0.0001	-0,015	<0.0001
BUN:Creat ratio	0,046	<0.0001	0,075	<0.0001	-0,008	<0.0001	0,021	<0.0001	0,027	<0.0001
Bilirubin (direct)	0,076	<0.0001	0,066	<0.0001	0,048	<0.0001				

Betas indicate change in effect size with the inclusion of the biomarker. Biomarkers are ordered by their beta-coefficients for the full WHAS data set, separately for individual slope with age and mortality. Negative coefficients are marked in red; coefficients not significant at $\alpha=0.05$ are in italics.