Article

The contribution of three dimensions of allostatic load to racial/ethnic disparities in poor/fair self-rated health

Alexis R. Santos-Lozada*, Jonathan Daw

Department of Sociology and Criminology, Population Research Institute, Pennsylvania State University, 211 Oswald Tower, University Park, PA 16802, USA

ARTICLE INFO

ABSTRACT

Objective: This study evaluates whether different dimensions of physiological dysregulation, modeled individually rather than additively mediate racial/ethnic disparities in self-reported health.

Methods: Using data from the National Health and Nutrition Examination Survey (2005–2010) and the Karlson, Hold, and Breen (KHB) mediation model, this paper explores what operationalization of biomarker data most strongly mediate racial/ethnic disparities in poor/fair self-rated health (SRH) among adults in the United States, net of demographic, socioeconomic, behavioral, and medication controls.

Results: Non-Hispanic blacks and Hispanics had significantly higher odds of reporting poor/fair self-rated health in comparison to non-Hispanic whites. Operationalizations of allostatic load that disaggregate three major dimensions of physiological dysregulation mediate racial/ethnic disparities strongly between non-Hispanic blacks and non-Hispanic whites, but not between Hispanics and non-Hispanic whites. Disaggregating these dimensions explains racial/ethnic disparities in poor/fair SRH better than the continuous score. Analyses on sex-specific disparities indicate differences in how individual dimensions of allostatic load contribute to racial/ethnic disparities in poor/fair SRH differently. All individual dimensions are strong determinants of poor/fair SRH for males. In contrast, for females, the only dimension that is significantly associated with poor/fair SRH is inflammation. For the analytic sample, additive biomarker scores fit the data as well or better than other approaches, suggesting that this approach is most appropriate for explaining individual differences. However, in sex-specific analyses, the interactive approach models fit the data best for men and women.

Conclusions: Future researchers seeking to explain racial/ethnic disparities in full or sex-stratified samples should consider disaggregating allostatic load by dimension.

Introduction

Biomarker data are widely used in population health research, especially in the study of the concept of allostatic load. Most previous research on this topic employs a continuous score approach which indicates how many of a participant's biomarkers exceed a given threshold that is either clinically or empirically determined. This continuous score is intended to capture multi-system physiological dysregulation by combining numerous biomarkers from different biological systems. However, previous research largely employs canonical methods for AL score construction without investigating what operationalization of the underlying biomarkers best serves their research purposes. The continuous score strategy has yielded many research insights, but we argue that a reexamination of this operationalization is timely and appropriate.

This paper reexamines the continuous score strategy by comparing it to alternative operationalizations of the same biomarkers by comparing their explanatory power for individual differences and racial/ethnic disparities in self-rated health (SRH) for a nationally representative sample of adults in the United States. We compare the continuous score measure to dichotomous variables indicating clinically significant biomarker values in the cardiovascular, metabolic, and inflammatory systems, and model the effects of dysregulation in each system both additively and interactively. To our knowledge there is no research investigating whether different specifications of allostatic load mediate racial/ethnic differences in SRH. To achieve this goal, identifying the appropriate operationalization of AL is critical to understanding individual differences and racial/ethnic disparities in this key health measure.

Background

Allostatic load: concept and measurement

Although the biomarkers used in allostatic load (AL) scores are biological measurements, individual and group differences therein are heavily linked to variations in the social environment. The social determinants of health framework argues that economic, early life, social context, environmental conditions, as well as individual characteristics and behaviors affect health outcomes such as morbidity, health status, functional limitations, healthcare expenditures and mortality (Macgregor, 1961). Exposure to unequal socioeconomic and environmental conditions, paired with individual characteristics, has been associated with homeostatic imbalance (McEwen, 1998), which triggers processes within the body that aim to correct this imbalance (McEwen, 1998; McEwen & Wingfield, 2003). The process of allostasis leads to the adaptation of the organism to these unequal conditions; the lasting effect of adaptation accumulates in the body through "wear and tear" (McEwen & Wingfield, 2003). Researchers have begun to reveal links between the concept of this "wear and tear", or allostatic load (AL), and a wide variety of health outcomes. AL scores are constructed from a variety of biomarkers to summarize the resulting burden of continuing internal processes which aim to attain or maintain stability within the body under stressful conditions (McEwen & Seeman, 1999).

Many AL studies and conceptual models consider three major dimensions of physiological dysregulation, using cardiovascular (CM), metabolic (MM, including anthropometric measures), and inflammatory (IM) biomarkers (McEwen, 1998; Doung, Bingham, Aldana, Chung & Summer, 2017; Morrison, Shenassa, Mendola, Wu & Schoendorf, 2013; Juster, McEwen & Lupien, 2010). Every approach begins with a set of biomarkers, which are then converted to a more informative value either by dichotomizing the underlying value compared to a clinically- or empirically-significant threshold, or by converting the values to a standardized distribution. These converted values are then summed together into a continuous AL score, typically without distinguishing between the biological systems involved, which may limit their explanatory power if each dimension does not contribute equally to health outcomes for individuals or these associations vary by race/ethnicity.

A recent review of the methods employed to construct allostatic load scores (Doung et al., 2017) indicates that researchers vary substantially in which biomarkers are used to construct these indexes of biological dysregulation. Although the number of biomarkers used for each score varies by study, with number of biomarkers considered ranging from 7 to 14, all of them include markers from the aforementioned dimensions. Despite the pervasive use of multiple biological dimensions in the construction of these continuous scores, relatively little research explores how similarly each system influences individual differences and racial/ethnic disparities in these health outcomes.

Racial/ethnic and sex disparities in self-rated health

In the United States (U.S.), racial/ethnic disparities are frequently documented for SRH, as higher proportions of non-Hispanic Blacks (NH Blacks) and Hispanics report poor or fair health when compared to non-Hispanic Whites (NH Whites) (Woo & Zajacova, 2016; Borrell & Dallo, 2008), a difference that remains strong even when models are adjusted for social status, access to healthcare services, and health behaviors (Lo, Howell & Cheng, 2013). Moreover, NH Blacks-NH White differences exist for the majority of health outcomes. NH Blacks have been found to have higher mortality rates (Levine, Foster & Fullilove, 2001), disability rates (Fuller-Thomson, Nuru-Jeter, Minkler & Guralnik, 2009; Hayward, Hummer, Chiu, González-González & Wong, 2014), lower life expectancy (Harper, MacLehose & Kaufman, 2014; Elo, Beltrán-Sánchez & Macinko, 2014), higher rates of engagement in risky health behaviors (Kawachi, Kennedy & Glass, 1999), and lower levels of engagement in exercise or healthy diets (August & Sorkin, 2011). In a study of the reliability of SRH measures, where respondents reported SRH on 2 occasions (about 1 month apart), NH Blacks were more likely than NH Whites to change their SRH answer and report worse health status (Zajacova & Dowd, 2011). Most recent approaches to understanding the NH Black-NH White gap in SRH have incorporated controls for period and cohorts (Beck, Finch, Lin, Hummer & Masters, 2014), wealth (Hajat, Kaufman, Rose, Siddiqi & Thomas, 2011), health conditions (Banerjee, Perry, Tran & Arafat, 2010), and contextual variables (Subramanian, Acevedo-Garcia & Osypuk, 2005; Bjornstrom & Kuhl, 2014) (i.e. residential segregation, percent NH Black within the county, etc.) but none of these have been able to eliminate the NH Black-NH White disparity.

The difference between NH Whites and Hispanics continues to puzzle researchers as the latter group has been found to have lower or similar mortality rates (Markides & Coreil, 1986), infant mortality risk (Hummer, Powers, Pullum, Gossman & Frisbie, 2007), poor/fair self-rated mental health (Santos-Lozada, 2016), self-reported hypertension among Hispanic-Whites (Borrell, 2009), and low birth-weights (Johnelle Sparks, 2009) when compared to NH Whites. This pattern has been termed the epidemiological paradox (Markides & Coreil, 1986) because it is inconsistent with these groups' respective socioeconomic positions in US society. SRH is one of the few outcomes where evidence that contradicts this paradox is present (Dubard & Gizlice, 2008; Viruell-Fuentes, Morenoff, Williams & House, 2011; Kandula, Lauderdale & Baker, 2007), however. The contrast between the usual pattern of Hispanic health advantage and poorer SRH has been referred to as the "Latino health puzzle" (Viruell-Fuentes et al., 2011). Numerous factors have been hypothesized to explain this puzzle, including language of interview (Dubard & Gizlice, 2008; Kandula et al., 2007), rating health based on different factors (Bzostek, Goldman & Pebley, 2007), socioeconomic and cultural influences (Kandula et al., 2007; Markides & Martin, 1979), and contextual effects (Patel, Eschbach, Rudkin, Peek & Markides, 2003), among others (Bzostek et al., 2007).

Furthermore, in the U.S., differences in health status by sex are well documented, as women tend to report higher poor/fair SRH when compared to males despite incorporating controls for demographic and socioeconomic characteristics (Prus, 2011). Because of this, a growing body of literature has started to pursue analyses of SRH and its determinants stratifying by sex. Two reasons for doing so are particularly salient here. First, men have lower odds of reporting poor/fair SRH than females but lower life expectancy (Gorman & Read, 2006; Zajacova, Huzurbazar & Todd, 2017; Oksuzyan et al. 2009; Case & Paxson, 2005). This pattern has been termed the male-female health-survival paradox and has been roughly translated to imply "men die, women suffer". Others have indicated that "females are sicker, but males die sooner" (Arber & Cooper, 1999). In light of gender differences in these later life outcomes, it is likely that allostatic load and its individual dimensions differentially explains individual and racial/ethnic differences in SRH by gender.

Second, poor/fair SRH predicts mortality better for males than for females (Hirve, Juvekar & Sambhudas, 2012; Ross, Masters & Hummer, 2012). Given that the predictive power of SRH varies by sex, it may be possible that SRH is capturing different elements of subjective health and this produces the differences in reporting SRH. Because differences exist in SRH reporting (i.e. male-female health survival paradox) and the difference in predictive power for subsequent mortality varies by sex, a sex-specific analyses is deemed both appropriate and essential to better understand the contribution of the dimensions of allostatic load to SRH. In summary, given significant differences by sex in self-rated health, biomarkers, and potentially the relationship between them, we pursue this sex-specific analysis to examine whether our findings are comparable to those found in the complete analytical sample or not.
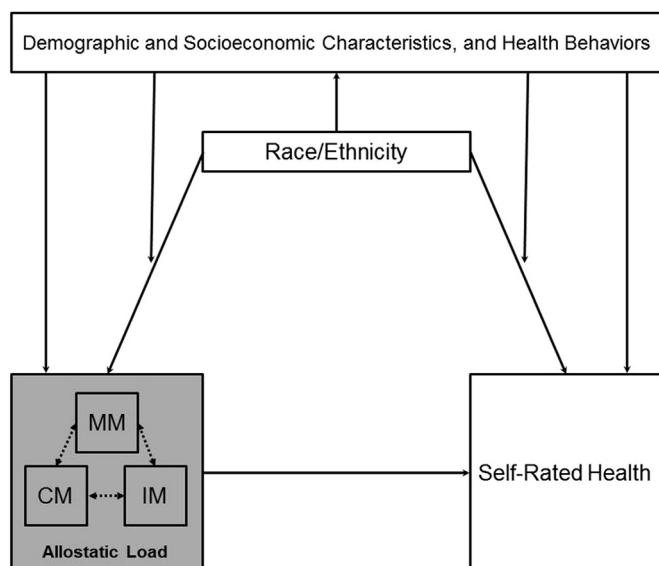
Fig. 1. Conceptual model. CM, cardiovascular markers; MM, metabolic markers; IM, inflammation markers (adapted from Daw (2017)).

The present study

Despite the extensive literature on racial/ethnic differences in poor/fair SRH, no previous study has explored whether allostatic load contribute similarly to racial/ethnic disparities in self-rated health status and whether differences exist in this contribution by sex. Moreover, previous research has not investigated the potentially *separate* contributions of the dimensions of allostatic load to these differences, nor the potential for *interactions* between them (in which the effect of one AL dimension depends on another). In other words, standard approaches to constructing allostatic load scales implicitly assume that the effects of each measure and dimension are uniform and independent. In contrast to this usual approach, we investigate whether the effect of one form of physiological dysregulation on self-rated health and racial disparities therein contributes separately or in a manner dependent on whether one also has other indicators of physiological dysregulation.

Fig. 1 illustrates the approach of this paper to racial/ethnic disparities in odds of reporting poor/fair SRH. Based on previous research discussed above, we hypothesize that race/ethnicity is associated with the odds of reporting fair/poor self-rated health, and that this association is potentially confounded by a wide range of demographic, socioeconomic, and health behavior characteristics, and medication use that are unequally distributed by race/ethnicity. Furthermore, race/ethnicity is associated with differential risk by allostatic load dimensions, which we hypothesize mediates the race-SRH association. Finally, we hypothesize that this mediation is potentially better explained by interactions between the three dimensions of allostatic load, as represented by the dashed lines connecting CM, MM, and IM in this figure. Thus, the objective of this paper is to determine whether disaggregated dimensions of allostatic load better explain individual and racial/ethnic patterns of poor/fair self-rated health, individually or interactively. Accordingly, this paper tests four hypotheses:

1. There are racial disparities in the prevalence of the dimensions of allostatic load, individually and in combination.
2. These disaggregated dimensions will more strongly mediate racial/ethnic disparities and explain individual differences in poor/fair SRH than their additive effects alone.
3. Allowing these disaggregated dimensions to interact will more strongly mediate racial/ethnic disparities and explain individual differences in poor/fair SRH than their independent effects.
4. The degree to which biomarkers mediate racial/ethnic disparities

and explain individual differences in poor/fair self-rated health will differ by sex.

Data and methods

Data

Data for this analysis come from the 2005–2010 survey years of the National Health and Nutrition Examination Survey (NHANES). We limit our investigation to the 2005-10 waves of NHANES for two reasons. First, NHANES no longer collected C-reactive protein data after the 2009-10 wave. Second, we sought a sample with sufficient sample size to test our hypotheses on the one hand, but would not be confounded with unmodeled period differences in the prevalence of these biomarkers and their effects over time.

The NHANES, conducted by the National Center for Health Statistics (NCHS), uses stratified, multistage probabilistic sampling to provide national estimates of health and nutritional status for the civilian, non-institutionalized population of the United States (Johnson, Paulose-Ram & Ogden, 2013). For this paper we used questionnaire data as well as clinically-assessed markers of physiological activity within the body. The analytical sample includes Non-Hispanic Whites, Non-Hispanic Blacks and Hispanics ages 25 and older at the time of interview. The survey years selected includes 12,757 individuals with valid information for the analyzed variables.

The exploratory data analysis for each biomarker revealed the existence of missing values. We employed a generalized regression-based methods to impute missing values in the biomarker data only, following the approach found in previous studies using NHANES (Howard and Sparks, 2016; Howard & Sparks, 2015; Howard & Sparks, 2016). These models incorporated controls for age, race/ethnicity, education, sex, marital status and income to predict a single value for each respondent with a missing value for each specific biomarker. Further analysis of biomarker data indicated no significant differences between initial and post-imputation measures of central tendency and dispersion. A detailed explanation of the non-imputed and imputed values distributions are presented in Table 1. Additionally, sensitivity analysis indicated that the imputation of missing cases did not substantively change the results reported. Missing data in other measures were handled either via listwise deletion or dummy variable adjustment (when so noted in the next section).

Measures

SRH is the dependent variable for this study, and is measured as a dichotomous variable indicating poor/fair SRH, following the usual practice (Subramanian et al., 2005; Manor, Matthews & Power, 2000; Acevedo-Garcia, Bates, Osypuk & McArdle, 2010). As a robustness check, we also fit all models specifying SRH as a continuous variable (not shown) and obtained comparable results to the dichotomous measure. *Allostatic load* is measured both additively and in three dimensions: CM, MM, and IM. CM indicators considered in this study include: diastolic blood pressure (mmHg), systolic blood pressure (mmHg) and pulse rate at 60 seconds. MM indicators considered include: total cholesterol (mg/dL), HDL cholesterol (mg/dL), triglycerides (mg/dL) and glycohemoglobin (%) and body mass index (BMI, measured as kg/m^2). Finally, IM indicators considered in this study include: albumin (g/dL) and C-reactive protein (mg/dL). AL components will be operationalized using the clinically derived cutoffs described in Table 1, such that the individual is designated to be at risk if they have biomarkers exceeding the listed thresholds. The clinical threshold cut-off points for each biomarker used for each operationalization of allostatic load (summation or dimensions) for each individual as discussed thoroughly in previous literature (Howard & Sparks, 2016; Crimmins, Kim & Seeman, 2009).

Table 1 presents the distribution of biomarkers with and without

Table 1
Weighted descriptive statistics for 10 allostatic load biomarkers with and without imputation, NHANES 2005–2010.

Biomarker	Without imputed values			With imputed values			Clinically based AL		
	n	Mean	S.E.	n	Mean	S.E.	Threshold	Prevalence without imputation (%)	Prevalence with imputation (%)
Cardiovascular markers									
Diastolic blood pressure (mmHg)	12,508	70.78	0.30	12,757	70.77	0.30	≥ 90	5.07	4.99
Systolic blood pressure (mmHg)	12,508	123.02	0.29	12,757	123.01	0.29	≥ 140	15.52	15.30
Pulse rate at 60 seconds	12,560	72.54	0.19	12,757	72.55	0.19	≥ 90	9.14	9.02
Metabolic markers									
Total cholesterol (mg/dL)	12,163	200.15	0.58	12,757	200.06	0.57	≥ 240	16.01	15.41
HDL cholesterol (mg/dL)	12,204	53.16	0.27	12,757	53.18	0.27	< 40	19.72	19.03
Triglycerides (mg/dL)	5,912	137.62	1.90	12,757	137.90	1.01	≥ 150	29.66	31.29
Glycohemoglobin (%)	12,279	5.60	0.02	12,757	5.61	0.02	≥ 6.4	8.02	7.87
Body Mass Index (kg/m^2)	12,606	29.02	0.10	12,757	29.02	0.10	≥ 30	36.57	36.38
Inflammation markers									
Albumin (g/dL)	12,165	4.25	0.01	12,757	4.25	0.01	< 3.8	5.83	5.61
C-reactive protein (mg/dL)	12,248	0.42	0.01	12,757	0.42	0.01	≥ 0.03	35.56	37.48

Survey Design: Sampling Unit = SDMVPSU, Stratum = SDMSTRA Weight = WTMEC6YR

Table 2
Descriptive statistics for overall population and by race/ethnicity, NHANES 2005–2010 (n = 12,757).

	Overall Population		Non-Hispanic White		Non-Hispanic Black		Hispanic	
	Mean/Prop.	95% C.I.	Mean/Prop.	95% C.I.	Mean/Prop.	95% C.I.	Mean/Prop.	95% C.I.
Poor/Fair Self-Rated Health	0.18	0.16–0.19	0.14	0.12–0.15	0.27	0.24–0.29	0.32	0.30–0.35
Allostatic Load Score (additive)	1.82	1.77–1.88	1.77	1.70–1.83	1.96	1.87–2.05	2.05	1.97–2.13
Allostatic Load Dimensions								
Cardiovascular	0.24	0.23–0.52	0.24	0.22–0.25	0.32	0.30–0.34	0.20	0.18–0.22
Metabolic	0.65	0.64–0.67	0.64	0.62–0.66	0.62	0.60–0.65	0.75	0.73–0.76
Inflammation	0.39	0.38–0.40	0.36	0.35–0.38	0.54	0.52–0.56	0.41	0.38–0.44
Race/Ethnicity								
Non-Hispanic White	0.76	0.72–0.80	–	–	–	–	–	–
Non-Hispanic Black	0.11	0.09–0.13	–	–	–	–	–	–
Hispanic	0.13	0.10–0.15	–	–	–	–	–	–
Age Group								
25–40 years	0.31	0.30–0.33	0.28	0.26–0.29	0.36	0.33–0.39	0.49	0.46–0.51
41–60 years	0.43	0.42–0.44	0.43	0.52-0.45	0.45	0.43–0.47	0.38	0.36–0.40
61 years and older	0.26	0.24–0.28	0.29	0.27–0.31	0.20	0.18–0.21	0.14	0.12–0.15
Education								
Less than High School	0.19	0.17–0.21	0.13	0.11–0.15	0.26	0.23–0.29	0.48	0.46–0.51
High School/Some College	0.54	0.53–0.56	0.56	0.54–0.58	0.58	0.55–0.60	0.41	0.39–0.43
College Degree or higher	0.27	0.25–0.29	0.31	0.28–0.34	0.17	0.15–0.19	0.11	0.09–0.13
Marital Status								
Never Married	0.11	0.10–0.12	0.09	0.08–0.10	0.23	0.21–0.25	0.13	0.11–0.15
Married	0.61	0.59–0.63	0.65	0.63–0.66	0.40	0.36–0.43	0.59	0.56–0.62
Divorced/Separated/Widowed	0.21	0.19–0.22	0.20	0.19–0.22	0.28	0.26–0.30	0.18	0.16–0.20
Cohabitating	0.07	0.06–0.08	0.06	0.05–0.07	0.09	0.08–0.11	0.10	0.08–0.11
Family income								
Less than $20,000	0.16	0.14–0.17	0.13	0.11–0.14	0.25	0.22–0.27	0.25	0.22–0.28
$20,000-$64,999	0.45	0.43–0.46	0.44	0.41–0.45	0.48	0.46–0.51	0.51	0.48–0.53
$65,000 or more	0.37	0.35–0.39	0.43	0.39–0.45	0.23	0.20–0.27	0.18	0.16–0.21
Don't know/Refused	0.03	0.02–0.03	0.02	0.02–0.03	0.04	0.03–0.05	0.06	0.05–0.07
Smoking								
Non-smoker	0.51	0.49–0.53	0.49	0.47–0.51	0.56	0.53–0.59	0.62	0.60–0.65
Current smoker	0.22	0.20–0.23	0.22	0.20–0.23	0.27	0.24–0.30	0.02	0.16–0.20
Former smoker	0.27	0.26–0.28	0.30	0.28–0.31	0.17	0.16–0.19	0.19	0.18–0.21
Drinking Habits								
Non-drinker	0.28	0.26–0.30	0.26	0.23–0.29	0.37	0.34–0.40	0.33	0.30–0.35
1 drink per week	0.45	0.43–0.46	0.44	0.42–0.46	0.42	0.40–0.44	0.52	0.50–0.54
Over 1 drink per week	0.27	0.25–0.29	0.30	0.27–0.32	0.21	0.19–0.23	0.15	0.14–0.17
Medication use								
Cholesterol	0.18	0.17–0.19	0.20	0.19–0.22	0.15	0.13–0.16	0.10	0.08–0.12
Diabetes	0.08	0.07–0.09	0.07	0.06–0.08	0.14	0.12–0.15	0.09	0.07–0.10
Heart	0.27	0.26–0.29	0.29	0.27–0.31	0.32	0.29–0.34	0.14	0.12–0.16
Unweighted n	12,757		6,658		2,657		3,442	

Survey Design: Sampling Unit = SDMVPSU, Stratum = SDMSTRA Weight = WTMEC6YR

imputed values by dimension, clinically determined thresholds, and the percentage above clinical thresholds within the analytic sample. The analysis of individual markers indicated no statistically significant differences between full-case and imputed values. Additionally, analyses of cases with information for all the biomarkers indicated that the exclusion of missing cases did not substantively change the results reported. The percentage exceeding clinical thresholds for each marker ranged between 4.99% (Diastolic blood pressure) and 37.48% (C-reactive protein).

Race/ethnicity was measured in three categories: NH Whites, NH Blacks and Hispanics of any race – members of other racial/ethnic groups were excluded from the analysis. *Sex* was measured dichotomously indicating whether the respondent was a male (reference category) or a female, according to self-report. *Age* was recoded into three categories: 25–40 years (reference), 41–60 years and 61 years and older, in order to flexibly model the functional form between age and self-rated health. No substantial differences were observed in the main associations explored in this article whenever age was specified as a continuous variable or using different categorical specifications (i.e. five or ten years age-groups, comparative models are included in Table S1 in Appendix 1). *Education* was measured by self-report and recoded into three categories: less than high school (reference), high school/ some college, or college degree or higher. *Marital status* was measured by self-reports and recoded into four categories: never married (reference), married, cohabitating, or divorced/separated/widowed. *Family income* was measured as a categorical variable from self-reports and recoded into four categories: less than $20,000 (reference), $20,000-$64,999, $65,000 or more, and don't know/refused. Analyses which eliminated don't know/refused observations from the analytical sample through listwise deletion yielded substantively identical results. We incorporated two health behaviors: smoking and drinking. Smoking was measured as a categorical variable indicating whether the respondent was a non-smoker (reference), current smoker, or former smoker. Drinking status was also measured in three categories indicating frequency of drinking reported by the respondent: non-drinker (reference), one drink per week, or more than one drink per week. Finally, we incorporate controls for medication use for cholesterol, diabetes and heart conditions. For cholesterol, we considered the use of antihyperlipidemic agents (Gu, Paulose-Ram, Burt & Kit, 2017); and for diabetes, we considered the use of antidiabetic agents. In the case of heart conditions, we identified respondents who reported use of cardiovascular agents; except for diuretics, vasodilators and pulmonary hypertension medication.

Statistical analysis

The analysis proceeded in three stages. First, Tables 2 and 3 describe the distribution of the variables included in this analysis, presented for the overall sample and separately and by race/ethnicity (Table 2), and by sex (Table 3). Second, the weighted percentage of the sample, overall and by race/ethnicity and sex, of all eight possible AL dimension combinations are presented in Tables 4 and 5, including chi-square tests for differences in the percentage with these outcome combinations.

We estimate the contribution of AL to racial/ethnic disparities in poor/fair SRH using a variety of operationalizations for AL, employing a mediation model described below. Using these clinically significant thresholds described above, we will model the effects of these indicators of allostatic load in four different ways, then compare how well each approach explains individual differences and racial/ethnic disparities in SRH:

1) Following research convention, we will construct a 10-point scale of AL, counting the number of biomarkers for which the respondent exceeds clinically significant thresholds. We refer to this as the *continuous AL score* approach.
2) To allow for the possibility that the effects of AL biomarkers vary by

sub-dimension (CM, MM, and IM), we will construct three indicator variables that equal one if at least one of the biomarkers for the dimension in question exceeds the clinically relevant threshold, and equals zero otherwise. We will then simultaneously estimate the associations of each with SRH. We refer to this as the *additive dummy variable* approach.
3) To allow for the possibility that the dummy variables for biological risk described in 2) have interdependent effects, we will combine them in two different ways:
   a. We will create an 8-category variable for each combination of these three dummy variables: no diseases, CM only, MM only, IM only, CM+MM, CM+IM, MM+IM, and CM+MM+IM. The effects of each category of biological risk will be compared to the reference category of no clinically significant biomarkers. We refer to this as the *categorical combinations* approach.
   b. We will specify their effects in three-interactions between CM,

Table 3
Descriptive statistics for overall population by sex, NHANES 2005–2010 (n = 12,757).

	Male		Female	
	Mean/ Prop.	95% C.I.	Mean/ Prop.	95% C.I.
Poor/Fair Self-Rated Health	0.17	0.16–0.18	0.18	0.17–0.20
Allostatic Load Score (additive)	2.00	1.93–2.06	1.66	1.60–1.71
Allostatic Load Dimensions				
Cardiovascular	0.23	0.22–0.25	0.25	0.24–0.26
Metabolic	0.75	0.73–0.76	0.56	0.54–0.58
Inflammation	0.32	0.30–0.34	0.46	0.44–0.47
Race/Ethnicity				
Non-Hispanic White	0.76	0.72–0.79	0.76	0.73–0.80
Non-Hispanic Black	0.11	0.09–0.12	0.12	0.10–0.14
Hispanic	0.14	0.11–0.16	0.12	0.09–0.14
Age Group				
25–40 years	0.33	0.31–0.35	0.30	0.28–0.31
41–60 years	0.43	0.42–0.45	0.42	0.41–0.44
61 years and older	0.24	0.22–0.25	0.28	0.26–0.30
Education				
Less than High School	0.19	0.17–0.21	0.18	0.17–0.20
High School/Some College	0.53	0.52–0.56	0.55	0.53–0.56
College Degree or higher	0.27	0.24–0.30	0.27	0.24–0.29
Marital Status				
Never Married	0.12	0.11–0.13	0.10	0.09–0.11
Married	0.65	0.64–0.67	0.57	0.55–0.59
Divorced/Separated/ Widowed	0.15	0.14–0.16	0.27	0.25–0.29
Cohabitating	0.08	0.07–0.09	0.06	0.05–0.07
Family income				
Less than $20,000	0.14	0.12–0.15	0.18	0.16–0.19
$20,000-$64,999	0.45	0.43–0.47	0.45	0.43–0.46
$65,000 or more	0.39	0.37–0.42	0.35	0.33–0.38
Don't know/Refused	0.03	0.02–0.03	0.03	0.02–0.03
Smoking				
Non-smoker	0.45	0.42–0.47	0.57	0.56–0.59
Current smoker	0.24	0.22–0.26	0.19	0.18–0.21
Former smoker	0.31	0.30–0.33	0.23	0.22–0.25
Drinking Habits				
Non-drinker	0.23	0.21–0.25	0.33	0.30–0.36
1 drink per week	0.42	0.40–0.44	0.48	0.45–0.50
Over 1 drink per week	0.35	0.33–0.37	0.19	0.22–0.25
Medication use				
Cholesterol	0.20	0.19–0.21	0.17	0.15–0.18
Diabetes	0.27	0.25–0.28	0.08	0.07–0.09
Heart	0.08	0.07–0.09	0.28	0.26–0.30
Unweighted n	6,430		6,327	

Survey Design: Sampling Unit = SDMVPSU, Stratum = SDMSTRA Weight = WTMEC6YR

Table 4
Cross-tablulation of allostatic load dimensions clusters by race/ethnicity and significance test, NHANES 2005–2010 (n = 12,757).

Dimension 1	Dimension 2	Dimension 3	All	Non-Hispanic White	Non-Hispanic Black	Hispanic	NH White-NH Black Chi-Square (p-value)	NH White-Hispanic Chi-Square (p-value)
None	–	–	21.94	23.23	18.85	16.91	9.37 (0.002)	17.99 (< 0.0001)
CM	–	–	4.20	4.40	4.87	2.37	1.01 (0.31)	11.82 (0.0006)
MM	–	–	26.60	27.42	15.66	31.56	1,036.72 (< 0.0001)	8.56 (0.003)
IM	–	–	6.40	6.14	9.61	5.06	21.72 (< 0.0001)	2.77 (0.096)
CM	MM	–	8.25	8.50	6.73	8.13	8.09 (0.005)	0.25 (0.619)
CM	IM	–	2.25	2.13	4.32	1.12	43.26 (< 0.0001)	13.42 (0.0002)
MM	IM	–	20.90	19.52	24.19	26.31	14.75 (0.0001)	34.16 (< 0.0001)
CM	MM	IM	9.45	8.66	15.78	8.53	56.78 (< 0.0001)	0.04 (0.835)
Unweighted n			12,757	6,658	2,657	3,442		

Survey Design: Sampling Unit = SDMVPSU, Stratum = SDMSTRA Weight = WTMEC6YR

MM, and IM, along with their additive effects. We refer to this as the *dummy variable interaction* approach.

Dummy categorical combinations (3a) and the variable interaction (3b) approaches will yield identical model fit statistics and predicted probabilities of poor/fair self-rated health for each racial/ethnic group, but different coefficients and statistical significance tests. These hypothesis tests differ subtly – the dummy variable interaction approach calculates the interactive association and tests the statistical significance of combinations of risk markers against that expected from their additive effects alone. The categorical combinations approach calculates the association and tests the statistical significance of combinations of risk markers against that associated with having no risk markers. We believe that both hypotheses tests yield valuable information. However, because these two approaches yield identical results for the mediation tests described below and model fit indicators, we present these results together in Table 6.

Tables 6 and 8 report the results from a series of logistic regression models, where the degree to which AL mediates the race/ethnicity-SRH relationship is estimated using the Karlson, Holm and Breen (KHB) method (Karlson, Holm & Breen, 2012; Breen, Karlson & Holm, 2013; Kohler, Karlson & Holm, 2011). These are estimated using the *-khb-* and *-logit-* commands in Stata/SE, version 14 (StataCorp, 2015). The KHB method allows the estimation of mediation with a categorical dependent variable (Daw, 2017). This method decomposes the association of race/ethnicity with poor/fair SRH into the direct association and the indirect association attributable to AL, which is expressed as a percentage of the total association. For each operationalization of AL, we estimate the KHB models with three different independent variable specifications – examining the effect of AL alone, adding demographic and socioeconomic covariates, and then adding health behaviors and medications use. All models incorporate complex sample design within the calculation of point estimates, odds ratios, and standard errors. The analysis incorporated sampling weights, stratification and sampling units as recommended by the guidelines published by the National Center for Health Statistics (Johnson et al., 2013). The resulting twelve

models (the result of four AL operationalizations multiplied by three independent variable specifications) are presented in Table 6 with six pieces of information apiece: for NH Black and Hispanics, the odds ratios obtained from the full and reduced model are presented, along with estimated the percentage of that racial/ethnic group's associations mediated by AL. Regression coefficients for each AL approach for the overall sample are presented in Table 7. In addition, Table 8 contains the mediation results and associations between each of the operationalizations of allostatic load described above for the sex-specific analyses. Corresponding Akaike Information Criterion (AIC) scores for the overall and sex-stratified analyses are presented in Table 9; lower AICs are indicative of better model fit for each model specification.

Results

Descriptive findings

Table 2 describes the key characteristics of the analytic sample, overall and by race/ethnicity. Weighted means or proportions are presented for each variable, with the corresponding 95% confidence interval (CI). 18% of the sample reported poor/fair SRH. The average count of biomarkers that exceed the clinically determined threshold is 1.82 (on a 0–10 scale). Of this sample, for the three dimensions of AL, 24% of the population had at least one biomarker exceeding clinically determined thresholds for cardiovascular markers, 65% for metabolic markers, and 39% for inflammation ones. NH Blacks (27%) and Hispanics (32%) had much higher proportions with poor/fair SRH in comparison to NH Whites (14%). NH Blacks and Hispanics had higher means of additive AL compared to NH Whites. Racial/ethnic differences in the prevalence for each dimension of AL were also found. NH Blacks had highest prevalence of CM and IM, whereas Hispanics had the highest prevalence in MM. NH Whites had significantly lower levels of prevalence for CM and IM in comparison to NH Blacks.

Table 3 describes the key characteristics of the analytic sample by sex. Again, weighted means or proportions are presented for each variable, with the corresponding 95% confidence interval (CI). Males

Table 5
Cross-tablulation of allostatic load dimensions clusters by sex and significance test, NHANES 2005–2010 (n = 12,757).

Dimension 1	Dimension 2	Dimension 3	Males	Females	Male-Female Chi-Square (p-value)
None	–	–	17.49	26.19	83.18 (< 0.0001)
CM	–	–	2.98	5.37	26.17 (< 0.0001)
MM	–	–	37.41	16.25	500.24 (< 0.0001)
IM	–	–	3.46	9.22	171.74 (< 0.0001)
CM	MM	–	10.00	6.58	24.54 (< 0.0001)
CM	IM	–	1.53	2.95	39.14 (< 0.0001)
MM	IM	–	18.40	23.30	26.08 (< 0.0001)
CM	MM	IM	8.74	10.14	5.67 (0.0172)
Unweighted n			6,430	6,327	

Survey Design: Sampling Unit = SDMVPSU, Stratum = SDMSTRA Weight = WTMEC6YR

Table 6
Odds ratios from logistic regression models, NHANES 2005–2010 (n = 12,757).

	Continuous AL Score			Additive Dummy Variables		
	Simple	+ Demog and SES	+ HB and Medication	Simple	+ Demog and SES	+ HB and Medication
Non-Hispanic White (reference group)	–	–	–	–	–	–
Non-Hispanic Black						
Reduced Odds Ratios	2.33 (2.07–2.64)	1.74 (1.53–1.98)	1.56 (1.36–1.79)	2.30 (2.04–2.59)	1.73 (1.52–1.97)	1.56 (1.36–1.79)
Full Odds Ratios	2.17 (1.93–2.45)	1.69 (1.49–1.93)	1.58 (1.37–1.81)	1.96 (1.74–2.22)	1.56 (1.37–1.79)	1.45 (1.26–1.67)
% Mediated	8.50	4.81	-2.04	18.85	18.27	16.15
Hispanics						
Reduced Odds Ratios	3.17 (2.82–3.55)	2.11 (1.85–2.40)	2.37 (2.07–2.72)	3.16 (2.82–3.54)	2.11 (1.85–2.40)	2.38 (2.08–2.73)
Full Odds Ratios	2.84 (2.53–3.18)	2.02 (1.77–2.30)	2.30 (2.01–2.63)	2.93 (1.03–1.12)	2.03 (1.79–2.31)	2.31 (2.01–2.64)
% Mediated	9.47	5.85	3.65	6.36	5.04	3.61
Inclusion of Allostatic Load	Yes	Yes	Yes	Yes	Yes	Yes
Demographic and Socioeconomic Controls?	No	Yes	Yes	No	Yes	Yes
Health Behaviors Controls?	No	No	Yes	No	No	Yes
Medication Controls?	No	No	Yes	No	No	Yes

	Categorical Combinations/Dummy Variable Interaction		
	Simple	+ Demog and SES	+ HB and Medication
Non-Hispanic White (reference group)	–	–	–
Non-Hispanic Black			
Reduced Odds Ratios	2.30 (2.04–2.60)	1.73 (1.52–1.97)	1.55 (1.36–1.79)
Full Odds Ratios	1.96 (1.74–2.22)	1.56 (1.37–1.79)	1.45 (1.26–1.67)
% Mediated	19.07	18.29	16.24
Hispanics			
Reduced Odds Ratios	3.16 (2.82–3.54)	2.11 (1.85–2.40)	2.38 (2.08–2.73)
Full Odds Ratios	2.93 (2.62–3.29)	2.03 (1.79–2.31)	2.31 (2.01–2.64)
% Mediated	6.55	5.01	3.70
Inclusion of Allostatic Load	Yes	Yes	Yes
Demographic and Socioeconomic Controls?	No	Yes	Yes
Health Behaviors Controls?	No	No	Yes
Medication Controls?	No	No	Yes

Note: 95% CIs are in parentheses.
Survey Design: Sampling Unit = SDMVPSU, Stratum = SDMSTRA Weight = WTMEC6YR

and females reported poor/fair SRH in similar level, 17% and 18% respectively. Males had a higher average count of biomarkers above the clinically determined threshold higher than females with 2.00 (95% C.I. 1.93–2.06) and 1.66 (95% C.I. 1.60–1.71), respectively. 23% of the male population had at least one biomarker exceeding clinically determined thresholds for cardiovascular markers, with corresponding values of 75% for metabolic markers and 32% for inflammation markers. Among females, 25% had at least one biomarker exceeding clinically determined thresholds for cardiovascular markers, with corresponding values of 56% for metabolic markers and 46% for

Table 7
Odds ratios derived from interactive modeling of allostatic load, NHANES 2005–2010 (n = 12,757).

	Continuous AL Score				Additive Dummy Variables		
	Simple	+Demog and SES	+ HB and Medication	Dimensions	Simple	+Demog and SES	+ HB and Medication
AL Score (0–10, continuous)	1.46 (1.41–1.51)	1.36 (1.31–1.42)	1.27 (1.22–1.32)	None (reference)	–	–	–
				CM	1.66 (1.48–1.86)	1.33 (1.17–1.51)	1.24 (1.09–1.41)
				MM	1.74 (1.54–1.98)	1.52 (1.33–1.73)	1.29 (1.12–1.46)
				IM	2.06 (1.85–2.30)	1.92 (1.71–2.15)	1.81 (1.60–2.03)
	Categorical Combinations				Dummy Variable Interactions		
Dimensions	Simple	+Demog and SES	+ HB and Medication	Dimensions	Simple	+Demog and SES	+ HB and Medication
None (reference)	–	–	–	None (reference)	–	–	–
CM	1.99 (1.45–2.73)	1.52 (1.09–2.13)	1.49 (1.06–2.09)	CM	1.99 (1.45–2.73)	1.52 (1.09–2.13)	1.49 (1.06–2.09)
MM	1.94 (1.60–2.35)	1.55 (1.27–1.90)	1.37 (1.12–1.67)	MM	1.94 (1.60–2.35)	1.56 (1.27–1.90)	1.37 (1.12–1.67)
IM	2.42 (1.86–3.14)	2.00 (1.53–2.62)	1.95 (1.48–2.56)	IM	2.42 (1.86–3.14)	2.00 (1.53–2.63)	1.95 (1.48–2.56)
CM + MM	3.03 (2.42–3.81)	2.00 (1.58–2.54)	1.64 (1.28–2.09)	CM*MM	0.79 (0.54–1.14)	0.85 (0.57–1.25)	0.81 (0.54–1.20)
CM + IM	3.62 (2.57–5.10)	2.32 (1.61–3.37)	2.26 (1.55–3.30)	CM*IM	0.75 (0.46–1.22)	0.76 (0.46–1.27)	0.78 (0.47–1.31)
MM + IM	3.82 (3.16–4.60)	2.97 (2.44–3.61)	2.44 (2.00–2.99)	MM*IM	0.81 (0.60–1.10)	0.95 (0.70–1.30)	0.92 (0.67–1.26)
CM + MM + IM	6.39 (5.16–7.90)	4.00 (3.18–5.04)	2.99 (2.36–3.78)	CM*MM*IM	1.42 (0.82–2.47)	1.37 (0.77–2.46)	1.31 (0.72–2.37)
Demographic and Socioeconomic Controls?	No	Yes	Yes		No	Yes	Yes
Health Behaviors Controls?	No	No	Yes		No	No	Yes
Medication Controls?	No	No	Yes		No	No	Yes

Note: 95% CIs are in parentheses. Plus signs (+) indicate categorical combinations. Asterisks (*) are formal statistial interactions.
Survey Design: Sampling Unit = SDMVPSU, Stratum = SDMSTRA Weight = WTMEC6YR

Table 8
Odds ratios from logistic regression models by sex, NHANES 2005–2010 (n = 12,757).

	Biomarkers Modeled Additively			Individual Dimensions	
	Male	Female		Male	Female
Non-Hispanic White (reference group)	–	–		–	–
Non-Hispanic Black					
Reduced Odds Ratios	1.37 (1.12–1.67)	1.76 (1.46–2.14)		1.35 (1.10–1.64)	1.76 (1.46–2.34)
Full Odds Ratios	1.51 (1.24–1.85)	1.68 (1.39–2.04)		1.29 (1.05–1.60)	1.63 (1.35–1.98)
% Mediated	-31.74	8.41		12.02	13.49
Hispanics					
Reduced Odds Ratios	2.04 (1.68–2.47)	2.77 (2.28–3.37)		2.05 (1.69–2.48)	2.78 (2.29–3.38)
Full Odds Ratios	1.95 (1.61–2.36)	2.73 (2.25–3.31)		1.95 (1.61–2.37)	2.74 (2.26–3.33)
% Mediated	6.51	1.55		6.41	1.48
AL Score (0–10, continuous)	1.31 (1.24–1.39)	1.22 (1.15–1.29)	None (reference)	–	
			CM	1.34 (1.11–1.62)	1.14 (0.95–1.37)
			MM	1.41 (1.13–1.75)	1.19 (0.99–1.43)
			IM	1.95 (1.65–2.31)	1.68 (1.42–1.98)

	Categorical Combinations			Interactions	
	Male	Female		Male	Female
Non-Hispanic White (reference group)	–	–		–	–
Non-Hispanic Black					
Reduced Odds Ratios	1.33 (1.09–1.63)	1.76 (1.46–2.14)		1.33 (1.09–1.63)	1.76 (1.46–2.14)
Full Odds Ratios	1.28 (1.04–1.59)	1.63 (1.34–1.98)		1.28 (1.04–1.59)	1.63 (1.34–1.98)
% Mediated	13.61	13.83		13.61	13.83
Hispanics					
Reduced Odds Ratios	2.05 (1.69–2.48)	2.78 (2.29–3.38)		2.05 (1.69–2.48)	2.78 (2.29–3.38)
Full Odds Ratios	1.94 (1.60–2.35)	2.75 (2.26–3.34)		1.94 (1.60–2.35)	2.75 (2.26–3.34)
% Mediated	7.40	1.31		7.40	1.31
None (reference)	–	–	None (reference)	–	–
CM	2.17 (1.28–3.67)	1.20 (0.77–1.88)	CM	2.17 (1.28–3.67)	1.20 (0.77–1.88)
MM	1.71 (1.26–2.32)	1.25 (0.93–1.67)	MM	1.71 (1.26–2.32)	1.25 (0.93–1.67)
IM	2.43 (1.50–3.96)	1.73 (1.23–2.43)	IM	2.43 (1.50–3.96)	1.73 (1.23–2.43)
CM + MM	2.45 (1.71–3.53)	1.11 (0.78–1.57)	CM*MM	0.66 (0.37–1.20)	0.74 (0.42–1.28)
CM + IM	4.64 (2.45–6.67)	1.52 (0.95–2.44)	CM*IM	0.88 (0.37–2.07)	0.73 (0.38–1.41)
MM + IM	3.52 (2.56–4.84)	1.88 (1.44–2.44)	MM*IM	0.85 (0.50–1.44)	0.87 (0.57–1.32)
CM + MM + IM	3.79 (2.61–5.51)	2.53 (1.86–3.45)	CM*MM*IM	0.86 (0.33–2.20)	2.08 (0.96–4.55)
Inclusion of Allostatic Load	Yes	Yes		Yes	Yes
Demographic and Socioeconomic Controls?	Yes	Yes		Yes	Yes
Health Behaviors	Yes	Yes		Yes	Yes
Medication	Yes	Yes		Yes	Yes

Note: 95% CIs are in parentheses. Plus signs (+) indicate categorical combinations. Asterisks (*) are formal statistial interactions.
Survey Design: Sampling Unit = SDMVPSU, Stratum = SDMSTRA Weight = WTMEC6YR

Table 9
Model fit measures for competing AL measurement approaches for overall population and by sex, NHANES 2005–2010 (n = 12,757).

	Continuous AL Score	Additive Dummy Variable	Categorical Combinations/ Dummy Variable Interaction
	AIC	AIC	AIC
Overall			
Simple	11,054	11,133	11,130
+Demographic and Socioeconomic Status	10,259	10,313	10,312
+HB and Medication	9,964	9,965	9,963
Fully Specified Models, Stratified by Sex			
Male	9,709	9,709	9,690
Female	10,161	10,153	10,140

Survey Design: Sampling Unit = SDMVPSU, Stratum = SDMSTRA Weight = WTMEC6YR.

inflammation markers. Thus, males had the highest metabolic prevalence of CM while females has the higher IM prevalence. The prevalence of at least one cardiovascular marker was similar by sex.

Table 4 presents the different combinations of prevalence in dimensions of AL by race/ethnicity. Racial/ethnic groups experience clustering of these dimensions of AL in qualitatively different ways. NH Whites include the highest percent of individuals with no clinically significant indicators of AL (23.23%), which is statistically significantly higher than the percentage for NH Blacks and Hispanics. However, among the large majority of all three racial groups that have at least one of these conditions, these patterns of inequality are more complex. For the group with the worst AL profile (with all three dimensions of AL), NH Whites and Hispanics have much lower and statistically equivalent rates of having all three biomarkers indicated (8.66% and 8.53% respectively vs. 15.78% for NH Blacks). For the three combinations in which a person has two out of three dimensions with at least one biomarker above clinical thresholds, each racial group has the highest prevalence in one combination apiece, with NH Whites showing the highest prevalence of the CM/MM combination (though this is not statistically significantly higher than Hispanics' prevalence), NH Blacks having the highest prevalence of the CM/IM combination, and Hispanics having the highest prevalence of the MM/IM combination. Similarly complex are the patterns of single biomarkers above clinical thresholds, as NH Blacks have the highest prevalence in two of the three (IM and CM, though not statistically significantly higher than whites' prevalence in the latter case), and Hispanics have the highest

prevalence of MM alone. Together, this evidence suggests that racial/ethnic disparities in the three dimensions of AL examined here are not unidirectional in their patterns by race/ethnicity, and that it is worthwhile to examine the contributions of each of these conditions to racial/ethnic disparities in self-rated health separately rather than assuming additive effects.

Table 5 presents the different combinations of prevalence in dimensions of AL by sex. Dimensions of AL cluster in qualitatively different ways by sex. Compared to males, a higher percentage of females show no clinically significant indicators of AL (26.19% vs 17.49%, a statistically significant difference). As in Table 4, other categories show complex patterns by sex. Females have significantly higher percentages for the CM-only and IM-only categories, while males have a higher percentage of MM-only outcomes. Turning to combinations of dimensions, a higher percentage of males are found in the CM/MM category, while females show higher percentages in all other combinations. Despite having a higher percentage of the population with no biomarkers above the clinically determined thresholds, females also are more likely than males to have at least one biomarker for all three dimensions (10.14% vs. 8.74% for males). Together, this evidence suggests that disparities in the three dimensions of AL examined here are not unidirectional in their patterns by sex, and that it is worthwhile to stratify our primary analyses of the contribution of AL to racial/ethnic disparities in SRH by sex.

Which AL measure best mediates the race/ethnicity-SRH relationship?

Are racial/ethnic differences in poor/fair SRH mediated by AL? Results presented in Table 6 indicate that non-Hispanic blacks and Hispanics experience increased risk of reporting poor/fair SRH in comparison to non-Hispanic Whites, which holds true net of demographic/socioeconomic controls, health behaviors and medication use. A small percentage of racial/ethnic differences are explainable as a function of continuous AL scores: 8.50% of the NH White-NH Black disparity and 9.47% of the NH White-Hispanic gap disparity is explained by this measure. In the fully specified model, racial/ethnic disparities are mediated -2.04% and 3.65% for NH Blacks and Hispanics in comparison to NH Whites, respectively. These results are substantively different when AL is operationalized separately or interactively by dimension. When AL is modeled using the additive dummy variables approach, the percent mediated for NH-Blacks in the simple model was 18.85%, while the percent for Hispanics was 6.36%. For the fully specified model the difference was mediated 16.15% for NH-Blacks and 3.61% for Hispanics.

Whenever AL is modeled interactively, either using the dummy variable interactions or categorical combinations approaches, the percent mediated for NH-Blacks in the simple model was 19.07% (higher), while the percent for Hispanics was 6.55% (lower). For the fully specified models, the difference is mediated 16.24% for non-Hispanic Blacks (higher) and 3.50% for Hispanics (lower). Thus, AL statistically accounts for these differences for NH Blacks more than for Hispanics, and differences are observed for the mediation effect by modeling approach. However, the continuous AL score may slightly accentuate the disparity, while other approaches provide evidence of a mediation of the NH White-NH Black gap. Additionally, a fraction of the reduction in gap mediated could also be attributed to the inclusion of additional covariates in the models.

The regression coefficients underlying these results are presented in Table 7. In the top-left quadrant of this table, the continuous AL score is significantly associated with the odds of reporting poor/fair self-rated health, regardless of the model specification. In the additive dummy variable approach, individuals have higher odds of reporting poor/fair SRH in comparison to those without any biomarker exceeding the threshold for each dimension of AL, the strongest effect being for the IM.

Both of the model results just discussed present two additive approaches to measuring the consequences of biological risk markers, but what happens when these dimensions are allowed to interact? In the bottom-left quadrant of Table 7, the categorical combinations results make clear that any combination of these dimension-specific indicators place respondents at elevated risk of reporting poor/fair self-rated health compared to those with no such indicators, regardless of model specification. What is unclear from these results is whether the effects of these three dimensions of biological risk are additive or interactive – in other words, whether those with multiple conditions are at higher risk of reporting poor/fair health because they have multiple dimensions of biological risk with independent effects, or whether the effects of combinations of those risks are greater than the sum of their parts. The dummy variable interaction approach in the bottom-right quadrant of Table 7 addresses this issue, and finds no evidence of statistical interactions between the three dimensions of AL in the full sample. Thus, we conclude that although respondents with multiple biomarkers of disease do have higher odds have reporting poor/fair self-rated health, in the full sample this elevated risk is additive, not multiplicative, in nature.

Do differences in AL mediate racial/ethnic differences in poor/fair SRH differently by sex? Results presented in Table 8 indicate that non-Hispanic blacks and Hispanics experience increased risk of reporting poor/fair SRH in comparison to non-Hispanic Whites, which holds true net of demographic/socioeconomic controls, health behaviors, and medication use for both males and females. A higher percentage of racial/ethnic differences are explainable as a function of continuous AL scores for males than for females: -31.74% of the NH White-NH Black disparity and 6.51% of the NH White-Hispanic gap for males is explained by this measure. The -31.74% figure indicates that in this model, NH Blacks are *advantaged* compared to whites in this modeling strategy. On the other hand, 8.41% of the NH White-NH Black disparity and 1.55% of the NH White-Hispanic gap is explained for females by the continuous AL score.

These results differ somewhat when the AL dimensions are disaggregated. When AL is modeled using the additive dummy variables approach, the percent mediated for NH-Blacks in the simple model was 12.02%, while the percent for Hispanics was 6.41% for males; and 13.49% and 1.48% for females, respectively. Whenever AL is modeled interactively, either using the dummy variable interactions or categorical combinations approaches, the percent mediated for NH-Black males was 13.61% (higher), while the percent for Hispanics was 7.40% (higher) for males; and 13.83% (higher) and 1.31% (lower) for females, respectively. Thus, AL statistically accounts for the differences in SRH for NH Blacks more than for Hispanics, and differences are observed for the mediation effect by modeling approach and by sex.

The regression coefficients underlying these results are also presented in Table 8. In the top-left quadrant of this table, the continuous AL score is statistically significantly associated with the odds of reporting poor/fair self-rated health, regardless of sex. In the additive dummy variable approach, males have higher odds of reporting poor/fair SRH in comparison to those without any biomarker exceeding the threshold for each dimension of AL. Females are only at higher odds of reporting the outcome whenever they have biomarkers exceeding the threshold in the IM dimension, but not so for the CM and MM dimensions. What happens when these dimensions are allowed to interact with models for each sex? In the bottom-left quadrant of Table 6, the categorical combinations results make clear that any combination of these dimension-specific indicators place respondents at elevated risk of reporting poor/fair self-rated health compared to those with no such indicators among males. On the other hand, in the models specified for females having being assigned in one group with at least one biomarker does not always imply higher odds of reporting poor/fair SRH, except for inflammation. Again, it remains unclear whether the effects of these three dimensions of biological risk are additive or interactive – and whether these effects differ by sex. The dummy variable interaction approach in the bottom-right quadrant of Table 6 addresses this issue,

and finds no evidence of statistical interactions between the three dimensions of AL, regardless of sex. Similarly to the analysis of the overall population, we conclude that although respondents with multiple biomarkers of biological dysregulation do have higher odds have reporting poor/fair self-rated health, this elevated risk is additive, not multiplicative, in nature.

Which AL measure best explains individual differences in SRH?

Thus far, all our analyses have been focused on the question of which operationalization of AL best mediates racial/ethnic disparities in SRH. Yet one question remains – which model best fits the data to model *individual* SRH outcomes? As Table 9 shows, the answer fundamentally depends on whether the analysis is stratified by sex. For the +HB and Medication models in the combined analysis, similar fit is achieved by all operationalization strategies. However, we obtain a very different answer when the analyses are stratified by sex. In those models, the Categorical Combinations/Dummy Variable Interaction approaches yield the best fit within both sex groups. This result suggests that sex is a confounder in this relationship. Therefore, we conclude that sex stratification is the appropriate model for these data. The answer to this question depends on whether we are exploring trends for the overall population or by sex. Given that differences are observed for sex in terms of both reporting the outcome and differences in the influences of the different dimensions, future analyses should approach this problem stratifying by sex. When we explore modeling strategies by sex then the Categorical Combinations/Dummy Variable Interactions are the most appropriate operationalization of the biomarker data for modeling individual differences in SRH.

Discussion and conclusion

Biomarker data are widely used in population health research, yet most of previous research on this topic uses the additive score approach to model health outcomes and contributions to individual and group differences in health. This paper examines the appropriateness of this approach by comparing it to alternative operationalizations. We conclude that modeling the dimensions of AL into the affected biological systems leads to greater mediation of racial/ethnic disparities in racial/ethnic disparities, and that allowing these dimensions to interact, by sex, also yields the best fit for our empirical models. We reached this conclusion by testing four hypotheses. The first hypothesis predicts that we will find racial/ethnic disparities in mean additive allostatic load scores and in the prevalence of individual dimensions of this score. This hypothesis is confirmed, in line with previous research.

The second hypothesis predicts that racial/ethnic disparities in self-rated health will be more strongly mediated, and individual differences better accounted for, when the three dimensions of allostatic load scores are disaggregated. This hypothesis was also confirmed for the racial/ethnic disparities sub-hypothesis, and is a novel contribution to the literature. The degree of the disparity by race/ethnicity varies by dimension of allostatic load, such that much larger disparities are found between NH Whites and NH Blacks for inflammation and cardiovascular markers, and between Hispanics and NH Whites for metabolic markers. Given that these markers may have divergent relationships with self-rated health, this is an important insight and leads to stronger mediation of the racial/ethnic disparity in this important health indicator. However, the degree to which individual differences in SRH are explained by these operationalizations of biomarker data depends on the model specification – in simpler specifications with fewer controls, the additive AL score outperforms the disaggregated dimensions, and all strategies yield approximately equal fit in the most fully specified models that we test. However, since models accounting for demographic, socioeconomic, health behaviors and medication use variables are common in this type of research, we argue that the disaggregated dimensions do not yield any loss of fit in appropriate model specifications.

The third hypothesis predicts that this racial/ethnic disparity will be even more strongly mediated when these dimensions are allowed to statistically interact, such that the effect of one dimension is dependent on the presence or absence of the other two dimensions. This hypothesis is not supported, as the amount of the disparity explained by these markers is approximately equal between Additive Dummy Variable and Categorical Combinations/Dummy Variable Interaction approaches in the combined sex analysis. Therefore, in the full analytical sample, the increased risk of fair/poor self-rated health increases with more dimensions of allostatic load, but the effect of individual dimensions does not meaningfully change as in combination. Rather, the risk is increased due to the individual effects of these markers, not their interactions.

The fourth hypothesis posits that these associations may vary by sex. This hypothesis is confirmed for the prediction of individual ratings of self-rated health, but not for the mediation of racial/ethnic disparities therein. Based on this finding, we argue that it is critical to disaggregate analyses by sex to explain individual differences in SRH, but this does not make a meaningful impact on the degree of mediation of these markers on racial/ethnic disparities in SRH.

Of course, like other research into racial/ethnic disparities in self-rated health, our results characterize patterns of associations between race/ethnicity, self-rated health, and biomarkers of physiological dysregulation. Appropriate caution should be taken when inferring causal relationships between these variables. Although we are not aware of any analytical approaches capable of differentiating causal and merely correlative relationships between these characteristics in these data, future research should seek to overcome this barrier to identify the causal contribution of physiological dysregulation to racial/ethnic disparities in self-rated health and other health outcomes.

Should the results of this analysis change how researchers model allostatic load biomarkers? We argue that they should. When modeling individual differences in self-rated health, interactive operationalizations of biomarker data yield clearly superior fit when disaggregating by sex, and equivalent fit in sex-pooled analyses with adequate controls. Therefore, we advise future analysts to both disaggregate their sample by sex and model the dimensions of AL independently.

Ethics approval

This article does not contain any studies with human participants or animals performed by any of the authors; as such ethics approval is not required. This is a secondary data analysis which relies on publicly available information.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ssmph.2017.11.007.

References

Macgregor, G. (1961). Social determinants of health practices. *American Journal of Public Health and the Nation's Health, 51*(11), 1709–1714.

McEwen, B. S. (1998). Stress, adaptation, and disease. Allostasis and allostatic load. *Annals of the New York Academy of Sciences, 840*, 33–44.

McEwen, B. S., & Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, 2–15.

McEwen, B. S., & Seeman, T. (1999). Protective and damaging effects of mediators of stress: Elaborating and Testing the concepts of allostasis and allostatic load. *Annals of the New York Academy of Sciences, 896*(1), 30–47.

Duong, M. T., Bingham, B. A., Aldana, P. C., Chung, S. T., & Summer, A. E. (2017). Variation in the calculation of allostatic load score: 21 Examples from NHANES. *Journal of Racial Ethnic Heal Disparities, 4*, 455–461.

Morrison, S., Shenassa, E. D., Mendola, P., Wu, T., & Schoendorf, K. (2013). Allostatic load may not be associated with chronic stress in pregnant women, NHANES 1999–2006. *Annals of Epidemiology, 23*(5), 294–297.

Juster, R. P., McEwen, B. S., & Lupien, S. J. (2010). Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience & Biobehavioral Reviews*, 2–16.

Woo, H., & Zajacova, A. (2016). Predictive strength of self-rated health for mortality risk among older adults in the United States: Does it differ by race and ethnicity? *Research on Aging,* **164027516**.

Borrell, L. N., & Dallo, F. J. (2008). Self-rated health and race among Hispanic and non-Hispanic adults. *Journal of Immigrant and Minority Health, 10*(3), 229–238. (Available from) ⟨http://www.ncbi.nlm.nih.gov/pubmed/17653864⟩.

Lo, C. C., Howell, R. J., & Cheng, T. C. (2013). Disparities in whites' versus blacks' self-rated health: Social status, health-care services, and health behaviors. *Journal of Community Health, 38*(4), 727–733.

Levine, R. S., Foster, J. E., Fullilove, R. E., et al. (2001). Black-white inequalities in mortality and life expectancy, 1933–1999: Implications for healthy people 2010. *Public Health Reports, 116*(5), 474–483.

Fuller-Thomson, E., Nuru-Jeter, A., Minkler, M., & Guralnik, J. M. (2009). Black-White disparities in disability among older Americans further untangling the role of race and socioeconomic status. *Journal of Aging and Health, 21*(5), 677–698.

Hayward, M. D., Hummer, R. A., Chiu, C. T., González-González, C., & Wong, R. (2014). Does the hispanic paradox in U.S. adult mortality extend to disability? *Population Research and Policy Review, 33*(1), 81–96.

Harper, S., MacLehose, R. F., & Kaufman, J. S. (2014). Trends in the black-white life expectancy gap among US States, 1990–2009. *Health Affair, 33*(8), 1375–1382.

Elo, I. T., Beltrán-Sánchez, H., & Macinko, J. (2014). The contribution of health care and other interventions to Black-White disparities in life expectancy, 1980–2007. *Population Research and Policy Review, 33*(1), 97–126.

Kawachi, I., Kennedy, B. P., & Glass, R. (1999). Social capital and self-rated health: A contextual analysis. *American Journal of Public Health, 89*(8), 1187–1193.

August, K. J., & Sorkin, D. H. (2011). Racial/ethnic disparities in exercise and dietary behaviors of middle-aged and older adults. *Journal of General Internal Medicine, 26*(3), 245–250.

Zajacova, A., & Dowd, J. B. (2011). Reliability of self-rated health in US Adults. *American Journal of Epidemiology, 174*(8), 977–983.

Beck, A. N., Finch, B. K., Lin, S. F., Hummer, R. A., & Masters, R. K. (2014). Racial disparities in self-rated health: Trends, explanatory factors, and the changing role of socio-demographics. *Social Science & Medicine, 104*, 163–177.

Hajat, A., Kaufman, J. S., Rose, K. M., Siddiqi, A., & Thomas, J. C. (2011). Long-term effects of wealth on mortality and self-rated health status. *American Journal of Epidemiology, 173*(2), 192–200.

Banerjee, D., Perry, M., Tran, D., & Arafat, R. (2010). Self-reported health, functional status and chronic disease in community dwelling older adults: Untangling the role of demographics. *Journal of Community Health, 35*(2), 135–141.

Subramanian, S. V., Acevedo-Garcia, D., & Osypuk, T. L. (2005). Racial residential segregation and geographic heterogeneity in black/white disparity in poor self-rated health in the US: A multilevel statistical analysis. *Social Science & Medicine, 60*(8), 1667–1679.

Bjornstrom, E. E. S., & Kuhl, D. C. (2014). A different look at the epidemiological paradox: Self-rated health, perceived social cohesion, and neighborhood immigrant context. *Social Science & Medicine, 120*, 118–125.

Markides, K. S., & Coreil, J. (1986). The health of Hispanics in the southwestern United States: An epidemiologic paradox. *Public Health Reports, 101*(3), 253–265.

Hummer, R. A., Powers, D. A., Pullum, S. G., Gossman, G. L., & Frisbie, W. P. (2007). Paradox found (again): Infant mortality among the Mexican-origin population in the United States. *Demography, 44*(3), 441–457.

Santos-Lozada, A. R. (2016). Self-rated mental health and race/ethnicity in the United States: Support for the epidemiological paradox. *PeerJ, 4*(e2508), 1–15.

Borrell, L. N. (2009). Race, ethnicity, and self-reported hypertension: Analysis of data from the national health interview survey, 1997–2005. *American Journal of Public Health, 99*(2), 313–319.

Johnelle Sparks, P. (2009). One size does not fit all: An examination of low birthweight disparities among a diverse set of racial/ethnic groups. *Maternal and Child Health Journal, 13*(6), 769–779.

Dubard, C. A., & Gizlice, Z. (2008). Language spoken and differences in health status, access to care, and receipt of preventive services among US hispanics. *American Journal of Public Health, 98*(11), 2021–2028.

Viruell-Fuentes, E. a, Morenoff, J. D., Williams, D. R., & House, J. S. (2011). Language of interview, self-rated health, and the other Latino health puzzle. *American Journal of Public Health, 101*(7), 1306–1313.

Kandula, N. R., Lauderdale, D. S., & Baker, D. W. (2007). Differences in self-reported health among Asians, Latinos, and non-hispanic Whites: The role of language and nativity. *Annals of Epidemiology, 17*, 191–198.

Bzostek, S., Goldman, N., & Pebley, A. (2007). Why do Hispanics in the USA report poor health? *Social Science & Medicine, 65*(5), 990–1003.

Markides, K. S., & Martin, H. (1979). Predicting self-rated health among the aged. *Research on Aging, 1*, 97–112.

Patel, K. V., Eschbach, K., Rudkin, L. L., Peek, M. K., & Markides, K. S. (2003). Neighborhood context and self-rated health in older Mexican Americans. *Annals of Epidemiology, 13*(9), 620–628. (Available from) ⟨http://www.ncbi.nlm.nih.gov/pubmed/14732301⟩.

Prus, S. G. (2011). Comparing social determinants of self-rated health across the United States and Canada. *Social Science & Medicine, 73*(1), 50–59. (Available from) ⟨http://www.sciencedirect.com/science/article/pii/S0277953611002474⟩.

Gorman, B. K., & Read, J. G. (2006). Gender disparities in adult health: An examination of three measures of morbidity. *Journal of Health Social Behaviour, 47*, 95–110.

Zajacova, A., Huzurbazar, S., & Todd, M. (2017). Gender and the structure of self-rated health across the adult life span. *Social Science & Medicine, 187*, 58–66.

Oksuzyan, A., Petersen, I., Stovring, H., Bingley, P., Vaupel, J. W., & Christensen, K. (2009). The male-female health-survival paradox: A survey and register study of the impact of sex-specific selection and information bias. *Annals of Epidemiology, 19*(7), 504–511.

Case, A., & Paxson, C. (2005). Sex differences in morbidity and mortality. *Demography, 42*(2), 189–214.

Arber, S., Cooper, H., S. A., H. C., Arber, S., & Cooper, H. (1999). Gender differences in health in later life: The new paradox? *Social Science & Medicine, 48*(1), 61–76.

Hirve, S., Juvekar, S., Sambhudas, S., et al. (2012). Does self-rated health predict death in adults aged 50 years and above in India? Evidence from a rural population under health and demographic surveillance. *International Journal of Epidemiol, 41*(6), 1719–1727.

Ross, C., Masters, R., & Hummer, R. (2012). Education and the gender gaps in health and mortality. *Demography, 49*(4), 1157–1183.

Daw, J. (2017). Contribution of four comorbid conditions to racial/ethnic disparities in mortality risk. *American Journal of Preventive Medicine, 52*(Supplement 1), S95–S102.

Johnson, C. L., Paulose-Ram, R., Ogden, C. L., et al. (2013). National health and nutrition examination survey: Analytic guidelines, 1999–2010. *Vital Health Statistcs, 2*(161), 1–24.

Howard, J. T., & Sparks, P. J. (2016). The effects of allostatic load on racial/ethnic mortality differences in the United States. *Population Research and Policy Review, 35*, 421–443.

Howard, J. T., & Sparks, P. J. (2015). The role of education in explaining racial/ethnic allostatic load differentials in the United States. *Biodemography and Social Biology, 61*(1), 18–39.

Howard, J. T., & Sparks, P. J. (2016). Does allostatic load calculation method matter? Evaluation of different methods and individual biomarkers functioning by race/ethnicity and educational level. *American Journal of Human Biology, 28*, 627–635. http://dx.doi.org/10.1002/ajhb.22843.

Manor, O., Matthews, S., & Power, C. (2000). Dichotomous or categorical response? Analysing self-rated health and lifetime social class. *International Journal of Epidemiology, 29*(1), 149–157.

Acevedo-Garcia, D., Bates, L. M., Osypuk, T. L., & McArdle, N. (2010). The effect of immigrant generation and duration on self-rated health among US adults 2003–2007. *Social Science & Medicine, 71*(6), 1161–1172.

Crimmins, E. M., Kim, J. K., & Seeman, T. E. (2009). Poverty and biological risk: The earlier 'aging' of the poor. *Journals of Gerontology. Series A: Biological Sciences and Medical Science, 64*(2), 286–292.

Gu, Q., Paulose-Ram, R., Burt, V., & Kit, B. (2017). Prescription cholesterol-lowering medication use in adults aged 40 and over: United States, 2003–2012. *NCHS Data Brief.* (Available from) ⟨http://www.cdc.gov/nchs/data/databriefs/db177.pdf⟩.

Karlson, K. B., Holm, A., & Breen, R. (2012). Comparing regression coefficients between same-sample nested models using logit and probit: A new method. *Sociological Methodology, 42*(1), 286–313.

Breen, R., Karlson, K. B., & Holm, A. (2013). Total, direct, and indirect effects in logit and probit models. *Sociological Methods & Research, 42*(2), 164–191.

U. Kohler K.B. Karlson A. Holm The Stata Journal Stata J **11** 3 2011 pp.420–438.

StataCorp (2015). *Stata statistical software: Release 14.* College Station, TX: StataCorp LP (2015).