**eAPPENDIX A: Measurement of life-course SES**

We selected eight indicators of early-life SES based on retrospective assessment and six indicators of adult SES (shown in **eTable 1)**.[[1](#_ENREF_1),[2](#_ENREF_2)] EFA with quartimin rotation was used to assess the dimensionality of our SES indicators. We used the root mean square error of approximation (RMSEA) and comparative fit index (CFI) to determine the smallest number of latent factors needed to explain the correlations among SES indicators. A RMSEA less than 0.06 and a CFI greater than 0.95 generally indicate a good fitting model.[[3](#_ENREF_3)] A three factor model with a RMSEA=0.06 and CFI=0.97 was initially identified. The first factor was measured by five indicators of early-life SES, including father’s and mother’s educational attainment, father’s occupational status, region of birth, and rural residence during childhood. Three indicators of early-life SES, including self-rated childhood SES, whether the respondent’s family moved for financial reasons during childhood, and whether there was a period during childhood when the respondent’s father was unemployed for several months measured the second factor. The third factor was measured by five adult SES indicators, including respondent’s educational attainment, respondent’s occupation, respondent’s labor force status, household income, and household wealth. The three-indicator factor for early-life SES was deleted because the indicators were, based on face validity, poor measures of the latent factor. The final two-factor model had a reasonable fit based on the RMSEA (RMSEA=0.07) and a good fit based on the CFI (CFI=0.98).

We next used CFA with maximum likelihood estimation to output factor scores for early-life and adult SES. The results of the CFA, including standardized model coefficients and standard errors are shown in **eFigure 1**. Approximately 47% of participants were missing data on at least one indicator of early-life SES. Missing values were handled with full information maximum likelihood (FIML) using the expectation maximization algorithm. This method uses all available data to estimate model parameters and has been shown, with ignorable missing data, to be unbiased and more efficient than other methods for handling missing data.[[4](#_ENREF_4)] We conducted sensitivity analyses comparing results using FIML to listwise deletion; results were qualitatively similar although FIML estimates were more precise. Factor analyses were conducted in Mplus version 5.

**eAPPENDIX B: Estimation of controlled direct effects with time varying confounding**

Our analytical objective was to estimate the extent to which the association between adult SES () and all-cause mortality (, where indexes the individual and indexes the wave of mortality assessment) was mediated by observed time-varying health behaviours (i.e., smoking, alcohol consumption, and physical inactivity), represented by , after accounting for: (i) a vector of baseline characteristics, , that included potentially confounding baseline covariates (i.e., age at enrolment, gender, race, and early-life SES) and (ii) potentially confounding time-varying covariates (i.e., self-report of doctor’s diagnosis of cancer, psychiatric disorder, stroke, heart disease, diabetes), represented by the vector .

Following the conventional regression approach to assessing mediation,[[5](#_ENREF_5)] in Model 6 we estimated the direct effect of SES on mortality not mediated by measured health behaviors or time-varying covariates through standard regression adjustment by fitting a log-linear model of the form:

 **[6]**

The coefficient in Model 6 estimates the direct effect of SES on mortality not through pathways involving measured health behaviors or time-varying health status provided that measured baseline and time-varying covariates suffice to control for confounding between (i) SES and mortality, (ii) health behaviors and mortality, and (iii) time-varying covariates and mortality. This conventional regression approach produces an unbiased estimate of the direct effect we wish to estimate only under very specific conditions.[[6](#_ENREF_6)] For example, if there is a consequence of SES that confounds the association between health behaviors and mortality (i.e., if time-varying health status acts simultaneously as a confounder of the association between health behaviors and mortality and mediator of the association between SES and mortality) then the controlled directed effect of SES on mortality not mediated through health behaviors cannot be identified using the conventional regression approach. Indeed, we find empirical evidence that health events predict health behaviors in the HRS. We ran multivariable models regressing each health behavior on lagged health conditions (cancer, psychiatric condition, diabetes, heart disease, stroke, self-rated health), controlling for age and gender. Results indicate that major chronic conditions, including diabetes, heart disease, and stroke, were associated with not smoking (**eTable 8)**, not consuming any alcohol (**eTable 9**), and not engaging in physical activity (**eTable 10**). In this situation, adjusting for time-varying health status using the conventional regression approach might block pathways linking SES to mortality and underestimate the direct effect of SES on mortality not through health behaviors. Alternatively, omitting these time-varying covariates and conditioning on health behaviors may induce bias due to collider stratification. By handling potential confounding by time-varying covariates through weighting rather than conditioning on covariates, marginal structural models (MSM) allow for identification of the direct effect of SES on mortality even in settings in which conventional approaches are biased, including when there is a consequence of SES that confounds the association between health behaviors and mortality.[[7](#_ENREF_7),[8](#_ENREF_8)]

In Model 7 we estimated the controlled direct effect of SES () on all-cause mortality () not through measured health behaviors , including smoking (S), alcohol consumption (D), and physical inactivity (Z), by fitting a stabilized inverse probability weighted marginal structural model, as described by VanderWeele and Robins.[[7](#_ENREF_7),[8](#_ENREF_8)] In brief, we fitted a weighted log-linear regression model of the form:



Potential confounding was accounted for by fitting the above model with stabilized inverse probability weights of the form , where:









The weight  accounts for measured confounding of the relation between SES and mortality and the weights , , and  account for measured confounding of the relation between smoking and mortality, alcohol consumption and mortality, and physical inactivity and mortality, respectively. We show in **eTable 3** that there is not evidence of interaction between SES and health behaviors in predicting mortality. If there is no interaction between SES and health behaviors, the coefficient  in the weighted model gives the direct effect of SES not through measured health behaviors provided that (i) the measured baseline confounders suffice to control for confounding between SES and mortality and (ii) the measured confounders, SES, and time-varying covariates suffice to control for confounding between health behaviors and mortality. Note that, unlike Model 6, the direct effect of SES estimated from Model 7 includes pathways mediated by time-varying health status.

Briefly, each participant in this model is weighted by the inverse of the probability that he had the mediating behaviors he actually had, given his SES, baseline covariates, and past history of behaviors and time-varying covariates. The denominator of  is the probability of receiving the quartile of SES the individual in fact received, conditional on baseline covariates. The denominator of  is the probability of smoking status reported by the individual at wave , conditional on quartiles of SES, baseline covariates, time-varying covariates, and smoking status at . The denominator of  is the probability of receiving the quartile of alcohol consumption the individual in fact received at wave , conditional on quartiles of SES, baseline covariates, time-varying covariates, smoking status at , and alcohol consumption at . The denominator of  is the probability of physical inactivity reported by the individual at wave , conditional on quartiles of SES, baseline covariates, time-varying covariates, smoking status at , alcohol status at , and physical inactivity at . Stabilizing the weights by including probabilities in the numerator results in more efficient estimation.[[8](#_ENREF_8)] Weights were estimated using logistic regressions for smoking and physical inactivity and ordinal logistic regression for quartiles of alcohol consumption. Predicted probabilities for the numerator and denominator were assigned based on the category of the outcome reported and divided to obtain stabilized weights. The distributions of the weights , , , and  are provided in **eTable 11.** We incorporated individual-level sample weights using the approach described by Brumback et al. (2010).[[9](#_ENREF_9)] In brief, the marginal structural model was weighted by the product of the four inverse probability weights and the sample weight, truncated at the 99th percentile to improve the precision of our estimates.

**eAPPENDIX C: Potential sources of selection bias**

We were concerned about two forms of selection: left censoring due to starting follow-up in 1998 (which could be attributable to mortality or dropout) and right censoring due to attrition between 1998 and 2008. With respect to left censoring, the bias is constrained by the fraction of the sample affected by mortality or attrition. HRS dropout between 1992 and 1998 was 16%.[[10](#_ENREF_10)] We incorporated respondent-level sample weights provided by HRS. These weights are constructed by first producing a Current Population Survey post-stratified household-level weight using the initial sampling probabilities and the birth years and race/ethnicity of household members and then constructing appropriate respondent-level weights.[[11](#_ENREF_11)] Applying the sample weights to our analyses of the total and direct effects should partially correct for bias introduced by differential attrition by socio-demographic characteristics between 1992 and 1998.

Additionally, we have considered the direction and magnitude of the potential bias introduced by selective mortality. Note that our study includes US birth cohorts 1931-1941. If we assume exposure was defined by age 25, then the relevant survival is survival between ages 25 and the year 1998. For the 1940 birth cohort, approximately 78% of those who survived to age 25 further survived to 1998 (age 58);[[12](#_ENREF_12)] of the 1930 birth cohort who survived to age 25, we estimate 54% survived to 1998 (age 68). To the extent that either our exposures or our hypothesized mediators influenced survival to 1998 or selection into the 1998 sample (due to either nonparticipation in 1992 or dropout between 1992 and 1998), this is likely to introduce bias. Further, this bias is of uncertain sign; although it would likely lead to an attenuation of the estimated effect of SES on post-1998 mortality, it would also likely lead to an attenuation of the estimated effect of the behavioral mediators on post-1998 mortality. The bias in the direct effect estimation is due to the joint influence of those two biases, making it difficult to predict its direction.

To explore the plausible magnitude of the bias introduced by selective mortality, we used the Chiba and VanderWeele (2011) framework comparing the treatment-outcome association observed among actual survivors to what would have been observed among the “always-survivor” subpopulation (individuals who would survive regardless of whether exposed to low SES or not).[[13](#_ENREF_13)] We assume the population includes “always-survivors”, “never survivors” (who would not survive regardless of exposure status), and “survive under non-exposure”(individuals who would die if exposed to low SES but survive is high SES). We assume nobody in the population would survive under low SES but die under high SES (i.e., monotonicity). We further assume that the “survive under non-exposure” individuals who died between age 25 and 1998 as a result of low SES had worse health and health behaviors than the “always survive” individuals. Chiba and VanderWeele provide a simple correction factor to be applied to the observed contrast among survivors:

Survivor bias (a) = E[*Y*1 | *A* = 1, *S* = 1] – E[*Y*1 | *A* = 0, *S* = 1].

This correction parameter is the counterfactual value of Y under exposure among those who were exposed and survived (i.e., always survivors) minus the counterfactual value of Y under exposure among those who were not exposed and survived (which is a combination of the “always survivors” and the “survive under non-exposure”). This parameter differs in proportion to the fraction of the population that is “survive only under non-exposure”.

Estimating this parameter requires estimating the fraction of the population that would survive only under non-exposure. The fraction of the population that died includes the “never survive” and the “survive only under non-exposure” who were in fact exposed. For the 1940 birth cohort, these two groupsconstitute 100-78%=22% of the population. Among this 22%, what fraction were “never survivors” and what fraction were “survive only under non-exposure”? A very conservative estimate of “never survivors” could be based on contemporary Japanese survival rates. The 2008 Japanese life tables indicated 7% mortality rate between ages 25 and 58, and a 16% mortality rate between ages 25 and 68.[[14](#_ENREF_14)] We think it is safe to say that if 7% of the contemporary Japanese population died between these ages, at least this fraction would have died regardless of their SES status from the 1940 US birth cohort (for whom the high SES group was much poorer than the contemporary Japanese population and the health care was much worse). So, let us assume that among the 1940 birth cohort, at least 7% of the population were never survivors, leaving 22%-7%=15% of the population at most as survive under non-exposure. For the 1930 birth cohort, we use a similar approach to estimate that 46%-15%=31% of the population could plausibly be in the “survive only if unexposed” category. This implies that if the true effect of exposure on smoking were null, but that 40% of the survive under non-exposure group would smoke (when not exposed) compared to only 20% smoking among the always-survivor group (when not exposed), we would find that the correction factor for the 1940 cohort was 0.4-(0.85\*0.4+0.15\*0.2)=0.03. For the 1930 cohort, the correction factor would be 0.4-(0.69\*0.4+0.31\*0.2)=0.062. This suggests that the size of the bias introduced by selective mortality is small compared to the size of the difference in health outcomes one must assume between the “survive under non-exposure” and “always survive” groups. Although we have no way of knowing the actual values of these parameters, we think it unlikely that they are large enough to introduce substantial bias.

**eAPPENDIX D: Potential sources of measurement error**

Our analyses were subject to several sources of potential measaurement error. First, there were changes in the measurement of physical activity in the HRS over time. Second, we were concerned about potential error in the measurement of health behaviors, particularly physical activity. Third, we measured alcohol consumption by the number of drinks respondents had on days they consumed alcohol in the last 3 months rather than the regularity of drinking. Fourth, we measured health behaviors at a unique time point rather than over the life-course. We discuss the implications of these potential sources of measurement error below.

*Change in physical activity measurement*

The response categories for physical inactivity changed in the HRS between waves 3-6 and 7-8, such that an individual who participated in physical activity 2 times/week would have been classified as physically inactive in waves 3-6, but active in waves 7-8. In order to assess the sensitivity of our results to this potential misclassification, we repeated analyses utilizing only measures of health behaviours through wave 6 and found that our inference was unchanged (**eTable 12**).

*Errors in the measurement of health behaviors*

Measurement of health behaviors in the HRS is not ideal. For example, smoking and physical activity were measured as binary variables in HRS and likely with error. It is possible that a greater proportion of the social gradient in mortality would have been explained had we been able to model these potential mediators using dose information. We were concerned about available measures of physical activity, in particular, because they do not distinguish between activities related to work and those related to household activity or exercise. Recent methods have been developed to test the robustness of the controlled direct effect when the mediator is measured with error. VanderWeele (2012) published on the role of measurement error in mediation analysis.[[15](#_ENREF_15)] Additionally, le Cessie and colleagues (2012) published a recent report that provides formulae for correcting the controlled direct effect for potential bias introduced by non-differential error in the measurement of the mediator.[[16](#_ENREF_16)]

We were curious about the sensitivity of our results to non-differential error in the measurement of health behaviors. To the best of our knowledge, sensitivity analyses have not yet been developed to correct controlled direct effects for errors in the measurement of multiple mediators, as in the case of our analysis. As such, we provide a simple example using physical inactivity as the mediator. Briefly, we can estimate corrected direct effects at various reliability ratios (the proportion of the variance in the mediator M\* measured with error explained by the mediator M measured without error) using the coefficients from two models: (1) a model for the controlled direct effect that regresses mortality on adult SES, physical inactivity, and time-fixed baseline covariates and (2) a model regressing physical inactivity on adult SES and time-fixed covariates. We show in **eTable 13** below an estimate of the observed controlled direct effect of adult SES not mediated by physical inactivity assuming no error in the measurement of physical inactivity (reliability ratio of 1.0), as well as corrected controlled direct effects at various reliability ratios from 0.9 to 0.5; confidence intervals were obtained by generating bootstrap distributions for each parameter, with 1000 re-samples.

Results confirm that errors in the measurement of the mediator (physical inactivity, in this case), result in overestimation of the controlled direct effect of adult SES on mortality (underestimation of the proportion of the total effect mediated). The magnitude of the bias is minimal for reliability ratios above 0.8; however, as the reliability ratio drops below 0.8, the observed direct effect assuming no measurement error becomes an increasingly greater overestimate of the corrected direct effect. For example, for a reliability ratio of 0.70, which translates to nearly one-third of the true variability in the measurement of physical inactivity being attributed to error, the controlled direct effect is RR=2.03 (compared to RR=2.38 for a reliability ratio of 1.0).

*Alternative measures of alcohol consumption*

We considered an alternative measure of alcohol consumption based on the regularity of drinking. In our main analyses we measured alcohol consumption by the number of drinks respondents had on days they consumed alcohol in the last 3 months, with non-drinkers coded as zero; this measure is intended to distinguish between non-drinkers, moderate drinkers (those who drink 1-2 or 3-4 drinks on drinking occasions), and problem/binge drinkers (respondents who consume 5+ drinks on drinking occasions). As an alternative, we recoded alcohol consumption to reflect the number of drinks consumed per week and categorized this according to the NIAAA guidelines[[17](#_ENREF_17)] for men and women—this is a three category variable (Men: 0 drinks/week, 0<drinks/week≤14 drinks/week, >14 drinks/week; Women: 0 drinks/week, 0<drinks/week≤7 drinks/week, >7 drinks/week). As seen in **eTable 14**, categorizing alcohol consumption based on the regularity of drinking seems to miss some of the elevated risk associated with heavier consumption relative to zero consumption that we observed using the prior categorization based on the number of drinks on drinking occasions. Additionally, as seen in **eTable 15**, the categorization of alcohol consumption based on the regularity of drinking also seems to obscure some of the social gradient in heavy drinking that we observed with the prior categorization (i.e., the elevated odds of heavy consumption relative to zero consumption comparing the lowest to highest social class). Thus, for our main results we categorized alcohol consumption based on the number of drinks consumed on drinking occasions.

*Lack of life-course measures*

We measured behaviors at a unique time point one wave prior to outcome assessment. However, it is likely that health behaviors may influence mortality in a cumulative fashion over the life-course. For example, it is intuitive that a 60 year-old respondent who began exercising at age 58 and another who has exercised continuously since age 25 may have different risks for mortality in the subsequent time interval. Similarly, epidemiologic evidence supports the conclusion that smoking has dose-response and cumulative effects on mortality and also that risks wane after quitting. Unfortunately, the HRS picks up respondents in their 50s and does not include a retrospective assessment of health behaviors in earlier years; therefore, we cannot capture much of this information. However, we are to some extent able to incorporate respondent’s history of health behaviors since enrollment, for example, by modeling health behaviors as a moving average. Moving averages incorporate prior exposure. Suppose we have two respondents with 5 visits each. Respondent A did not report smoking at any visit, whereas respondent B smoked for the first 3 visits and then quit at wave 4. If we assign health behaviors based on the unique value at time t-1), both respondents would have a value of zero for visit 5 based on their visit 4 smoking status. In contrast, with the moving average, respondent A would have a value of zero for visit 5, whereas respondent B would have a value of 0.75, reflecting prior smoking history over all prior waves up to and including wave t-1. To test for whether incorporating prior history influenced our conclusions, we created moving averages of current smoking, alcohol consumption, and physical activity based on all prior waves up to and including wave t-1 and assessed the controlled direct effects of SES on mortality comparing results from models with: (1) behaviors modeled categorically at time t-1 (**as shown in Table 4**); (2) behaviors as a continuous moving average over all prior waves until t-1 (**eTable 16**); and (3) behaviors as a continuous moving average over all prior waves until t-1, modeled using restricted cubic splines (**eTable 17**). As seen in the tables below, estimates were very similar and conclusions the same in all cases, suggesting our results were not sensitive to measurement of health behaviors at a unique point versus a moving average. There are several potential explanations for these results. First, health behaviors at the prior wave may be a reasonable measure of cumulative exposure to the extent that these measures are correlated over time. Second, as noted earlier, our measure of cumulative health behaviors using the moving average does not incorporate health behavior over the entire life course from early-adulthood and still misses a potentially important exposure window.

**References**

1. Braveman PA, Cubbin C, Egerter S, et al. Socioeconomic status in health research: one size does not fit all. *JAMA*. 2005;294:2879-88.

2. Shavers VL. Measurement of socioeconomic status in health disparities research. *J Natl Med Assoc.* 2007;99:1013-23.

3. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling.* 1999;6:1-55.

4. Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct Equ Modeling.* 2001;8:430-57.

5. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51:1173-82.

6. Kaufman JS, Maclehose RF, Kaufman S. A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. *Epidemiol Perspect Innov*. 2004;1:4.

7. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. 2009;20:18-26.

8. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11:550-60.

9. Brumback BA, Bouldin ED, Zheng HW, Cannell MB, Andresen EM. Testing and estimating model-adjusted effect-measure modification using marginal structural models and complex survey data. *Am J Epidemiol.* 2010;172:1085-91.

10. Health and Retirement Study. Sample sizes and response rates [Internet]. 2011 [Accessed December 20, 2011]. Available from: <http://hrsonline.isr.umich.edu/sitedocs/sampleresponse.pdf>.

11. Health and Retirement Study. Sampling weights: Revised for tracker 2.0 and beyond [Internet]. N.D. [Accessed March 9, 2013]. Available from: <http://hrsonline.isr.umich.edu/sitedocs/wghtdoc.pdf>.

12. Centers for Disease Control, National Center for Health Statistics. United States life tables, 2006 [Internet]. 2010. [Accessed March 9, 2013]. Available from: <http://www.cdc.gov/nchs/data/nvsr/nvsr58/nvsr58_21.pdf>.

13. Chiba Y, VanderWeele TJ. A simple method for principal strata effects when the outcome has been truncated due to death. *Am J Epidemiol*. 2011;173: 745-51.

14. Ministry of Health, Labour and Welfare, Statistics and Information Department. Life tables for Japan 2008 [Internet]. N.D. [Accessed March 9, 2013]. Available from: <http://www.mhlw.go.jp/za/0723/b09/b09.pdf>.

15. VanderWeele TJ, Valeri L, Ogburn EL. The role of measurement error and misclassification in mediation analysis: mediation and measurement error. *Epidemiology*. 2012;23:561-4.

16. le Cessie S, Debeij J, Rosendaal FR, Cannegieter SC, Vandenbroucke JP. Quantification of bias in direct effects estimates due to different types of measurement error in the mediator. *Epidemiology.* 2012;23:551-60.

17. National Institute on Alcohol Abuse and Alcoholism. Rethinking Drinking (NIH Publication No. 10-3770) [Internet]. 2010. [Accessed March 9, 2013]. Available from: <http://pubs.niaaa.nih.gov/publications/RethinkingDrinking/Rethinking_Drinking.pdf>.

**eTable 1.** Summary of indicators used to measure SES in early-life and adulthood, Health and Retirement Study

|  |  |  |  |
| --- | --- | --- | --- |
| **Indicatora** | **Yearb** | **Description** | **Latent factorc** |
| *Indicators of early-life SES* | | | |
| Father’s educational attainment (years) | 1992 | 4 indicators (0-7, 8-11, 12, >12) | Early-life SES |
| Mother’s educational attainment (years) | 1992 | 4 indicators (0-7, 8-11, 12, >12) | Early-life SES |
| Father’s occupation | 1998 | Dichotomous (manual/unskilled service, professional/white collar) | Early-life SES |
| Birth in the southern US | 1992 | Dichotomous (yes/no) | Early-life SES |
| Rural residence during childhood | 1992 | Dichotomous (yes/no) | Early-life SES |
| Self rated childhood SES | 1998 | 3 indicators (varied/poor, about average, pretty well off financially) | Not included in final model |
| Family moved for financial reasons | 1998 | Dichotomous (yes/no) | Not included in final model |
| Father unemployed for several months | 1998 | 3 indicators (no, yes/never lived with father) | Not included in final model |
| *Indicators of adult SES* | | | |
| Respondent’s educational attainment | 1992 | 4 indicators (<HS, HS/GED, some college, college+) | Adult SES |
| Respondent’s longest job occupation | 1992 | Dichotomous (manual/unskilled service, professional/white collar) | Adult SES |
| Respondent’s labor force status | 1992 | Dichotomous (works FT/PT/retired, unemployed/disabled/not in labor force) | Adult SES |
| Household income | 1992 | Split into quartiles (indicators) | Adult SES |
| Household wealth | 1992 | Split into quartiles (indicators) | Adult SES |
| Adult height | 1992 | Split into quartiles (indicators) | Not included in final model |

a Hypothesized indicators of SES selected from the Health and Retirement Study

b Year of assessment

c Indicates the latent factor measured by each indicator in the final two-factor model

**eTable 2.** Summary of health behavioral indicators in the Health and Retirement Study

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Indicator** | **Yearb** | **HRS survey assessment** | **Response categories** | **Coding for analysis** |
| Smoking | 1998-2006 | *“Do you smoke cigarettes now?”* | Binary | Yes/No |
| Alcohol consumption | 1998-2006 | *In the last three months, on the days you drink, about how many drinks do you have?* | Continuous | Drinks on drinking occasions coded as 0 (non-drinkers), 1-2, 3-4, or 5+ drinks |
| Physical inactivitya | 1998-2002 | *“On average over the last 12 months have you participated in vigorous physical activity or exercise three times a week or more? By vigorous physical activity, we mean things like sports, heavy housework, or a job that involves physical labor.”* | Binary | Physical inactivity was defined as engaging in vigorous physical activity less than 3 times per week |
|  | 2004-2006 | *“How often do you take part in sports or activities that are vigorous, such as running or jogging, swimming, cycling, aerobics or gym workout, tennis, or digging with a spade or shovel?”* | Categorical (more than once a week, once a week, one to three times a month, or hardly ever or never) | Physical inactivity was defined as engaging in vigorous physical activity less than once per week |

a Response categories for physical inactivity changed in the HRS between 2002 (Wave 6) and 2004 (Wave 7), such that an individual who participated in physical activity 2 times/week would have been classified as physically inactive prior to 2004, but active after 2004. In order to assess the sensitivity of our results to this potential misclassification, we repeated analyses utilizing only measures of health behaviours through wave 6 and found that our inference was unchanged (**eTable 6**)

b Year of assessment

**eTable 3** Coefficients, Standard Errors, and p-values from Direct Effects Models Regressing Mortality on SES, Health Behavior(s), and Cross-product Terms Representing Interaction Between SES and Health Behavior(s), Health and Retirement Study, n=8037

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Model 2a** | | **Model 3b** | | **Model 4c** | | **Model 5d** | | **Model 6e** | | **Model 7f** | |
|  | **Betag (SE)** | **p-value** | **Beta (SE)** | **p-value** | **Beta (SE)** | **p-value** | **Beta (SE)** | **p-value** | **Beta (SE)** | **p-value** | **Beta (SE)** | **p-value** |
| SESh |  |  |  |  |  |  |  |  |  |  |  |  |
| Q1: highest SES | 0.00 | -- | 0.00 | -- | 0.00 | -- | 0.00 | -- | 0.00 | -- | 0.00 | -- |
| Q2 | 0.21 (0.14) | 0.120 | 0.16 (0.14) | 0.246 | 0.10 (0.25) | 0.699 | 0.04 (0.27) | 0.883 | -0.03 (0.27) | 0.904 | 0.86 (0.40) | 0.034 |
| Q3 | 0.46 (0.14) | <0.001 | 0.31 (0.14) | 0.027 | 0.45 (0.25) | 0.070 | 0.10 (0.26) | 0.710 | -0.07 (0.27) | 0.799 | 0.72 (0.37) | 0.054 |
| Q4: lowest SES | 1.00 (0.14) | <0.001 | 0.83 (0.14) | <0.001 | 0.94 (0.25) | <0.001 | 0.67 (0.26) | 0.011 | 0.31 (0.27) | 0.246 | 1.11 (0.48) | 0.019 |
| Current smoker |  |  |  |  |  |  |  |  |  |  |  |  |
| No | 0.00 | -- |  |  |  |  | 0.00 | -- | 0.00 | -- | 0.00 | -- |
| Yes | 0.81 (0.21) | <0.001 |  |  |  |  | 0.69 (0.21) | <0.001 | 0.59 (0.21) | 0.005 | 0.66 (0.44) | 0.134 |
| Alcohol |  |  |  |  |  |  |  |  |  |  |  |  |
| 0 drinks |  |  | 0.00 | -- |  |  | 0.00 | -- | 0.00 | -- | 0.00 | -- |
| 1-2 |  |  | -0.70 (0.20) | <0.001 |  |  | -0.64 (0.20) | 0.001 | -0.48 (0.20) | 0.018 | -0.47 (0.35) | 0.180 |
| 3-4 |  |  | -0.41 (0.35) | 0.244 |  |  | -0.52 (0.35) | 0.139 | -0.38 (0.36) | 0.284 | -1.42 (0.47) | 0.003 |
| 5+ |  |  | 0.35 (0.60) | 0.562 |  |  | -0.12 (0.57) | 0.837 | 0.10 (0.54) | 0.851 | -4.24 (0.76) | <0.001 |
| Physically inactivei |  |  |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  | 0.00 | -- | 0.00 | -- | 0.00 | -- | 0.00 | -- |
| Yes |  |  |  |  | 1.11 (0.20) | <0.001 | 1.01 (0.20) | <0.001 | 0.81 (0.21) | <0.001 | 1.39 (0.30) | <0.001 |
| SES\*smoking |  |  |  |  |  |  |  |  |  |  |  |  |
| SES Q2\*smoke | -0.19 (0.28) | 0.216 |  |  |  |  | -0.15 (0.28) | 0.577 | -0.12 (0.28) | 0.653 | -0.89 (0.75) | 0.238 |
| SES Q3\*smoke | -0.00 (0.25) | 0.986 |  |  |  |  | 0.02 (0.24) | 0.385 | -0.10 (0.26) | 0.693 | -0.02 (0.55) | 0.970 |
| SES Q4\*smoke | -0.31 (0.24) | 0.193 |  |  |  |  | -0.21 (0.24) | 0.385 | -0.17 (0.25) | 0.504 | -0.74 (0.51) | 0.147 |
| SES\*alcohol consumption |  |  |  |  |  |  |  |  |  |  |  |  |
| SES Q2\*1-2 |  |  | -0.16 (0.30) | 0.599 |  |  | -0.11 (0.30) | 0.705 | -0.14 (0.30) | 0.633 | 0.36 (0.55) | 0.513 |
| SES Q2\*3-4 |  |  | -0.28 (0.51) | 0.582 |  |  | -0.23 (0.50) | 0.655 | -0.19 (0.51) | 0.714 | 0.36 (0.75) | 0.637 |
| SES Q2\*5+ |  |  | -0.50 (0.81) | 0.533 |  |  | -0.06 (0.79) | 0.942 | -0.23 (0.78) | 0.765 | 5.74 (1.04) | <0.001 |
| SES Q3\*1-2 |  |  | 0.51 (0.28) | 0.074 |  |  | 0.50 (0.28) | 0.077 | 0.48 (0.28) | 0.091 | 0.31 (0.45) | 0.496 |
| SES Q3\*3-4 |  |  | 0.39 (0.47) | 0.407 |  |  | 0.37 (0.47) | 0.422 | 0.38 (0.48) | 0.430 | 0.84 (0.76) | 0.270 |
| SES Q3\*5+ |  |  | 0.02 (0.69) | 0.977 |  |  | 0.21 (0.65) | 0.742 | 0.16 (0.62) | 0.791 | 4.87 (0.98) | <0.001 |
| SES Q4\*1-2 |  |  | 0.23 (0.32) | 0.458 |  |  | 0.19 (0.31) | 0.546 | 0.25 (0.31) | 0.425 | -0.72 (0.52) | 0.170 |
| SES Q4\*3-4 |  |  | 0.00 (0.50) | 0.994 |  |  | -0.03 (0.48) | 0.949 | 0.20 (0.49) | 0.679 | 0.32 (1.04) | 0.760 |
| SES Q4\*5+ |  |  | -0.28 (0.67) | 0.679 |  |  | -0.08 (0.65) | 0.904 | -0.12 (0.63) | 0.850 | 4.17 (1.05) | <0.001 |
| SES\*physical inactivity |  |  |  |  |  |  |  |  |  |  |  |  |
| SES Q2\*inactive |  |  |  |  | 0.08 (0.28) | 0.769 | 0.10 (0.29) | 0.740 | 0.01 (0.29) | 0.964 | -1.13 (0.46) | 0.015 |
| SES Q3\*inactive |  |  |  |  | 0.01 (0.28) | 0.963 | 0.05 (0.28) | 0.853 | -0.09 (0.28) | 0.755 | -0.87 (0.45) | 0.053 |
| SES Q4\*inactive |  |  |  |  | -0.09 (0.27) | 0.740 | -0.03 (0.27) | 0.911 | -0.16 (0.27) | 0.569 | -0.69 (0.50) | 0.178 |

Abbreviations: SE, standard error; SES, socioeconomic status

**a** Model 2: SES adjusted for time-fixed covariates (age, sex, race, early-life SES) + smoke, lagged 1 visit

**b** Model 3: SES adjusted for time-fixed covariates (age, sex, race, early-life SES) + drink, lagged 1 visit

**c** Model 4: SES adjusted for time-fixed covariates (age, sex, race, early-life SES) + physical inactivity, lagged 1 visit

**d** Model 5: SES adjusted for time-fixed covariates (age, sex, race, early-life SES) + smoke + drink + physical inactivity, lagged 1 visit

**e** Model 6: Model 5 + time varying confounders (i.e., cancer, psychiatric disorder, stroke, heart disease, diabetes, self rated health), lagged 2 visits

**f** Model 7: Results from marginal structural model fit with inverse probability weights, where each participant in the model is weighted by the inverse of the probability that he had the mediating behaviors he actually had, given his SES, baseline covariates, and past history of behaviors and time-varying covariates

g Beta coefficients represent log risk ratios

h Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles (Q1-Q4) of adult SES

i Between waves 4 in 1998 and 6 in 2002, respondents were asked if they engaged in vigorous physical activity 3 or more times per week for the past 12 months; physical inactivity was defined as engaging in vigorous physical activity less than 3 times per week. For waves 7 in 2004 and 8 in 2006 respondents were asked if they engaged in vigorous physical activity more than once per week, once per week, 1 to 3 times per month, or never. For these last two waves, physical inactivity was defined as engaging in vigorous physical activity less than once per week.

**eTable 4.** Role of Time-varying Health Behaviors in “Explaining” Association Between Education and Mortality, Health and Retirement Study, n=8037a

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Model 1**b  **RR (95%CI)** | **Model 2**c  **RR (95%CI)** | **Model 3**d  **RR (95%CI)** | **Model 4**e  **RR (95%CI)** | **Model 5**f  **RR (95%CI)** | **Model 6**g  **RR (95%CI)** | **Model 7**h,i,j  **RR (95%CI)** |
| Educational attainment |  |  |  |  |  |  |  |
| College or greater | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| At least some college | 1.14 (0.87, 1.48) | 1.06 (0.81, 1.39) | 1.06 (0.81, 1.38) | 1.10 (0.85, 1.44) | 1.00 (0.77, 1.30) | 0.86 (0.66, 1.11) | 1.00 (0.61, 1.63) |
| High school/GED | 1.21 (0.95, 1.54) | 1.12 (0.88, 1.42) | 1.08 (0.84, 1.38) | 1.14 (0.90, 1.45) | 0.98 (0.77, 1.25) | 0.85 (0.67, 1.08) | 0.97 (0.61, 1.54) |
| <High school | 1.78 (1.40, 2.25) | 1.57 (1.24, 1.99) | 1.51 (1.19, 1.93) | 1.57 (1.24,1.99) | 1.29 (1.00, 1.61) | 0.90 (0.71, 1.15) | 1.10 (0.70, 1.74) |
| Current smoker |  |  |  |  |  |  |  |
| No |  | 1.00 |  |  | 1.00 | 1.00 | 1.00 |
| Yes |  | 2.05 (1.75, 2.39) |  |  | 1.88 (1.61, 2.20) | 1.75 (1.50, 2.05) | 1.40 (0.96, 2.02) |
| Alcohol |  |  |  |  |  |  |  |
| 0 drinks |  |  | 1.00 |  | 1.00 | 1.00 | 1.00 |
| 1-2 |  |  | 0.53 (0.42, 0.65) |  | 0.57 (0.46, 0.70) | 0.68 (0.55, 0.85) | 0.62 (0.44, 0.89) |
| 3-4 |  |  | 0.66 (0.47, 0.93) |  | 0.60 (0.43, 0.84) | 0.75 (0.53, 1.05) | 0.33 (0.19, 0.59) |
| 5+ |  |  | 1.19 (0.79, 1.79) |  | 0.93 (0.63, 1.38) | 1.08 (0.72, 1.62) | 1.68 (0.72, 3.90) |
| Physically inactivek |  |  |  |  |  |  |  |
| No |  |  |  | 1.00 | 1.00 | 1.00 | 1.00 |
| Yes |  |  |  | 3.17 (2.62, 3.84) | 2.91 (2.41, 3.53) | 2.15 (1.77, 2.61) | 1.94 (1.39, 2.72) |

Abbreviations: RR, risk ratio; CI, confidence interval

a All models include respondent-level sample weights

b Model 1: education adjusted for time-fixed covariates (age, sex, race, early-life SES)

c Model 2: Model 1 + smoke, lagged 1 visit

d Model 3: Model 1 + alcohol consumption, lagged 1 visit

e Model 4: Model 1 +physical inactivity, lagged 1 visit

f Model 5: Model 1 + smoke + drink + physical inactivity, lagged 1 visit

g Model 6: Model 5 + time varying confounders (i.e., cancer, psychiatric disorder, stroke, heart disease, diabetes, self rated health), lagged 2 visits

h Model 7: Results from marginal structural model fit with inverse probability weights, where each participant in the model is weighted by the inverse of the probability that he had the mediating behaviors he actually had, given his education, baseline covariates, and past history of behaviors and time-varying covariates

i Combined weights (the product of individual sample weights and inverse probability weights) were truncated at the 99th percentile

j Health behaviors were included as independent variables in weight models and effects of health behaviors on mortality are not interpretable

k Between waves 4 in 1998 and 6 in 2002, respondents were asked if they engaged in vigorous physical activity 3 or more times per week for the past 12 months; physical inactivity was defined as engaging in vigorous physical activity less than 3 times per week. For waves 7 in 2004 and 8 in 2006 respondents were asked if they engaged in vigorous physical activity more than once per week, once per week, 1 to 3 times per month, or never. For these last two waves, physical inactivity was defined as engaging in vigorous physical activity less than once per week.

**eTable 5**. Role of Time-varying Health Behaviors (including BMI as a Proxy for Nutrition) in “Explaining” Association Between SES and Mortality, Health and Retirement Study, n=8037a

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Model 1**b  **RR (95%CI)** | **Model 2**c  **RR (95%CI)** | **Model 3**d  **RR (95%CI)** | **Model 4**e  **RR (95%CI)** | **Model 5**f  **RR (95%CI)** | **Model 6**g  **RR (95%CI)** |
| SESh |  |  |  |  |  |  |
| Q1: highest SES | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Q2 | 1.23 (0.98, 1.55) | 1.19 (0.94, 1.49) | 1.13 (0.90, 1.43) | 1.18 (0.94, 1.48) | 1.29 (1.02, 1.62) | 1.11 (0.88, 1.40) |
| Q3 | 1.75 (1.40, 2.19) | 1.61 (1.30, 2.01) | 1.53 (1.22, 1.93) | 1.58 (1.27, 1.97) | 1.83 (1.46, 2.29) | 1.37 (1.10, 1.72) |
| Q4: lowest SES | 2.84 (2.25, 3.60) | 2.48 (1.96, 3.14) | 2.39 (1.88, 3.05) | 2.37 (1.87, 3.00) | 2.93 (2.31, 3.70) | 1.90 (1.49, 2.41) |
| Current smoker |  |  |  |  |  |  |
| No |  | 1.00 |  |  |  | 1.00 |
| Yes |  | 1.94 (1.66, 2.26) |  |  |  | 1.63 (1.39, 1.92) |
| Alcohol |  |  |  |  |  |  |
| Abstain |  |  | 1.00 |  |  | 1.00 |
| 1-2 |  |  | 0.57 (0.46, 0.71) |  |  | 0.58 (0.47, 0.72) |
| 3-4 |  |  | 0.69 (0.49, 0.97) |  |  | 0.61 (0.44, 0.85) |
| 5+ |  |  | 1.16 (0.77, 1.73) |  |  | 0.89 (0.61, 1.31) |
| Physically inactivei |  |  |  |  |  |  |
| No |  |  |  | 1.00 |  | 1.00 |
| Yes |  |  |  | 3.05 (2.52, 3.69) |  | 2.92 (2.40, 3.54) |
| Body mass index (BMI) |  |  |  |  |  |  |
| BMI<25 |  |  |  |  | 1.00 | 1.00 |
| 25≤BMI<30 |  |  |  |  | 0.59 (0.50, 0.70) | 0.63 (0.54, 0.75) |
| BMI≥30 |  |  |  |  | 0.61 (0.50, 0.73) | 0.60 (0.50, 0.73) |

Abbreviations: RR, risk ratio; CI, confidence interval; SES, socioeconomic status; BMI, body mass index

a All models include respondent-level sample weights

b Model 1: SES adjusted for time-fixed covariates (age, sex, race, early-life SES)

c Model 2: Model 1 + smoke, lagged 1 visit

d Model 3: Model 1 + alcohol consumption, lagged 1 visit

e Model 4: Model 1 + physical inactivity, lagged 1 visit

f Model 5: Model 1 + BMI, lagged 1 visit

g Model 6: Model 1 + smoke + drink + physical inactivity + BMI, lagged 1 visit

h Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles (Q1-Q4) of adult SES

i Between waves 4 in 1998 and 6 in 2002, respondents were asked if they engaged in vigorous physical activity 3 or more times per week for the past 12 months; physical inactivity was defined as engaging in vigorous physical activity less than 3 times per week. For waves 7 in 2004 and 8 in 2006 respondents were asked if they engaged in vigorous physical activity more than once per week, once per week, 1 to 3 times per month, or never. For these last two waves, physical inactivity was defined as engaging in vigorous physical activity less than once per week.

**eTable 6.** Weighted Percentage Distribution of Health Behaviors in 1998by SES, Health and Retirement Study, n=8037a

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **SES**a | | | | |
|  | *Q1: highest SES* | *Q2* | *Q3* | *Q4: lowest SES* | *TOTAL* |
| Current smoker |  |  |  |  |  |
| No | 86.7 | 82.7 | 76.3 | 71.0 | 80.1 |
| Yes | 13.3 | 17.3 | 23.7 | 29.0 | 19.9 |
| Alcohol consumption |  |  |  |  |  |
| 0 drinks | 49.8 | 62.9 | 73.1 | 83.4 | 65.3 |
| 1-2 | 41.5 | 28.0 | 18.1 | 8.8 | 26.0 |
| 3-4 | 7.8 | 7.0 | 5.6 | 4.2 | 6.4 |
| 5+ | 0.9 | 2.1 | 3.1 | 3.6 | 2.3 |
| Physically inactiveb |  |  |  |  |  |
| No | 54.5 | 51.8 | 51.8 | 36.5 | 48.6 |
| Yes | 45.5 | 48.2 | 48.2 | 63.5 | 51.4 |

Abbreviations: SES, socioeconomic status

a  Weighted by respondent-level sample weights

b Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles (Q1-Q4) of adult SES

c Between waves 4 in 1998 and 6 in 2002, respondents were asked if they engaged in vigorous physical activity 3 or more times per week for the past 12 months; physical inactivity was defined as engaging in vigorous physical activity less than 3 times per week. For waves 7 in 2004 and 8 in 2006 respondents were asked if they engaged in vigorous physical activity more than once per week, once per week, 1 to 3 times per month, or never. For these last two waves, physical inactivity was defined as engaging in vigorous physical activity less than once per week.

**eTable 7.** Role of Time-varying Health Behaviors in “Explaining” Association Between SES and Mortality, Risk Difference Models, Health and Retirement Study, n=8037a

|  |  |  |
| --- | --- | --- |
|  | **Model 1**b  **RD (95%CI)** | **Model 7**c,d  **RD (95%CI)** |
| SESh |  |  |
| Q1: highest SES | 0.000 | 0.000 |
| Q2 | 0.005 (0.001, 0.009) | 0.006 (-0.005, 0.017) |
| Q3 | 0.014 (0.009, 0.019) | 0.007 (-0.002, 0.016) |
| Q4: lowest SES | 0.035 (0.029, 0.042) | 0.017 (0.003, 0.032) |
| Current smoker |  |  |
| No |  | 0.000 |
| Yes |  | 0.015 (-0.003, 0.033) |
| Alcohol |  |  |
| 0 drinks |  | 0.000 |
| 1-2 |  | -0.010 (-0.018, -0.002) |
| 3-4 |  | -0.013 (-0.021, -0.004) |
| 5+ |  | 0.032 (-0.035, 0.099) |
| Physically inactivee |  |  |
| No |  | 0.000 |
| Yes |  | 0.020 (0.011, 0.028) |

Abbreviations: RD, risk difference; CI, confidence interval; SES, socioeconomic status

a All models include respondent-level sample weights

b Model 1: SES adjusted for time-fixed covariates (age, sex, race, early-life SES)

c Model 7: Results from marginal structural model fit with inverse probability weights, where each participant in the model is weighted by the inverse of the probability that he had the mediating behaviors he actually had, given his SES, baseline covariates, and past history of behaviors and time-varying covariates

d Health behaviors were included as independent variables in weight models and effects of health behaviors on mortality are not interpretable

e Between waves 4 in 1998 and 6 in 2002, respondents were asked if they engaged in vigorous physical activity 3 or more times per week for the past 12 months; physical inactivity was defined as engaging in vigorous physical activity less than 3 times per week. For waves 7 in 2004 and 8 in 2006 respondents were asked if they engaged in vigorous physical activity more than once per week, once per week, 1 to 3 times per month, or never. For these last two waves, physical inactivity was defined as engaging in vigorous physical activity less than once per week.

**eTable 8.** Association Between Health Conditions Measured at the Prior Visit and Current Smoking, Adjusted for Age and Gender, Health and Retirement Study, n=8037a

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Risk Ratio** |  | |
| **95%CI** | |
| RR cancer (vs. not) | 1.03 | 0.89 | 1.20 |
| RR psychiatric condition (vs. not) | 1.23 | 1.09 | 1.37 |
| RR diabetes (vs. not) | 0.66 | 0.58 | 0.76 |
| RR heart disease (vs. not) | 0.87 | 0.77 | 0.97 |
| RR stroke (vs. not) | 1.16 | 0.98 | 1.39 |
| RR self-rated health very good (vs. excellent) | 1.25 | 1.09 | 1.43 |
| RR self-rated health good (vs. excellent) | 1.64 | 1.42 | 1.89 |
| RR self-rated health fair (vs. excellent) | 1.96 | 1.69 | 2.29 |
| RR self-rated health poor (vs. excellent) | 2.34 | 1.96 | 2.79 |

a All models include respondent-level sample weights

**eTable 9.** Association Between Health Conditions Measured at the Prior Visit and any Alcohol Consumption, Adjusted for Age and Gender, Health and Retirement Study, n=8037a

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Risk Ratio** |  | |
| **95%CI** | |
| RR cancer (vs. not) | 1.20 | 1.10 | 1.31 |
| RR psychiatric condition (vs. not) | 0.84 | 0.76 | 0.93 |
| RR diabetes (vs. not) | 0.67 | 0.60 | 0.75 |
| RR heart disease (vs. not) | 1.00 | 0.92 | 1.07 |
| RR stroke (vs. not) | 0.70 | 0.58 | 0.85 |
| RR self-rated health very good (vs. excellent) | 0.89 | 0.84 | 0.94 |
| RR self-rated health good (vs. excellent) | 0.76 | 0.71 | 0.82 |
| RR self-rated health fair (vs. excellent) | 0.57 | 0.52 | 0.62 |
| RR self-rated health poor (vs. excellent) | 0.40 | 0.34 | 0.47 |

a All models include respondent-level sample weights

**eTable 10.** Association Between Health Conditions Measured at the Prior Visit and Vigorous Physical Activity, Adjusted for Age and Gender, Health and Retirement Study, n=8037a

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Risk Ratio** |  | |
| **95%CI** | |
| RR cancer (vs. not) | 0.97 | 0.91 | 1.04 |
| RR psychiatric condition (vs. not) | 0.84 | 0.78 | 0.90 |
| RR diabetes (vs. not) | 0.83 | 0.78 | 0.88 |
| RR heart disease (vs. not) | 0.94 | 0.89 | 1.00 |
| RR stroke (vs. not) | 0.74 | 0.64 | 0.85 |
| RR self-rated health very good (vs. excellent) | 0.81 | 0.78 | 0.84 |
| RR self-rated health good (vs. excellent) | 0.66 | 0.63 | 0.69 |
| RR self-rated health fair (vs. excellent) | 0.48 | 0.44 | 0.51 |
| RR self-rated health poor (vs. excellent) | 0.30 | 0.26 | 0.34 |

a All models include respondent-level sample weights

**eTable 11.** Distribution of weights a, b, c, and d

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Weight** | **Mean (SD)** | **Min, Max** | **1%, 99%** | **IQR (25%, 75%)** |
|  | 1.00 (1.08) | 0.23, 21.53 | 0.23, 4.69 | 0.73, 0.95 |
|  | 1.00 (1.34) | 0.14, 24.54 | 0.15, 7.33 | 0.75, 0.89 |
|  | 1.00 (0.97) | 0.01, 14.70 | 0.05, 5.23 | 0.59, 0.87 |
|  | 1.00 (0.52) | 0.38, 7.10 | 0.47, 2.78 | 0.68, 1.22 |

Abbreviations: SD, standard deviation

a Weight accounting for the relation between SES and mortality

b Weight accounting for the relation between smoking and mortality

c Weight accounting for the relation between alcohol consumption and mortality

d Weight accounting for confounding of the relation between physical inactivity and mortality

**eTable 12**. Role of Time-varying Health Behaviors in “Explaining” Association Between SES and Mortality, with Analyses Restricted to Survey Waves with Consistent Measurement of Physical Inactivity, Health and Retirement Study, n=7906a

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Model 1**b  **RR (95%CI)** | **Model 2**c  **RR (95%CI)** | **Model 3**d  **RR (95%CI)** | **Model 4**e  **RR (95%CI)** | **Model 5**f  **RR (95%CI)** |
| SESg |  |  |  |  |  |
| Q1: highest SES | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Q2 | 1.18 (0.87, 1.60) | 1.14 (0.84, 1.54) | 1.10 (0.81, 1.49) | 1.16 (0.86, 1.57) | 1.05 (0.77 1.43) |
| Q3 | 1.82 (1.37, 2.43) | 1.68 (1.26, 2.24) | 1.61 (1.20, 2.16) | 1.73 (1.30, 2.30) | 1.43 (1.07, 1.92) |
| Q4: lowest SES | 3.06 (2.27, 4.12) | 2.69 (1.99, 3.65) | 2.61 (1.92, 3.55) | 2.67 (1.98, 3.61) | 2.10 (1.54, 2.86) |
| Current smoker |  |  |  |  |  |
| No |  | 1.00 |  |  | 1.00 |
| Yes |  | 1.91 (1.55, 2.35) |  |  | 1.86 (1.51, 2.29) |
| Alcohol |  |  |  |  |  |
| Abstain |  |  | 1.00 |  | 1.00 |
| 1-2 |  |  | 0.52 (0.39, 0.69) |  | 0.54 (0.41, 0.73) |
| 3-4 |  |  | 0.75 (0.49, 1.14) |  | 0.67 (0.44, 1.02) |
| 5+ |  |  | 0.61 (0.30, 1.22) |  | 0.47 (0.24, 0.93) |
| Physically inactiveh |  |  |  |  |  |
| No |  |  |  | 1.00 | 1.00 |
| Yes |  |  |  | 2.42 (1.95, 3.00) | 2.26 (1.82, 2.81) |

Abbreviations: RR, risk ratio; CI, confidence interval; SES, socioeconomic status

a All models include respondent-level sample weights

b Model 1: SES adjusted for time-fixed covariates (age, sex, race, early-life SES)

c Model 2: Model 1 + smoke, lagged 1 visit

d Model 3: Model 1 + alcohol consumption, lagged 1 visit

e Model 4: Model 1 +physical inactivity, lagged 1 visit

f Model 6: Model 1 + smoke + drink + physical inactivity, lagged 1 visit

g Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles (Q1-Q4) of adult SES

h Between waves 4 in 1998 and 6 in 2002, respondents were asked if they engaged in vigorous physical activity 3 or more times per week for the past 12 months; physical inactivity was defined as engaging in vigorous physical activity less than 3 times per week.

**eTable 13.** Sensitivity Analyses for the Controlled Direct Effect of SES not Mediated Through Physical Inactivity, with Error in Measurement of Physical Inactivity Represented by the Reliability Ratio from 1.0 to 0.5, Health and Retirement Study, n=8037a

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Reliability ratio=1.0** | | | | **Reliability ratio=0.90** | | | | **Reliability ratio=0.80** | | | | **Reliability ratio=0.70** | | | | **Reliability ratio=0.60** | | | | **Reliability ratio=0.50** | | | |
|  | **RR** | | **95%CI** | | **RR** | | **95%CI** | | **RR** | | **95%CI** | | **RR** | | **95%CI** | | **RR** | | **95%CI** | | **RR** | | **95%CI** | |
| SESa |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |
| Q1: highest SES |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |
| Q2 | 1.18 | 0.90 | | 1.43 | 1.17 | 0.89 | | 1.42 | 1.15 | 0.87 | | 1.40 | 1.13 | 0.86 | | 1.37 | 1.10 | 0.83 | | 1.34 | 1.06 | 0.80 | | 1.29 |
| Q3 | 1.59 | 1.21 | | 1.91 | 1.55 | 1.18 | | 1.86 | 1.50 | 1.15 | | 1.81 | 1.44 | 1.10 | | 1.74 | 1.37 | 1.04 | | 1.66 | 1.28 | 0.96 | | 1.55 |
| Q4: lowest SES | 2.38 | 1.79 | | 2.91 | 2.28 | 1.71 | | 2.79 | 2.17 | 1.62 | | 2.66 | 2.03 | 1.52 | | 2.49 | 1.86 | 1.38 | | 2.29 | 1.64 | 1.21 | | 2.04 |

Abbreviations: RR, risk ratio; CI, confidence interval; SES, socioeconomic status

a All models include respondent-level sample weights

b Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles (Q1-Q4) of adult SES

**eTable 14.** Association Between Alternative Measure of Alcohol Consumptiona and Mortality 1998-2008, Adjusted for Age and Sex and Weighted by Individual Sample Weights, Health and Retirement Study, n=8037

|  |  |  |
| --- | --- | --- |
|  | **Risk ratio** | |
| **RR death vs. non-death** | **95%CI** |
| Alcohol consumption |  |  |
| No consumption | 1.00 | -- |
| Moderate consumption | 0.49 | 0.40, 0.59 |
| Heavy consumption | 0.81 | 0.60, 1.10 |

a For men: 0 drinks/week (no consumption), 0<drinks/week≤14 drinks/week (moderate consumption), >14 drinks/week (heavy consumption); For women: 0 drinks/week (no consumption), 0<drinks/week≤7 drinks/week (moderate consumption), >7 drinks/week (heavy consumption)

**eTable 15.** Association between SES and Alternative Measure of Alcohol Consumption,b Adjusted for Age and Sex and Weighted by Individual Sample Weights, Health and Retirement Study, n=8037

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Alcohol consumption** | | | | |
| **RR moderate drinks vs. 0 drinks** | **95%CI** | **RR heavy drinks vs. 0 drinks** | **95%CI** |  |
| SESa |  |  |  |  |  |
| Q1: highest SES | 1.00 | -- | 1.00 | -- |  |
| Q2 | 0.71 | 0.65, 0.76 | 0.87 | 0.71, 1.06 |  |
| Q3 | 0.53 | 0.49, 0.58 | 0.63 | 0.51, 0.78 |  |
| Q4: lowest SES | 0.32 | 0.28, 0.36 | 0.48 | 0.37, 0.62 |  |

Abbreviations: RR, risk ratio; CI, confidence interval; SES, socioeconomic status

a Factor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles (Q1-Q4) of adult SES

b For men: 0 drinks/week (no consumption), 0<drinks/week≤14 drinks/week (moderate consumption), >14 drinks/week (heavy consumption); For women: 0 drinks/week (no consumption), 0<drinks/week≤7 drinks/week (moderate consumption), >7 drinks/week (heavy consumption)

**eTable 16.** Role of Time-varying Health Behaviors, Modeled as a Continuous Moving Average of all Values up to and Including t-1, in “Explaining” the Association Between SES and Mortality, Health and Retirement Study, n=8037a

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Model 1**b  **RR (95%CI)** | **Model 2**c  **RR (95%CI)** | **Model 3**d  **RR (95%CI)** | **Model 4**e  **RR (95%CI)** | **Model 5**f  **RR (95%CI)** | **Model 6**g  **RR (95%CI)** | **Model 7**h,i  **RR (95%CI)** |
| SESj |  |  |  |  |  |  |  |
| Q1: highest SES | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Q2 | 1.22 (0.96, 1.54) | 1.17 (0.92, 1.47) | 1.22 (0.97, 1.54) | 1.20 (0.95, 1.51) | 1.14 (0.91, 1.44) | 0.98 (0.78, 1.23) | 1.12 (0.73, 1.70) |
| Q3 | 1.70 (1.37, 2.13) | 1.55 (1.24, 1.93) | 1.70 (1.36, 2.12) | 1.58 (1.27, 1.97) | 1.44 (1.15, 1.79) | 1.08 (0.86, 1.35) | 1.38 (0.93, 2.03) |
| Q4: lowest SES | 2.61 (2.10, 3.23) | 2.26 (1.82, 2.80) | 2.59 (2.09, 3.21) | 2.15 (1.73, 2.66) | 1.89 (1.53, 2.35) | 1.18 (0.93, 1.49) | 1.64 (1.12, 2.39) |

Abbreviations: RR, risk ratio; CI, confidence interval; SES, socioeconomic status

a All models include respondent-level sample weights

bModel 1: SES adjusted for time-fixed covariates (age, sex, race, early-life SES)

cModel 2: Model 1 + smoke, lagged 1 visit

dModel 3: Model 1 + alcohol consumption, lagged 1 visit

eModel 4: Model 1 +physical inactivity, lagged 1 visit

fModel 5: Model 1 + smoke + drink + physical inactivity, lagged 1 visit

gModel 6: Model 5 + time varying confounders (i.e., cancer, psychiatric disorder, stroke, heart disease, diabetes, self rated health), lagged 2 visits

hModel 7: Results from marginal structural model fit with inverse probability weights, where each participant in the model is weighted by the inverse of the probability that he had the mediating behaviors he actually had, given his SES, baseline covariates, and past history of behaviors and time-varying covariates

iHealth behaviors were included as independent variables in weight models and effects of health behaviors on mortality are not interpretable

jFactor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles (Q1-Q4) of adult SES

kBetween waves 4 in 1998 and 6 in 2002, respondents were asked if they engaged in vigorous physical activity 3 or more times per week for the past 12 months; physical inactivity was defined as engaging in vigorous physical activity less than 3 times per week. For waves 7 in 2004 and 8 in 2006 respondents were asked if they engaged in vigorous physical activity more than once per week, once per week, 1 to 3 times per month, or never. For these last two waves, physical inactivity was defined as engaging in vigorous physical activity less than once per week.

**eTable 17.** Role of Time-varying Health Behaviors in “Explaining” Association Between SES and Mortality, Health and Retirement Study, n=8037; *after incorporating sample weights in all models, measuring health behaviors as a continuous moving average of all values up to and including t-1 modeled using cubic splines, and truncating the product of the inverse probability of treatment weights (Model 7) at the 99th percentile*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Model 1**a  **RR (95%CI)** | **Model 2**b  **RR (95%CI)** | **Model 3**c  **RR (95%CI)** | **Model 4**d  **RR (95%CI)** | **Model 5**e  **RR (95%CI)** | **Model 6**f  **RR (95%CI)** | **Model 7**g,h  **RR (95%CI)** |
| SESi |  |  |  |  |  |  |  |
| Q1: highest SES | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Q2 | 1.22 (0.96, 1.54) | 1.16 (0.92, 1.47) | 1.17 (0.92, 1.47) | 1.20 (0.95, 1.51) | 1.10 (0.87, 1.39) | 0.96 (0.76, 1.20) | 1.10 (0.73, 1.66) |
| Q3 | 1.70 (1.37, 2.13) | 1.53 (1.23, 1.91) | 1.58 (1.26, 1.98) | 1.57 (1.26, 1.96) | 1.35 (1.08, 1.69) | 1.04 (0.83, 1.31) | 1.29 (0.87, 1.92) |
| Q4: lowest SES | 2.61 (2.10, 3.23) | 2.25 (1.81, 2.79) | 2.35 (1.89, 2.94) | 2.14 (1.73, 2.65) | 1.76 (1.41, 2.20) | 1.14 (0.90, 1.44) | 1.58 (1.08, 2.30) |

Abbreviations: RR, risk ratio; CI, confidence interval; SES, socioeconomic status

aModel 1: SES adjusted for time-fixed covariates (age, sex, race, early-life SES)

bModel 2: Model 1 + smoke, lagged 1 visit

cModel 3: Model 1 + alcohol consumption, lagged 1 visit

dModel 4: Model 1 +physical inactivity, lagged 1 visit

eModel 5: Model 1 + smoke + drink + physical inactivity, lagged 1 visit

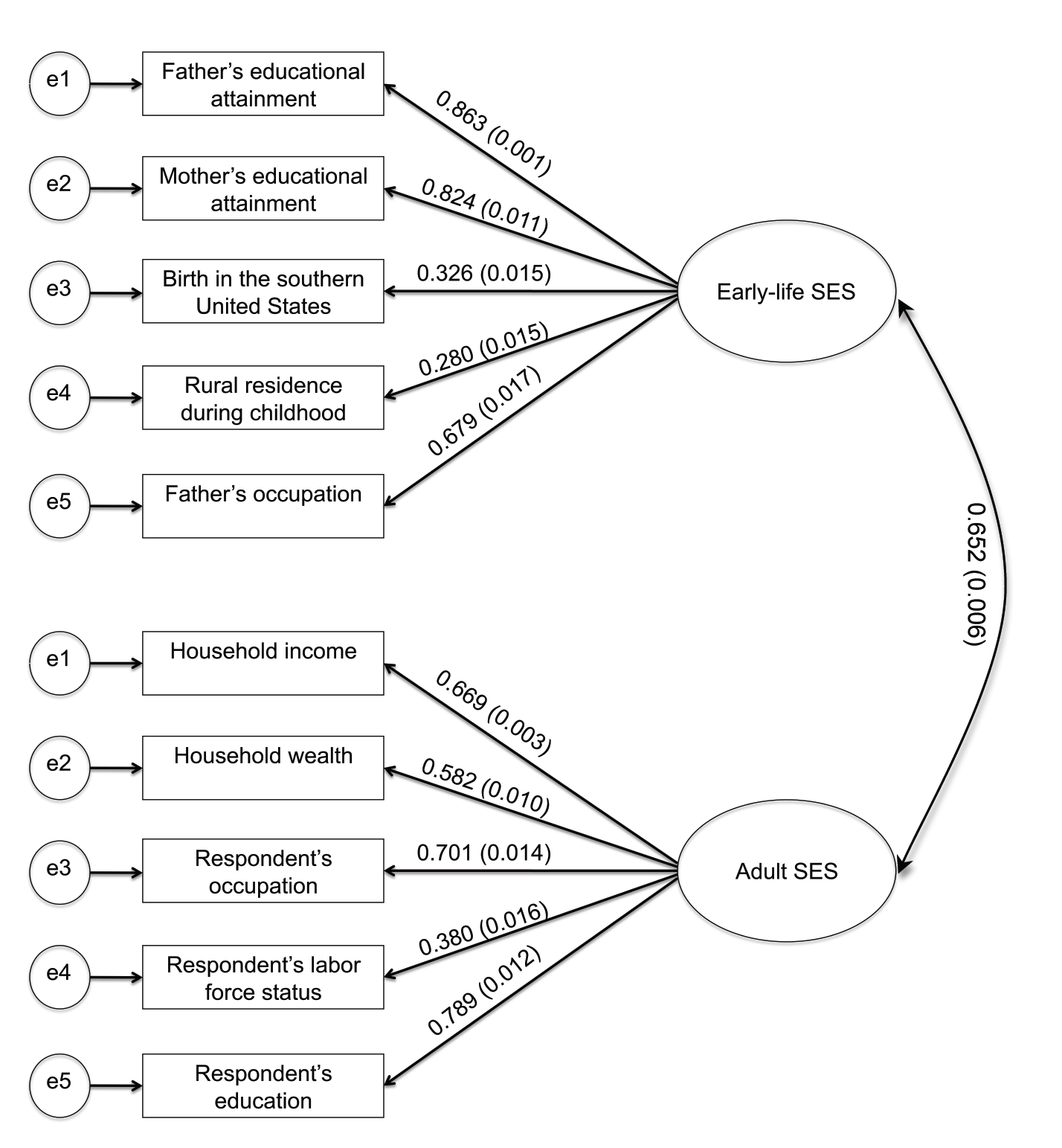
fModel 6: Model 5 + time varying confounders (i.e., cancer, psychiatric disorder, stroke, heart disease, diabetes, self rated health), lagged 2 visits

gModel 7: Results from marginal structural model fit with inverse probability weights, where each participant in the model is weighted by the inverse of the probability that he had the mediating behaviors he actually had, given his SES, baseline covariates, and past history of behaviors and time-varying covariates

hHealth behaviors were included as independent variables in weight models and effects of health behaviors on mortality are not interpretable

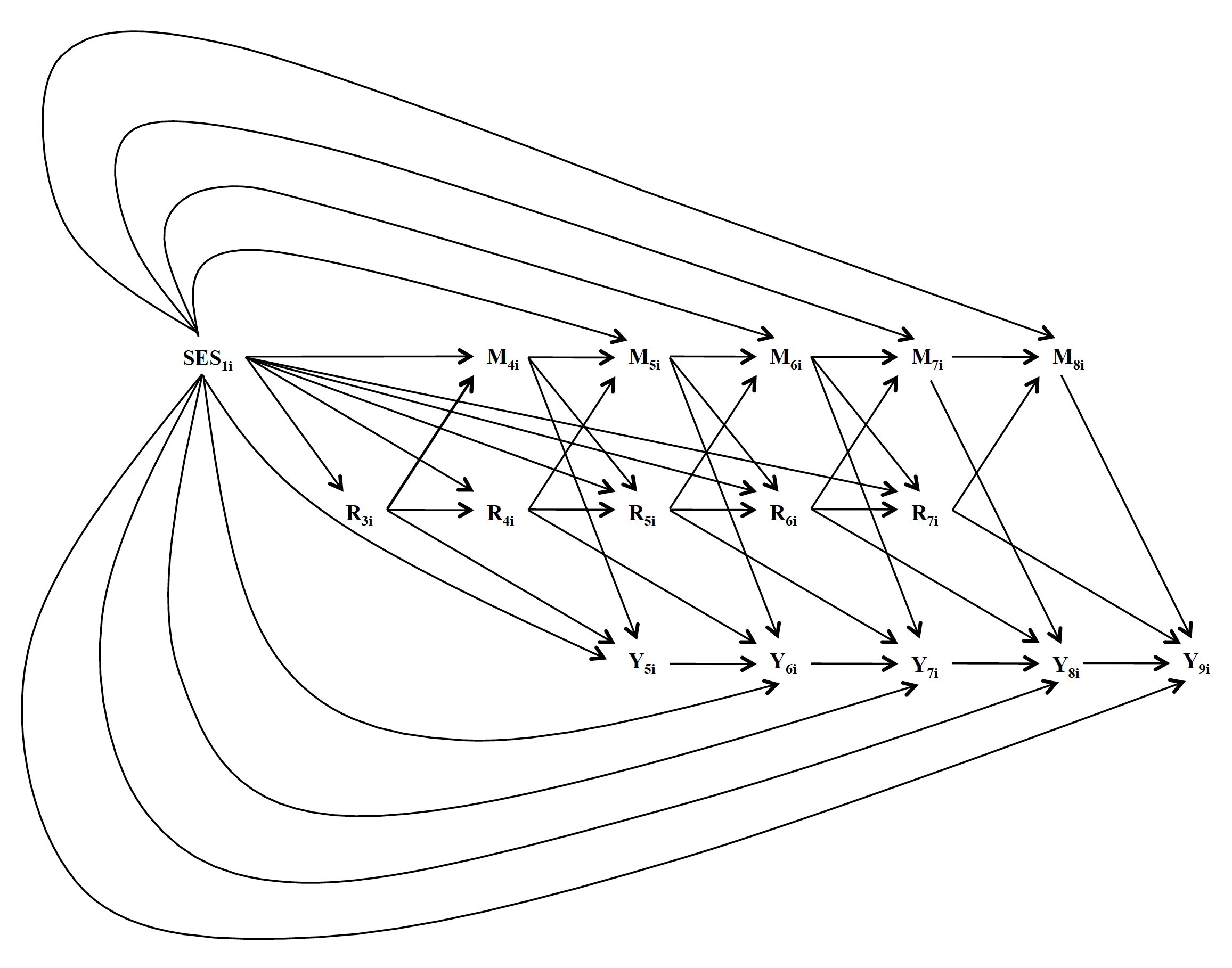
iFactor scores from confirmatory factor analysis with maximum likelihood estimation, split into quartiles (Q1-Q4) of adult SES

**eFigure 1.** Standardized model coefficients and standard errors (in parentheses) from confirmatory factor analysis for measurement of early-life and adult SES; n=9,760 Health and Retirement Study enrollees followed from 1992-2006a



a Coefficients and standard errors standardized using both latent variable and observed variable variances

**eFigure 2.** Diagram Demonstrating Temporal Ordering of SES, Health Behaviors, and Mortality

****

*Notes*

For individual *i* at wave t (from wave 1 in 1992 to wave 9 in 2008): Y represents all-cause mortality; the vector M represents time-varying health behaviors (i.e., smoking, alcohol consumption, and physical inactivity); and the vector R represents potential time-varying confounders (i.e., cancer, psychiatric disorder, stroke, heart disease, diabetes, self-rated health)