SUPPLEMENTARY MATERIAL

for

Predicting Cognitive Impairment and Dementia: A Machine Learning Approach

Damaris Aschwanden[1], Stephen Aichele[2,3], Paolo Ghisletta[2,4,5], Antonio Terracciano[1], Matthias Kliegel[2,5], Angelina R. Sutin[1], Justin Brown[1], & Mathias Allemand[6,7]

[1] Florida State University, Tallahassee, USA
[2] Faculty of Psychology and Educational Sciences, University of Geneva, Switzerland
[3] Colorado State University, Fort Collins, USA
[4] Swiss Distance University Institute, Switzerland
[5] Swiss National Centre of Competence in Research LIVES – Overcoming vulnerability: Life course perspectives, Universities of Lausanne and of Geneva, Switzerland
[6] University of Zurich, Zurich, Switzerland
[7] University Research Priority Program Dynamics of Healthy Aging, University of Zurich, Switzerland

Correspondence should be addressed to Damaris Aschwanden, Department of Geriatrics, College of Medicine, Florida State University, 1115 West Call Street, Tallahassee, FL 32306, USA. E-mail address: damaris.aschwanden@med.fsu.edu.

S1: Background information of methods

The first section of the supplementary material (S1) contains information about the methods. Figure S1 (see below) displays the percentage of cognitive impairment and dementia per age group across follow-ups.

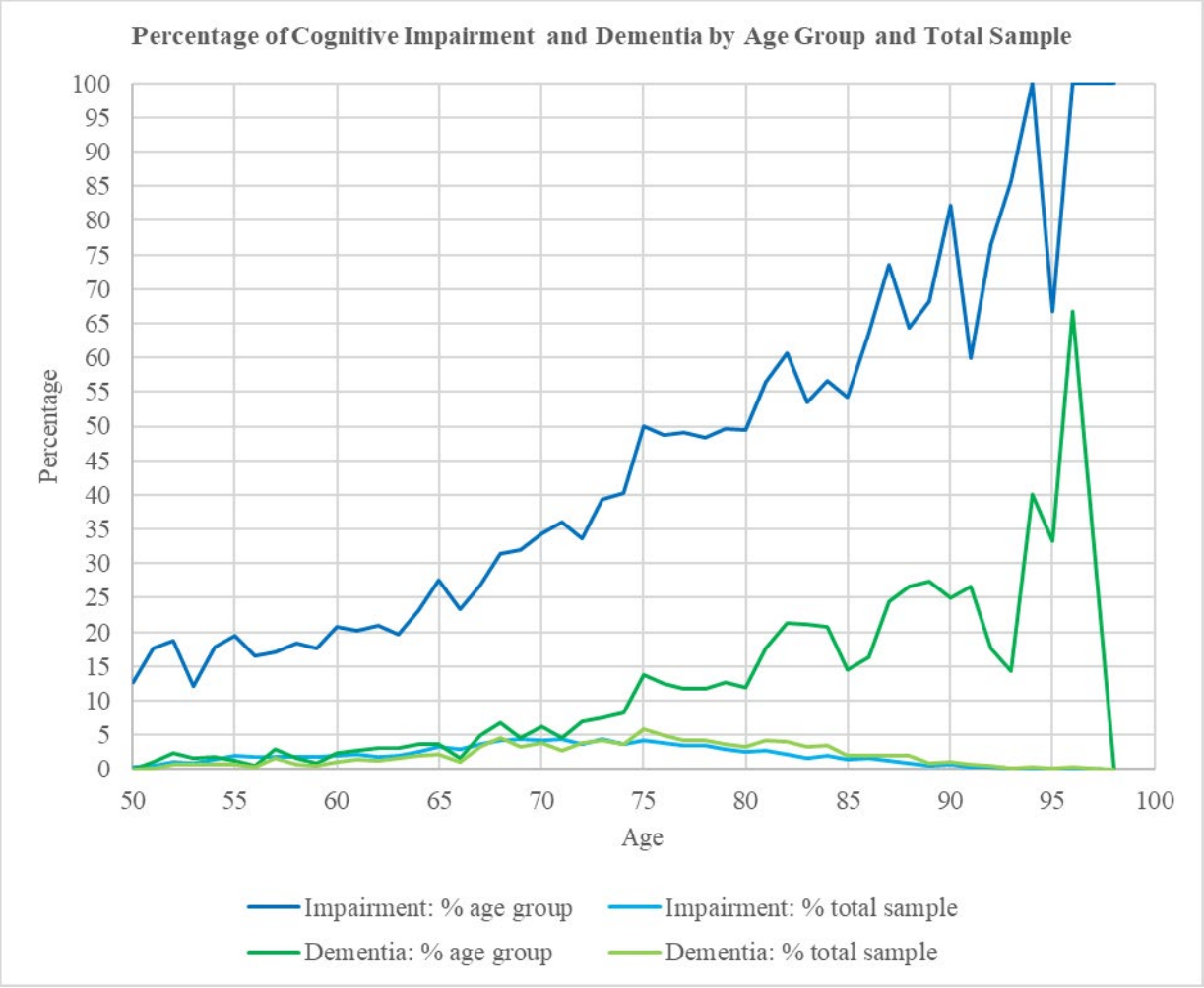

Figure S1. Across the follow-up interval, n = 3,119 participants (31.3%) developed cognitive impairment and n = 622 participants (6.2%) developed dementia. This figure illustrates the percentage of cognitive impairment and dementia per age group and per total sample. Percentage (in %) is shown on the y-axis. Age (at baseline) is shown on the x-axis and ranged from 50-98 years at baseline. Regarding participants' own age group, the percentage of onset of cognitive impairment or dementia increased over time. However, concerning the total sample, older participants showed lower percentage of onset of cognitive impairment or dementia. This can be interpreted as a selection effect (survival of the fittest): Older participants at baseline have already demonstrated their cognitive fitness (before study entry) and are thus at lower risk for cognitive impairment compared to younger participants at baseline. Previous research suggests that despite an age-related decline in cognitive ability, dementia-free survival is thought to increase with advanced age [51]. The onset of cognitive impairment is estimated to have a 3-year delay for each decade of life from 90 years onwards, followed by a rapid cognitive decline toward the end of life [52].

S1.1 Description of predictors

We aimed to include a comprehensive list of risk factors identified in previous literature and available in the present data. In total, the set of predictors included 52 variables: demographic (10), psychosocial (9), health (26), and biomarkers/polygenic scores (7).

S1.1.1 Demographic variables. Demographic predictors were as follows: (a) age; (b) gender *(1 = male, 0 = female)*; (c) education *(in years)*; (d) race *(African American = 1 and other non-White = 1 compared to White = 0)*; (e) ethnicity *(non-Hispanic = 0, Hispanic = 1)*; (f) income (total income in the year prior to interview wave); (g) wealth (i.e., assets minus debts); (h) marital status *(1 = married, 2 = separated/divorced; 3 = widowed, 4 = never married)*; (i) work *(1 = yes, 0 = no)*; and (j) type of home *(1 = assisted such as senior housing or retirement center, 0 = not assisted)*.

S1.1.2 Psychosocial variables. The HRS assessed an extensive battery of personality and self-variables. The present study analyzed nine psychosocial variables (conscientiousness, openness, extraversion, agreeableness, emotional distress, life satisfaction, positive affect, purpose in life, and social contact). Emotional distress was a composite variable obtained by averaging scores from nine standardized negative emotion variables (i.e., neuroticism, hostility, anxiety, negative affect, hopelessness, pessimism, depression, loneliness and perceived constraints; all inter-correlated at least $r > .3$).

S1.1.3 Personality. The Big Five personality traits were assessed using the Midlife Development Inventory (MIDI; [53]). Participants were asked how well 26 adjectives described them on a Likert-type scale ranging from 1 (*not at all*) to 4 (*a lot*). The MIDI includes five items for conscientiousness (e.g., organized), four for neuroticism (e.g., moody), seven for openness (e.g., imaginative), five for extraversion (e.g., outgoing), and five for agreeableness (e.g., helpful).

S1.1.4 Self-variables. Hostility was measured using the Cook Medley Hostility Inventory [54]. Five items (e.g., "I think most people would lie in order to get ahead") were rated on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*).

Anxiety was measured with five items from the Beck Anxiety Inventory [55]. Participants reported on items (e.g., "I had fear of the worst happening"), rating how often they felt that way during the past week on a scale from 1 (*never*) to 4 (*most of the time*).

Perceived constraints were measured with five items from the Perceived Constraints on Personal Control Scale [56]. Items such as "What happens in my life is often beyond my control" were rated on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*).

Hopelessness was assessed using the Hopelessness Scale [57]. Participants reported on four items (e.g., "I feel it is impossible for me to reach the goals that I would like to strive for") on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*).

Loneliness was measured using three items [58]. Participants were asked whether they feel left out, feel isolated, or lack companionship. Items were rated on a scale from 1 (*often*) to 3 (*hardly ever or never*), reverse coded and averaged.

Depression was measured with a 9-item version of the Center for Epidemiological Studies Depression Scale [59]. Participants reported whether or not they experienced depressive symptoms (e.g., feeling depressed) during last week. Items were reverse coded (if needed) and averaged, with higher scores indicating the presence of more symptoms.

Positive and negative affect were assessed by the Positive and Negative Affect Schedule [60]. Participants rated 6 (in 2006) and 12 (in 2008) items (e.g., enthusiastic, ashamed) on a scale ranging from 1 (*very much*) to 5 (*not at all*). Scores were thus standardized ($M = 0$, $SD = 1$)

within wave before combining them as baseline in order to account for the difference in number of items.

Optimism and *pessimism* were measured with each three items from the revised version of the Life Orientation Test [61]. Items (e.g., "In uncertain times, I usually expect the best" for optimism; "If something can go wrong for me it will" for pessimism) were rated on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*).

Life satisfaction was assessed using five items from the Satisfaction With Life Scale [62]. Items (e.g., "In most ways my life is close to ideal") were rated on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*) in 2006 and from 1 (*strongly disagree*) to 7 (*strongly agree*) in 2008. Scores were thus standardized ($M = 0$, $SD = 1$) within wave before combining them as baseline in order to account for this difference in response scale.

Purpose in life was drawn from the Ryff Measures of Psychological Well-being [63]. Seven items (e.g., "I have a sense of direction and purpose in my life") were rated on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*).

Furthermore, we created a *social contact* index consisting of information about the extent to which respondents are in contact with their social networks (children, family, and friends). Nine items were used to assess how often participants (a) meet up (include both arranged and chance meetings), (b) speak on the phone, and (c) write or email. The items were rated on a scale from 1 (*three or more times a week*) to 6 (*less than once a year or never*), reverse coded and averaged.

S1.1.5 Health. We analyzed 26 health predictors. *Subjective health*, *childhood health*, and *hearing* were each assessed with a single measure scaled from 1 (*excellent*) to 5 (*poor*). We reverse coded these items, so that higher scores indicated better health and hearing. Participants also indicated (yes/no) as to whether they wear a *hearing aid*, take *sleep medications*, and whether they have ever experienced *childhood traumas* or *lifetime traumas* (e.g., being in a major fire, flood, earthquake, or other natural disaster; [64]).

We included three variables of BMI: *BMI at baseline*, *highest BMI ever* and *change of BMI* (i.e., BMI slope). Given that there was no measure for midlife obesity in the data, highest BMI ever and BMI slope seemed the best possible alternatives. Weight and height were self-reported. Weight (in pounds) was divided by the square of height (in inches) x 702 (lb/in^2×702) to calculate BMI. For highest BMI ever, the self-reported highest weight ever and baseline height were used for calculation. BMI slope was calculated using information from the earliest available measurement occasion in HRS - e.g., 1992 - until 2 years before year of detection of cognitive impairment/dementia or the last year of cognitive testing for unimpaired individuals. Furthermore, *waist circumference* (in inches) was also included.

Further, participants indicated (yes/no) as to whether they have a physician diagnosis of *hypertension*, *diabetes*, *stroke*, *cancer,* and *heart disease* (e.g., heart attack, coronary heart disease, angina). Likewise, they reported (yes/no) whether they currently drink *alcohol* and engage in *physical activity*. We included four physical activity variables, one for *mild activity* ("sports or activities that are mildly energetic, such as vacuuming, laundry, home repairs"), one for *moderate activity* ("sports or activities that are moderately energetic, such as gardening, cleaning the car, walking at a moderate pace, dancing, floor or stretching exercises"), and one for *vigorous activity* ("sports or activities that are vigorous, such as running or jogging, swimming, cycling, aerobics or gym workouts, tennis, or digging with a spade or shovel"). For each of these, participants reported the degree to which they engaged in that type of activity and we coded them as 0 (*never*) and 1 (*1-3 times per month to everyday*). We also created a *total physical activity* variable (consisting of mild, moderate and vigorous activity) to indicate whether participants

were physically inactive or active. For *smoking*, we included a binary variable whether participants have ever smoked from 1992 to 2008.

Functional limitations were measured with 10 yes-no items (e.g., ″Because of a health problem, do you have any difficulty with lifting or carrying weights over 10 pounds, like a heavy bag of groceries?"). Items were averaged, with higher scores indicating more functional limitations.

Hand grip strength was measured in kilograms using a Smedley spring type hand dynamometer. The measure was conducted with the respondent standing and holding the dynamometer at a 90-degree angle. Two measurements were taken on each hand alternating between the left and right hand. The best score of the two attempts for the dominant hand was used in the analyses.

S1.1.6 Biomarkers/polygenic scores. We analyzed seven cardiovascular, inflammation, and metabolic biomarkers: total cholesterol, high-density lipoprotein, Cystatin C, C-reactive protein, and hemoglobin A1c. The assessment of these biomarkers from a dried blood spot has been described in detail by Crimmins et al. [65]. Polygenic scores for Alzheimer's Disease were obtained from a subsample (2013) and refer to APOE risk status without and with ε4 risk variant.

S1.1.7 Preliminary data conditioning of predictors. We removed outliers to avoid analyzing data that were nonsensical or potentially erroneous: BMI variables were truncated by +/- 3 *SD*, wealth was truncated if the mean was exceeded by 3 *SD*, and income was truncated at 300,000 USD. Furthermore, wealth, emotional distress, and hemoglobin A1C were positively skewed and thus log-transformed. All continuous variables were z-transformed.

S1.2 Methods: Statistical analyses

The goal of this study was to examine the relative and combined importance of 52 predictors of risk for cognitive impairment and dementia in survival analyses. Using biyearly outcomes in survival analysis may seem rather rough at a first glance, but onset of cognitive impairment and dementia is slow [66] and does not occur instantaneously, making it difficult to specify the exact time point of occurrence. Survival analysis still offers a powerful tool to investigate important research questions that revolve around the occurrence of events that are less well-defined than mortality per se (i.e., date of death) (see also [67–69]). Age-at-onset was chosen as time scale in the survival analyses [23].

S1.2.1 Random forest survival analysis (RFSA). RFSA is a nonparametric statistical technique that is related to classification and regression trees [14–16]. A random survival forest is computed by binary decision trees which can be used for selecting the most important variables that are linked with survival time. The random survival forest is formed by an ensemble of independent decision trees. To grow these decision trees, bootstrapping and random node splitting are applied [69]. The trees consist of nodes (i.e., predictor variables and corresponding split values) and branches (i.e., pathways linking nodes). Nodes that are close to the root (i.e., start point) of the tree are considered as stronger predictors than nodes that are farther away from the root. The results are pooled across multiple trees, where each tree is derived from randomly sampled subsets of observations and predictors (i.e., built-in cross-validation [15]). The trees are not estimated on the entire sample of individuals, across all variables. Rather, trees are different from each other in terms of the randomly chosen individuals and variables that are included. Thus, while the results of any single tree may be quite specific to that subsample of observations and predictors, the results aggregated across all trees reflect the most robust associations between predictors and outcome, thereby minimizing the chances of overfitting. In essence, RFSA

attributes participants to subgroups according to the associations between the outcome (here, cognitive impairment or dementia) and all the predictors. The analysis maximizes the differences among the subgroups while, at the same time, minimizing the differences among the individuals within any given group.

By aggregating estimates of predictor-outcome strength across all trees, the importance of each predictor (i.e., relative influence) can be derived, resulting in a variable importance (VIMP) ranking. VIMP ranking has been shown to be more stable than stepwise variable selection for logistic regression, which is known to be affected by order effects (cf. [16,70]). Furthermore, predictors that influence a given outcome primarily through their higher-order interactions with other variables may show a high VIMP ranking (because RFSA implicitly accounts for such complex effects) but appear non-significant in standard regression approaches (where explicit modeling of such effects may not be possible due to the number of predictors or sample size) [18].

S1.2.2 Cox PH survival analysis. The Cox PH model [17] is the most popularly used survival regression approach, and it estimates relationships between predictors and survival time based on a "hazard function". It assumes that the predictors have a multiplicative effect on the hazard and that this effect is constant over time [17]. Under this assumption, exponentiated regression coefficients for each predictor can be interpreted in terms of the hazard ratio (HR) associated with a 1-unit change in the given predictor.

S2-S7: Summary of sensitivity analyses

The goal of the sensitivity analyses was to check the robustness of the variable importance (VIMP) rankings derived by the RFSA. We ran a series of five sensitivity analyses with the first subsample that was used for the main RFSA reported in the manuscript. We then ran a sixth sensitivity analysis by using the second subsample - that was used for the Cox PH survival analysis reported in the manuscript-, for the RFSA.

First, we tested whether the ranks were robust if we excluded the top predictor (i.e., African American for cognitive impairment and BMI slope for dementia) given the substantial magnitude on the risk of cognitive impairment and dementia. Second, we only included participants who were aged >65 years because the risk to develop cognitive impairment and dementia increases with age, usually occurring in people aged over 65 years [71]. The sample sizes were n = 3,101 for cognitive impairment and n = 3,151 for dementia. Third, we only included participants with more than two years of time-to-detection. Since the temporal processes of the disease's onset are unclear, it might be that two years of time-to-detection are too short to investigate predictor-outcome associations. If ranks show to be unstable, the reverse causality hypothesis could be a possible explanation. The sample sizes were n = 4,220 for cognitive impairment and n = 4,658 for dementia. Fourth, missing data were imputed with 20 iterations (vs. 5 iterations in main RFSA). Fifth, we ran the analyses with a different R package; that is "party" using the function "cForest" [72]. Lastly, we ran the analyses with the second subsample to compare the VIMP ranking between the subsamples.

We compared the VIMP ranking between main results and each sensitivity analysis, considering the 15 strongest predictors. If a predictor showed up in the VIMP ranking of both the main and sensitivity analyses, it was counted as an overlap.

The results for cognitive impairment are summarized in Table S2 (sensitivity analyses 1-4) and Table S3 (sensitivity analysis 5). The findings for dementia are displayed in Table S4 (sensitivity analyses 1-4) and Table S5 (sensitivity analysis 5). Table S6 presents the results for sensitivity analysis 6 for cognitive impairment and dementia. In addition, Pearson correlations were performed to examine the associations between the relative importance (I_{rel}) of main and sensitivity analyses; that is, to evaluate the overlaps more quantitatively (Table S7).

Across all sensitivity analyses, the overlaps ranged from 7/15 to 14/15 variables. The correlation coefficients ranged from r =.67 to r = 1.00 between the relative importance of main and sensitivity analyses within cognitive status using the R package "randomForestSRC"; indicating strong associations. It should be noted that the associations were less consistent using the R package "cForest", leading to negative correlation coefficients. However, caution should be applied when interpreting the variable importance for survival forests of "cForest" given its feature is experimental. Based on these sensitivity analyses, we conclude that the VIMP ranking for cognitive impairment and dementia seems consistent.

Table S2

Results of sensitivity analyses (SA) 1-4 for cognitive impairment.

	Main results		SA1: Without top predictor		SA2: >65y		SA3: Survival time >2y		SA4: Iterations = 20	
Rank	Predictor	I_{rel}	Predictor	I_{rel}	Predictor	I_{rel}	Predictor	I_{rel}	Predictor	I_{rel}
1	African American	1.00	Wealth	1.00	Education	1.00	African American	1.00	African American	1.00
2	Wealth	0.59	Education	0.94	African American	0.78	Education	0.47	Education	0.61
3	Education	0.57	BMI Slope	0.50	Age at Baseline	0.75	Wealth	0.38	Wealth	0.54
4	BMI Slope	0.33	Emotional Distress	0.43	Smoking	0.37	Subjective Health	0.32	BMI Slope	0.32
5	Subjective Health	0.27	Subjective Health	0.40	Wealth	0.33	Emotional Distress	0.20	Subjective Health	0.27
6	Emotional Distress	0.27	Smoking	0.26	Subjective Health	0.28	Hispanic	0.19	Emotional Distress	0.26
7	Hispanic	0.22	Grip Strength	0.24	Marital Status	0.21	BMI Slope	0.15	Hispanic	0.22
8	Grip Strength	0.17	Childhood Traumas	0.22	Hispanic	0.21	Grip Strength	0.12	Grip Strength	0.17
9	Childhood Traumas	0.16	Hispanic	0.21	Emotional Distress	0.18	Smoking	0.12	Childhood Traumas	0.17
10	Smoking	0.15	Marital Status	0.20	Income	0.15	Cystatin C	0.11	Marital Status	0.13
11	Marital Status	0.13	Cystatin C	0.19	Grip Strength	0.14	Childhood Traumas	0.11	Smoking	0.12
12	Social Contact	0.11	Income	0.18	BMI Slope	0.13	Cholesterol	0.10	Cystatin C	0.11
13	Cystatin C	0.10	Polygenic without APOE ε4	0.17	Cholesterol	0.11	Funct. limitations	0.07	Income	0.09
14	Income	0.08	Cholesterol	0.16	Cystatin C	0.11	Work	0.07	Work	0.08
15	Cholesterol	0.07	Social Contact	0.15	Work	0.11	Social Contact	0.06	Social Contact	0.08

Note. Funct. limitations = Functional limitations. Sensitivity analyses were conducted in subsample 1 (n = 4,990) with 1,000 trees per random forest and 5 iterations (except Sensitivity analysis 4). Sensitivity analysis 1 (SA1) excluded the top predictor (i.e., African American); Sensitivity analysis 2 (SA2) excluded participants aged 50-64 years; Sensitivity analysis 3 (SA3) excluded participants with time-to-detection of 2 years; and Sensitivity analysis 4 (SA4) increased the number of iterations to 20. The top 15 predictors of each sensitivity analysis were compared with the main results. For SA1, 14/14 predictors overlap. For SA2, 13/15 predictors overlap. For SA3, 13/15 predictors overlap. For SA4, 14/15 predictors overlap.

Table S3

Sensitivity analysis 5 for cognitive impairment.

	Rank	
Predictor	cForest	randomForestSRC
Age	1	52
Education	2	3
Smoking	3	10
Work	4	16
Emotional Distress	5	6
African American	6	1
Marital Status	7	11
Wealth	8	2
Subjective Health	9	5
Grip Strength	10	8
Childhood Traumas	11	9
Alcohol	12	26
BMI	13	48
Cancer	14	25
Cystatin C	15	13

Note. Sensitivity analysis 5 (SA5) was conducted in subsample 1 (*n* = 4,990) with 1,000 trees per random forest and 5 iterations. The ranks of the strongest 15 predictors were compared using “cForest” vs. "randomForestSRC". For SA5, 10/15 predictors overlap.

Table S4

Results of sensitivity analyses (SA) 1-4 for dementia.

	Main results		SA1: Without top predictor		SA2: >65y		SA3: Survival time >2y		SA4: Iterations = 20	
Rank	Predictor	I_{rel}	Predictor	I_{rel}	Predictor	I_{rel}	Predictor	I_{rel}	Predictor	I_{rel}
1	BMI Slope	0.42	Diabetes	0.39	Diabetes	0.90	BMI Slope	0.37	BMI Slope	0.47
2	Emotional Distress	0.40	Emotional Distress	0.33	African American	0.83	Subjective Health	0.22	Emotional Distress	0.45
3	Diabetes	0.34	African American	0.24	Education	0.62	Funct. Limitations	0.16	Diabetes	0.32
4	African American	0.33	Hispanic	0.23	BMI Slope	0.50	Hemoglobin A1C	0.15	Hemoglobin A1C	0.29
5	Childhood Traumas	0.23	Hemoglobin A1C	0.20	Life Satisfaction	0.48	African American	0.13	African American	0.27
6	Hemoglobin A1C	0.20	Life Satisfaction	0.19	Hemoglobin A1C	0.39	Emotional Distress	0.12	Childhood Traumas	0.23
7	Education	0.19	Childhood Traumas	0.18	Emotional Distress	0.39	Diabetes	0.12	Subjective Health	0.21
8	Life Satisfaction	0.17	Subjective Health	0.15	Stroke	0.34	Life Satisfaction	0.10	Hispanic	0.19
9	Hispanic	0.15	Optimism	0.15	Childhood Traumas	0.30	Hispanic	0.10	Education	0.18
10	Childhood Health	0.15	Education	0.14	Funct. Limitations	0.27	Purpose in Life	0.10	Childhood Health	0.18
11	Funct. Limitations	0.13	Purpose in Life	0.12	Purpose in Life	0.25	Childhood Traumas	0.09	Wealth	0.17
12	Subjective Health	0.10	Childhood Health	0.12	Smoking	0.24	Lifetime Traumas	0.09	Life Satisfaction	0.16
13	Stroke	0.09	Funct. Limitations	0.12	Hispanic	0.23	Positive Affect	0.07	Lifetime Traumas	0.15
14	Social Contact	0.08	Income	0.11	Grip Strength	0.17	Childhood Health	0.06	Funct. Limitations	0.12
15	Activity Total	0.07	Stroke	0.11	Subjective Health	0.17	Stroke	0.05	Activity Total	0.12

Note. Funct. limitations = Functional limitations. Sensitivity analyses were conducted in subsample 1 (n = 4,990) with 1,000 trees per random forest and 5 iterations (except Sensitivity analysis 4). Sensitivity analysis 1 (SA1) excluded the top predictor (i.e., BMI slope); Sensitivity analysis 2 (SA2) excluded participants aged 50-64 years; Sensitivity analysis 3 (SA3) excluded participants with time-to-detection of 2 years; and Sensitivity analysis 4 (SA4) increased the number of iterations to 20. The top 15 predictors of each sensitivity analysis were compared with the main results. For SA1, 12/14 predictors overlap. For SA2, 12/15 predictors overlap. For SA3, 12/15 predictors overlap. For SA4, 13/15 predictors overlap.

Table S5

Sensitivity analysis 5 for dementia.

	Rank	
Predictor	cForest	randomForestSRC
Age	1	52
Smoking	2	50
Marital Status	3	25
Emotional Distress	4	2
Subjective Health	5	12
Diabetes	6	3
Alcohol	7	20
Education	8	7
Work	9	24
Cystatin C	10	18
Life Satisfaction	11	8
Hypertension	12	27
Childhood Traumas	13	5
Hearing	14	17
BMI Slope	15	1

Note. Sensitivity analysis 5 (SA5) was conducted in subsample 1 (n = 4,990) with 1,000 trees per random forest and 5 iterations. The ranks of the strongest 15 predictors were compared using “cForest” vs. "randomForestSRC". For SA5, 7/15 predictors overlap.

Table S6

Results of sensitivity analysis 6: VIMP comparison for cognitive impairment (Part A) and dementia (Part B).

	A: Impairment				B: Dementia			
	Main results (subsample 1)		SA6 (subsample 2)		Main results (subsample 1)		SA6 (subsample 2)	
Rank	Predictor	I_{rel}	Predictor	I_{rel}	Predictor	I_{rel}	Predictor	I_{rel}
1	African American	1.00	African American	1.00	BMI Slope	0.42	African American	1.00
2	Wealth	0.59	Emotional Distress	0.78	Emotional Distress	0.40	BMI Slope	0.76
3	Education	0.57	Wealth	0.71	Diabetes	0.34	Emotional Distress	0.53
4	BMI Slope	0.33	Education	0.71	African American	0.33	Wealth	0.37
5	Subjective Health	0.27	BMI Slope	0.63	Childhood Traumas	0.23	Cholesterol	0.32
6	Emotional Distress	0.27	Subjective Health	0.47	Hemoglobin A1C	0.20	Childhood Traumas	0.30
7	Hispanic	0.22	Hispanic	0.22	Education	0.19	Subjective Health	0.26
8	Grip Strength	0.17	Grip Strength	0.21	Life Satisfaction	0.17	Income	0.26
9	Childhood Traumas	0.16	Smoking	0.19	Hispanic	0.15	Education	0.19
10	Smoking	0.15	Cystatin C	0.15	Childhood Health	0.15	Childhood Health	0.19
11	Marital Status	0.13	Income	0.12	Funct. Limitations	0.13	Social Contact	0.17
12	Social Contact	0.11	Funct. limitations	0.12	Subjective Health	0.10	Diabetes	0.17
13	Cystatin C	0.10	Childhood Traumas	0.08	Stroke	0.09	Purpose in Life	0.15
14	Income	0.08	Highest BMI	0.08	Social Contact	0.08	High Density Lipoprotein	0.12
15	Cholesterol	0.07	Marital Status	0.07	Activity Total	0.07	BMI	0.11

Note. Funct. limitations = Functional limitations. Main results were conducted in subsample 1 (n = 4,990), whereas sensitivity analysis 6 (SA6) was ran in subsample 2 (n = 4,989) with 1,000 trees per random forest and 5 iterations. The top 15 predictors were compared. For cognitive impairment, 13/15 predictors overlap. For dementia, 9/15 predictors overlap.

Table S7

Correlations between relative variable importance of main and sensitivity analyses.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Impairment															
1. Main															
2. SA1	.97**														
3. SA2	.58**	.50**													
4. SA3	.97**	.92**	.52**												
5. SA4	1.00**	.97**	.59**	.98**											
6. SA5	-.10	-.07	.66**	-.17	-.10										
7. SA6	.93**	.91**	.53**	.89**	.93**	-.06									
Dementia															
9. Main	.57**	.48**	-.15	.59**	.57**	-.66**	.61**	.13							
10. SA1	.46**	.39**	-.21	.52**	.47**	-.71**	.49**	-.01	.96**						
11. SA2	.62**	.48**	.14	.65**	.63**	-.43**	.61**	.29*	.85**	.84**					
12. SA3	.45**	.42**	-.32*	.49**	.46**	-.80**	.49**	-.04	.93**	.93**	.74**				
13. SA4	.57**	.54**	-.15	.57**	.56**	-.64**	.64**	.13	.97**	.95**	.82**	.93**			
14. SA5	-.18	-.10	.55**	-.25	-.19	.94**	-.11	.34*	-.56**	-.60**	-.33*	-.70**	-.56**		
15. SA6	.80**	.65**	.23	.79**	.80**	-.29*	.83**	.59**	.75**	.65**	.69**	.64**	.73**	-.28*	

Note. SA = Sensitivity analyses. Sensitivity analysis 1 (SA1) excluded the top predictor from the main analyses; Sensitivity analysis 2 (SA2) excluded participants aged 50-64 years; Sensitivity analysis 3 (SA3) excluded participants with survival time of 2 years; Sensitivity analysis 4 (SA4) increased the number of iterations to 20; Sensitivity analysis 5 (SA5) compared variable importance ranks using the R packages "cForest" vs. "randomForestSRC"; Sensitivity analysis 6 (SA6) compared the variable importance ranks obtained from subsample 1 vs. subsample 2.

The correlation coefficients ranged from r =.67 to r = 1.00 between the relative importance of main and sensitivity analyses using the R package "randomForestSRC"; indicating strong associations. It should be noted that the associations were less consistent using the R package "cForest" (SA5), leading to negative correlation coefficients. However, caution should be applied when interpreting the variable importance for survival forests of "cForest" given its feature is experimental.

*p < .05, ** p < .01.