STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract Multiple Vulnerabilities to Health Disparities and Incident Heart Failure Hospitalization in the REasons for Geographic and Racial differences in Stroke (REGARDS) cohort study.
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found Page 2 The objective of this study was to determine the cumulative effect of SDV to health disparities on incident HF hospitalization. Using the REasons for Geographic and Racial differences in (REGARDS) cohort study, we studied 25,790 participants without known HF and followed them 10+ years. Our primary outcome was an incident HF hospitalization through 12/31/2016. Guided by the Healthy People 2020 framework for social determinants of health, we examined 10 potential SDVs. We retained SDVs associated with incident HF hospitalization ($p<0.10$) and created a SDV count (0, 1, 2, 3+). Using the count, we estimated Cox proportional hazard models to examine associations with incident HF hospitalization, adjusting for potential confounders. Models were stratified by age (45-64, 65-74, and 75+ years) because past reports suggest greater disparities in HF incidence at younger ages. Participants were followed for a median of 10.1 years (IQR 6.5, 11.9). Black race, low educational attainment, low annual household income, zip code poverty, poor public health infrastructure, and lack of health insurance were associated with incident HF hospitalization. In adjusted models, among those 45-64 years, compared to having no SDV, having 1 SDV (HR 1.85; 95% CI 1.12-3.05), 2 SDV (HR 2.12; 95% CI 1.28-3.50) and 3+ SDVs (HR 2.45; 95% CI 1.48-4.04) were significantly associated with incident HF hospitalization (*p* for trend 0.001). We observed no significant associations for older individuals.
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported Page 1 Heart failure is a common chronic disease among older Americans. As there is no cure for HF, preventing its onset is of public health interest. Studies have identified predictors of incident HF including older age, male gender, cardiovascular disease (CVD) risk factors (e.g., diabetes, hypertension, smoking), and obesity. Recent studies investigated the effects of socio-demographic factors on incident HF and found that Black race, low education, low income, and neighborhood deprivation predict HF incidence. Racial disparities in HF are well-established; the prevalence of HF is greater among Blacks compared to Non-Hispanic Whites. Blacks develop HF at younger ages and have a 50% higher HF incidence at earlier ages than Whites. Racial difference in incident HF are partially attributed to a greater burden of CVD risk factors among Blacks compared to Whites. However, additional social determinants of health disparities may play an important role. Low educational attainment, low annual income, living in an area with relatively few healthcare services, and lacking health insurance put individuals at risk for incident HF.
Objectives	3	State specific objectives, including any prespecified hypotheses

		Pages 2, 4 The objective of this study was to determine the cumulative effect of SDV to health disparities on incident HF hospitalization. We hypothesized that as a person's number of SDV rose, the risk of incident HF hospitalization would rise. Because of prior observations of disparities for HF being greatest at younger ages, we hypothesized that the effect of a rising burden of SDV on risk of incident HF hospitalization would be greatest among individuals <65 years.
Methods		
Study design	4	Present key elements of study design early in the paper Pages 4-5 Using the REasons for Geographic and Racial differences in (REGARDS) cohort study, we studied 25,790 participants without known HF and followed them 10+ years. We assembled a cohort at risk for HF, using HF-related medications to exclude individuals with suspected HF at baseline.
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection Page 4 REGARDS is a national, prospective, longitudinal cohort study evaluating racial and geographic disparities in stroke mortality. REGARDS recruited 30,239 community-dwelling, English-speaking individuals ≥45 years of age from 2003-2007 and is following participants for 10+ years.
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants Pages 4-5 To assemble a cohort at risk for HF, we used HF-related medications to exclude individuals with suspected HF at baseline. The approach to determining suspected HF using medications was internally validated among a subgroup of REGARDS participants for whom Medicare claims were available.
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed Pages 4-5 To assemble a cohort at risk for HF, we used HF-related medications to exclude individuals with suspected HF at baseline. HF-related medications included: digoxin in the absence of atrial fibrillation, angiotensin-converting enzyme inhibitor/angiotensin receptor blocker plus beta-blocker in the absence of hypertension; carvedilol; spironolactone; loop diuretics including furosemide, bumetanide, or torsemide; and/or a combination of hydralazine and nitrates.
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable **Outcome:** An incident HF hospitalization through 12/31/2016.

Primary Exposure Variable: Guided by the Healthy People 2020 framework for social determinants of health, we examined 10 potential SDVs. We evaluated SDVs from 5 domains of the framework: 1) education (<high school); 2) economic stability (<$35,000 annual household income); 3) neighborhood/built environment (living in a zip code with >25% of residents living below the Federal poverty line, and living in a rural area as defined by rural urban commuting area codes 9 and 10); 4) health and healthcare (living in a Health Professional Shortage Area [HPSA], lacking health insurance, and living in a US state with poor public health infrastructure); and 5) social and community (Black race, social isolation).

Covariates: To understand the mechanisms leading to associations between SDV and incident HF hospitalization, we sequentially adjusted for variables reflecting 1) socio-demographics, 2) medical conditions, 3) functional status, 4) health behaviors, and 5) physiologic variables.

Data sources/ measurement

8*

For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group

Pages 4-5

REGARDS Study: REGARDS is a national, prospective, longitudinal cohort study evaluating racial and geographic disparities in stroke mortality. REGARDS recruited 30,239 community-dwelling, English-speaking individuals ≥45 years of age from 2003-2007 and is following participants for 10+ years. At enrollment, REGARDS participants completed a baseline computer assisted telephone interview, which ascertained sociodemographic information and medical history. Participants also received an in-home physical exam during which blood and urine samples were obtained, an electrocardiogram was performed, and a medication inventory was done through pill bottle review.

Incident HF Hospitalizations: REGARDS participants were contacted by phone to ascertain CVD outcomes every six-months. CVD events including incident HF hospitalizations were adjudicated by experts using a structured form,[22] based on signs and symptoms of HF collected from chart-level data obtained from the hospital.

Covariates: Socio-demographics included age at baseline, gender, and Southeastern region (stroke belt/buckle, defined as North Carolina, South Carolina, Georgia, Tennessee, Mississippi, Alabama, Louisiana and Arkansas; or non-stroke belt). Medical conditions included history of high blood pressure (self-report of hypertension diagnosis, use of antihypertensive medications, or blood pressure ≥140/90 mm Hg at the baseline in-home visit reflecting hypertension guidelines at the time of the observation period), high cholesterol (self-reported diagnosis, total cholesterol >=240 or low density lipoprotein (LDL) cholesterol >160 mg/dL or high density lipoprotein (HDL) <40), diabetes (use of diabetes medications or insulin, or fasting blood glucose >126 mg/dL, or non-fasting glucose >200 mg/dL). Use of antihypertensive medications, statins, and insulin were included separately. Functional status was assessed with the Physical Component Summary (PCS) and the Mental Component Summary (MCS) scores. Health behaviors included smoking (currently vs. not), alcohol use (risky drinking based on sex-specific National Institute on Drug Abuse cut points vs. others), physical activity (enough activity to work up a sweat on most days of the week vs. others), and adherence to the Mediterranean diet using the

		Mediterranean diet score. Physiologic variables included body mass index, systolic and diastolic blood pressure, total cholesterol, high density lipoprotein cholesterol, log transformed high sensitivity c-reactive protein, log transformed urinary albumin-to-creatinine ratio, and estimated glomerular filtration rate using the CKD-Epi equation.
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at We reached a baseline sample of 30,239 participants; we excluded 496 participants due to loss of follow-up. We further excluded participants with suspected HF at baseline, or if information necessary to determine an event was missing. We finally reached a sample size of 25,790. (Please refer to the exclusion cascade outlined in Supplementary Figure 1.)
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why Pages 5-4, 7-8 **SDV to health disparities:** We created a SDV count (0, 1, 2, 3+) and described characteristics of our HF-free cohort within SDV count categories. We assessed multicollinearity among SDVs using variance inflation factors (VIF). Sub-groups: To assess for effect modification by age, we tested interactions between SDV count and three age subgroups in an overall model: <65, 65-74, and 75+ years. Since the interaction term was significant (joint test p<0.0001), we present age-stratified results. Using Kaplan Meier plots, we depicted the cumulative risk of HF by SDV count by age group. Using the log-rank test, we assessed the equal incident HF hospitalization rates by SDV count for each age group.
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding Page 7 We estimated Cox models to examine the effect of the SDV count on incident HF hospitalization, by age. First, a crude model examined the association between SDV count and incident HF. Second, a minimally-adjusted Cox model adjusted for age and gender. Finally, we added covariates in groups: 1) socio-demographics, 2) medical conditions and medications, 3) functional status, 4) health behaviors, and 5) physiologic variables. We calculated adjusted hazard ratios (aHR) and 95% confidence intervals (95% CI) for each estimate. To reduce the effect of missing data, we performed multiple imputation by chained equations on covariates that were missing.
		(*b*) Describe any methods used to examine subgroups and interactions Page 6 To assess for effect modification by age, we tested interactions between SDV count and three age subgroups in an overall model: <65, 65-74, and 75+ years. Since the interaction term was significant (joint test p<0.0001), we present age-stratified results. Using Kaplan Meier plots, we depicted the cumulative risk of HF by SDV count by age group. Using the log-rank test, we assessed the equal incident HF hospitalization rates by SDV count for each age group.
		(*c*) Explain how missing data were addressed Page 7 To reduce the effect of missing data, we performed multiple imputation by chained

equations on covariates that were missing.

(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed

Participants were censored if loss to follow-up

(*e*) Describe any sensitivity analyses

Page 8,10

We examined death as a competing risk, which resulted in sub-distribution HR estimates that were nearly identical to the main analysis.

We conducted a competing risk survival analysis fitting Fine and Gray's sub-distribution hazard model, where death from any cause was considered a competing event.

Results

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed Pages 4-5 To assemble a cohort at risk for HF, we used HF-related medications to exclude individuals with suspected HF at baseline. The approach to determining suspected HF using medications was internally validated among a subgroup of REGARDS participants for whom Medicare claims were available. HF-related medications[21] included: digoxin in the absence of atrial fibrillation, angiotensin-converting enzyme inhibitor/angiotensin receptor blocker plus beta-blocker in the absence of hypertension; carvedilol; spironolactone; loop diuretics including furosemide, bumetanide, or torsemide; and/or a combination of hydralazine and nitrates. We excluded participants with: 1) missing data on self-reported atrial fibrillation, 2) on baseline medication use; and 3) participants with HF hospitalizations between the baseline CATI and in-home visit. Compared to Medicare claims, the negative predictive value was >95%. HF hospitalization rates was 27 per 1,000 person-years (PYs) among individuals with suspected HF versus 4 per 1,000 PYs among those without suspected HF.
		(b) Give reasons for non-participation at each stage Page 5 We excluded participants with: 1) missing data on self-reported atrial fibrillation, 2) on baseline medication use; and 3) participants with HF hospitalizations between the baseline CATI and in-home visit. (Please refer to the exclusion cascade outlined in Supplementary Figure 1.)
		(c) Consider use of a flow diagram Supplementary Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders Page 7 We included 25,790 participants without suspected HF at baseline. The study sample included 13,487 (52%) participants aged 45-64 years; 8,214 (32%) aged 65-74 years; and 4,089 (16%) aged 75+ years. Among individuals with no missingness on SDVs who were 45-64 years, individuals with a greater number of SDVs were female, had hypertension and diabetes, had worse physical well-being, were smokers, and lived in the Southeast (see Table 1). We observed similar characteristics among individuals with more SDVS in the two older age groups. (Please refer to Supplemental Table 1 and 2)
		(b) Indicate number of participants with missing data for each variable of interest Page 8 We observed missing data for some covariates, with the largest proportions of missing information were for Mediterranean diet scores (28%) and annual household income (12%). Missingness for the rest of the variables was <6%. The degree of missingness observed in

		our study was less than established thresholds of 50%.
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount) Median follow-up of 10.1 years (SD=3.3)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time

Pages 8-10

Over a median follow-up of 10.1 years (SD=3.3), we observed 1,109 incident HF hospitalizations. Of these, 276 hospitalizations (25%) occurred among participants aged 45-64 years, 441 (40%) occurred among 65-74 years, and 392 (35%) occurred among 75+ years. Median [IQR] time to HF hospitalization was 6.0 [3.0-9.2] for those aged 45-64 years, 6.0 [2.8-8.7] for 65-74 years, and 5.6 [3.2-9.2] for 75+ years. Kaplan Meier survival curves are shown in Figure 1. The log rank test p-value was <0.0001 for differences in the cumulative incidence of HF hospitalization among SDV groups for those aged 45-64 and 65-74, but not for those aged 75+ (p=0.59). Age-adjusted incidence rates of HF hospitalizations per 1,000 PYs by SDV groups and age are shown in eFigure 3. HF hospitalization incidence was lowest among those <65 years and highest among those 75+ years. In the <65 and 65-74 year old groups, HF hospitalization incidence increased with each additional SDV. For individuals with 3+ SDVs compared to no SDVs, HF hospitalization incidence was nearly 7 times higher in <65 years old stratum, and 2 times higher among those 65-74 years old. In the 75+ group, the highest HF hospitalization incidence was observed for individuals with 2 SDVs.

Adults 45-64 years: In unadjusted models (eFigure 4), we observed significant associations between SDV count and incident HF hospitalization. HRs increased in a graded fashion with each additional SDV (1 SDV: 2.69; 95% CI 1.61-4.49; 2 SDVs: 4.13, 95% CI 2.47-6.91; 3+ SDVs: 7.16; 95% CI 4.39-11.67, *p* for trend <.0001). In models adjusting for age at baseline and gender (Figure 2), we continued to observe graded, statistically significant HRs for 1 SDV (2.72; 95% CI 1.63-4.54), 2 SDVs (4.25; 95% CI 2.54-7.13), and 3+ SDVs (7.41; 4.53-12.12) compared to 0 SDVs (*p* for trend <0.0001). In fully adjusted models (Figure 3), adjusted HRs for the association between SDV count and incident HF hospitalization were attenuated but remained significant. Compared to having 0 SDV, having 1 SDV had aHR 1.85 (95% CI 1.12-3.05), having 2 SDV (aHR 2.12; 95% CI 1.28-3.51) and 3+ SDVs (aHR 2.45; 95% CI 1.48-4.04) were significantly associated with incident HF hospitalization (*p* for trend <0.0001).

Adults 65-74 years: In unadjusted Cox models (eFigure 4), statistically significant associations were observed for individuals with 2 SDV (1.43; 1.04, 1.95) and 3+ SDV (1.72; 1.27-2.34) compared to 0 SDV, *p* for trend <.0001. In age and gender adjusted models (Figure 2), HRs remained significant for 2 SDVs (1.54; 95% CI 1.12-2.11) and 3+ SDVs (1.91; 95% CI 1.41-2.61) compared to 0 SDV (*p* for trend <0.0001). In fully adjusted Cox models (Figure 3), we did not observe significant associations between number of SDVs and incident HF hospitalization (*p* for trend=0.986).

Adults 75+ years: Among the 554 adults with no SDV, we observed 39 incident HF hospitalizations. Among the 2,930 adults with 1 or more SDV, we observed 228 incident HF hospitalizations. None of the crude HRs for the 75+ year group were statistically significant, *p* for trend=0.602 (eFigure 4). In age and gender adjusted models (Figure 2), the

		only statistically significant minimally adjusted HR was for individuals with 2+ SDV (1.45; 95% CI 1.03-2.03) compared to 0 SDV, *p* for trend=0.211. We observed no significant associations between SDV count and incident HF hospitalization with fully-adjusted HRs near 1.0 (*p* for trend =0.379). (Please see Figure 1. Kaplan-Meier Survival Curves by Age Strata)
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included Page 2 Black race, low educational attainment, low annual household income, zip code poverty, poor public health infrastructure, and lack of health insurance were all significantly associated with incident HF hospitalization at p<0.10.
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses Page 7 We used stratified analyses by three age groups (45-64, 65-74, and 75+ years). As a sensitivity analysis, we conducted a competing risk survival analysis fitting Fine and Gray's sub-distribution hazard model, where death from any cause was considered a competing event.
Discussion		
Key results	18	Summarise key results with reference to study objectives Page 8 As the number of SDVs increased, the risk of incident HF hospitalization increased among adults <65 years of age. Page 9 Among adults 45-64 years, In unadjusted models (eFigure 4), we observed significant associations between SDV count and incident HF hospitalization. HRs increased in a graded fashion with each additional SDV (1 SDV: 2.69; 95% CI 1.61-4.49; 2 SDVs: 4.13, 95% CI 2.47-6.91; 3+ SDVs: 7.16; 95% CI 4.39-11.67, *p* for trend <.0001). In models adjusting for age at baseline and gender (Figure 2), we continued to observe graded, statistically significant HRs for 1 SDV (2.72; 95% CI 1.63-4.54), 2 SDVs (4.25; 95% CI 2.54-7.13), and 3+ SDVs (7.41; 4.53-12.12) compared to 0 SDVs (*p* for trend <0.0001). In fully adjusted models (Figure 3), adjusted HRs for the association between SDV count and incident HF hospitalization were attenuated but remained significant. Compared to having 0 SDV, having 1 SDV had aHR 1.85 (95% CI 1.12-3.05), having 2 SDV (aHR 2.12; 95% CI 1.28-3.51) and 3+ SDVs (aHR 2.45; 95% CI 1.48-4.04) were significantly associated with incident HF hospitalization (*p* for trend <0.0001). Among adults 54-74 years, in unadjusted Cox models (eFigure 4), statistically significant associations were observed for individuals with 2 SDV (1.43; 1.04, 1.95) and 3+ SDV (1.72; 1.27-2.34) compared to 0 SDV, *p* for trend <.0001. In age and gender adjusted models (Figure 2), HRs remained significant for 2 SDVs (1.54; 95% CI 1.12-2.11) and 3+ SDVs (1.91; 95% CI 1.41-2.61) compared to 0 SDV (*p* for trend <0.0001).

		Page 10 Among the 554 adults 75+ years, after adjusting for age and gender (Figure 2), the only statistically significant minimally adjusted HR was for individuals with 2+ SDV (1.45; 95% CI 1.03-2.03) compared to 0 SDV, *p* for trend=0.211. (Please see Figure 3. Fully Adjusted Estimates for Associations Between Socially Determined Vulnerabilities and Incident Heart Failure)
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias Page 13 An observational design limits our ability to draw causal inferences. Additionally, demographic and medical history variables were self-reported. Because participants were followed prospectively until they experienced an incident HF hospitalization event, SDVs were captured at the baseline survey so we were unable to examine the effects of time-varying SDVs (e.g., insurance status). We used incident HF hospitalization as a proxy for incident HF but recognize that some incident HF cases are diagnosed in the outpatient setting. Finally, the suspected HF-free cohort was internally validated with Medicare data, which is an imperfect gold standard for HF. We cannot corroborate the same operating characteristics would be observed in commercial claims data.
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence Page 13 Our study suggests that among individuals <65, the cumulative burden of SDV is an important risk factor for incident HF hospitalization that rises with each additional SDV to health disparities. This effect was not explained by CVD risk factors and confounders. Similar patterns were not observed for individuals 65+ years. While our findings should be confirmed in cohorts with larger samples of younger adults, the number of SDVs in individuals <65 years may be a simple and novel strategy to identify individuals at increased risk for incident HF hospitalization.
Generalisability	21	Discuss the generalisability (external validity) of the study results Page 13 We cannot corroborate the same operating characteristics would be observed in commercial claims data.
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based Page 14 This research project is supported by cooperative agreement U01 NS041588 co-funded by the National Institute of Neurological Disorders and Stroke (NINDS) and the National Institute on Aging (NIA), National Institutes of Health, Department of Health and Human Service. This work is also supported by R01 HL80477 from the National Heart Lung and Blood Institute (NHLBI), National Institutes of Health, Department of Health and Human Service. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NINDS, NIA or NHLBI. Representatives of the NINDS were involved in the review of the manuscript but were not directly involved in the collection, management, analysis or interpretation of the data.

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.