

SUPPLEMENTARY APPENDICES:

Supplementary Appendix 1. Number of deaths, person-years of exposures, and health state transitions, ages 50-89

Characteristic	All (N=14,804)	Men (N=6,657)	Women (N=8,147)
Deaths, number	4,305	2,292	2,013
Person-years of follow-up	153,991	66,534	87,457
Health State Transitions, number			
Non-disabled to disabled	5,795	2,363	3,432
Non-disabled to dead	2,503	1,435	1,068
Disabled to non-disabled	3,871	1,508	2,363
Disabled to dead	1,802	857	945

Note: Disability is defined as the presence of at least one activity of daily living (ADL) limitations

Source: U.S. Health and Retirement Survey, 1998-2012

Supplementary Appendix 2. Regression and Matrix Population Models

On a person-year file, we estimated the following multinomial regression model to generate age-and sex-specific transition probabilities across the states non-disabled, disabled, and dead. The models took the form: $\log\left(\frac{p_{ij}(x,t)}{p_{ii}(x,t)}\right) = a_{ij} + b_{1ij}Age + b_{2ij}Age^2 + c_{ij}Female + \sum s_{ijk}Smoke + \sum o_{ijk}Obese + \sum d_{ijk}Alcohol + e_{ij}Z$; where $p_{ij}(x,t)$ is the annual probability of transitioning from state i to state j (for all $i \neq j$); a_{ij} is the intercept; Age is age over the follow-up measured in single years; Smoke, Obese, and Alcohol represent k dichotomous (0/1) categories for each level of the behavioral factor with the low-risk category excluded; e_{ij} represents coefficients for other control variables Z (education and race/ethnicity).

The matrix population model calculations are in the spirit of Hunter and Caswell 2005⁴⁸ but simplify to standard Markov chain techniques for the case of the present study. Based on our multinomial logit model, we estimate transition probabilities for each health state $s \in \{healthy, unhealthy, dead\}$, and annual age classes between ages 50 and 110. The Markov transition matrix is then given by:

$$\tilde{M} = \begin{pmatrix} p_{hh} & p_{uh} & 0 \\ p_{hu} & p_{uu} & 0 \\ p_{hd} & p_{ud} & 1 \end{pmatrix}$$

where each of the matrix blocks \mathbf{p}_{hh} , \mathbf{p}_{hu} , \mathbf{p}_{uh} and \mathbf{p}_{uu} is 61 x 61 (the number of age classes) and has non-zero elements on the first subdiagonal only. The non-zero entries contain probabilities of moving from health state i to health state j when making a particular age transition. \mathbf{p}_{hd} and \mathbf{p}_{ud} are 1 x 61 row vectors. Columns of $\tilde{\mathbf{M}}$ sum to one. The fundamental matrix is given by

$$\mathbf{F} = (\mathbf{I} - \mathbf{M})^{-1} = \begin{pmatrix} \mathbf{q}_{hh} & \mathbf{q}_{uh} \\ \mathbf{q}_{hu} & \mathbf{q}_{uu} \end{pmatrix}$$

where \mathbf{I} is the identity matrix and \mathbf{M} the submatrix of $\tilde{\mathbf{M}}$ that contains only the transient states:

$$\mathbf{M} = \begin{pmatrix} \mathbf{p}_{hh} & \mathbf{p}_{uh} \\ \mathbf{p}_{hu} & \mathbf{p}_{uu} \end{pmatrix}$$

Each entry of \mathbf{F} indicates the expected duration of being in a particular age-health combination before dying. Summing the entries of the first column of \mathbf{q}_{hh} , for example, yields the healthy life expectancy for healthy subjects aged 50. Summing the first column of \mathbf{q}_{hu} yields the unhealthy life expectancy for healthy subjects aged 50, and adding the two numbers gives the total life expectancy for healthy subjects aged 50. We repeat the same calculations for initially unhealthy subjects (using matrix blocks \mathbf{q}_{uh} and \mathbf{q}_{uu}) and use estimates of healthy and unhealthy population proportions in order to obtain overall (weighted) estimates of life expectancies (healthy, unhealthy, total). Initial proportions were derived from the first

available observation for 50-54 year olds using HRS survey waves between 1992 and 2012.

While the multinomial logit model is estimated only once, the above calculations are done separately for each population subgroup studied in this paper. To produce transition probabilities for the total population, we set all covariates to their grand mean and then compute the predicted probabilities from the estimated multinomial logit model equation. Transition probabilities for LRB (Low Risk Behavior) 1 (non-obese never-smokers) were produced by assigning a value of 0 to the underweight, obese, current smoker, and former smoker variables and setting the normal/overweight and never-smoker categories to 1. The remaining covariates are set to their grand mean. LRB 2 (non-obese, never-smoker, moderate drinker) transition probabilities were produced by following the same rules as for LRB 1, except that the never drinker and the heavy drinker categories are set to 0 and the moderate drinker category to 1.

Numbers for the mean age of the first transition to the disabled state are predictions of the Markov model. Markov transition probabilities is the only information used. Note that this method of calculation is different from incidence rate-type calculations. The Markov model-based numbers are obtained as follows: First, probabilities of healthy 50-year olds of becoming disabled for the first time after N years are calculated: the probability of becoming disabled after one year,

the probability of staying healthy for one year and then becoming disabled, the probability of staying healthy for two years and then becoming disabled, and analogously for all ages in the age range of the model. By doing so, we are ignoring transitions from non-disabled to dead states. Since the latter transitions have positive probabilities, our calculated transition probabilities from non-disabled to disabled states sum to a number smaller than one. In other words, the Markov model predicts that a fraction of the population never makes a transition from disability-free to disabled; this fraction of the population never leaves the disability-free state before dying. Therefore, we rescale the probabilities appropriately such that they do sum to one, and then use them as weights for the ages in the age range of the model, which results in the mean-age estimates shown in the paper.

Supplementary Appendix 3. Coefficients estimated from multinomial regression models

Characteristic	Men				Women			
	Initial State: Non-Disabled		Initial State: Disabled		Initial State: Non-Disabled		Initial State: Disabled	
	to Disabled	to Dead	to Non-Disabled	to Dead	to Disabled	to Dead	to Non-Disabled	to Dead
Age in years	0.96**	1.09***	0.99	1.08**	0.95***	1.02	1.00	1.07**
Age squared	1.0023***	1.0004	0.9992*	0.9996	1.0029***	1.0015***	0.9991**	0.9999
Educational attainment (Ref: Less than high school)								
High school / some college	0.58***	0.69*	0.83	0.62*	0.72**	0.36***	1.20	0.37***
College graduate+	0.41***	0.52***	1.03	0.82	0.74	0.43***	0.77	0.35**
Race/ethnicity (Ref: Non-Hispanic White)								
Non-Hispanic black	1.13*	1.15*	0.90	1.01	1.37***	1.30***	0.84***	0.89
Hispanic	1.13	0.75**	0.90	0.92	1.40***	0.67***	0.81***	0.52***
Non-Hispanic other race	1.01	0.99	0.88	0.89	1.25*	0.81	1.05	0.65*
BMI (Ref: BMI 18.5-29.9)								
Underweight (BMI<18.5)	2.39	19.34***	2.00	71.62***	1.63	1.60	1.42	17.42***
Obese (BMI≥30)	1.25	0.98	0.70*	0.67	2.26***	1.42	0.59***	0.73
Cigarette smoking (Ref: Never Smoker)								
Current	0.88	2.57***	1.12	1.49	1.54***	2.51***	1.27	1.28
Former	1.58**	3.25***	0.60*	1.17	1.35*	3.20***	0.94	1.06
Alcohol consumption (Ref: Moderate Drinker)								
Non-Drinker	1.69***	2.26***	0.44***	1.38	2.22***	0.93	0.56**	0.98
Heavy Drinker	0.93	1.98*	0.54*	1.28	1.52	1.23	1.26	1.15

* p<.05; ** p<.01; *** p<.001

Source: U.S. Health and Retirement Survey, 1998-2012

Supplementary Appendix 4. Sample distribution at baseline wave in 1998, ages 50-74

Characteristic	All (N=14,804)	Men (N=6,657)	Women (N=8,147)
Health state, %			
Non-Disabled	88.5 [87.7-89.2]	89.6 [88.6-90.5]	87.5 [86.5-88.5]
Disabled	11.5 [10.8-12.3]	10.4 [9.5-11.4]	12.5 [11.5-13.5]
BMI Category, %			
Non-Obese	72.8 [71.7-73.8]	74.3 [72.9-75.6]	71.5 [70.0-72.9]
Obese (BMI \geq 30)	25.9 [24.9-26.9]	25.2 [23.9-26.5]	26.5 [25.1-27.9]
Cigarette smoking, %			
Never	38.0 [36.6-39.4]	27.8 [26.2-29.4]	47.0 [45.1-49.0]
Former	41.4 [40.2-42.7]	50.7 [49.0-52.4]	33.2 [31.7-34.8]
Current	20.6 [19.4-21.9]	21.5 [20.2-23.0]	19.7 [18.3-21.3]
Alcohol Consumption, %			
Non/Irregular Drinker	65.8 [64.3-67.4]	56.6 [54.9-58.3]	74.0 [72.0 75.9]
Moderate Drinker	27.2 [25.7-28.7]	32.2 [30.7-33.9]	18.7 [17.1 20.3]
Heavy Drinker	7.0 [6.5-7.4]	11.2 [10.3-12.0]	7.3 [6.4 8.4]
Low Risk Behavioral Profiles (LRB), %			
LRB 1	27.1 [26.1-28.2]	20.0 [18.8-21.3]	33.3 [31.9-34.8]
LRB 2	7.0 [6.3-7.8]	6.7 [5.9-7.6]	7.3 [6.4-8.2]
High-Risk Behavioral Profiles			
Obese + Ever Smokers	15.4 [14.5-16.2]	17.5 [16.4-18.7]	13.4 [12.4-14.5]
Obese + Ever Smokers + Non-Moderate Drinkers	11.8 [11.1-12.6]	12.9 [11.9-13.9]	10.9 [9.9-11.9]

Note: Non/Irregular drinkers are individuals who either report not drinking alcoholic beverages or report drinking alcohol, but report doing so zero or less than one day per week. Moderate drinkers are those that drink <14 drinks/week [men] and <7 drinks/week [women]. Heavy drinkers are those that drink \geq 14 drinks/week [men] and \geq 7 drinks/week [women]. LRB 1 is non-obese never smokers. LRB 2 is non-obese, never smokers, and moderate alcohol consumers. Estimates reflect HRS provided sample weights. 95% confidence intervals shown in parentheses.

Data Source: Author analysis of U.S. Health and Retirement Survey, 1998

Supplementary Appendix 5. Percentage distribution of educational attainment by behavioral profile status at baseline in 1998.

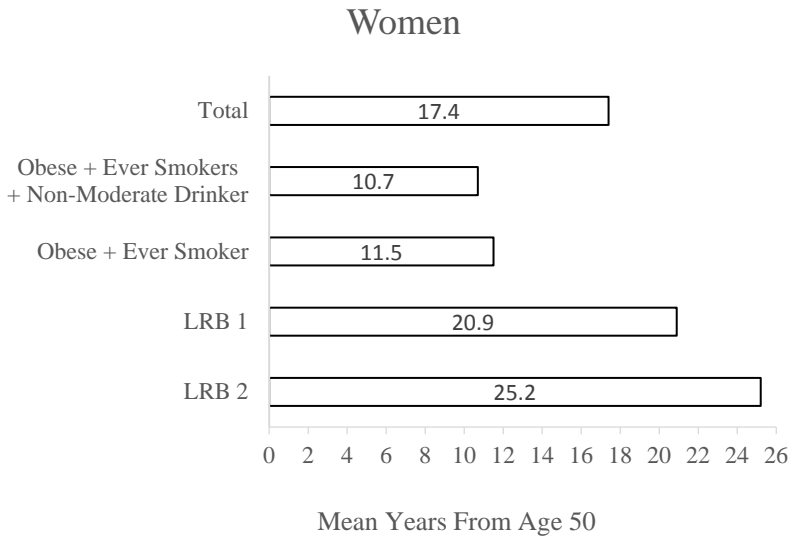
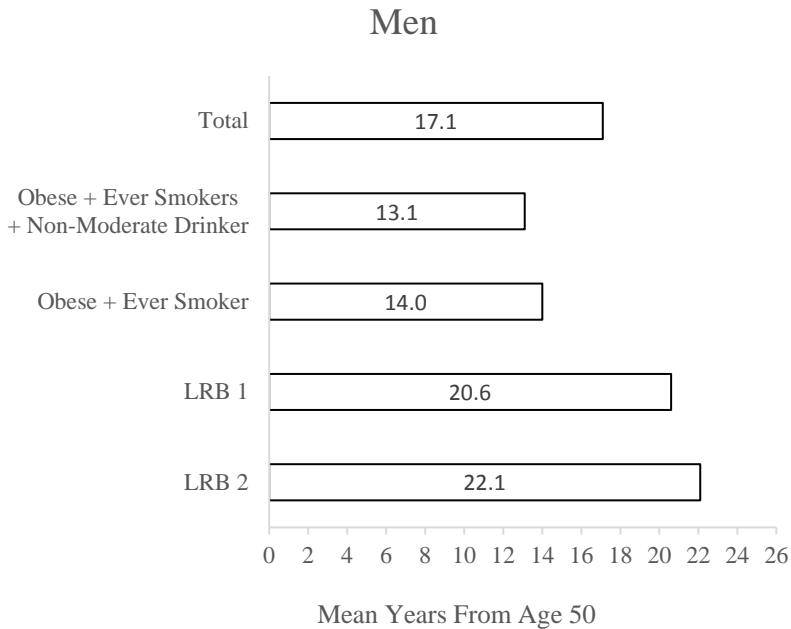
Subpopulation	Distribution of Educational Attainment for each subpopulation (row sums to 100%)		
	Less than High School	High School Degree/GED or Some College	College Graduate
Men (N=6,657)			
Total	20.7	53.1	26.2
LRB 1	15.5	43.3	41.2
LRB 2	7.2	37.3	55.5
Women (N=8,147)			
Total	21.9	61.2	16.9
LRB 1	18.7	61.7	19.6
LRB 2	8.9	59.8	31.3

Source: U.S. Health and Retirement Survey, 1998

LRB (Low Risk Behavior) 1: Non-obese, never smokers

LRB 2: Non-obese, never smokers, moderate alcohol consumers

Supplementary Appendix 6. Mean age of first disability incidence since age 50 by behavioral profile.



Notes: Total is entire analytical sample from HRS. Obesity is defined as BMI \geq 30.0 kg/m². Ever smoker includes former and current smokers. Moderate drinking are individuals who drink and consume <14 drinks/week (men) and <7 drinks/week (women). Heavy drinking is \geq 14 drinks/week (men) and \geq 7 drinks/week (women).

Data Source: Author analysis of U.S. Health and Retirement Study (1998-2012)

Supplementary Appendix 7. Health Expectancies at age 50.

Subpopulation	Men						Women					
	TLE	95% CI	DFLE	95% CI	DLE	95% CI	TLE	95% CI	DFLE	95% CI	DLE	95% CI
Total Population	27.7	[26.9-29.0]	23.7	[23.0-24.6]	4.0	[3.6-4.6]	31.4	[30.7-32.5]	25.7	[25.1-26.4]	5.8	[5.3-6.5]
Individual Behavioral Factors												
Non-Obese	27.8	[26.9-29.0]	24.3	[23.4-25.1]	3.6	[3.2-4.2]	32.0	[31.2-33.1]	27.0	[26.4-27.7]	5.0	[4.5-5.7]
Obese	27.8	[26.6-29.5]	22.0	[21.1-23.2]	5.8	[5.1-7.0]	30.5	[29.4-32.1]	22.2	[21.3-23.2]	8.3	[7.5-9.5]
Never Smokers	32.2	[31.2-34.2]	27.6	[26.7-28.8]	4.6	[3.9-5.9]	34.5	[33.5-35.9]	28.3	[27.5-29.0]	6.2	[5.6-7.5]
Ever Smokers	25.7	[24.8-26.9]	22.0	[21.2-22.9]	3.8	[3.3-4.3]	28.3	[27.5-29.4]	23.0	[22.4-23.8]	5.3	[4.9-6.0]
Moderate Drinker	30.8	[29.8-32.5]	26.5	[25.7-27.4]	4.3	[3.6-5.4]	35.5	[34.0-37.9]	29.7	[28.8-30.9]	5.7	[4.7-7.8]
Heavy Drinker	28.1	[26.4-30.7]	23.9	[22.6-25.4]	4.2	[3.2-5.8]	33.4	[30.9-36.9]	27.4	[25.9-29.2]	6.0	[4.4-8.9]
Non-Drinker	25.5	[24.6-26.7]	21.6	[20.8-22.6]	3.9	[3.4-4.5]	30.2	[29.4-31.2]	24.3	[23.7-25.0]	5.9	[5.4-6.6]
Low-Risk Behavioral Profile												
LRB 1	32.3	[31.2-34.2]	28.2	[27.3-29.2]	4.1	[3.4-5.5]	34.9	[33.9-36.4]	29.5	[28.8-30.3]	5.4	[4.8-6.6]
LRB 2	34.9	[33.5-37.0]	30.3	[29.3-31.5]	4.6	[3.5-6.5]	38.6	[36.8-41.5]	32.9	[32.0-34.2]	5.6	[4.4-8.2]
High-Risk Behavioral Profile												
Obese + Ever Smoker + Non-Moderate Drinker	23.8	[22.6-25.5]	18.5	[17.5-19.7]	5.3	[4.6-6.3]	26.3	[25.1-27.8]	18.4	[17.4-19.6]	8.0	[7.2-9.0]
Obese + Ever Smoker	25.8	[24.6-27.6]	20.3	[19.4-21.4]	5.5	[4.8-6.6]	27.4	[26.2-28.9]	19.6	[18.7-20.8]	7.8	[7.0-8.8]

Note: Expectancies in years since age 50. 95% Confidence Intervals shown in parenthesis.

Source: U.S. Health and Retirement Survey, 1998-2012

TLE: Total Life Expectancy

DFLE: Disability-free Life Expectancy

DLE : Disabled Life Expectancy

LRB (Low Risk Behavior) 1: Non-obese, never smokers

LRB 2: Non-obese, never smokers, moderate alcohol consumers