

Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging

Teresa E. Seeman^{*†}, Bruce S. McEwen[‡], John W. Rowe[§], and Burton H. Singer[¶]

^{*}Division of Geriatrics, University of California School of Medicine, 10945 Le Conte Avenue, Suite 2339, Los Angeles, CA 90095-1687; [‡]Laboratory for Neuroendocrinology, The Rockefeller University, New York, NY 10021; [§]Aetna, Inc, Hartford, CT 06156; and [¶]Office of Population Research, Princeton University, Princeton, NJ 08544

Contributed by Burton H. Singer, February 13, 2001

Allostatic load (AL) has been proposed as a new conceptualization of cumulative biological burden exacted on the body through attempts to adapt to life's demands. Using a multisystem summary measure of AL, we evaluated its capacity to predict four categories of health outcomes, 7 years after a baseline survey of 1,189 men and women age 70–79. Higher baseline AL scores were associated with significantly increased risk for 7-year mortality as well as declines in cognitive and physical functioning and were marginally associated with incident cardiovascular disease events, independent of standard socio-demographic characteristics and baseline health status. The summary AL measure was based on 10 parameters of biological functioning, four of which are primary mediators in the cascade from perceived challenges to downstream health outcomes. Six of the components are secondary mediators reflecting primarily components of the metabolic syndrome (syndrome X). AL was a better predictor of mortality and decline in physical functioning than either the syndrome X or primary mediator components alone. The findings support the concept of AL as a measure of cumulative biological burden.

In elderly populations, comorbidity in the form of multiple co-occurring chronic conditions is the norm rather than the exception. For example, in the U.S. 61% of women and 47% of men age 70–79 report two or more chronic conditions. These figures rise to 70% of women and 53% of men age 80–89 with 2+ chronic conditions (1). No single form of comorbidity occurs with high frequency, but rather a multiplicity of diverse combinations are observed (e.g., osteoarthritis and diabetes, colon cancer, coronary heart disease, depression, and hypertension). This diversity underscores the need for an early warning system of biomarkers that can signal early signs of dysregulation across multiple physiological systems.

One response to this challenge was the introduction of the concept of allostatic load (AL) (2–4) as a measure of the cumulative physiological burden exacted on the body through attempts to adapt to life's demands. The ability to successfully adapt to challenges has been referred to by Sterling and Eyer (5) as allostasis. This notion emphasizes the physiological imperative that, to survive, “an organism must vary parameters of its internal milieu and match them appropriately to environmental demands” (5). When the adaptive responses to challenge lie chronically outside of normal operating ranges, wear and tear on regulatory systems occurs and AL accumulates.

An initial operationalization of the notion of AL (6) used assessments of 10 biological parameters reflecting functioning of the hypothalamic-pituitary-adrenal (HPA) axis, sympathetic nervous system, cardiovascular system, and metabolic processes. Higher scores on a summary numerical measure of AL were shown to predict four major health outcomes: incident cardiovascular disease (CVD), decline in physical functioning, decline in cognitive functioning, and mortality (6). Although the summary AL measure was a significant predictor of outcomes, none of the individual components exhibited strong predictive capacity. This finding suggests that risk for the above outcomes is related to the overall impact of dysregulation across multiple

regulatory systems. In addition, examination of the most commonly occurring pairs and triples of factors contributing to overall AL scores clearly indicated that there were multiple pathways of relatively comparable frequency through which individuals achieved higher AL scores. This finding is consistent with the idea that a cumulative measure of “biological dysregulation” across multiple systems may act as an early warning indicator of multiple health outcomes. It also emphasizes the point that there are multiple pathways to comorbidity and mortality.

Despite these positive indications, some important caveats are associated with previous analyses. First, the predictive ability of AL was assessed over a relatively short time period; i.e., 2.5 years. Second, six of the 10 components of the measure of AL are standard indicators of syndrome X, the metabolic syndrome that has been related to increased risk of CVD (7). These include levels of seated systolic blood pressure (SBP) and diastolic blood pressure (DBP), high density lipoproteins (HDL), and total cholesterol, glycosylated hemoglobin, and abdominal obesity [i.e., high waist-hip ratio (WHR)]. Thus, the question arises as to whether the extant operationalization of AL is something more than an alternative measure of syndrome X. Third, there is a cascade of events through which AL develops, with chemical messengers such as cortisol (C), epinephrine (EPI), and norepinephrine (NE) playing the role of primary mediators (4). Previous studies give no indication of the potential predictive value of such primary mediators, nor do they clarify how they contribute to a summary measure of AL beyond the syndrome X components.

The purposes of this paper are to investigate: (i) the predictive value of a summary measure of AL over a period of 7 years beyond baseline survey of an elderly population; (ii) the predictive value of syndrome X parameters, which are all secondary mediators in the cascade that leads to cumulating AL, for outcomes other than CVD, such as decline in physical and cognitive functioning, and (iii) the kinds of outcomes for which a set of primary mediators are good predictors in their own right and those for which they contribute, above and beyond syndrome X parameters, to the predictive capacity of a summary measure of AL.

Methods

Data for these analyses come from the MacArthur Successful Aging Study, a longitudinal study of relatively high functioning men and women aged 70–79. As described in greater detail elsewhere (8), subjects were subsampled on the basis of age and both physical and cognitive functioning from three community-

Abbreviations: AL, allostatic load; CVD, cardiovascular disease; CI, confidence interval; C, cortisol; NE, norepinephrine; EPI, epinephrine; DBP, diastolic blood pressure; SBP, systolic blood pressure; HDL, high density lipoproteins; HPA, hypothalamic-pituitary-adrenal; WHR, waist-hip ratio; DHEA-S, dihydroepiandrosterone sulfate; OR, odds ratio.

[†]To whom reprint requests should be addressed. E-mail: tseeman@mednet.ucla.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

based cohorts in Durham, NC, East Boston, MA, and New Haven, CT that were part of the Established Populations for Epidemiological Studies of the Elderly. Age was restricted to 70–79 years, and age-eligible men and women were screened on the basis of four criteria of physical functioning and two criteria of cognitive functioning to identify those functioning in the top third of the age group.

Of the 4,030 age-eligible men and women, a cohort of 1,313 subjects met all screening criteria; 1,189 (90.6%) agreed to participate and provided informed consent. Baseline data collection was completed between May 1988 and December 1989 and included a 90-min, face-to-face interview as well as blood samples and 12-h overnight urine; 80.3% agreed to provide blood samples and 85.8% consented to provide urine samples.

The cohort was reinterviewed in 1991 and again in 1996; with reassessments of all measures included in the baseline interview. Outcomes examined here are based on the 1996 follow-up (data collection from October 1995 to February 1997). The mean interval from the first follow-up was 85 months (i.e., 7 years; SD = 7 months). Of the 1,118 individuals alive at the end of data collection in 1991, 202 had died by the time of the 1996 follow-up (18%); of the remaining 916 subjects, 722 completed face-to-face interviews (79%), 107 (11.7%) had proxy-partial interviews, 44 (4.8%) refused, and 43 (4.7%) were alive but couldn't be contacted.

AL. Our measure of AL, described previously (6), is designed to summarize levels of physiological activity across a range of regulatory systems pertinent to disease risks (2–4). These parameters included: SBP and DBP (indices of cardiovascular activity); WHR (an index of more chronic levels of metabolism and adipose tissue deposition, thought to be influenced by increased glucocorticoid activity) (9); serum HDL and total cholesterol (markers whose levels are known to influence the development of atherosclerosis, increased risk being seen with higher levels in the case of total cholesterol and lower levels in the case of HDL); blood plasma levels of glycosylated hemoglobin (Hb_{A1c}; an integrated measure of glucose metabolism over several months time); serum dihydroepiandrosterone sulfate (DHEA-S; a functional HPA axis antagonist) (10); 12-h urinary C excretion (an integrated measure of 12-h HPA axis activity); and 12-h urinary NE and EPI excretion levels (integrated indices of 12-h sympathetic nervous system activity).

SBP and DBP were calculated as the average of the second and third of three seated blood pressure readings (11). WHR was calculated based on waist circumference (measured at its narrowest point between the ribs and iliac crest) and hip circumference (measured at the maximal buttocks) (12). Blood samples for assays of HDL cholesterol, total cholesterol, glycosylated hemoglobin, and DHEA-S were obtained by phlebotomists who visited subjects' homes the morning after their home interview to collect the blood sample and an overnight urine sample (see below). Although subjects were not required to fast, most blood samples were taken early in the morning before subjects had eaten. A total of 9 cm³ of blood was drawn in serum-separator tubes, and another 10 cm³ was drawn into heparinized tubes. Samples were processed and frozen within 4 h of blood drawing. Blood in the serum-separator tubes was allowed to clot, and the tube was centrifuged at 15 relative centrifugal force for 10 min in a refrigerated (4°C) centrifuge. The sera were sent to Nichols Laboratories (San Juan Capistrano, CA) for measurements of HDL and total cholesterol, and DHEA-S. A total of 2 cm³ of blood from a heparinized tube was removed after mixing for assays of glycosylated hemoglobin (Hb_{A1c}), assayed by affinity chromatography methods (13).

Subjects also completed an overnight urine collection from 8 p.m. on the evening after their home interview to 8 a.m. the next morning. The overnight collection protocol serves to min-

Table 1. AL: Criteria for contributions from individual biological parameters

Highest quartile	
SBP	(≥ 148 mm Hg)
DBP	(≥ 83 mm Hg)
WHR	(≥ 0.94)
Ratio total cholesterol/HDL	(≥ 5.9)
Glycosylated hemoglobin	(≥ 7.1%)
Urinary C	(≥ 25.7 μg/g creatinine)
Urinary NE	(≥ 48 μg/g creatinine)
Urinary EPI	(≥ 5 μg/g creatinine)
Lowest quartile	
HDL cholesterol	(<= 37 mg/dl)
DHEA-S	(<= 350 ng/ml)

imize the potential confounding effects of physical activity as subjects generally spent this time at home (and much of that time in bed). Thus, individual differences in overnight excretion of C as well as NE and EPI provide integrated measures of differences in more “steady-state” operating levels of the HPA axis and sympathetic nervous system (i.e., estimates of “basal,” nonstimulated levels of activity). All samples were sent to Nichols Institute immediately after collection for assays of NE, EPI, and C. Determinations were made by HPLC (14, 15). Interassay variation was 8–9% for catecholamines regardless of range; interassay variation for C was 1% in low range (<50 μg/dl) and 8% in higher range. Results for each of the three outcomes are reported as μg (NE, EPI, or C)/g creatinine to adjust for body size.

For each of the 10 indicators of AL, subjects were classified into quartiles based on the distribution of scores in the baseline cohort. AL was measured by summing the number of parameters for which the subject fell into the highest risk quartile (i.e., top quartile for all parameters except HDL cholesterol and DHEA-S for which membership in the lowest quartile corresponds to highest risk). Use of the top/bottom quartile criterion to define contributions to higher AL from these various biological parameters reflects a data-driven partitioning of the sample. Membership in the upper/lower quartile represents a quantitative way of classifying those exposed to more extreme levels of system activity relative to the rest of the population and thus potentially at greater risk of disease pathology. Table 1 indicates the actual criterion cutoff for each of the 10 indices. Other methods of summarizing the data, including averaging z scores and use of other criterion cut points also were examined in earlier analyses and yielded comparable results (6).

Contributions of syndrome X and non-X (i.e., primary mediator) components of AL were examined by dividing the total AL score into two component parts: one reflecting the sum of the syndrome X components (i.e., SBP, DBP, HDL and total cholesterol, Hb_{A1c}, WHR) and the other reflecting the sum of the non-syndrome X components (i.e., urinary C, NE, EPI, and DHEA-S).

Health Outcomes. All-cause mortality. Deaths were identified through contact with next of kin at the time of subsequent interviews (1991 and 1996), ongoing local monitoring of obituary notices, and National Death Index searches.

Incident CVD (1988–1996). Incident CVD events were assessed through self-reports of new, doctor-diagnosed myocardial infarction, stroke, hypertension, or diabetes between 1988 and 1996. Reports of such newly diagnosed conditions at the time of the 1991 and 1996 follow-up interviews were compared with annual interview data from Established Populations for the Epidemiologic Study of the Elderly interviews for the previous 5 years (for Durham subjects) and 10 years (for New Haven and

East Boston subjects) to confirm that the condition had not been previously reported. Although medical records were not available for further confirmation, self-reports of such conditions have been found to agree with medical records for these conditions with percent agreement of 0.86 or higher and kappas of 0.70 or higher (16). Because of the small number of cases for individual diseases, we used a summary measure reflecting the occurrence of any incident CVD (i.e., doctor-diagnosed myocardial infarction, hypertension, diabetes, or stroke). Of the total 161 incident CVD events reported between 1988 and 1995, there were 10 incident reports of a diagnosis of hypertension in subjects who also had high blood pressure at baseline as a contributor to their AL score. To reduce potential confounding between baseline AL and subsequent reports of incident hypertension, these 10 cases were deleted from the count of incident CVD events.

Cognitive and physical performance. Detailed, performance-based assessment of both cognitive and physical performance were available at both baseline (1988) and the final follow-up (1996), permitting examination of both cross-sectional and longitudinal associations with AL.

Cognition. Measures of cognition were designed to assess those aspects of higher cortical function that are necessary to perform complex cognitive activities. Measures included six tasks: language, abstraction, spatial ability (copying), delayed spatial recognition, delayed recall of confrontation naming items, and delayed recall of a story. A measure of overall cognitive performance was created by summing scores on the six tasks (17).

Physical performance. A summary measure of physical performance, based on five separate tests of physical ability, was used. The five abilities assessed included timed measures of balance, gait, chair stands, foot taps, and manual ability. Scores for each of the subscales were scaled to indicate the proportion of the best possible score that a subject achieved, i.e., rescaled scores range from 0 (worst possible performance) to 1 (best possible performance), and a summary measure of performance was created by adding the scaled subscale scores for the respective years (range each year = 0 to 5) (18).

Covariates. Multivariate analyses included age, sex, race, income, and education along with summary measures of baseline health status. Age was measured as of the date of the baseline interview (range = 70 to 79). Race was coded African-American versus Caucasian. Education was measured as less than high school vs. high school or higher. Annual household income was measured in \$10,000 increments to \$30,000+; final analyses use a dichotomous classification: <\$10,000 versus \$10,000 or more. Summary measures of baseline prevalence of CVD (i.e., myocardial infarction, stroke, diabetes, high blood pressure) and of other conditions (cancer, broken hip, or other bones) were developed based on self-reports of doctor-diagnosed conditions. These measures were included in analyses of cognitive and physical functioning and of mortality.

Analyses. Logistic regression models were used to test for associations between AL and incident CVD and mortality. For the continuous outcome measures of cognitive and physical functioning, linear regression models were used to assess associations between AL and changes in performance. Comparison of the 720 individuals with complete baseline data to those missing either AL or covariate data indicated that those with complete data did not differ in terms of their age, education, income, burden of chronic conditions, or incidence of CVD; mortality was only marginally higher among those excluded from the analyses (25% vs. 21%, $P = 0.08$). Those with complete data were more likely to be male and to exhibit somewhat better cognitive and physical performance abilities at baseline, although the actual differences in performance were relatively minor (data not shown).

To assess associations between AL and changes in cognitive

and physical functioning, 1996 function was regressed on AL scores, controlling for baseline function. With inclusion of baseline performance scores, the analysis is modeling residualized change, i.e., change adjusted for baseline level of physical performance (19) and coefficients for other variables provide estimates of the impact of these factors on change in performance from baseline. Models were examined controlling for possible confounding by socio-demographic and/or health status characteristics. Similar multivariable logistic (rather than proportional hazards) regression models were examined with respect to incident and recurrent CVD and mortality outcomes because information on time to event was not available. Possible nonlinear risk relationships for AL were examined by including a series of four dummy variables reflecting increasing levels of AL (levels 1–2, 3–4, 5–6, 7+) with scores of zero as the reference group.

To assess the question of whether AL is something more than simply an alternative measure of syndrome X, models were run, replacing the summary AL score with the two separate subscales measuring syndrome X components and non-X components. SAS version 6.12 was used for all analyses (20).

Results

The sample was 49% male, 19% non-white and showed considerable socio-economic diversity. Forty percent of the subjects reported annual household incomes of less than \$10,000/year whereas 14% reported annual incomes of \$30,000 or more. Educational attainment also varied from less than eighth grade (17%), to high school (25.6%), to some college or more (22%). At baseline, subjects reported on average slightly more than one chronic condition; 56% reported prevalent CVD and 26% reported no prevalent chronic conditions.

Over the following 7.5 years, there were 153 deaths among the 720 subjects with complete baseline data and a total of 151 incident CVD events among the 582 subjects free of baseline CVD. Average declines in physical performance as measured by our five-point scale were equivalent to approximately 1 SD ($\bar{x} = -0.49$), whereas average declines in cognitive performance based on the 90-point scale were equivalent to approximately half of a SD ($\bar{x} = -3.7$). The actual range of change scores for both performance domains was quite wide: changes in physical performance ranged from -3.35 to $+0.62$ (maximum possible range: -5 to $+5$), changes in cognitive performance ranged from -49 to $+21$ (maximum possible range = -89 to $+89$).

AL and 7-Year Outcomes. Logistic regression analyses revealed significantly increased risk of mortality with higher AL scores, independent of age, sex, ethnicity, education, income, and baseline morbidity [odds ratio (OR) = 1.23; 95% confidence interval (CI) = 1.09, 1.39]. Use of dummy-coded indicator variables for categories of increasing AL revealed an approximately monotonic increase in relative risk with increasing AL. As shown in Fig. 1a, those with scores of 1–2 experienced 1.67 greater mortality compared with those scoring zero on AL (95% CI = 0.71, 4.59); those scoring 3–4 on AL experienced 2.45 times greater mortality (95% CI = 1.04, 6.77) whereas those at the highest levels of AL (i.e., 7+) experience the greatest relative risk for mortality (OR = 6.42; 95% CI = 1.36, 32.12).

Logistic regression models for incident CVD revealed a marginally significant association between baseline AL and subsequent incident CVD events (OR = 1.12; 95% CI = 0.99, 1.27; $P = 0.06$). Examination of the individual dummy-coded indicators for increasing levels of AL revealed a generally increasing pattern of risk for CVD with increasing AL although none of the individual comparisons reached statistical significance (see Fig. 1b).

Linear regression analyses of changes in physical functioning also revealed that higher AL scores were associated with greater

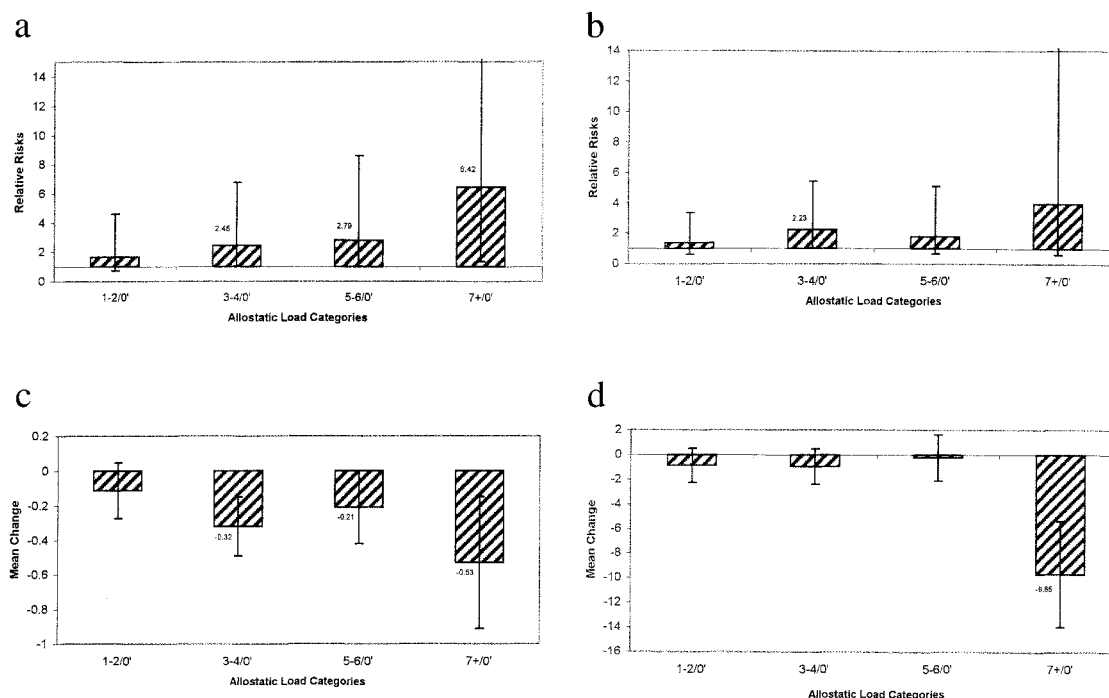


Fig. 1. Associations between baseline AL and 7-year health outcomes. Effects reaching statistical significance ($P < 0.05$) are indicated by numbers at top of bars. Indicated are 95% CIs for each estimated relative risk. (a) Mortality risks. (b) Incident CVD risks. (c) Changes in physical functioning. (d) Changes in cognitive functioning.

average decline ($b = -0.06, P = 0.02$). As shown in Fig. 1c, those with highest AL scores (i.e., those with scores of 7+) experienced the greatest declines and, although not strictly monotonic, there was a general trend toward smaller declines moving from higher to lower AL groups.

For cognitive declines, linear regression models again revealed a significant negative effect of AL ($b = -0.58, P = 0.03$). As shown in Fig. 1d, this effect appears to reflect the significantly greater declines among those with the highest AL scores.

Cumulative AL versus Syndrome X Risk. We next considered the question of whether these observed associations reflect the cumulative impact of the 10 components or were simply reflections of the effects of a specific subset of measures such as those associated with syndrome X. To test this hypothesis, models were examined that included two separate scores: one score reflected the number of syndrome X characteristics that were present (i.e., high SBP and DBP, high ratio of total cholesterol to HDL, high WHR, and high Hb_{A1c}), and the other score reflected the number of non-X characteristics that were present (i.e., high urinary C, NE, EPI, and low DHEA-S).

The possibility that each of these subscales might contribute independently to health risk was supported by evidence indicating considerable nonoverlap between these two scores. For example,

among those with a 0 score on the syndrome X subscale, over 70% had scores of 1 or more on the non-X subscale; likewise, among those with 0 scores on the non-X subscale, 80% had a score of 1 or more on the syndrome X subscale. Analyses including the two subscale scores provide an indication of the independent contribution of each subset of biological measures (i.e., an assessment of the extent to which the non-X components of AL are contributing independently to the prediction of our four health outcomes).

Analyses of mortality revealed significant effects for both syndrome X and non-X risk scores (for syndrome X: OR = 1.18, $P = 0.03$, 95% CI = 1.02, 1.36; and for non-X: OR = 1.36, $P = 0.002$, 95% CI = 1.12, 1.66). The pattern of mortality risks across the levels of syndrome X and non-X scores based on dummy-coded indicators for each of the levels of syndrome X and non-X is shown in Table 2. For non-X scores, there is a pattern of increasing mortality as one moves from lower to higher non-X scores; as also indicated in Table 2, each of these groups experiences significantly increased mortality as compared with those scoring 0 on the non-X risk factors. By contrast, syndrome X scores of 1, 2, and 3 are associated with similar and nonsignificant increases in mortality risk (OR ≈ 1.7) when compared with those with scores of 0. Scores of 4 or more on syndrome X, however, are associated with significantly greater risk (OR = 2.47; 95% CI = 1.28, 2.66)

Table 2. Mortality and cardiovascular risks associated with syndrome X and non-X factors

Syndrome X & non-X scores	Mortality		Incident CVD	
	Syndrome X risks OR	Non-X risks OR	Syndrome X risks OR	Non-X risks OR
1/0	1.68 (0.96, 3.01)	1.9* (1.21, 3.02)	1.59 (0.93, 2.75)	1.55* (1.0, 2.44)
2/0	1.65 (0.91, 3.06)	2.01* (1.14, 3.51)	1.64 (0.92, 2.96)	1.21 (0.68, 2.14)
3/0	1.54 (0.77, 3.07)	2.69* (1.30, 5.45)	1.65 (0.85, 3.2)	1.34 (0.64, 2.72)
4+ /0	2.47* (1.19, 5.13)		1.71 (0.78, 3.68)	

AL = 0 is reference group in all analyses. *, $P < 0.05$.

Table 3. Changes in physical and cognitive performance associated with syndrome X and non-X factors

Syndrome X & non-X scores	Change in physical function		Change in cognitive function	
	Syndrome X risks <i>b</i>	Non-X risks <i>b</i>	Syndrome X risks <i>b</i>	Non-X risks <i>b</i>
1/0	-0.09	-0.11*	1.03*	1.22*
2/0	-0.11	-0.03	-0.17	-1.43*
3/0	-0.4*	-0.34*	-2.69*	-0.58
4+/ —	-0.15	—	-0.15	—

*, $P < 0.05$.

Analyses of incident CVD events revealed a stronger, positive association between the syndrome X scores and incident CVD than was the case for the non-X scores, although even the syndrome X scores had only a marginally significant effect (OR = 1.14; 95% CI = 0.98, 1.32; $P = 0.08$). Examination of the patterns of risk for specific levels of syndrome X and non-X revealed an unexpectedly flat pattern of risk for syndrome X scores of 1 or more (see Table 2); in no case did the estimated relative odds of incident CVD achieve statistical significance. A somewhat similar pattern was seen for the non-X scores, although in the latter case, the estimated relative odds of an incident CVD event did achieve marginal significance for those with scores of 1 versus 0 (OR = 1.55; $P = 0.06$).

Linear regression modeling of declines in physical performance revealed only a marginal association for syndrome X ($b = -0.062$, $P = 0.07$) and no significant association for non-X ($b = -0.066$, $P = 0.13$). However, as shown in Table 3, analyses based on the dummy-coded indicators for levels of these subscales suggested that these linear estimates may be inaccurate as there appear to be threshold effects for both syndrome X and non-X scores—the largest declines in performance being evident among those with highest or next highest scores on syndrome X and/or non-X scales. This pattern is consistent with the previously observed risks for decline in physical performance associated with the highest AL scores (i.e., scores of 7+). Indeed, the stronger overall effect of AL appears to reflect more modest effects of both syndrome X and non-X factors.

Linear regression models for cognitive performance revealed a significant, negative effect for syndrome X risk scores ($b = -0.61$, $P = 0.04$), with a nonsignificant effect for non-X risk scores ($b = -0.46$, $P = 0.25$). Analyses based on the dummy-coded indicators did not reveal a consistent pattern of effects (see Table 3).

Discussion

A primary goal of these analyses was to provide a more thorough examination of the concept of AL as a measure of cumulative biological burden. A first set of analyses confirmed that overall AL was a significant predictor of 7.5-year health outcomes in this cohort of older adults as it had been in earlier analyses of 2.5-year outcomes (6). Higher baseline AL scores were associated with significantly greater risk for mortality as well as for declines in cognitive and physical functioning; and a marginally significant effect was also seen for incident CVD. The risk relationship was generally linear for mortality but appeared to reflect more of a threshold pattern with respect to risk for decline in cognitive and physical functioning, where only those at the highest levels of AL experienced significantly greater declines in functioning.

A second set of analyses, focusing on the contributions of the syndrome X and non-X components of AL, indicated that for two of our outcomes, mortality and physical functioning, both subscales appear to contribute to subsequent risks. In the case of mortality, both the syndrome X and non-X components of AL contributed significantly to risk. In the case of physical functioning, although neither of the two subscales was significantly

related to patterns of change in its own right, each was marginally and negatively related to risk for decline, and the summary measure of AL, combining these subscale effects, did exhibit a significant association with subsequent risks for decline in physical functioning.

The other two outcomes examined here, incident CVD and cognitive functioning, each provided evidence suggesting that the syndrome X components are largely responsible for the observed effects of AL—a finding that is not surprising in the case of incident CVD but more unexpected with respect to risks for cognitive decline. This pattern of results provides somewhat mixed support for the general hypothesis that the more comprehensive summary measure of AL would offer the best prediction of outcomes. The evidence for this is strongest with respect to mortality and physical functioning but less so for incident CVD and cognition.

Limitations of the data presented here are important to consider in evaluating the findings. Two features of the MacArthur Study of Successful Aging database are worth noting. First, the analyses use secondary data (a convenience sample) and are thus constrained by the available biological data (a set of measures that were not collected specifically to assess AL). Longitudinal data collection explicitly designed to assess AL lies in the future. Thus, our inability to consistently demonstrate significant contributions from non-syndrome X factors may be due, at least in part, to the fact that the available data did not provide as comprehensive coverage of these latter factors as was true for the syndrome X characteristics. One might speculate, for example, that stronger evidence for AL as a predictor of both cognition and CVD, and indeed as a predictor of all of the outcomes, would have been found had our measure of AL included a more comprehensive assessment of parameters of dysregulation in all important regulatory systems. Thus, for example, inclusion of information on missing elements such as measures of hemostatic or immune/inflammatory processes, which have been linked to CVD (21, 22) and/or cognition (J. Weaver, M. Albert, T. Harris, J.W.R., and T.E.S., unpublished work), might improve the ability of our measure of AL to predict outcomes.

Clearly, the measure of AL used in these analyses should be seen as an initial, and only partial, assessment of the concept. AL has been conceptualized as a marker of cumulative biological dysregulation across all major regulatory systems; contributing factors include dysregulation in both system state (or usual) levels of activity and in system dynamics. Although our measure of AL included information on metabolism, cardiovascular activity and both HPA and sympathetic nervous system activity, it does not include information on such things as inflammatory markers, clotting factors, or immune function. Component measures are also snapshots of levels of activity; there is no information on system dynamics.

A second potential limitation of the MacArthur study is the fact that our cohort was selected at baseline to represent the top third of those aged 70–79 in terms of their physical and cognitive functioning. As we have demonstrated in previous analyses (6),

the resulting cohort has a more restricted range of AL scores than do groups representing those individuals who did not meet criteria for enrollment into the cohort (i.e., those with greater impairments). Thus, although our cohort does exhibit a range of scores on AL that are associated with significant differences in risk for major health outcomes, it is possible that because of our relatively more restricted range of baseline AL scores we have underestimated the impact of AL on health risk. Notably, despite these limitations, our initial summary measure does provide evidence to support the value of a more cumulative approach to assessment of biological burden. This evidence is particularly strong for two of our outcomes, namely, mortality and declines in physical performance.

Further evaluation of the concept of AL will require attention to a number of issues related to measurement. First, what would be the *a priori* set of component measures needed to comprehensively assess AL (i.e., which regulatory systems and which parameters of those regulatory systems, including both system dynamics and system state/level assessments)? As noted above, our existing measure does not include assessments of hemostatic or inflammatory parameters, or assessment of parasympathetic activity; nor does it encompass assessments of system dynamics. A second issue concerns the procedures for developing optimal operational measures of AL from these component measures. In this vein, alternatives to our initial, summated measure where each component had equal weight and “higher risk” was uniformly designated by values in the uppermost or lowermost quartile need to be explored. Examples of alternatives include use of recursive partitioning (23) to identify nonlinear patterns of association between the various components of AL that are associated with greater versus lesser health risk. In addition, parameters associated with different biological systems should be weighted differentially according to the health outcomes to be predicted. Construction of such weighting systems based on canonical correlation analyses has been done (A. Karlamangla, T.E.S., B.H.S., J.W.R., and B.S.M., unpublished work).

Finally, the concept of AL postulates that a more comprehensive, cumulative measure of biological wear and tear will

provide a better index of health risk over the life course, better that is, in comparison to more traditional approaches that assess biological risks in terms of individual parameters. Further development of this concept, however, requires clearer delineation of the cascade of events (both within and across systems) through which AL develops, including specification of the relationships among regulatory systems and the pattern of cumulating dysregulation (e.g., what are most frequently the initial markers of dysregulation and how does this ripple outward to other systems and parameters?). Along these lines, a reformulation of AL that more explicitly addresses the cascade of events by delineating three levels of events that are seen as contributing to AL (24) has been developed. At the first level are primary mediators, e.g., chemical messengers such as C, NE, and EPI. A second level of events is represented by what we refer to as primary effects, e.g., cellular events, like enzymes, receptors, ion channels, or structural proteins induced genomically or phosphorylated via second messenger systems, that are regulated as part of allostasis by the primary mediators. A third level is represented by secondary mediators, e.g., more integrated processes such as blood pressure, heart rate, and metabolic profiles that are the outcome of effects of one or more primary mediators. The final outcome of this cascade of effects would be the actual disease outcomes.

The analyses completed to date suggest that the concept of AL offers considerable insight into the cumulative risks to health from biological dysregulation across multiple regulatory systems. Further specification of the roles of primary and secondary mediators is likely to offer greater insight into the cascade of events leading from allostasis to AL. This should allow us to more sharply delineate the health consequences of cumulating AL.

Work on this article was supported by National Institute on Aging Grants AG-00586 and AG 17056 (to T.E.S.) and by the MacArthur Research Network on Successful Aging and the MacArthur Research Network on Socioeconomic Status and Health through grants from the John D. and Catherine T. MacArthur Foundation.

- Jaur, L. & Stoddard, S. (1999) *Charitbook on Women and Disability in the U.S.* (U. S. National Institute on Disability and Rehabilitation Research, Washington, DC).
- McEwen, B. S. & Stellar, E. (1992) *Arch. Int. Med.* **153**, 2093–2101.
- McEwen, B. S. (1998) *N. Eng. J. Med.* **338**, 171–179.
- McEwen, B. S. & Seeman, T. E. (1999) *Ann. N.Y. Acad. Sci.* **896**, 30–47.
- Sterling, P. & Eyer, J. (1988) in *Handbook of Life Stress, Cognition, and Health*, eds. Fisher, S. & Reason, J. (Wiley, New York), pp. 631–651.
- Seeman, T. E., Singer, B., Rowe, J. W., Horwitz, R. & McEwen, B. S. (1997) *Arch. Int. Med.* **157**, 2259–2268.
- Reaven, G. M. (1988) *Diabetes* **27**, 1595–1607.
- Berkman, L. F., Seeman, T. E., Albert, M., Blazer, D., Kahn, R., Mohs, R., Finch, C., Schneider, E., Cotman, C., McClearn, G., et al. (1993) *J. Clin. Epidemiol.* **46**, 1129–1140.
- Bjorntorp, P. (1987) *Acta Med. Scand.* **723**, 121–134.
- Svec, S. & Lopez, S. A. (1989) *Lancet* **ii**, 1335–1336.
- Hypertension Detection and Follow-up Program Cooperative Group (1978) *J. Chronic Dis.* **31**, 651–667.
- Lohman, T. G., Roche, A. F. & Martorell, R. (1988) *Anthropometric Standardization* (Human Kinetics Books, Champaign, IL).
- Little, R. R., England, J. D., Wiedmyer, H. M. & Coldstein, D. E. (1983) *Clin. Chem.* **29**, 1080–1082.
- Canalis, E., Reardon, G. & Caldarella, A. (1982) *Clin. Chem.* **28**, 2418–2420.
- Krstulovic, A. M. (1983) *J. Chromatogr.* **22**, 1–34.
- Bush, T. L., Miller, S. R., Golden, A. L. & Hale, W. E. (1989) *Am. J. Pub. Health* **79**, 1554–1556.
- Inouye, S., Albert, M., Mohs, R. C., Sun, K. & Berkman, L. F. (1993) *J. Gerontol.* **48**, M146–M151.
- Seeman, T. E., Charpentier, P. & Berkman, L. F. (1994) *J. Gerontol.* **49**, M97–M108.
- Kendall, M. G. & Stuart, A. (1986) *The Advanced Theory of Statistics* (Charles Griffin, London), Vol. 2.
- SAS Institute (1996) *SAS/STAT Guide for Personal Computers* (SAS Institute, Cary, NC), Version 6.
- Tracy, R. (1995) *Arterioscler. Thromb. Vasc. Biol.* **15**, 1269–1279.
- Ridker, P., Cushman, M., Stampfer, M., Tracy, R. & Hennekens, C. (1997) *N. Engl. J. Med.* **336**, 973–979.
- Zhang, H. & Singer, B. (1999) *Recursive Partitioning in the Health Sciences* (Springer, New York).
- McEwen, B. S. & Seeman, T. E. (1999) *Ann. N.Y. Acad. Sci.* **896**, 30–47.