Loneliness, Dementia Status, and Their Association with All-Cause Mortality Among Older US Adults

Michael F. Georgescu[a,1], May A. Beydoun[a,1], Jason Ashe[a], Christian A. Maino Vieytes[a],
Hind A. Beydoun[b,c], Michele K. Evans[a,3] and Alan B. Zonderman[a,3]
[a]*Laboratory of Epidemiology and Population Sciences, National Institute on Aging, Intramural Research Program, NIA/NIH/IRP, Baltimore, MD, USA*
[b]*U.S. Department of Veterans Affairs, VA National Center on Homelessness Among Veterans, Washington, DC, USA*
[c]*Department of Management, Policy, and Community Health, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA*

Abstract.

Background: Loneliness, dementia, and mortality are interconnected.

Objective: We aimed at understanding mediating pathways and interactions between loneliness and dementia in relation to mortality risk.

Methods: The study tested bi-directional relationships between dementia, loneliness, and mortality, by examining both interactions and mediating effects in a large sample of older US adults participating in the nationally representative Health and Retirement Study. Out of $\leq$ 6,468 older participants selected in 2010, with mean baseline age of 78.3 years and a follow-up time up to the end of 2020, 3,298 died at a rate of 64 per 1,000 person-years (P-Y). Cox proportional hazards and four-way decomposition models were used.

Results: Algorithmically defined dementia status (yes versus no) was consistently linked with a more than two-fold increase in mortality risk. Dementia status and Ln(odds of dementia) were strongly related with mortality risk across tertiles of loneliness score. Loneliness z-score was also linked to an elevated risk of all-cause mortality regardless of age, sex, or race or ethnicity, and its total effect (TE) on mortality was partially mediated by Ln(odds of dementia), z-scored, ($\leq$40% of the TE was a pure indirect effect). Conversely, a small proportion (<5%) of the TE of Ln(odds of dementia), z-scored, on mortality risk was explained by the loneliness z-score.

Conclusions: In sum, dementia was positively associated with all-cause mortality risk, in similar fashion across loneliness score tertiles, while loneliness was associated with mortality risk. TE of loneliness on mortality risk was partially mediated by dementia odds in reduced models.

Keywords: Aging, Alzheimer's disease, cohort studies, dementia, loneliness, mortality

---

[1]These authors contributed equally to this work.

[2]MFG and MAB had full access to the data used in this manuscript and completed all the statistical analyses.

[3]Co-senior authors.

*Correspondence to: May A. Beydoun, PhD, MPH, NIH Biomedical Research Center, National Institute on Aging, IRP, 251 Bayview Blvd, Suite 100, 04B118, Baltimore, MD 21224, USA. Fax: +1 410 558 8236; E-mail: baydounm@mail.nih.gov.

INTRODUCTION

Worldwide, around 50 million people have dementia and this is projected to increase to 152 million in 2050, with a greater increase in low-and middle-income countries, with Alzheimer's disease remaining the most common sub-type [1]. Additionally, 4.7% individuals aged 60 years and older are projected to be affected with dementia, with 4.6–7.7 million additional cases worldwide per year [2]. Dementia is characterized as a progressive loss of global functions, with an emphasis on memory loss and on deficits in executing everyday life activities [3]. The World Health Organization (WHO) recognized the importance of addressing dementia globally through a global action plan responding to dementia over the period of 2017–2025, highlighting key action areas in health and social systems [4]. Despite recent advancements in drug development [5] there are limited treatment options. Therefore, increasing research on dementia prevention and risk reduction remains critical, along with identifying modifiable lifestyle and risk factors linked to dementia risk.

In 2020, the Lancet commission identified potentially modifiable risk factors that accounted for 40% of dementia risk; these include education, hearing loss, traumatic brain injuries, hypertension, alcohol use, obesity, smoking, depression, social isolation, physical inactivity, air pollution, and diabetes [6]. The same study also highlighted social contact as a protective factor against dementia, with the potential to improve late-life cognitive functioning [6]. In 2020 the National Academies of Science, Engineering and Medicine report identified how social isolation and loneliness are risk factors for depression, anxiety, cognitive deficits, dementia, poor cardiovascular health, and early mortality [7]. The same report also summarized the pathways which link social isolation and loneliness with health outcomes ending with risk for mortality; along with mediators that are behavioral (e.g., lifestyle, medial adherence), psychological (e.g., stress), and biological (blood pressure, gene expression) suggesting that social connection influences morbidity and mortality [7]. Several cross-sectional studies have identified feelings of loneliness to be predictive to dementia onset [8, 9] one even concluding feelings of loneliness to be more predictive than social isolation [10]. Recent studies have also shown dementia and poor cognitive performance to be predictive of higher mortality risk [11, 12]. At the same time, loneliness in older adults has recently become a public health concern

associated with cardiovascular disease [13, 14], psychiatric morbidities [15, 16], hospitalization [17, 18], and mortality [19, 20]. A recent systematic review and meta-analysis of 144 studies identified loneliness to be associated with adverse health outcomes, especially on mental health and overall well-being, with a stronger effect on men [21]. Another recent meta-analysis of 36 studies not only concluded that loneliness was associated with mortality but also a small social network size was also associated with mortality [22], providing a possible objective aspect to addressing subjective feelings of loneliness in older adults. Nevertheless, most studies have either examined the association between loneliness and mortality (e.g., [19, 22–27]) or loneliness and dementia (e.g., [8, 28–35]).

Given that loneliness, dementia, and mortality are interconnected, understanding the pathways, along with the potential interaction between loneliness and dementia in predicting mortality, provides a relevant area of research to be developed. To our knowledge, no known study has comprehensively examined the mediating and moderating role of loneliness, dementia, and mortality risk among older adults using a nationally representative sample. Therefore, we test whether dementia and loneliness are related to mortality risk, independently of potential confounders and across sex and race groups, and their 2-way interactions. We then simultaneously test how loneliness and dementia odds interact to determine mortality risk and bidirectional mediation of "loneliness"-mortality and "dementia"-mortality through dementia status and loneliness, respectively, using four-way decomposition models.

MATERIALS AND METHODS

Database

Repeated every two years since 1992, the Health and Retirement Study (HRS) collects data on a nationally representative sample of US individuals over the age of 50 years [36, 37]. The National Institute on Aging (grant number U01AG009740) and the Social Security Administration fund the prospective cohort study. The core data contains a long list of variables that are obtained every two years from all HRS participants, covering a variety of domains related to health and retirement. This core data is made available in a curated manner in the RAND longitudinal dataset used in this study. In 2006, the Enhanced

Face-to-Face Interview (EFTF) was launched for the HRS to incorporate physical (i.e., performance tests & anthropometric measures), biological (i.e., blood and saliva samples), and psychosocial (i.e., self-administered psychosocial questionnaires) measures. Throughout the primary sampling units (PSUs), nearly half of homes with one or more living respondent were included for an EFTF interview.

From the 2006 HRS wave, a random half sample of respondents were used for EFTF interviews. In 2008, the second half of the sample was examined. Likewise, for 2010 new cohort families were assigned at random to one of these two groups, beginning in 2010 or 2012 for EFTF data collection. Ensuring the same request was sent to both household members, the household level was chosen as the sample. New spouses of respondents who were marked for an EFTF interview were asked to do the same. In coupled families, both members were chosen. A few EFTF sample participants were not required to complete physical or biomarker measures. Those respondents a) were questioned by a proxy, b) were in a nursing home, or c) refused face-to-face interview but agreed to a phone interview. In this study, EFTF psychosocial data was used for the years 2006 and 2008 combined.

Standard protocol approvals, registrations, and patient consents

All procedures followed the ethical standards of the institutional or regional committee on human experimentation. Approval was obtained from the committee of human subjects at a local universities Institutional Review Board (IRB). For the current prospective longitudinal cohort analysis, approval was obtained from the parent IRB-approved study, from a government agencies research program, while the ethics board determined that participant consent was not required or waived.

Study sample

A total of 43,561 HRS particpants (1992–2018) were included in the initial RAND file for the longitudinal study (randhrs1992_2018v2). Based on the participant flowchart (Supplementary Figure 1), in 2010, those who were included and alive were 22,034. Among them, those aged > 50y were 20,695. Additonally, those with data on dementia status or probability in 2010 were 7,677. In the final sample, we accounted for missing data on loneliness, including 6,468 participants > 50y in 2010 with dementia status or probability and loneliness data who were followed up until 2020 for mortality from all causes using the HRS latest tracker file available (trk2020tr_r). The mean age in 2010 was 78y. Around 5.5% of this sample reported being in a nursing home over the past 2 years.

All-cause mortality

The present study had the outcome variable all-cause mortality between 2010 and 2020, which was documented from 2010 for both the loneliness exposure and dementia status or probability measurement. Deaths were identified in the HRS using data linkage from the population registry and interviews [38]. Variables to 2020 (verson 2) were used from the tracker file. During the follow-up period from 2010–2020, 6,468 participants were used in our final sample. Among those who died from any cause with a known date of death (month and year), 3,298 were counted. Time-to-death or censoring was calculated from time elapsed (measured in years) between age in 2010 and age at death or age at censoring.

Dementia status and probability

The data file (*hrsdementia_2021_1109.sas7bdat*) which is publicly available includes dementia probabilities and classifications projected for 2000–2016 HRS respondents 60 + in 2006. Several new algorithms were developed. Those included a modified version from Hurd and colleagues (Modified Hurd Model) who originally developed the model, along with a new expert-informed logistic model (Expert Algorithm) and the Least Absolute Shrinkage and Selection Operator (LASSO) algorithm were trained and tested using HRS data along with data from all four waves of the Aging, Demographics, and Memory Study (ADAMS; http://hrsonline.isr. umich.edu/index.php?p=shoavail&iyear=XB). This process achieved 77–83% sensitivity, 92–94% specificity, and 90–92% accuracy overall in out-of-sample performance. The Supplementary Material and Supplementary Methods File contain further information regarding these algorithms. In summary, two forms of these variables were available, one being dichotomized as 0 = no dementia versus 1 = dementia and another being a predicted probability from each of the models, which was then transformed into Ln(odds of dementia).

Loneliness score

Loneliness was operationalized using 3 items from 2006 and 2008, each measured on a Likert scale of 1–3. These items were KLB020A, KLB020B, KLB020 C, LLB020A, LLB020B, and LLB020 C, which were reverse coded to reflect greater loneliness with higher scores and can be obtained from the Leave-Behind Questionnaires section LB, for the 2006 wave of data (public use data, wave K). The three items from 2006 were combined with those in 2008 creating a total loneliness score from 1–3 with 3 indicating greater loneliness. The general formulation of these items was as follows: In 2006: Item 1: "How often do you feel you lack companionship?"; Item 2: "How often do you feel left out?"; and Item 3: "How often do you feel isolated from others?"

In 2008: Item 1: "You lack companionship?"; Item 2: "Left out?"; and Item 3: "Isolated from others?". Initially scaled from 1–3, items were combined together by taking the mean if at most 1 item is missing out of 3 [39, 40]. This resulted in a score ranging between 1 and 3, which was then multiplied by 3. Subtracting 3 from the resulting product, we then obtained a potential range of 0 to 6, with "0" reflecting no or little loneliness and "6" reflecting "extreme loneliness". Details and code are provided in the Supplementary Methods 2 and in (https://hrs.isr.umich.edu/sites/default/files/meta/2006/core/codebook/h06lb_ri.htm). The loneliness score ordered data was also split into thirds or tertiles using the final selected sample (T1, T2, and T3). Furthermore, also within the final selected sample, a z-score of the loneliness total score was estimated with a unit increase denoting a 1 standard deviation increase in that score. In summary, three forms of the loneliness exposure were used alternatively: (A) The raw form ranging from 0–6; (B) The loneliness score tertiles; (C) The loneliness standardized z-score calculated within the final selected sample, as deemed most appropriate for each type of analysis.

Covariates

Socio-demographic characteristics

The following socio-demographic variables were considered: sex (male, female), age measured in 2010, self-reported race/ethnicity (White versus Black/Hispanic/Other), marital status (never married, married/partnered, separated/divorced, widowed), education (no degree, GED, high school graduate, some college, college degree or higher) work status

(working, not working) and total wealth (in US dollars) (<25,000, 25,000–124,999, 125,000–299,000, 300,000+).

Lifestyle factors

Factors falling in the category of lifestyle were also measured in 2010, including smoking status (never smoker, past smoker, current smoker) and frequency of moderate/vigorous exercise (never, 1–4 times per month, >1 times per week).

Health characteristics

The following health characteristics measured in 2010 were considered: self-rated health (excellent/very good/good versus fair/poor), self reported weight and height measures along with existing cardiometabolic risk factors and chronic illness determined by a physician and depressive symptoms measured by an 8-item Centers for Epidemiologic Studies-Depression (CES-D) scale total score. To enhance the sample size, multiple imputation was used with these variables.

Statistical analysis

Data taken from the HRS were analyzed with Stata 18.0 (StataCorp, College Station TX) [41] and sampling design complexity was accounted for [42] using PSU, stratifications, and sampling weights. Multiple imputation using chained equations [43, 44] were used for covariates only, specifically with 5 imputations and 10 iterations. Stata commands *mi register*, *mi stset*, *mi svyset*, *mi impute*, *mi passive* and *mi estimate* were used among others. Missingness ranged between 0% and 1.3% for all variables that were imputed, with up to 270 observations imputed out of over 21,000 in the case of BMI. To estimate population means, proportions, and regression coefficients we used the survey (*svy*) command. Taylor series linearization were used to adjust standard errors [42]. To compare means and proportions of primary variables across sex and race groups, svy:reg and svy:mlogit commands were used, stratifying on sex and race.

Age at entry > 50 years (i.e., delayed entry) until last age when the event of concern or censoring (loss to follow-up or end of follow-up) occurred is how we defined time-to-event (in years). We defined loss to follow-up as the age respondents were not present for a wave but were alive and responded in the previous wave. Only respondents alive in 2010 were included

in the final sample analysis. The time variable used in this analysis was the number of years elapsed between age at the end of 2010 (age at entry), and age at death, censoring or by December 31, 2020. Across the loneliness tertiles and dementia status groups, Kaplan-Meier survival curves were used, along with a log-rank test to evaluate differences in survival functions.

We implemented race- and sex-stratified Cox proportional hazards (PH) models to evaluate the relationship between loneliness, dementia, and all-cause mortality using both reduced (Model 1) and fully adjusted models (Model 2). Age in 2010, sex and race were included in Model 1 while Model 2 further adjusted for education, total wealth, marital status, smoking status, physical exercise, self-rated health, body mass index categories, cardiometabolic risk groups, and CES-D total scores. The main exposures were loneliness as a continuous measure and each "dementia probability" [Ln(odds) (Pr/(1-Pr) transformed; Hurd, Expert and LASSO] (Analysis A) for models 1 and 2. Loneliness and Ln(odds of dementia)'s associations with mortality risk by sex and race were tested for heterogeneity effects by adding 2-way interactions between each exposure and sex or race. Binary dementia status variables and loneliness tertiles as main exposures were applied with a similar analysis (Analysis B).

To test the association of Ln(odds of dementia), Analysis A) or dementia status (Analysis B) and mortality risk across loneliness tertiles, an additional set of models were used (reduced—i.e., Model 1—and fully adjusted—i.e., Model 2) with a similar analytic process as before. Using Cox PH models, each dementia exposure was entered separately with all-cause mortality of all causes as the outcome of interest, while stratifying by loneliness tertile. Loneliness tertiles were tested for heterogeneity in an unstratified model with 2-way interaction terms for each exposure of interest, considering loneliness tertile was an ordinal variable.

Next, we used four-way decomposition to decompose TE into four distinctive components: (1) no interaction or mediation, or controlled direct effect (CDE); (2) interaction is present rather than mediation or reference interaction (INTREF); (3) both an interaction and mediation are present, or INTMED; and (4) mediation is present but not an interaction or Pure Indirect Effect (PIE) (see the Supplementary Methods File) and a summary in Table 1. These models were recently operationalized in Stata using the *Med4way* command [45]

[https://github.com/anddis/med4way]. We considered three alternative potential mediator/moderators, the Ln(odds of dementia) (i.e., Hurd, expert, and LASSO), in the total effect (TE) of loneliness on all-cause mortality risk between 2010 and 2020 and evaluated these relationships using four-way decomposition models. In each of the four-way decomposition models, standardized z-scores for exposures and mediator/moderators were utilized. Mediation and interaction effects of the TE of loneliness z-score on all-cause mortality were tested, along with three Ln(odds of dementia), z-scored, measures utilized as different potential mediator/moderators. In these four-way decomposition models, Cox PH regression was estimated for the exposure/mediator/moderator versus all-cause mortality outcome equation, and OLS regression for each mediator/moderator as predicted by the exposure of interest. Main four-way decomposition models were applied to the total sample, using a reduced set of exogeneous covariates, namely age in 2010, sex and race; with additional models tested stratifying by sex and by race, separately. Consideration of the full model was also used as a secondary analysis. In a similar analytic approach, a four-way decomposition for the TE of Ln(odds of dementia), z-scored, on mortality risk [Ln(odds of dementia, z-scored)-mortality risk analysis] was decomposed using the loneliness z-score as a potential mediator/moderator. As a sensitivity analysis, the TE of loneliness z-score and Ln(odds of dementia), z-scored, in four-way decomposition models were compared with addition of each exogenous variable aside from those included in the reduced model, to test confounding effects of those variables. A type I error of 0.05 was set for all analyses, except for the 2-way interaction terms whereby 0.10 was set [46]. To address multiple testing, we adjusted the type I error using familywise Bonferroni correction accounting for multiplicity in dementia algorithms[47], assuming all other analyses to be independent hypotheses. Thus, the type I error was reduced to 0.017 for main effects and 0.033 for interaction terms, similar to what was done in previous studies (e.g., [48–50]).

RESULTS

Of the 6,468 participants that were selected in the final analytic sample, 2,718 were male, 3,564 were female, 5,135 were White and 1,147 were "Black/Hispanic/Other ethnic groups" adults, with

Table 1
Summary of four-way decomposition models

Parameter	Equation	Interpretation	Sources
Total Effect, TE	$TE = CDE + INT_{ref} + INT_{med} + PIE$	The total effect of the exposure on the outcome, which is a sum of four partitioned components that explain the total effect or variance between the exposure and the outcome.	[45]
General form of the model	$E[Y\|(a, m, c)] = \theta_0 + \theta_{1a} + \theta_{2m} + \theta_{3a} * m + \theta_c c$ $E[M\|(a, c)] = \beta_0 + \beta_1 a + \beta_c^T c$	The two regression models required to estimate the individual effect components that go into the sum for the total effect. In the first equation, we regress the outcome and condition on the exposure, mediators, and the vector of confounders. In the second equation, we regress the mediators and condition on the exposure and confounders.	[45]
Controlled Direct Effect, CDE	$E[CDE\|c] = \theta_1(a - a')$	The effect of the exposure on the outcome, not related to interaction or mediation.	[45]
Reference Interaction, INTref	$E[INT_{ref}\|c] = \theta_3(\beta_0 + \beta_1 a' + \beta_c^T c)(a - a')$	The interaction effect of the exposure on the outcome in the presence of the mediator, when the presence of the outcome is not necessary for the presence of the mediator.	[45]
Mediated Interaction, INTmed	$E[INT_{med}\|c] = \theta_3 \beta_1(a - a')(a - a')$	The interaction effect of the exposure on the outcome in the presence of the mediator, when the presence of the outcome is necessary for the presence of the mediator.	[45]
Pure Indirect Effect, PIE	$E[PIE\|c] = (\theta_2 \beta_1 + \theta_3 \beta_1 a')(a - a')$	The effect of the mediator on the outcome when the exposure is necessary for the presence of the mediator.	[45]
Final equation form, Cox PH model	$\lambda(t\|x, v, z) = \lambda_0(t) + \theta_1 x + \theta_2 v + \theta_3 x * v + \theta_z^T z$ $E[V\|(x.z)] = \beta_0 + \beta_1 x' + \beta_z^T z$	The two regression models required for the computation of the four components describing the total effect of the exposure on the outcome. In the first equation, the log hazard at time t is conditioned on the exposure, mediator, and confounders (using a Cox proportional hazards model). In equation two we implement ordinary least squares and regress the mediator on the exposure and confounders.	[45]

Details are provided in the technical appendix: loneliness-dementia-four-way-decomposition-models.pdf

	Overall	Males	Females	NHW	NW	Sex Differences	Sex Differences p	Race Differences	Race Differences p
	Mean/%±SE	Mean/%±SE	Mean/%±SE	Mean/%±SE	Mean/%±SE	Coefficient/±SE		Coefficient/±SE	
	N=6,468	N=2,718	N=3,564	N=5,135	N=1,147	N=6,282		N=6,282	
Sociodemographic									
Sex:									
Male	42.4±0.6	100.0	–	43.0±0.6	37.7±1.7	–	–	–0.2±0.1	0.003
Female	57.6±0.6	–	100.0	57.0±0.6	62.3±1.7	–	–	–	–
Age (y):									
Mean±SE	78.3±0.1	77.7±0.1	78.7±0.1	78.4±0.1	77.4±0.2	1.0±0.1	<0.001	–1.1±0.3	0.001
Race/Ethnicity:									
Non-Hispanic White	88.3±0.8	89.6±1.0	87.3±0.9	100.0	–	–	–	–	–
Black/Hispanic/Other	11.7±0.8	10.4±1.0	12.7±0.9	–	100.0	–	0.003	–0.2±0.1	–
Non-Hispanic Black	6.9±0.5	5.9±0.5	7.6±0.5	–	–	0.3±0.1	<0.001	–	–
Hispanic	4.9±0.7	4.6±0.8	0.5±0.8	–	–	0.1±0.1	0.294	–	–
Other	0.0	0.0	0.0	0.0	0.0	–	–	–	–
Education:									
No degree	18.5±0.8	18.7±1.0	18.4±1.0	14.8±0.7	46.0±2.3	–0.3±0.1	<0.001	1.5±0.1	<0.001
GED	4.3±0.4	5.6±0.5	3.3±0.4	4.3±0.4	4.2±0.7	–0.8±0.1	<0.001	0.3±0.2	0.137
High School graduate	35.2±07	29.0±0.5	39.8±0.8	37.5±0.7	25.3±1.7	–	–	–	–
Some college	21.0±0.6	19.0±0.9	22.6±0.9	21.7±0.7	16.0±1.1	–0.1±0.1	0.067	0.1±0.1	0.538
College degree or higher	20.9±1.0	27.8±1.6	15.9±0.8	22.6±1.1	8.5±1.1	–0.9±0.1	<0.001	–0.6±0.2	<0.001
Marital status:									
Never married	4.0±0.4	4.2±0.5	3.9±0.6	3.8±0.4	5.9±1.1	0.6±0.2	0.003	0.8±0.2	0.001
Married/ Partnered	55.0±0.7	74.9±0.9	40.4±0.8	56.9±0.8	40.6±1.4	–	–	–	–
Separated/ Divorced	8.9±0.4	7.3±0.6	10.1±0.6	8.0±0.4	15.2±1.0	0.9±0.1	<0.001	1.0±0.2	<0.001
Widowed	32.1±0.7	13.7±0.6	45.7±0.9	31.3±0.8	38.2±1.6	1.8±0.1	<0.001	0.6±0.1	<0.001
Working status:									
No Working	86.9±0.5	82.4±0.8	90.2±0.6	86.6±0.6	89.2±1.0	–	–	–	–
Working	13.1±0.5	17.6±0.8	9.8±0.6	13.4±0.6	10.8±1.0	–0.7±0.1	<0.001	–0.2±0.1	0.050

(*Continued*)

Table 2
(*Continued*)

	Overall	Males	Females	NHW	NW	Sex Differences	Sex Differences p	Race Differences	Race Differences p
Total wealth ($):									
<25,000	37.6 ± 1.0	25.3 ± 1.1	46.6 ± 0.1	33.8 ± 1.1	66.3 ± 1.7	0.9 ± 0.0	<0.001	–1.3 ± 0.1	<0.001
25,000–124,999	56.4 ± 0.8	66.2 ± 0.9	49.1 ± 1.0	59.7 ± 1.0	31.6 ± 1.7	–	–	–	–
125,000–299,999	5.0 ± 0.5	7.2 ± 0.9	3.4 ± 0.4	5.4 ± 0.6	2.0 ± 0.5	–0.5 ± 0.1	<0.001	–0.4 ± 0.3	0.261
300,000–649,999	0.9 ± 0.2	1.2 ± 0.2	0.7 ± 0.2	1.0 ± 0.2	0.1 ± 0.1	–0.2 ± 0.3	0.508	–2.1 ± 1.0	0.050
≥650,000	0.2 ± 0.1	0.1 ± 0.1	0.2 ± 0.1	0.2 ± 0.1	0.0	1.0 ± 0.6	0.086	–21.0 ± 0.4	<0.001
Lifestyle									
Smoking status:									
Never smoker	43.8 ± 0.8	30.4 ± 1.0	53.5 ± 1.0	43.4 ± 0.9	46.8 ± 1.7	1.0 ± 0.1	<0.001	0.2 ± 0.1	0.055
Past smoker	50.1 ± 0.6	63.1 ± 1.0	40.6 ± 0.8	50.7 ± 0.8	45.5 ± 2.0	–	–	–	–
Current Smoker	6.1 ± 0.5	6.5 ± 0.5	5.9 ± 0.6	5.9 ± 0.5	7.7 ± 1.1	0.3 ± 0.1	0.004	0.4 ± 0.2	0.047
Physical activity:									
Never	27.7 ± 0.7	21.6 ± 1.0	32.1 ± 0.8	27.2 ± 0.7	31.6 ± 1.6	0.6 ± 0.1	<0.001	0.3 ± 0.1	<0.001
1–4 times per month	24.5 ± 0.9	25.1 ± 1.0	24.0 ± 1.0	24.1 ± 1.0	27.5 ± 1.4	0.1 ± 0.1	0.025	0.3 ± 0.1	0.001
>1 times per week	47.9 ± 0.9	53.2 ± 1.4	43.9 ± 0.9	48.8 ± 1.0	40.9 ± 1.4	–	–	–	–
Health related									
Self-rated health:									
Excellent/very good/ good	73.5 ± 0.8	73.0 ± 1.0	72.9 ± 0.9	76.1 ± 0.7	54.5 ± 1.8	–	–	–	–
Fair/poor	26.5 ± 0.8	27.0 ± 1.0	26.1 ± 0.9	23.9 ± 0.7	45.5 ± 1.8	–0.0 ± 0.1	0.347	1.0 ± 0.1	<0.001
Body mass index (kg/m2):									
<25	36.1 ± 0.8	29.0 ± 1.2	41.4 ± 0.9	37.1 ± 0.8	28.7 ± 1.8	0.7 ± 0.1	<0.001	–0.3 ± 0.1	0.001
25–29.9	39.1 ± 0.7	46.1 ± 1.2	34.0 ± 0.9	39.0 ± 0.8	40.6 ± 1.3	–	–	–	–
≥30	24.7 ± 0.6	24.9 ± 1.0	24.6 ± 0.8	24.0 ± 0.7	30.7 ± 1.6	0.3 ± 0.1	<0.001	0.2 ± 0.1	0.020
Number of chronic conditions[c]									
0	20.8 ± 0.4	18.7 ± 0.8	22.4 ± 0.8	22.0 ± 0.5	15.7 ± 1.1	0.2 ± 0.1	0.045	–0.3 ± 0.1	0.000
1–2	66.2 ± 0.5	66.0 ± 1.1	66.4 ± 0.8	66.0 ± 0.6	68.1 ± 1.4	–	–	–	–
≥3	12.9 ± 0.4	15.4 ± 0.7	11.1 ± 0.5	12.5 ± 0.5	16.2 ± 1.3	–0.3 ± 0.1	<0.001	0.2 ± 0.1	0.032
CES-D total score	1.3 ± 0.0	1.0 ± 0.3	1.5 ± 0.0	1.2 ± 0.0	1.8 ± 0.1	0.4 ± 0.0	<0.001	0.6 ± 0.1	<0.001
Loneliness[d]	1.3 ± 0.0	1.2 ± 0.0	1.4 ± 0.0	1.3 ± 0.0	1.6 ± 0.1	0.2 ± 0.0	<0.001	0.3 ± 0.1	<0.001

MORTALITY RATE, 2010–2020:

#death per 1000 P-Y with 95% CI	64 (62–66)	74(70–77)	57 (54–60)	65 (63–68)	58 (53–63)	–	–	–	–
Dementia Status									
Hurd:									
No	86.3 ± 0.6	87.6 ± 0.8	85.4 ± 0.7	87.0 ± 0.7	81.4 ± 1.5	–	–	–	–
Yes	13.7 ± 0.6	12.4 ± 0.8	14.6 ± 0.7	13.0 ± 0.7	18.6 ± 1.5	0.2 ± 8.5	0.031	0.4 ± 0.1	0.001
Expert:									
No	84.0 ± 0.5	85.1 ± 0.8	83.2 ± 0.6	84.5 ± 0.5	80.7 ± 1.3	–	–	–	–
Yes	16.0 ± 0.5	14.9 ± 0.8	16.8 ± 0.6	16.0 ± 0.5	19.3 ± 1.3	14.3 ± 7.6	0.064	0.3 ± 0.1	0.007
Lasso:									
No	83.5 ± 0.6	85.2 ± 0.8	82.3 ± 0.6	84.6 ± 0.7	75.7 ± 1.4	–	–	–	–
Yes	16.4 ± 0.6	14.8 ± 0.8	17.7 ± 0.6	15.4 ± 0.7	24.3 ± 1.4	0.2 ± 0.1	0.002	0.6 ± 0.1	<0.001
Dementia Probability									
Hurd:									
Mean ± SE	0.094 ± 0.003	0.086 ± 0.004	0.100 ± 0.004	0.087 ± 0.004	0.146 ± 0.009	0.013 ± 0.005	0.009	0.059 ± 0.010	<0.001
Expert:									
Mean ± SE	0.133 ± 0.004	0.122 ± 0.005	0.141 ± 0.005	0.126 ± 0.004	0.180 ± 0.010	0.019 ± 0.006	0.003	0.053 ± 0.011	<0.001
Lasso:									
Mean ± SE	0.133 ± 0.004	0.117 ± 0.005	0.144 ± 0.004	0.127 ± 0.004	0.174 ± 0.008	0.027 ± 0.005	<0.001	0.046 ± 0.009	<0.001

Abbreviations: LASSO = Least Absolute Shrinkage and Selection Operator; NHW = Non-Hispanic White; SE = Standard Error. [a]Values are means ± SE column percentages for overall and fixed sample characteristics, across sex and race/ethnicity groups, accounting for sampling design complexity. All covariates are measured in 2010. [b]Based on linear or multinomial logit models using sex or race as main predictors for both continuous and categorical variables, accounting for sampling design complexity. [c]Number of chronic conditions among hypertension, diabetes, heart problems and stroke. [d]Loneliness score had a mean of 1.33 and a standard deviation of 1.56 in the unweighted and unimputed sample. Loneliness tertiles 1, 2 and 3, had means ± SD of 0.00 ± 0.00 (n = 2,902), 1.45 ± 0.50 (n = 2,010), and 3.68 ± 1.00 (n = 1,556) in the unweighted and unimputed sample, N = 6,468.

an average age of 78.3 years. The mortality rate in this sample was 64 per 1000 person-years (P-Y), overall (Supplementary Figure 1), with a median survival time of 9.9 years and a maximum of 11.2 years. Weighted proportions and means $\pm$ SE of key characteristics are presented in Table 2, overall, by sex and by racial minority status groups. Male older adults were significantly more likely than female older adults to be married/parternered (74.9% versus 40.4%, $p < 0.001$). They also were more likely to be currently working (17.6% versus 9.8%), to have a college degree or higher (27.8% versus 15.9%), and less likely to have a total wealth ($)<25,000 (25.3% versus 46.6%). In terms of lifestyle factors, females had a higher proportion of never smokers (53.5% versus 30.4%), though their phyiscal activity level was significantly lower, with 43.9% among females versus 53.2% among males reporting > 1 times/week of physical exercise. BMI within the normal range was a more likely outcome among male adults (46.1% versus 34.0%, $p < 0.001$). While self-rated health did not differ across sex; male older adults reported ≥ 3 cardiometabolic factors and conditions to a greater extent than their female counterparts (15.4% versus 11.1%, $p < 0.001$). In contrast, female older adults had on average a higher level of depressive symptoms compared to male older adults (mean $\pm$ SE:

1.5 ± 0.0 versus 1.0 ± 0.0, $p < 0.001$). Similarly, loneliness scores, proportion with dementia based on the 3 algorithms, and dementia probabilities were higher among female older adults compared to their male counterparts.

In terms of racial disparities, "Black/Hispanic/ Other ethnic groups" older adults had a greater likelihood of dementia based on all 3 algorithms, higher scores on loneliness and depressive symptoms, and significantly higher proportion reporting fair/poor self-rated health or ≥ 3 chronic condition compared to their White counterparts. Furthermore, lower level of physical exercise, current smoking and lower SES in terms of education and income were also more prevalent among "Black/Hispanic/Other ethnic groups" older adults compared to their White counterparts.

Figure 1 presents survival probabilities across loneliness score tertiles and by dementia status (3 algorithmic definitions). Loneliness score tertiles were directly associated with a higher mortality rate, with a dose-response relationship. Notably, dementia status was consistently and strongly associated with all-cause mortality risk.

Table 3 presents Cox PH models that included "Loneliness score" as the main ordinal exposure and the outcome being all-cause mortality risk, displaying
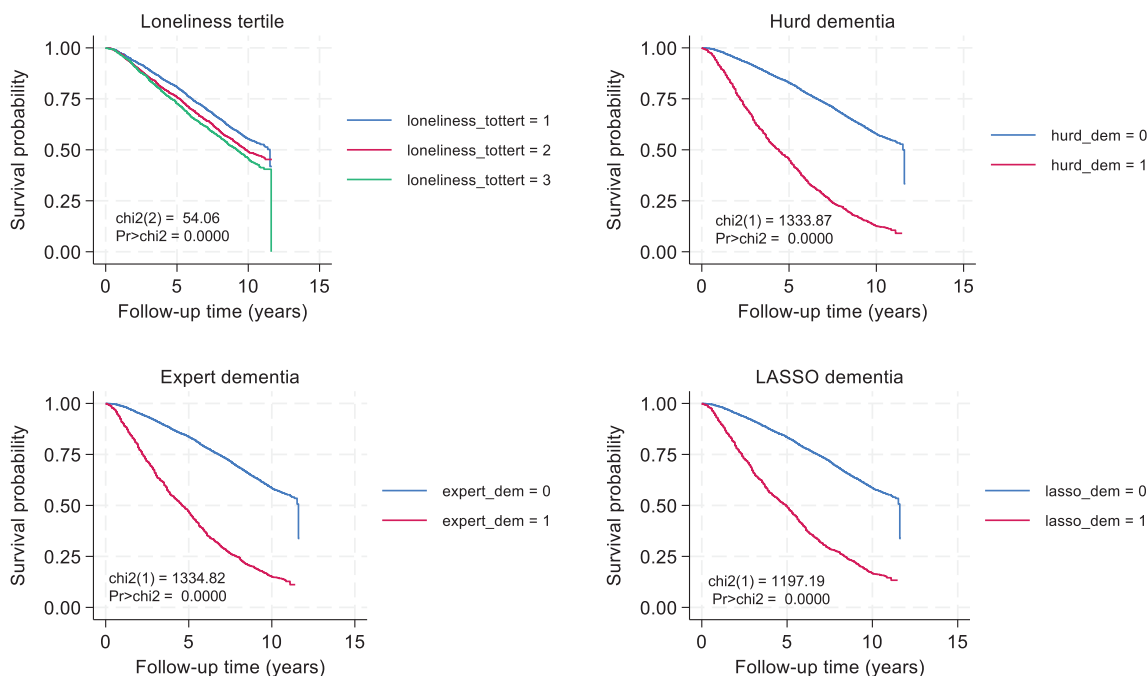

Fig. 1. Loneliness score tertile, dementia status, and all-cause mortality: K-M survival curves. *Abbreviations:* chi2 = Chi-square test, Log-rank test; K-M = Kaplan Meier; dem = Dementia; tert = Tertile.

Table 3

Loneliness, [Ln(odds of dementia), Analysis A] or dementia status (Analysis B) and all-cause mortality: Cox PH models, overall, by sex and by race, HRS 2010–2020[a,b]

	Overall	Male	Female	White	Black/Hispanic/Other	p_{sex}[f]	p_{race}[f]
	β±SE	β±SE	β±SE	β±SE	β±SE		
Unweighted N	$N = 6,450$	$N = 2,707$	$N = 3,553$	$N = 5,116$	$N = 1,144$		
Analysis A							
Reduced models							
Model 1A: Loneliness	0.082 ± 0.012[d]	0.099 ± 0.020[d]	0.070 ± 0.014[d]	0.083 ± 0.013[d]	0.074 ± 0.026[c]	0.217	0.843
Model 1B: Hurd dementia	0.092 ± 0.011[d]	0.130 ± 0.009[d]	0.084 ± 0.012[d]	0.092 ± 0.012[d]	0.099 ± 0.015[d]	0.003	0.711
Model 1C: Expert dementia	0.164 ± 0.008[d]	0.166 ± 0.011[d]	0.164 ± 0.010[d]	0.168 ± 0.009[d]	0.132 ± 0.018[d]	0.270	0.145
Model 1D: LASSO dementia	0.205 ± 0.010[d]	0.200 ± 0.014[d]	0.211 ± 0.014[d]	0.207 ± 0.011[d]	0.185 ± 0.025[d]	0.688	0.543
Full models	$N = 6,188$	$N = 2,555$	$N = 3,443$	$N = 4,919$	$N = 1,079$		
Model 2A: Loneliness	0.012 ± 0.013	0.012 ± 0.021	0.009 ± 0.019	0.005 ± 0.015	0.043 ± 0.032	0.583	0.758
Model 2B: Hurd dementia	0.107 ± 0.010[d]	0.115 ± 0.016[d]	0.104 ± 0.012[d]	0.109 ± 0.011[d]	0.098 ± 0.033[c]	0.212	0.297
Model 2C: Expert dementia	0.105 ± 0.009[d]	0.102 ± 0.013[d]	0.111 ± 0.012[d]	0.110 ± 0.010[d]	0.080 ± 0.022[c]	0.438	0.038
Model 2D: LASSO dementia	0.148 ± 0.012[d]	0.143 ± 0.017[d]	0.157 ± 0.015[d]	0.152 ± 0.013[d]	0.120 ± 0.035[c]	0.625	0.138
Analysis B							
Reduced models	$N = 6,450$	$N = 2,707$	$N = 3,553$	$N = 5,116$	$N = 1,144$		
Model 1A: Loneliness	0.153 ± 0.025[d]	0.188 ± 0.038[d]	0.123 ± 0.033[d]	0.152 ± 0.028[d]	0.163 ± 0.059[c]	0.180	0.817
Model 1B: Hurd dementia	0.705 ± 0.049[d]	0.567 ± 0.084[d]	0.822 ± 0.066[d]	0.729 ± 0.051[d]	0.549 ± 0.146[d]	0.171	0.245
Model 1C: Expert dementia	0.692 ± 0.060[d]	0.591 ± 0.080[d]	0.776 ± 0.074[d]	0.716 ± 0.064[d]	0.518 ± 0.138[d]	0.215	0.219
Model 1D: LASSO dementia	0.658 ± 0.042[d]	0.557 ± 0.073[d]	0.747 ± 0.064[d]	0.691 ± 0.050[d]	0.443 ± 0.119[c]	0.338	0.078
Full models	$N = 6,188$	$N = 2,555$	$N = 3,443$	$N = 4,919$	$N = 1,079$		
Model 2A: Loneliness	0.031 ± 0.025	0.038 ± 0.041	0.018 ± 0.040	0.021 ± 0.029	0.077 ± 0.065	0.535	0.899
Model 2B: Hurd dementia	0.428 ± 0.050[d]	0.363 ± 0.090[d]	0.497 ± 0.070[d]	0.453 ± 0.053[d]	0.360 ± 0.171	0.990	0.193
Model 2C: Expert dementia	0.371 ± 0.063[d]	0.316 ± 0.092[d]	0.435 ± 0.076[d]	0.391 ± 0.066[d]	0.327 ± 0.166	0.876	0.174
Model 2D: LASSO dementia	0.402 ± 0.044[d]	0.335 ± 0.083[d]	0.490 ± 0.066[d]	0.431 ± 0.054[d]	0.280 ± 0.132	0.843	0.116

Abbreviations: HRS = Health and Retirement Study; LASSO = Least Absolute Shrinkage and Selection Operator; PH = Proportional hazards; SE = Standard Error. [a]Values are β±SE, with β representing Ln HR from Cox PH model for each exposure-outcome relationship. Cox PH models were conducted overall and across sex or race/ethnicity groups, accounting for sampling weights and sampling design complexity in multiple imputed data. [b]All covariates are measured in 2010 unless stated otherwise. Analysis A included continuous forms of dementia probability (Ln(odds of dementia) and ordinal Loneliness score, while Analysis B included binary dementia status (3 algorithms) and Loneliness tertile. [c]$p < 0.010$; [d]$p < 0.001$ for null hypothesis that β = 0. [e]Reduced models (Models 1A-1D) were adjusted for age in 2010, sex and race/ethnicity. Full models (Model 2A-2D) further adjusted the reduced model by all covariates described under the Covariates section, including socio-demographic, socio-economic, lifestyle and health-related factors. [f]p-value associated with 2-way interaction term between sex or race and the main exposure (loneliness score or dementia odds), in a model not stratified by sex or race.

adjusted Ln(HR) with their SEs on multiple imputed data. In these Cox PH models, a unit increase on the "loneliness" score was associated with a 1.09-fold increased risk of all-cause mortality [HR = 1.09, 95% CI: 1.06–1.10, $p < 0.001$] adjusted only for age, sex and race/ethnicity (Table 3). This association was detected among all sex and race sub-groups, without any significant heterogeneity. Nevertheless, upon adjustment for other covariates, including socio-economic, lifestyle and health-related factors, the "Loneliness" score was no longer associated with mortality risk, with an overall HR = 1.01 (0.99–1.02) per unit increase ($p > 0.05$). In models with the loneliness score tertile as the exposure of interest (*Analysis B*), each tertile increase in the loneliness score measure was associated with a 16% increase in mortality risk overall and specifically in the reduced model. In the full model, a tertile increase in the loneliness score was not associated with mortality risk ($p > 0.05$), overall or within any of the sub-groups.

In Table 3, we included three different algorithms to estimate the odds of dementia, which were then entered into a Cox PH model for all-cause mortality risk, using a similar modeling approach to as for the "Loneliness" exposures. The Ln(odds of dementia) were consistently associated with higher mortality risk, with up to 22% higher risk for the LASSO algorithm (reduced model). Upon further covariate adjustment, the association between dementia odds and mortality risk was attenuated for most algorithms, even though the relationship retained its statistical significance at type I error of 0.05. In all these models, a unit increase in the Ln(odds of dementia) corresponded to dementia probability moving from 0% to 73%. Similar patterns were observed for the binary dementia exposures (0 = no, 1 = yes), whereby having a probable dementia diagnosis based on each algorithm was linked to approximately a two-fold increased risk of all-cause mortality in the reduced model, an association that was markedly reduced in the full model though retaining statistical significance (HR~1.5), with no detectable sex or racial differences.

Table 4 displays findings from Cox PH models for the relationship between Ln(odds of dementia) and all-cause mortality, stratified by loneliness score tertile. In the fully adjusted and reduced models, there was no heterogeneity detected across loneliness levels, of the association between dementia and mortality, regardless of the algorithm or the definition of the dementia exposure. Thus, dementia status and Ln(odds of dementia) were both associated with mortality risk with comparable strength across loneliness score tertiles.

Four-way decomposition of the TE of loneliness, z-scored, on mortality risk through Ln(odds of dementia), z-scored, is presented in Table 5, by displaying findings from both a reduced and a fully adjusted model and for all 3 dementia algorithms. In the reduced models, the TE was > 0 and statistically significant for all algorithms and strata, indicating that loneliness was associated with increased risk of mortality independent of age, sex and race/ethnicity. This TE was partially mediated by Ln(odds of dementia), z-scored, with PIE explaining about 21–40% of the TE, depending on the algorithm, indicating that a significant proportion of the TE of loneliness on mortality risk is accounted for pure mediation without interaction. This pattern was observed across all socio-demographic strata. Nevertheless, the TE was markedly attenuated when other exogenous variables were accounted for [e.g., from +0.15 ± 0.020 ($p < 0.001$) to 0.027 ± 0.021 ($p = 0.43$), for Hurd algorithm], including other socio-demographic, socio-economic, lifestyle and health-related factors. The PIE remained statistically significant among female and White participants between the fully adjusted and the reduced models, explaining a large portion of the TE (>20% up to 100% of TE).

Four-way decomposition of the TE of Ln(odds of dementia), z-scored, on mortality risk through the loneliness z-score is shown in Supplementary Table 1. First, looking at the Hurd dementia algorithm and specifically the reduced model, the positive TE of Ln(odds of dementia), z-scored, on mortality (TE:0.411 ± 0.020, $p < 0.001$) consisted mainly of a controlled direct effect, CDE (+0.389 ± 0.020, $p < 0.001$). Nevertheless, around 3.6% of this TE was mediated through a positive association between Ln(odds of dementia), z-scored, and mortality risk (PIE:+0.015 ± 0.003, $p < 0.001$), another 1.1% was due to both mediation and interaction (INTMED:+0.005 ± 0.002, $p = 0.028$). After adjustment for other potentially confounding factors, the TE of loneliness z-score retained its statistical significance ($p < 0.001$), as did the CDE for all 3 algorithms, while PIE became markedly atenuated ($p > 0.05$).

A sensitivity analysis was carried out to determine the most influential potential confounder in each of the four-way decomposition models, starting from the reduced model. When loneliness z-score was consid-

Table 4

[Ln(odds of dementia), Analysis A] or dementia status (Analysis B) and all-cause mortality across loneliness tertile: Cox PH models, HRS 2010–2020[a]

	T1	T2	T3	$p_{loneliness}$ [f]
	β±SE	β±SE	β±SE	
Analysis A				
Unweighted N	$N = 2,850$	$N = 1,933$	$N = 1,477$	
Reduced models				
Model 1A: Hurd dementia	**0.080 ± 0.014**[d]	**0.132 ± 0.015**[d]	**0.095 ± 0.011**[d]	0.237
Model 1B: Expert dementia	**0.161 ± 0.013**[d]	**0.176 ± 0.015**[d]	**0.141 ± 0.016**[d]	0.771
Model 1C: LASSO dementia	**0.211 ± 0.018**[d]	**0.211 ± 0.019**[d]	**0.178 ± 0.020**[d]	0.311
Full models	$N = 2,794$	$N = 1,849$	$N = 1,402$	
Model 2A: Hurd dementia	**0.105 ± 0.019**[d]	**0.111 ± 0.020**[d]	**0.111 ± 0.023**[d]	0.965
Model 2B: Expert dementia	**0.094 ± 0.018**[d]	**0.129 ± 0.018**[d]	**0.099 ± 0.023**[d]	0.991
Model 2C: LASSO dementia	**0.146 ± 0.023**[d]	**0.168 ± 0.024**[d]	**0.145 ± 0.030**[d]	0.787
Analysis B				
Unweighted N	$N = 2,850$	$N = 1,933$	$N = 1,477$	
Reduced models				
Model 1A: Hurd dementia	**0.712 ± 0.090**[d]	**0.760 ± 0.086**[d]	**0.580 ± 0.118**[d]	0.509
Model 1B: Expert dementia	**0.654 ± 0.090**[d]	**0.815 ± 0.096**[d]	**0.542 ± 0.104**[d]	0.596
Model 1C: LASSO dementia	**0.684 ± 0.075**[d]	**0.704 ± 0.101**[d]	**0.506 ± 0.099**[d]	0.285
Full models	$N = 2,794$	$N = 1,849$	$N = 1,402$	
Model 2A: Hurd dementia	**0.446 ± 0.118**[d]	**0.487 ± 0.106**[d]	**0.390 ± 0.131**[c]	0.793
Model 2B: Expert dementia	**0.353 ± 0.100**[d]	**0.547 ± 0.107**[d]	0.285 ± 0.123	0.755
Model 2C: LASSO dementia	**0.426 ± 0.089**[d]	**0.528 ± 0.111**[d]	0.298 ± 0.121	0.505

Abbreviations: HRS = Health and Retirement Study; LASSO = Least Absolute Shrinkage and Selection Operator; PH = Proportional hazards; SE = Standard Error. [a]Values are β±SE, with β representing Ln(HR) from Cox PH model for each exposure-outcome relationship. Cox PH models were conducted overall and across sex or race/ethnicity groups, accounting for sampling weights and sampling design complexity in multiple imputed data. [b]All covariates are measured in 2010 unless stated otherwise. Analysis A included continuous forms of dementia probability (Ln(odds of dementia); while Analysis B included binary dementia status (3 algorithms). [c]$p < 0.010$; [d]$p < 0.001$ for null hypothesis that β = 0. [e]Reduced models (Models 1A-1C) were adjusted for age in 2010, sex and race/ethnicity. Full models (Model 2A-2C) further adjusted the reduced model by all covariates described under the Covariates section, including socio-demographic, socio-economic, lifestyle and health-related factors. [f]p-value associated with 2-way interaction term between loneliness tertile entered as an ordinal variable and the main exposure Ln(odds of dementia), in a model not stratified by loneliness score tertile.

Table 5
Loneliness z-score and all-cause mortality: four-way decomposition models by Ln(odds of dementia), z-scored, overall, by sex and by race

Y = All-cause mortality; X = Loneliness z-score	TE β±SE	p	CDE β±SE	p	INTREF β±SE	p	INTMED β±SE	p	PIE β±SE	p
Overall										
Reduced Model 1A-C (N = 6,468)										
M = Hurd	+0.153 ± 0.020	<0.001	+0.113 ± 0.020	<0.001	0.004 ± 0.003	0.255	+0.003 ± 0.002	0.023	+0.032 ± 0.004	<0.001
M = Expert	+0.155 ± 0.021	<0.001	+0.086 ± 0.019	<0.001	+0.002 ± 0.006	0.701	+0.003 ± 0.002	0.165	+0.062 ± 0.006	<0.001
M = LASSO	+0.152 ± 0.021	<0.001	+0.093 ± 0.018	<0.001	+0.000 ± 0.005	0.927	+0.003 ± 0.002	0.233	+0.057 ± 0.006	<0.001
Full Model 2A-C (N = 6,079)										
M = Hurd	0.027 ± 0.021	0.197	0.017 ± 0.021	0.432	–0.000 ± 0.001	0.402	–0.000 ± 0.001	0.605	+0.011 ± 0.003	0.001
M = Expert	0.028 ± 0.021	0.193	0.019 ± 0.021	0.380	–0.001 ± 0.001	0.313	–0.000 ± 0.001	0.582	+0.010 ± 0.003	0.002
M = LASSO	0.027 ± 0.021	0.203	0.016 ± 0.021	0.461	–0.001 ± 0.001	0.338	–0.000 ± 0.001	0.599	+0.012 ± 0.003	<0.001
Male										
Reduced Model 1A-C (N = 2,778)										
M = Hurd	+0.176 ± 0.031	<0.001	+0.136 ± 0.029	<0.001	–0.001 ± 0.005	0.808	+0.002 ± 0.003	0.526	+0.040 ± 0.007	<0.001
M = Expert	+0.189 ± 0.033	<0.001	0.094 ± 0.029	0.001	+0.017 ± 0.010	0.096	+0.011 ± 0.005	0.025	+0.067 ± 0.010	<0.001
M = LASSO	+0.182 ± 0.032	<0.001	+0.108 ± 0.029	<0.001	+0.004 ± 0.008	0.62	+0.006 ± 0.004	0.149	+0.064 ± 0.010	<0.001
Full model 2A-C (N = 2,578)										
M = Hurd	0.026 ± 0.033	0.427	0.015 ± 0.033	0.653	0.002 ± 0.003	0.577	0.001 ± 0.001	0.499	0.009 ± 0.006	0.120
M = Expert	0.025 ± 0.033	0.443	0.015 ± 0.033	0.636	0.001 ± 0.003	0.706	0.000 ± 0.001	0.598	0.008 ± 0.005	0.115
M = LASSO	0.024 ± 0.033	0.461	0.015 ± 0.033	0.641	–0.001 ± 0.001	0.544	0.000 ± 0.001	0.599	0.009 ± 0.006	0.111
Female										
Reduced Model 1A-C (N = 3,690)										
M = Hurd	+0.132 ± 0.026	<0.001	+0.094 ± 0.025	<0.001	+0.005 ± 0.004	0.257	+0.003 ± 0.002	0.060	+0.030 ± 0.005	<0.001
M = Expert	+0.131 ± 0.027	<0.001	+0.073 ± 0.025	0.004	–0.004 ± 0.007	0.563	+0.000 ± 0.003	0.893	+0.062 ± 0.008	<0.001
M = LASSO	+0.131 ± 0.027	<0.001	+0.077 ± 0.025	<0.001	–0.001 ± 0.007	0.863	+0.001 ± 0.003	0.643	+0.054 ± 0.008	<0.001
Full model 2A-C (N = 3,501)										
M = Hurd	+0.024 ± 0.028	0.395	+0.010 ± 0.028	0.715	+0.001 ± 0.002	0.521	–0.001 ± 0.001	0.342	+0.014 ± 0.004	0.001
M = Expert	+0.024 ± 0.028	0.380	+0.014 ± 0.028	0.623	–0.001 ± 0.001	0.488	–0.001 ± 0.001	0.338	+0.012 ± 0.004	0.004
M = LASSO	+0.023 ± 0.028	0.414	+0.020 ± 0.028	0.732	–0.001 ± 0.001	0.373	–0.001 ± 0.001	0.425	+0.015 ± 0.004	0.001

White
Reduced Model 1A-C (N = 5,289)

M = Hurd	+0.159 ± 0.022	0.000	+0.117 ± 0.021	0.000	+0.004 ± 0.003	0.158	+0.005 ± 0.002	0.007	+0.033 ± 0.004	0.000
M = Expert	+0.160 ± 0.023	0.000	+0.089 ± 0.021	0.000	+0.001 ± 0.006	0.896	+0.003 ± 0.003	0.227	+0.066 ± 0.007	0.000
M = LASSO	+0.157 ± 0.023	0.000	+0.096 ± 0.021	0.000	−0.001 ± 0.005	0.874	+0.003 ± 0.002	0.304	+0.059 ± 0.007	0.000
Full model 2A-C (N = 4,987)										
M = Hurd	+0.014 ± 0.023	0.552	+0.001 ± 0.023	0.973	+0.002 ± 0.002	0.227	−0.001 ± 0.001	0.134	+0.012 ± 0.004	0.001
M = Expert	+0.014 ± 0.023	0.550	+0.002 ± 0.023	0.925	+0.000 ± 0.001	0.985	−0.001 ± 0.001	0.133	+0.013 ± 0.004	0.001
M = LASSO	+0.013 ± 0.023	0.569	−0.000 ± 0.023	0.988	+0.001 ± 0.001	0.443	−0.002 ± 0.001	0.092	+0.014 ± 0.004	0.000
Black/Hispanic/Other										
Reduced Model 1A-C (N = 1,179)										
M = Hurd	+0.120 ± 0.046	0.009	+0.104 ± 0.044	0.017	−0.015 ± 0.016	0.348	−0.003 ± 0.004	0.445	+0.034 ± 0.010	0.000
M = Expert	+0.131 ± 0.047	0.006	+0.067 ± 0.044	0.123	+0.010 ± 0.018	0.591	+0.004 ± 0.005	0.469	+0.050 ± 0.013	0.000
M = LASSO	+0.129 ± 0.047	0.006	+0.071 ± 0.043	0.094	+0.006 ± 0.017	0.743	+0.002 ± 0.005	0.594	+0.049 ± 0.013	0.000
Full model 2A-C (N = 1,092)										
M = Hurd	+0.062 ± 0.051	0.228	+0.033 ± 0.053	0.535	+0.018 ± 0.013	0.162	+0.003 ± 0.002	0.294	+0.008 ± 0.006	0.186
M = Expert	+0.061 ± 0.051	0.236	+0.047 ± 0.052	0.364	+0.013 ± 0.010	0.218	+0.000 ± 0.002	0.893	+0.001 ± 0.006	0.892
M = LASSO	+0.058 ± 0.051	0.258	+0.034 ± 0.052	0.516	+0.017 ± 0.010	0.080	+0.002 ± 0.002	0.401	+0.005 ± 0.006	0.374

Abbreviations: CDE = Controlled Direct Effect; HRS = Health and Retirement Study; INTMED = Mediated Interaction; INTREF = Reference Interaction; LASSO = Least Absolute Shrinkage and Selection Operator; PIE = Pure Indirect Effect; SE = Standard Error; SD = Standard Deviation. [a]Cox PH regression models with mortality as the main outcome and loneliness score (z-scored, 1SD˜ = 1.56 unit increase) as the exposure. [Ln(odds of dementia), z-scored] using Hurd, expert and LASSO algorithms were potential mediators/moderators allowed to interact with the main exposure, sample size N = 6,079 (full model) and 6,468 (reduced model), four-way decomposition analysis. 1 SD for each Ln transformed odds of dementia was 3̃.8, 3.0, and 2.3, for Hurd, Expert and LASSO algorithms, respectively. [b]Exogenous variables are the ones included in Table 2, Models 1A-2D, as covariates for the reduced and full models, respectively. See Covariates section for detail.

ered as the exposure of interest (as in Table 5), the TE on mortality risk was most influenced by self-rated health and the CES-D score, for all dementia algorithms. When Ln(odds of dementia), z-scored, was the exposure of interest (as in Supplementary Table 1), the most influential potential confounders were physical activity, self-rated health and the CES-D score. Those attenuations in the TE on mortality risk are shown in more detail in Supplementary Figure 2.

DISCUSSION

Summary of findings

This is to our knowledge, the first study to examine bi-directional associations of dementia, loneliness, and mortality, while testing both interactions and mediating effects in a large sample of older US adults using the nationally representative Health and Retirement Study. Among 6,468 individuals with mean age of 78.3 years in 2010 followed up until 2020, we found that female older adults had higher loneliness scores, dementia prevalence, and dementia probability than male older adults. Black/Hispanic/Other ethnic groups among older adults had higher scores on loneliness and depressive symptoms, a greater likelihood of dementia, and a significantly higher proportion of fair or poor self-rated health or chronic conditions than their White counterparts. Among key findings, probable dementia was consistently linked to a two-fold increased mortality risk. The presence of and Ln(odds of dementia) were equally strongly related with mortality risk across tertiles of the loneliness score. The study also found that the loneliness z-score was linked to an elevated risk of mortality regardless of age, sex, or race or ethnicity, and that its TE on mortality was partially mediated by Ln(odds of dementia), z-scored, with up to 40% of TE being a pure indirect effect. In contrast, a very small proportion (<5%) of the TE of Ln(odds of dementia), z-scored, on mortality risk was explained by the loneliness z-score.

Previous studies

Dementia and mortality

The mechanisms underlying these connections between poor cognitive performance assessed in mid-[51] to late-life [52] and an increased risk of death are poorly understood. The intricate interactions between genetic and environmental influences throughout life,

which may also affect health and mortality trajectories, are most likely the source of cognitive function. In addition to the social and physical consequences, problems in cognitive function may indicate underlying biological abnormalities or hereditary issues that go beyond neurodegenerative illnesses. Importantly, a meta-analysis of over 60 studies revealed that cognitive impairment, including overt dementia, was linked to a higher risk of death from all causes [53]. In a systematic study, dementia was also linked to higher mortality among COVID-19 patients [54]. It is also well known that the mortality rates of various dementias vary. Particularly, it has been demonstrated that dementias other than Alzheimer's disease have a higher death rate [55]. Further research is needed to disentangle stages of dementia, cognitive performance, and their mechanistic pathways to mortality risk.

Loneliness and mortality

This study's findings contribute to existing research that demonstrated loneliness remains a risk factor for mortality in older adults, irrespective of suboptimal lifestyle-related behaviors, underlying biological aging processes, and chronic health problems often associated with natural aging [22]. Prior research has shown that even in European and Asian countries where people tend to live longer than in the US, loneliness increases the risk of mortality [23], underscoring the need for targeted interventions for this age group [24]. Whereas many reports have found that feelings of loneliness were robust predictors of excess risk of death, others have found additional possible pathways wherein these associations are explained by other psychosocial factors or chronic conditions [19, 25]. A study using HRS data from 2002 to 2008 found that the associations of loneliness and increased mortality risk were influenced by health outcomes such as depressive symptoms, self-rated overall health, and functional capabilities [26]. Because older adults are more likely to experience cognitive and physical decline and other aging-related neurocognitive disorders that are risk conditions for mortality, additional work is needed to understand the interplay between loneliness and dementia [27].

Loneliness and dementia

Studies show that feelings of loneliness may be a precursor to diminished cognitive function and

dementia in older adults [28–30]. A systematic review including ten studies with adult samples of mean ages 60 years and older reported that loneliness was associated with cognitive decline [31]. Although studies' neuropsychological test batteries differ, and broad cognitive domains appear to be affected [32], the causal mechanisms remain unclear. In much of this literature, a wide array of measurements has been used to conceptualize loneliness, including the loss of a spouse or partner, living alone and lack of social connections, and perceptions of feeling alone [33]. Loneliness can affect memory recall, executive function (planning, visual cognition), and underlying neuroendocrine and genetic mechanisms that affect brain chemistry and matter [8, 34, 35]. Our study contributes to current body of work that not only might loneliness affect mortality risk by way of poorer health outcomes [21], but dementia may be a plausible pathway as well. Future research should explore how the progression of cognitive decline and impairment influences the strength of these associations.

Strengths and limitations

Our study has several strengths. First, this is the first study to explore the potential bi-directional mediational effects between loneliness and dementia in determining mortality risk. It is also the first to do so in a nationally representative study of older individuals with up to 14 years of follow-up. We may test a wide range of hypotheses using a wealth of data from the HRS while considering any potential confounding effects of unrelated variables. The current study additionally included cutting-edge methodologies, such as multiple imputation applied to variables, Cox proportional hazards models accounting for the complexity of the sampling design, and four-way decomposition models to simultaneously examine mediation and interaction. Our study does, however, have several important limitations. First, even though date of death was known, the analysis contained some measurement error because the date of birth's precision was only at the month and year levels. Second, a small number of questions that were not verified against other measures, like a full scale with many items, were used to measure loneliness. Third, even though some of these algorithms were based on the comprehensive ADAMS sub-study, measurement error in the algorithmically defined dementia result could not be evaluated. Given the observational nature of the study, selection biases and lingering confounding cannot be ruled out. In fact, we were not

able to adjust for diet quality indices or dietary behavior given the limited data available in the HRS on a sub-sample and with few waves available compared with the core data. In addition, we did not include alcohol use among potential exogenous variables, given the strong suggestion of it being a consequence rather than a cause of loneliness. Finally, we cannot rule out the possibility of exposure-induced mediator-outcome confounding given that many of the assumptions of the four-way decomposition models are not empirically testable (see Supplementary Methods File) [56]. We also acknowledge that, to our knowledge, *med4way* does not account for this potential [56].

Conclusions

In sum, dementia was positively associated with mortality risk, in similar fashion across loneliness score tertiles, while loneliness was associated with mortality risk in the overall sample. The TE of loneliness on mortality risk was partially mediated by dementia probability, particularly when models were only adjusted for age, sex, and race/ethnicity. Future research should further investigate the pathway explaining how loneliness, dementia, and mortality are related to each other. Our findings underscore the potential importance of intervening both on loneliness and dementia to reduce mortality risk.

AUTHOR CONTRIBUTIONS

Michael F. Georgescu (Study concept, plan of analysis, assistance with data management and statistical analysis; literature search and review, write-up of parts of the manuscript, revision of the manuscript); May A. Beydoun (plan of analysis, data management, statistical analysis, literature search and review, write-up of the manuscript, revision of the manuscript); Jason Ashe (Plan of analysis, literature search and review, write-up of parts of the manuscript, revision of the manuscript); Christian A. Maino Vieytes (methodology, literature search and reviews, write-up of parts of the manuscript, revision of the manuscript); Hind A. Beydoun (Literature search and reviews, write-up of parts of the manuscript, revision of the manuscript); Michele K. Evans (Data acquisition, write-up of parts of the manuscript, revision of the manuscript); Alan B. Zonderman (Data acquisition, write-up of parts of the manuscript, revision of the manuscript).

ACKNOWLEDGMENTS

FUNDING

CONFLICT OF INTEREST

The authors have no conflict of interest to report.

DATA AVAILABILITY

While the data is owned by the University of Michigan Ann Arbor and HRS is public use data, this work is owned and funded by the National Institute on Aging at the NIH. Scripts used in this analysis will be made available on a GitHub repository: baydounm/HRS_Loneliness_dementia_mortality (github.com). For additional information please contact the corresponding author by e-mail contact at baydounm@mail.nih.gov.

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/JAD-231359.

REFERENCES

[1] World Health Organization (2019) Risk reduction of cognitive decline and dementia: WHO guidelines.

[2] Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP (2013) The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimers Dement* **9**, 63-75 e62.

[3] Roman GC (2002) Defining dementia: Clinical criteria for the diagnosis of vascular dementia. *Acta Neurol Scand Suppl* **178**, 6-9.

[4] Cahill S (2020) WHO's global action plan on the public health response to dementia: Some challenges and opportunities. *Aging Ment Health* **24**, 197-199.

[5] Tolea MI, Ezzeddine R, Camacho S, Galvin JE (2023) Emerging drugs for dementia with Lewy bodies: A review of Phase II & III trials. *Expert Opin Emerg Drugs* **28**, 167-180.

[6] Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, Brayne C, Burns A, Cohen-Mansfield J, Cooper C, Costafreda SG, Dias A, Fox N, Gitlin LN, Howard R, Kales HC, Kivimaki M, Larson EB, Ogunniyi A, Orgeta V, Ritchie K, Rockwood K, Sampson EL, Samus Q, Schneider LS, Selbaek G, Teri L, Mukadam N (2020) Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* **396**, 413-446.

[7] Donovan NJ, Blazer D (2020) Social isolation and loneliness in older adults: Review and commentary of a national academies report. *Am J Geriatr Psychiatry* **28**, 1233-1244.

[8] Akhter-Khan SC, Tao Q, Ang TFA, Itchapurapu IS, Alosco ML, Mez J, Piers RJ, Steffens DC, Au R, Qiu WQ (2021) Associations of loneliness with risk of Alzheimer's disease dementia in the Framingham Heart Study. *Alzheimers Dement* **17**, 1619-1627.

[9] Zhou Z, Wang P, Fang Y (2018) Loneliness and the risk of dementia among older Chinese adults: Gender differences. *Aging Mental Health* **22**, 519-525.

[10] Holwerda TJ, Deeg DJ, Beekman AT, van Tilburg TG, Stek ML, Jonker C, Schoevers RA (2014) Feelings of loneliness, but not social isolation, predict dementia onset: Results from the Amsterdam Study of the Elderly (AMSTEL). *J Neurol Neurosurg Psychiatry* **85**, 135-142.

[11] Zhu Z, Liao H (2020) Impact of cognitive impairment and systemic vascular comorbidities on risk of all-cause and cardiovascular mortality: National Health and Nutrition Examination Survey 1999 to 2002. *Int J Cardiol* **300**, 255-261.

[12] Holm H, Bachus E, Jujic A, Nilsson ED, Wadstrom B, Molvin J, Minthon L, Fedorowski A, Nagga K, Magnusson M (2020) Cognitive test results are associated with mortality and rehospitalization in heart failure: Swedish prospective cohort study. *ESC Heart Fail* **7**, 2948-2955.

[13] Paul E, Bu F, Fancourt D (2021) Loneliness and risk for cardiovascular disease: Mechanisms and future directions. *Curr Cardiol Rep* **23**, 68.

[14] Vigorito C, Giallauria F (2018) Loneliness, social isolation and risk of cardiovascular disease in the English Longitudinal Study of Ageing. *Eur J Prev Cardiol* **25**, 1384-1386.

[15] Domenech-Abella J, Mundo J, Haro JM, Rubio-Valera M (2019) Anxiety, depression, loneliness and social network in the elderly: Longitudinal associations from The Irish Longitudinal Study on Ageing (TILDA). *J Affect Disord* **246**, 82-88.

[16] Hawkley LC, Cacioppo JT (2010) Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Ann Behav Med* **40**, 218-227.

[17] Prince JD, Oyo A, Mora O, Wyka K, Schonebaum AD (2018) Loneliness among persons with severe mental illness. *J Nerv Mental Dis* **206**, 136-141.

[18] Parmar M, Ma RM, Attygalle S, Mueller C, Stubbs B, Stewart R, Perera G (2022) Associations between loneliness and acute hospitalisation outcomes among patients receiving mental healthcare in South London: A retrospective cohort study. *Soc Psychiatry Psychiatr Epidemiol* **57**, 397-410.

[19] Luo Y, Waite LJ (2014) Loneliness and mortality among older adults in China. *J Gerontol B Psychol Sci Soc Sci* **69**, 633-645.

[20] Teguo MT, Simo-Tabue N, Stoykova R, Meillon C, Cogne M, Amieva H, Dartigues JF (2016) Feelings of loneliness and living alone as predictors of mortality in the elderly: The PAQUID Study. *Psychosom Med* **78**, 904-909.

[21] Park C, Majeed A, Gill H, Tamura J, Ho RC, Mansur RB, Nasri F, Lee Y, Rosenblat JD, Wong E, McIntyre RS (2020) The effect of loneliness on distinct health outcomes: A comprehensive review and meta-analysis. *Psychiatry Res* **294**, 113514.

[22] Schutter N, Holwerda TJ, Comijs HC, Stek ML, Peen J, Dekker JJM (2022) Loneliness, social network size and mortality in older adults: A meta-analysis. *Eur J Ageing* **19**, 1057-1076.

[23] Yu B, Steptoe A, Chen LJ, Chen YH, Lin CH, Ku PW (2020) Social isolation, loneliness, and all-cause mortality in patients with cardiovascular disease: A 10-year follow-up study. *Psychosom Med* **82**, 208-214.

[24] Poganik JR, Gladyshev VN (2023) We need to shift the focus of aging research to aging itself. *Proc Natl Acad Sci U S A* **120**, e2307449120.

[25] Beller J (2023) Loneliness and mortality: The moderating effect of positive affect. *Appl Psychol Health Well Being* **15**, 49-65.

[26] Luo Y, Hawkley LC, Waite LJ, Cacioppo JT (2012) Loneliness, health, and mortality in old age: A national longitudinal study. *Soc Sci Med* **74**, 907-914.

[27] Tilvis RS, Kähönen-Väre MH, Jolkkonen J, Valvanne J, Pitkala KH, Strandberg TE (2004) Predictors of cognitive decline and mortality of aged people over a 10-year period. *J Gerontol A Biol Sci Med Sci* **59**, 268-274.

[28] Qiao L, Wang G, Tang Z, Zhou S, Min J, Yin M, Li M (2022) Association between loneliness and dementia risk: A systematic review and meta-analysis of cohort studies. *Front Hum Neurosci* **16**, 899814.

[29] Cardona M, Andrés P (2023) Are social isolation and loneliness associated with cognitive decline in ageing? *Front Aging Neurosci* **15**, 1075563.

[30] Freak-Poli R, Wagemaker N, Wang R, Lysen TS, Ikram MA, Vernooij MW, Dintica CS, Vernooij-Dassen M, Melis RJF, Laukka EJ, Fratiglioni L, Xu W, Tiemeier H (2022) Loneliness, not social support, is associated with cognitive decline and dementia across two longitudinal population-based cohorts. *J Alzheimers Dis* **85**, 295-308.

[31] Boss L, Kang DH, Branson S (2015) Loneliness and cognitive function in the older adult: A systematic review. *Int Psychogeriatr* **27**, 541-553.

[32] Lara E, Caballero FF, Rico-Uribe LA, Olaya B, Haro JM, Ayuso-Mateos JL, Miret M (2019) Are loneliness and social isolation associated with cognitive decline? *Int J Geriatr Psychiatry* **34**, 1613-1622.

[33] Duan D, Dong Y, Zhang H, Zhao Y, Diao Y, Cui Y, Wang J, Chai Q, Liu Z (2017) Empty-nest-related psychological distress is associated with progression of brain white matter lesions and cognitive impairment in the elderly. *Sci Rep* **7**, 43816.

[34] Luchetti M, Terracciano A, Aschwanden D, Lee JH, Stephan Y, Sutin AR (2020) Loneliness is associated with risk of cognitive impairment in the Survey of Health, Ageing and Retirement in Europe. *Int J Geriatr Psychiatry* **35**, 794-801.

[35] Pecchinenda A, Yankouskaya A, Boccia M, Piccardi L, Guariglia C, Giannini AM (2024) Exploring the relationship between perceived loneliness and subjective cognitive decline in older individuals. *Aging Ment Health* **28**, 73-82.

[36] McGrath R, Vincent BM, Hackney KJ, Robinson-Lane SG, Downer B, Clark BC (2020) The longitudinal associations of handgrip strength and cognitive function in aging Americans. *J Am Med Dir Assoc* **21**, 634-639.e1.

[37] Hunter JC, Handing EP, Casanova R, Kuchibhatla M, Lutz MW, Saldana S, Plassman BL, Hayden KM (2018) Neighborhoods, sleep quality, and cognitive decline: Does where you live and how well you sleep matter? *Alzheimers Dement* **14**, 454-461.

[38] Weir DR (2016) Validating mortality ascertainment in the Health and Retirement Study. https://hrs.isr.umich.edu/publications/biblio/9022

[39] Mazza GL, Enders CK, Ruehlman LS (2015) Addressing item-level missing data: A comparison of proration and full information maximum likelihood estimation. *Multivariate Behav Res* **50**, 504-519.

[40] Beydoun MA, Beydoun HA, Fanelli-Kuczmarski MT, Weiss J, Georgescu MF, Meirelles O, Lyall DM, Evans MK, Zonderman AB (2023) Pathways explaining racial/ethnic and socio-economic disparities in dementia incidence: The UK Biobank study. *Aging (Albany NY)* **15**, 9310-9340.

[41] STATA (2019) Statistics/Data Analysis: Release 16.0. Stata Corporation, Texas.

[42] NCHS (2010) Office of Analysis and Epidemiology, Public-use Third National Health and Nutrition Examination Survey Linked Mortality File. http://www.cdc.gov/nchs/data_access/data_linkage/mortality/nhanes3_linkage.htm

[43] Lee KJ, Carlin JB (2010) Multiple imputation for missing data: Fully conditional specification versus multivariate normal imputation. *Am J Epidemiol* **171**, 624-632.

[44] Van Buuren S (2007) Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* **16**, 219-242.

[45] Discacciati A, Bellavia A, Lee JJ, Mazumdar M, Valeri L (2019) Med4way: A Stata command to investigate mediating and interactive mechanisms using the four-way effect decomposition. Oxford University Press.

[46] Selvin S (2004) *Statistical analysis of epidemiologic data*, Oxford University Press.

[47] Hochberg Y, Tamhane AC (1987) *Multiple comparison procedures*, Wiley, New York.

[48] Beydoun MA, Canas JA, Dore GA, Beydoun HA, Rostant OS, Fanelli-Kuczmarski MT, Evans MK, Zonderman AB (2016) Serum uric acid and its association with longitudinal cognitive change among urban adults. *J Alzheimers Dis* **52**, 1415-1430.

[49] Beydoun MA, Hossain S, Fanelli-Kuczmarski MT, Beydoun HA, Canas JA, Evans MK, Zonderman AB (2018) Vitamin D status and intakes and their association with cognitive trajectory in a longitudinal study of urban adults. *J Clin Endocrinol Metab* **103**, 1654-1668.

[50] Beydoun MA, Noren Hooten N, Weiss J, Beydoun HA, Georgescu M, Freeman DW, Evans MK, Zonderman AB (2023) GDF15 and its association with cognitive performance over time in a longitudinal study of middle-aged urban adults. *Brain Behav Immun* **108**, 340-349.

[51] Sabia S, Guéguen A, Marmot MG, Shipley MJ, Ankri J, Singh-Manoux A (2010) Does cognition predict mortality in midlife? Results from the Whitehall II cohort study. *Neurobiol Aging* **31**, 688-695.

[52] Bosworth HB, Siegler IC (2002) Terminal change in cognitive function: An updated review of longitudinal studies. *Exp Aging Res* **28**, 299-315.

[53] Dewey ME, Saz P (2001) Dementia, cognitive impairment and mortality in persons aged 65 and over living in the community: A systematic review of the literature. *Int J Geriatr Psychiatry* **16**, 751-761.

[54] July J, Pranata R (2021) Prevalence of dementia and its impact on mortality in patients with coronavirus disease 2019: A systematic review and meta-analysis. *Geriatr Gerontol Int* **21**, 172-177.

[55] Liang CS, Li DJ, Yang FC, Tseng PT, Carvalho AF, Stubbs B, Thompson T, Mueller C, Shin JI, Radua J, Stewart R, Rajji TK, Tu YK, Chen TY, Yeh TC, Tsai CK, Yu CL, Pan CC, Chu CS (2021) Mortality rates in Alzheimer's disease and non-Alzheimer's dementias: A systematic review and meta-analysis. *Lancet Healthy Longev* **2**, e479-e488.

[56] Vanderweele TJ, Vansteelandt S, Robins JM (2014) Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology* **25**, 300-306.

ONLINE SUPPLEMENTARY MATERIALS

Loneliness, dementia status and their association with all-cause mortality among older US adults

Georgescu et al.

SUPPLEMENTARY METHODS 1: ALGORITHMICALLY DEFINED DEMENTIA OUTCOMES

The algorithms estimate a predicted dementia probability using different combinations of sociodemographic characteristics, health and physical functioning variables, social engagement indicators, and cognitive indicators (i.e. cognition test item scores and proxy-reports of cognition) to classify dementia status using race/ethnicity-specific probability thresholds. Each algorithm was designed to minimize differences in prediction accuracy across race/ethnicity groups, with pairwise differences of 3 percentage points for sensitivity and 5 percentage points for specificity, and is thus suitable for use in race/ethnicity disparities research.

This data file (hrsdementia_20211109.sas7bdat) was created using the 2018 RAND V1 HRS longitudinal file ("randhrs1992_2018v1") and core HRS data; code for reproducing this dataset is available in the following Github repository, and is dated 2021_1109.

Note the small differences in the probabilities and classifications for all years in this dataset compared to the previously distributed dataset (hrsdementia_20191028.sas7bdat) due to differences in the source data. This previously distributed dataset covered 2000 to 2014, and was created using the 2014 RAND HRS longitudinal V2 file ("randhrs1992_2014v2") and core HRS data; code for reproducing this prior version of the dataset is available in the following Github repository and is dated 2019_0529: https://github.com/powerepilab/AD_algorithm_development.

SUPPLEMENTARY METHODS 2: LONELINESS SCALE AND SCORING

2006:

How often do you feel you lack companionship?

How often do you feel left out?

How often do you feel isolated from others?

2008:

You lack companionship?

Left out?

Isolated from others?

STATA code:

*****LONELINESS DATA, 2006 & 2008: MERGE AND CREATE THE LONELINESS VARIABLE**

2006

*KLB020A "How often do you feel you lack companionship?"

*KLB020B "How often do you feel left out?"

*KLB020C "How often do you feel isolated from others?"

2008

*LLB020A "You lack companionship?"

*LLB020B "Left out?"

*LLB020C "Isolated from others?"

*LLB020D "That you are "in tune" with the people around you?"

*LLB020E "Alone?"

*LLB020F "That there are people you can talk to?"

*LLB020G "That there are people you can turn to?"

*LLB020H "That there are people who really understand you?"

*LLB020I "That there are people you feel close to?"

*LLB020J "Part of a group of friends?"

*LLB020K "That you have a lot in common with the people around you?"

```
use H06LB_R,clear
destring HHID, replace
destring PN, replace


capture drop HHIDPN
egen HHIDPN = concat(HHID PN)


destring HHIDPN, replace
sort HHIDPN


save H06LB_Rfin, replace



use H08LB_R,clear
destring HHID, replace
destring PN, replace


capture drop HHIDPN
egen HHIDPN = concat(HHID PN)


destring HHIDPN, replace
sort HHIDPN


save H08LB_Rfin, replace
```

```
use H06LB_Rfin,clear

merge HHIDPN using H08LB_Rfin


save H06LB_R_H08LB_R, replace


keep HHIDPN KLB020A KLB020B KLB020C LLB020A LLB020B LLB020C

save Loneliness_data2006_2008, replace


use Loneliness_data2006_2008,clear


tab1 KLB020A KLB020B KLB020C LLB020A LLB020B LLB020C


save Loneliness_data2006_2008, replace



*****************************Create a loneliness
variable*********************************************************
 use Loneliness_data2006_2008, clear



**Your lack of companionship? 1: Often, 2: Some of the time, 3: Hardly ever or never**

*LLB020A & KLB020A merged*


**2006**

tab KLB020A, missing

capture drop lackcomp_2006

gen lackcomp_2006 = KLB020A


replace lackcomp_2006=1 if KLB020A==3
```

replace lackcomp_2006=2 if KLB020A==2

replace lackcomp_2006=3 if KLB020A==1

tab1 lackcomp_2006

2008

tab LLB020A, missing

capture drop lackcomp_2008

gen lackcomp_2008 = LLB020A

replace lackcomp_2008=1 if LLB020A==3

replace lackcomp_2008=2 if LLB020A==2

replace lackcomp_2008=3 if LLB020A==1

tab1 lackcomp_2008

Left out? 1: Often, 2: Some of the time, 3: Hardly ever or never

LLB020B & KLB020B merged

2006

tab KLB020B, missing

capture drop leftout_2006

gen leftout_2006 = KLB020B

replace leftout_2006=1 if KLB020B==3

replace leftout_2006=2 if KLB020B==2

replace leftout_2006=3 if KLB020B==1

tab1 leftout_2006

2008

tab LLB020B, missing

capture drop leftout_2008

gen leftout_2008 = LLB020B

replace leftout_2008=1 if LLB020B==3

replace leftout_2008=2 if LLB020B==2

replace leftout_2008=3 if LLB020B==1

tab1 leftout_2008

Isolated from others? 1: Often, 2: Some of the time, 3: Hardly ever or never

LLB020C & KLB020C merged

2006

tab LLB020C, missing

capture drop Iso_2006

gen Iso_2006 = KLB020C

replace Iso_2006=1 if KLB020C==3

replace Iso_2006=2 if KLB020C==2

replace Iso_2006=3 if KLB020C==1

tab1 Iso_2006

```
**2008**
tab LLB020C, missing
capture drop Iso_2008
gen Iso_2008 = LLB020C


replace Iso_2008=1 if LLB020C==3
replace Iso_2008=2 if LLB020C==2
replace Iso_2008=3 if LLB020C==1


tab1 Iso_2008



capture rename HHIDPN, lower



*********Combined 2006-2008 items********************
capture drop lackcomp_2006_2008
gen lackcomp_2006_2008=.
replace lackcomp_2006_2008=lackcomp_2006 if lackcomp_2006~=.
replace lackcomp_2006_2008=lackcomp_2008 if lackcomp_2008~=.



capture drop leftout_2006_2008
gen leftout_2006_2008=.
replace leftout_2006_2008=leftout_2006 if leftout_2006~=.
replace leftout_2006_2008=leftout_2008 if leftout_2008~=.


capture drop Iso_2006_2008
gen Iso_2006_2008=.
replace Iso_2006_2008=Iso_2006 if Iso_2006~=.
```

replace Iso_2006_2008=Iso_2008 if Iso_2008~=.

total loneliness score, scaled from 1 - 3

capture drop loneliness_tot_miss

egen loneliness_tot_miss=rowmiss(lackcomp_2006_2008 leftout_2006_2008 Iso_2006_2008)

tab loneliness_tot_miss

capture drop loneliness_tot_rmean

egen loneliness_tot_rmean=rmean(lackcomp_2006_2008 leftout_2006_2008 Iso_2006_2008) if loneliness_tot_miss<1

capture drop loneliness_tot

gen loneliness_tot=loneliness_tot_rmean*3-3

su loneliness_tot, detail

histogram loneliness_tot

graph save loneliness_tot.gph,replace

sort hhidpn

save, replace

e-Table S1. Ln(odds of dementia), z-scored and all-cause mortality: four-way decomposition models by Loneliness z-score, overall, by sex and by race

Y=All-cause mortality; M=Loneliness z-score	TE β±SE	P	CDE β±SE	P	INTREF β±SE	P	INTMED β±SE	P	PIE β±SE	P
Overall										
Reduced Model 1A-C (N=6,468)										
X=Hurd	+0.411±0.020	<0.001	+0.389±0.020	<0.001	0.002±0.002	0.314	0.005±0.002	0.028	0.015±0.003	<0.001
X=Expert	+0.709±0.037	<0.001	+0.687±0.037	<0.001	0.001±0.002	0.777	0.006±0.005	0.195	0.016±0.004	<0.001
X=LASSO	+0.665±0.033	<0.001	+0.646±0.033	<0.001	0.000±0.002	0.998	0.004±0.004	0.269	0.015±0.003	<0.001
Full model 2A-C (N=6,079)										
X=Hurd	+0.514±0.058	<0.001	+0.515±0.058	<0.001	0.000±0.001	0.979	-0.002±0.004	0.578	0.001±0.002	0.439
X=Expert	+0.376±0.038	<0.001	+0.376±0.038	<0.001	-0.000±0.000	0.961	-0.001±0.002	0.560	0.001±0.001	0.392
X=LASSO	+0.415±0.040	<0.001	+0.415±0.040	<0.001	-0.000±0.000	0.966	-0.001±0.002	0.576	0.001±0.001	0.466
Male										
Reduced Model 1A-C (N=2,778)										
X=Hurd	+0.487±0.037	<0.001	+0.467±0.037	<0.001	-0.002±0.001	0.002	0.003±0.005	0.573	+0.020±0.005	<0.001
X=Expert	+0.768±0.062	<0.001	+0.732±0.061	<0.001	-0.003±0.002	0.246	0.020±0.009	0.026	+0.018±0.006	0.002
X=LASSO	+0.708±0.054	<0.001	+0.683±0.054	<0.001	-0.002±0.001	0.062	0.010±0.007	0.171	+0.018±0.005	0.001
Full model 2A-C (N=2,578)										
X=Hurd	+0.607±0.093	<0.001	+0.608±0.094	<0.001	-0.004±0.005	0.391	0.003±0.004	0.507	+0.001±0.002	0.663
X=Expert	+0.395±0.057	<0.001	+0.395±0.057	<0.001	-0.002±0.004	0.552	0.001±0.002	0.606	+0.001±0.001	0.648
X=LASSO	+0.442±0.060	<0.001	+0.442±0.060	<0.001	-0.002±0.004	0.548	0.001±0.002	0.608	+0.001±0.001	0.652
Female										
Reduced Model 1A-C (N=3,690)										
X=Hurd	+0.384±0.025	<0.001	0.362±0.024	<0.001	+0.005±0.003	0.170	0.005±0.003	0.067	+0.012±0.003	0.001
X=Expert	+0.690±0.047	<0.001	0.678±0.047	<0.001	-0.001±0.004	0.799	0.000±0.006	0.948	+0.013±0.005	0.005
X=LASSO	+0.647±0.042	<0.001	0.632±0.042	<0.001	0.000±0.003	0.976	0.002±0.005	0.672	+0.013±0.004	0.003
Full model 2A-C (N=3,501)										
X=Hurd	+0.473±0.075	<0.001	0.480±0.076	<0.001	-0.003±0.003	0.234	-0.006±0.006	0.335	+0.001±0.003	0.715

	Est±SE	p	Est±SE	p	Est±SE	p	Est±SE	p	Est±SE	p
X=Expert	+0.381±0.051	<0.001	+0.385±0.052	<0.001	-0.002±0.002	0.266	-0.003±0.003	0.331	+0.001±0.002	0.626
X=LASSO	+0.415±0.054	<0.001	+0.419±0.055	<0.001	-0.002±0.002	0.364	-0.003±0.004	0.419	+0.001±0.003	0.732
White										
Reduced Model 1A-C (N=5,289)										
X=Hurd	+0.418±0.022	0.000	+0.395±0.022	<0.001	+0.002±0.002	0.299	+0.007±0.003	0.008	+0.015±0.003	0.000
X=Expert	+0.726±0.041	0.000	+0.703±0.040	<0.001	-0.000±0.001	0.741	+0.006±0.005	0.272	+0.017±0.004	0.000
X=LASSO	+0.665±0.035	0.000	+0.646±0.035	<0.001	-0.001±0.001	0.519	+0.004±0.004	0.354	+0.016±0.004	0.000
Full model 2A-C (N=4,987)										
X=Hurd	+0.538±0.064	0.000	+0.541±0.064	<0.001	+0.004±0.004	0.277	-0.007±0.005	0.135	+0.000±0.002	0.973
X=Expert	+0.404±0.042	0.000	+0.405±0.042	<0.001	+0.003±0.002	0.261	-0.004±0.002	0.134	+0.000±0.002	0.925
X=LASSO	+0.435±0.044	0.000	+0.437±0.044	<0.001	+0.003±0.003	0.226	-0.005±0.003	0.093	-0.000±0.002	0.988
Black/Hispanic/Other ethnic groups										
Reduced Model 1A-C (N=1,179)										
X=Hurd	+0.409±0.057	0.000	+0.408±0.063	<0.001	-0.009±0.009	0.338	-0.006±0.008	0.448	+0.016±0.007	0.035
X=Expert	+0.630±0.088	0.000	+0.605±0.089	<0.001	+0.007±0.011	0.552	+0.007±0.009	0.462	+0.012±0.008	0.135
X=LASSO	+0.667±0.091	0.000	+0.646±0.093	<0.001	+0.004±0.010	0.693	+0.005±0.009	0.589	+0.012±0.008	0.110
Full model 2A-C (N=1,092)										
X=Hurd	+0.436±0.139	0.002	+0.398±0.137	0.004	+0.025±0.020	0.206	+0.011±0.010	0.298	+0.003±0.005	0.564
X=Expert	+0.294±0.088	0.001	+0.277±0.088	0.002	+0.017±0.014	0.222	+0.000±0.004	0.893	+0.000±0.002	0.894
X=LASSO	+0.333±0.099	0.001	+0.301±0.098	0.002	+0.026±0.016	0.098	+0.005±0.007	0.402	+0.001±0.003	0.592

Abbreviations CDE=Controlled Direct Effect; HRS=Health and Retirement Study; INTMED=Mediated Interaction; INTREF= Reference Interaction; LASSO= Least Absolute Shrinkage and Selection Operator; PIE=Pure Indirect Effect; SE=Standard Error.

[a] Cox PH regression models with mortality as the main outcome and Ln(odds of dementia), z-scored, using Hurd, expert and LASSO algorithms as the main alternate exposures. 1 SD for each Ln transformed odds of dementia was ~3.8, 3.0, and 2.3, for Hurd, Expert and LASSO algorithms, respectively. Loneliness score (z-scored, 1SD~=1.56 unit increase) was a potential mediators/moderators allowed to interact with the main exposure, sample size N=6,079 (full model) and 6,468 (reduced model), four-way decomposition analysis.

[b] Exogenous variables are the ones included in Table 2, Models 1A-2D, as covariates for the reduced and full models, respectively. See Covariates section for detail.

Supplementary Methods File

Contents

    ** Please note the text from this file is taken and only slightly modified from the supplementary file our team created for doi: 10.1016/j.isci.2023.108526.[1]

1 Dementia Probabilities and Scoring Methods

Ascertainment of dementia status is a costly and time-consuming diagnostic venture that makes it impractical to implement in survey and longitudinal studies.[2] This phenomenon has long been recognized and is the basis of the development of several indices and algorithms for dementia status probability and classification by incorporating available information on cognitive status, behavioral, and sociodemographic characteristics. In our analysis we implemented three established algorithms for computing predicted probabilities of dementia:

- The Modified Hurd Algorithm[3]

- The Expert Algorithm[4]

- The LASSO Algorithm[4]

We now describe these methods in detail.

The Hurd Algorithm

1.0.1 *Dementia Probability Probit Models*

The Hurd algorithm relies on an ordered probit model where the response is a three-category factor (*dementia*, *cognitive impairment but not dementia* (*CIND*), and *aging normally*–i.e., $y_i \in j = 1, 2, 3$).[5]. The model was trained on a subset of HRS subjects from the Aging, Demographics, and Memory Study (ADAMS) who underwent an in-home clinical examination for dementia.[3] The probit model used cognitive status data from a modified version of the Telephone Interview for Cognitive Status (TICS) survey that is widely used in epidemiological studies for cognitive testing.[6] TICS data came from self-respondents while study subjects who had their answers provided by a proxy respondent used the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE), a 16-item questionnaire that assesses the subject's memory and functional independence.[7] Notably, in an effort to eschew the issue of reverse causality and confounding, the model did not use any variables that may have been affected by the onset of dementia. Instead, the training model included sociodemographic variables, variables summarizing activity of Daily Living (ADL), those summarizing Instrumental Activity of Daily Living (IADL) limitations, and measures of cognitive functioning captured in HRS surveys. Given the different surveys on cognitive functioning completed by self-respondents and proxy respondents, there were separate probability models (where the probability of the ith subject was denoted, Y_i) fit for each of these two study groups (Equations 1.0.0 and 1.0.1, respectively).

$$y_i^* = +\mathbf{x}_i^t\boldsymbol{\beta}_1 + \mathbf{TICS}_i^t\boldsymbol{\beta}_2 + \Delta\mathbf{TICS}_i^t\boldsymbol{\beta}_3 + \epsilon_i \qquad (1.0.0)$$

where $\mathbf{x}_i^t\boldsymbol{\beta}_1$ was a vector of covariates (listed in Table 1) multiplied by a vector of coefficients in $\boldsymbol{\beta}_1$, $\mathbf{TICS}_i^t\boldsymbol{\beta}_2$ was the vector of responses to the TICS items used in the model multiplied by another vector of coefficients, and $\Delta\mathbf{TICS}_i^t\boldsymbol{\beta}_3$ was a vector comprised of the changes in the responses to the TICS items, between the previous two HRS waves, multiplied by a third coefficients vector.

Table S1. Variables in **x**.
Age (<75, 75-79, 80-84, 85-89, ≥ 90)
Sex
Education level
ADL Limitations
IADL Limitations

Table S2. Variables in **TICS**.
Date Recall
Backward counting 20
Serial 7
Scissor recall
Cactus recall
Current President
Immediate word recall
Delayed word recall

Finally, the model included a residual, ϵ_i, modeled from a normal distribution ($e_i \sim N(0, 1)$). The model for the proxy respondents was similar:

$$y_i* = \mathbf{x}_i^t\boldsymbol{\beta}_1 + \mathbf{IQCODE}_i^t\boldsymbol{\beta}_2 + \Delta\mathbf{IQCODE}_i^t\boldsymbol{\beta}_3 + \mathbf{pTICS}\boldsymbol{\beta}_4 + PP\beta_5 + \epsilon_i \quad (1.0.1)$$

where $\mathbf{IQCODE}_i$ is the score from the IQCODE survey in the previous HRS wave, $\Delta\mathbf{IQCODE}_i$ is the change in the IQCODE score between the two previous HRS waves, $\mathbf{pTICS}$ is the vector of items from the TICS assessment occurring two waves prior to the ADAMS assessment (for subjects who were not proxy at that wave), and PP is a binary indicator variable indicating whether the subject was proxy two waves prior to the ADAMS assessment.

In both equations, the linear predictor is used to estimate y_i*, a latent (and therefore unobserved) outcome. A note: the missing indicator method was used for addressing the missingness patterns across self-respondents and proxy respondents and for substituting missing values in the cognitive variables, limitations in ADLs, and limitations in IADLS.[8]

$$y_i = \begin{cases} dementia, & \text{if } y* \leq c_1 \\ CIND, & \text{if } c_1 \leq y* \leq c_2 \\ aging\ normally, & \text{if } y* > c_2 \end{cases}$$

We do some algebraic manipulation to arrive at a closed-form expression (under the assumption of the normally distributed error residual) for computing probabilities of observing each of the $J = 3$ levels of the ordinal outcome variable:

For $j = 1$:
$$Pr[y_i = 1] = P[y_i^* \leq c_1]$$
$$= Pr[\boldsymbol{x}_i\boldsymbol{\beta} + \epsilon_i \leq c_1]$$
$$= Pr[\epsilon_i \leq c_1 - \boldsymbol{x}_i\boldsymbol{\beta}]$$
$$= \Phi(\epsilon_i \leq c_1 - \boldsymbol{x}_i\boldsymbol{\beta})$$

For $j = 2$:
$$Pr[y_i = 2] = P[c_1 \leq y* \leq c_2]$$
$$= Pr[c_1 \leq \boldsymbol{x}_i\boldsymbol{\beta} + \epsilon_i \leq c_2]$$
$$= Pr[c_1 - \boldsymbol{x}_i\boldsymbol{\beta} \leq \epsilon_i \leq c_2 - \boldsymbol{x}_i\boldsymbol{\beta}]$$
$$= \Phi(c_2 - \boldsymbol{x}_i\boldsymbol{\beta}) - \Phi(c_1 - \boldsymbol{x}_i\boldsymbol{\beta})$$

For $j = 3$:
$$Pr[y_i = 3] = P[y_i^* > c_2]$$
$$= Pr[\boldsymbol{x}_i\boldsymbol{\beta} + \epsilon_i > c_2]$$
$$= Pr[\epsilon_i > c_2 - \boldsymbol{x}_i\boldsymbol{\beta}]$$
$$= 1 - \Phi(\epsilon_i \leq c_2 - \boldsymbol{x}_i\boldsymbol{\beta})$$

where Φ is the cumulative distribution function(CDF) for the standard normal distribution and we use the short-hand $\boldsymbol{x_i\beta}$ to indicate the fixed part of the models in 1.0 and 1.1 The model itself, and therefore the estimation of y_i^* proceeds with maximum likelihood estimation and we rely on software for these computations.

1.0.2   *Classification*

Using the models in 1.0.0 and 1.0.1, the predicted probabilities are computed as outlined above in **1.0.1**. Classification of dementia status into one of the three categories indicated above proceeds via the "maximum likelihood" approach whereby the category with the highest probability for a given subject is the category assigned.

For more details and data on sensitivity and specificity of the classification method, we refer the readers to the original publication by Hurd et al.[3]

1.1   *The Expert Algorithm*

1.1.1   *Probability Model*

The Expert algorithm is similar to the Hurd model in that it is a regression-based algorithm for classifying subjects into dementia categories. Rather than implementing a probit model, the Expert algorithm specifies a logit model with a set of covariates chosen by two investigators with "expert" knowledge on risk factors for dementia.[4] The model was again trained on ADAMS participants and modeled the probability of dementia vs no dementia as a function of covariates and a number of interaction terms. The model is provided in Equation 1.1.0.

$$logit[Pr(Y_i = dementia)] = +\mathbf{x}_i^t\boldsymbol{\beta}_1 + \mathbf{pCOG}_i^t\boldsymbol{\beta}_2 + \mathbf{sCOG}_i^t\boldsymbol{\beta}_3 + \mathbf{int}_i^t\boldsymbol{\beta}_4 \quad (1.1.0)$$

Table S3. Variables in **x**.

Age
Sex
Race
Ethnicity
Education level
Respondent stats (Proxy, Self)
ADL Limitations
IADL Limitations
Self-reported health status
Diabetes history
Partnered/married
Time spent volunteering in last year
Time spent helping others in last year

Table S4. Variables in **COG** vectors.

sCOG (self-respondent cognition variables)
Date Recall
Backward counting 20
Serial 7
Scissor recall
Cactus recall
Current President
Current Vice President
Immediate word recall
Delayed word recall

pCOG (proxy-respondent cognition variables)
Proxy-rated memory
Proxy assessment of Jorm symptoms
IQCODE average score

Table S5. Interaction terms in **int**.[4]

Race x Education
Ethnicity x Education
Race x Serial 7
Ethnicity x Serial 7
Race x Cactus recall
Ethnicity x Cactus recall
Race x Scissor recall
Ethnicity x Scissor recall
Race x President recall
Ethnicity x President recall
Race x Vice President recall
Ethnicity x Vice President recall
Respondent status x Age
Respondent status x Education
Respondent status x Sex
Respondent status x ADL and IADL
Respondent status x Self-reported health status
Respondent status x Age
Date recalls x ADL
Object, President, Vice President recall x ADLs Age x IADL
Serial 7 x Self-reported health status
Serial 7 x Education
Word recall x Self-reported health status

1.1.2 *Classification*

Assignment of dementia status proceeded by adopting a a range of probability thresholds and using 10-fold cross validation to estimate sensitivity, specificity, and accuracy to gauge whether a global threshold or race-specific thresholds should be adopted.[4] The results of their analysis demonstrated that race-specific thresholds of 0.19, 0.25, and 0.27 (for non-Hispanic White, non-Hispanic Black, and Hispanics, respectively) maximized those estimates.

1.2 *The LASSO Algorithm*

The LASSO algorithm relies on the specification of a logistic regression with the LASSO penalization. Equation 1.2.0 demonstrates the formulation of the model. LASSO and penalized regression, in general, comprise a class of regression models whose foremost goal is coefficient shrinkage, as an approach to model selection. The LASSO penalty term is introduced for logistic regression case is introduced into the likelihood function used to fit the model (Equation 1.2.0)

$$-2[\ell(\beta)] = -2\sum_{i=1}^{n} log(Pr(Y_i|X_i, \beta) + \lambda||\beta||_1]) \tag{1.2.0}$$

The -2log-likelihood includes the -2log-likelihood function for a logistic regression model plus an additional penalty term, $\lambda||\beta||_1]$, which is the product of the L1 norm of the coefficients and a tuning parameter, λ, that is unknown and conventionally optimized via K-fold cross-validation. The goal then becomes to minimize the the objective function in Equation 1.2.0 to find the optimal values for the coefficients, $\boldsymbol{\beta}$, and the tuning parameter, λ.

The variables entered into the LASSO model, to model the probability of dementia status, are listed in table S6.

Table S6. Variables entered into the LASSO model.[4]
Sex
Hispanic
Less Than High School (LTHS)
LTHS x non-Hispanic Black
LTHS x Hispanic
Self-reported health status
At least 1 ADL
At least 1 IADL
Age x IADL
Post-diagnosis of diabetes
Social engagement
Proxy respondent status
Serial 7 score
Serial 7 score x non-Hispanic Black
Serial 7 score x Hispanic
Serial 7 score x LTHS
Serial 7 score x self-reported health status
Immediate word recall score
Delayed word recall score
Serial 7 score x non-Hispanic Black
Immediate word recall x self-reported health status
Delayed word recall x self-reported health status
Date recall
Date recall x at least 1 IADL
Name recall
Name recall x non-Hispanic Black
Name recall x Hispanic
Name recall x at least 1 IADL
IQCODE (centered at 5)
Proxy memory score(centered at 5)
Jorm symptoms
Proxy status x age
Proxy status x LTHS
Proxy status x self-reported health status
Proxy status x male
Proxy status x at least 1 ADL
Proxy status x at least 1 IADL

Classification of dementia status then proceeded as described above in **1.1.2**.

2 Four-Way Decomposition Models

2.1 *Background*

Let Y denote the response, A the exposure (in this loneliness–continuous), a its realized value (the loneliness index score), M the mediator (the dementia probabilities/scores–transformed using $\log[\frac{Pr}{1-Pr}]$), m its realized value (obtained using the Hurd, expert, or LASSO algorithms described above), and $\boldsymbol{c}$ the vector of confounders. Under the assumption of no confounding, namely:

$$\text{i. } Y_{am} \coprod A|C$$

$$\text{ii. } Y_{am} \coprod M|(A,C)$$

$$\text{iii. } M_a \coprod A|C$$

$$\text{iv. } Y_{am} \coprod A * |C$$

that the effect of A on Y is unconfounded conditional on C (i), that the effect of M on Y is unconfounded conditional on (A,C), (ii), the effect of A on M is unconfounded conditional on C (iii), and that any mediator outcome confounders are not affected by the exposure (iv), we can partition the sources of the total effect of the exposure, A, on the outcome, Y, into four components (Equation 2.1.0).[9]

$$TE = CDE + INT_{ref} + INT_{med} + PIE \qquad (2.1.0)$$

The total effect of the exposure, A, on the outcome, Y, is a sum of several components (i.e., sources of variation) that include a *controlled direct effect* (CDE–i.e., the effect of A on Y not due to any interaction or mediation), *the reference interaction* (INT_{ref}–i.e., the effect of interaction only), *the mediated interaction* (INT_{med}–i.e., the effect of interaction and mediation), and *the purely indirect effect* (PIE–i.e., the indirect effect only). Descriptions of the interpretations of these components are provided below[9]:

- *controlled direct effect*: the exposure's effect on the outcome in the absence of the mediator.

- *the reference interaction*: the exposure affects the outcome in the presence of the mediator and the presence of the exposure is not required for the presence of the mediator.

- *the mediated interaction*: the exposure affects the outcome in the presence of the mediator and the presence of the exposure is required for the presence of the mediator.

- *the purely indirect effect*: also called the "mediated main effect" and reflects that the mediator causes the outcome in the absence of the exposure but the exposure is required for the mediator to become present.

We refer readers to some of the original literature describing the four-way decomposition model.[9]

2.2 Four-Way Decomposition: Implementation

Two regression models are used to estimate the components of the partitioned model in Equation 2.1.0 The parameter estimates from these models are then used for the computation of the components. We refer the readers to supporting literature describing the *med4way* command in Stata.[10] We provide some explanation of it here. We assume no unmeasured confounding (as detailed above in **2.1**) and estimate, on average, the four components of the model. The first regression model we fit is for the expectation of Y conditioned on the exposure, mediator, and confounders (Equation 2.2.0). The second is a model for the expectation of M conditioned on the exposure and confounders (Equation 2.2.1):

$$\text{E}[Y|(a, m, \boldsymbol{c})] = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 a * m + \boldsymbol{\theta}_c c \tag{2.2.0}$$

$$\text{E}[M|(a, \boldsymbol{c})] = \beta_0 + \beta_1 a + \boldsymbol{\beta}_c^T \boldsymbol{c} \tag{2.2.1}$$

The parameter estimates from these models are then harnessed for the computation of estimates of the four component sources of variation for the total effect (TE):

$$\text{E}[CDE|c] = \theta_1(a - a')$$

$$\text{E}[INT_{ref}|c] = \theta_3(\beta_0 + \beta_1 a' + \boldsymbol{\beta}_c^T \boldsymbol{c})(a - a')$$

$$\text{E}[INT_{med}|c] = \theta_3 \beta_1 (a - a')(a - a')$$

$$\text{E}[PIE|c] = (\theta_2 \beta_1 + \theta_3 \beta_1 a')(a - a')$$

where a and a' are arbitrary values of the exposure such that $a - a'$ reflects a one-unit difference (or some other meaningful difference) in the exposure.[11] We also note that *med4way* can be generalized to and accomodate other types of regression models and distributions (e.g., binomial, log-binomial, Poisson, negative binomial, Weibull, Cox, etc.).[10] In our analysis, $\text{E}[Y|(a, m, c)]$ and $\text{E}[M|(a, c)]$ are specified as follows:

$$\lambda(t|x, v, \boldsymbol{z}) = \lambda_0(t) + \theta_1 x + \theta_2 v + \theta_3 x * v + \boldsymbol{\theta}_z^T \boldsymbol{z} \tag{2.2.2}$$

$$\text{E}[V|(x, \boldsymbol{z})] = \beta_0 + \beta_1 x + \boldsymbol{\beta}_z^T \boldsymbol{z} \tag{2.2.3}$$

where, in Equation 2.2.2, we model the log hazard at time t as a function of x, the loneliness index score, v, the dementia log odds (computed from one of the Hurd, Expert, or LASSO algorithms), and $\boldsymbol{z}$, the vector of confounders/covariates discussed in the manuscript. In Equation 2.2.3, we model the expectation of the mediator, V, as a function of the loneliness index and the other covariates using ordinary least squares.

References

1. Beydoun, M. A. *et al.* Hospital-treated prevalent infections, the plasma proteome and incident dementia among UK older adults. *iScience* **26** (2023).

2. Gianattasio, K. Z., Wu, Q., Glymour, M. M. & Power, M. C. Comparison of Methods for Algorithmic Classification of Dementia Status in the Health and Retirement Study. eng. *Epidemiology (Cambridge, Mass.)* **30,** 291–302. doi:`10.1097/EDE.0000000000000945` (Mar. 2019).

3. Hurd, M. D., Martorell, P., Delavande, A., Mullen, K. J. & Langa, K. M. Monetary Costs of Dementia in the United States. en. *New England Journal of Medicine* **368,** 1326–1334. doi:`10.1056/NEJMsa1204629` (Apr. 2013).

4. Gianattasio, K. Z., Ciarleglio, A. & Power, M. C. Development of Algorithmic Dementia Ascertainment for Racial/Ethnic Disparities Research in the US Health and Retirement Study. eng. *Epidemiology (Cambridge, Mass.)* **31,** 126–133. doi:`10.1097/EDE.0000000000001101` (Jan. 2020).

5. McKelvey, R. D. & Zavoina, W. A statistical model for the analysis of ordinal level dependent variables. *The Journal of Mathematical Sociology* **4,** 103–120. doi:`10.1080/0022250X.1975.9989847`. eprint: `https://doi.org/10.1080/0022250X.1975.9989847` (1975).

6. Brandt, J., Spencer, M. & Folstein, M. *Telephone Interview for Cognitive Status* en. Institution: American Psychological Association. Sept. 2022. doi:`10.1037/t28542-000`.

7. Jorm, A. F. The Informant Questionnaire on cognitive decline in the elderly (IQCODE): a review. *International psychogeriatrics* **16,** 275–293 (2004).

8. Groenwold, R. H. *et al.* Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. *Cmaj* **184,** 1265–1269 (2012).

9. VanderWeele, T. J. A unification of mediation and interaction: a four-way decomposition. *Epidemiology (Cambridge, Mass.)* **25,** 749 (2014).

10. Discacciati, A., Bellavia, A., Lee, J. J., Mazumdar, M. & Valeri, L. *Med4way: a Stata command to investigate mediating and interactive mechanisms using the four-way effect decomposition* 2019.

11. VanderWeele, T. J. & Tchetgen Tchetgen, E. J. Mediation analysis with time varying exposures and mediators. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **79,** 917–938 (2017).

Supplementary Figure 1. Participant flowchart

```
┌─────────────────────────────────────────────────────┐
│ HRS all sample RAND longitudinal file (1992-2018): N=43,561 │
└─────────────────────────────────────────────────────┘
        │                    ┌──────────────────────┐
        │         ──────────▶│ Not included in 2010 │
        │                    │ N=21,527             │
        ▼                    └──────────────────────┘
┌─────────────────────────────────────────────────────┐
│ Alive and included in HRS in 2010                   │
│ N=22,034                                             │
└─────────────────────────────────────────────────────┘
        │                    ┌──────────────────────┐
        │         ──────────▶│ ≤50y in 2010         │
        │                    │ N=1,339              │
        ▼                    └──────────────────────┘
┌─────────────────────────────────────────────────────┐
│ Alive and included in HRS in 2010, >50y in 2010     │
│ N=20,695                                             │
└─────────────────────────────────────────────────────┘
        │                    ┌──────────────────────┐
        │         ──────────▶│ No dementia status   │
        │                    │ data in 2010         │
        │                    │ N=13,018             │
        ▼                    └──────────────────────┘
┌─────────────────────────────────────────────────────┐
│ Alive and included in HRS in 2010, >50y in 2010,    │
│ with dementia status data in 2010                   │
│ N=7,677; Age(mean) in 2010=65.7y                    │
└─────────────────────────────────────────────────────┘
        │                    ┌──────────────────────┐
        │         ──────────▶│ No dementia status in│
        │                    │ 2010 or loneliness   │
        │                    │ data (2006-2008)     │
        │                    │ N=1,209              │
        ▼                    └──────────────────────┘
┌─────────────────────────────────────────────────────┐
│ Alive and included in HRS in 2010, >50y in 2010,    │
│ with dementia status data in 2010 and loneliness    │
│ data in 2006-2008                                   │
│ N=6,468; 3,273 deaths (2010-2020); MR=64/1000 P-Y   │
└─────────────────────────────────────────────────────┘
```
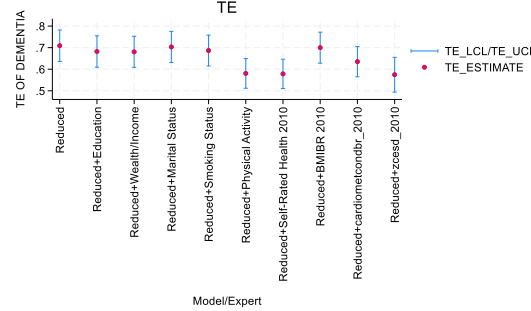
HRS, Health and Retirement Study; MR, mortality rate

Supplementary Figure 2. Total effects of loneliness and Ln(odds of dementia), z-scored, 3 algorithms: Four-way decomposition models incrementally adjusted for exogenous variables
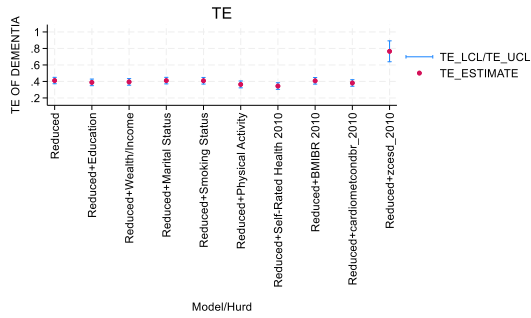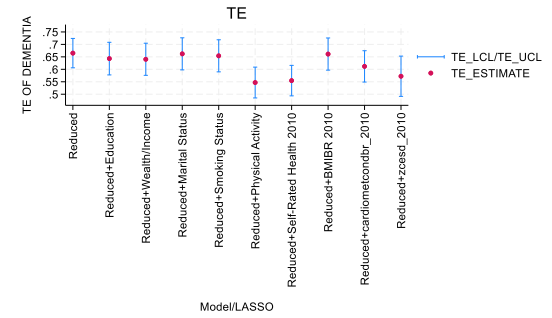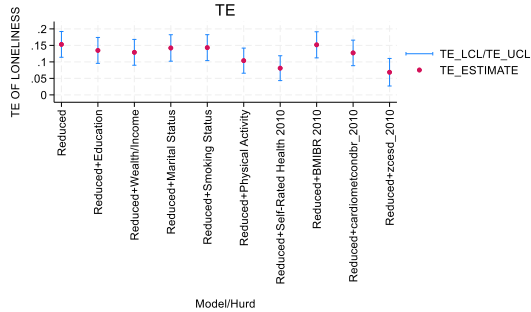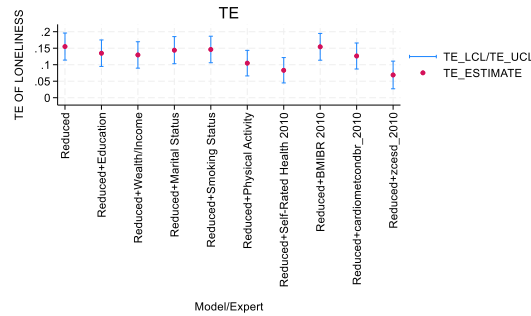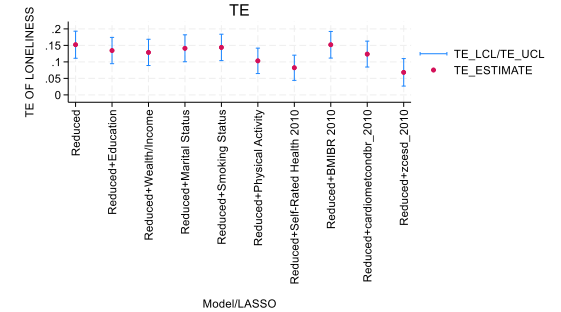

(A.1.) X= Ln(odds of dementia), z-scored: Hurd algorithm, M=Loneliness z-score
(A.2.) X= Ln(odds of dementia) z-scored: Expert algorithm, M=Loneliness z-score
(A.3.) X= Ln(odds of dementia) z-scored: LASSO algorithm, M=Loneliness z-score
(B.1.) M= Ln(odds of dementia) z-scored: Hurd algorithm, X=Loneliness z-score
(B.2.) M= Ln(odds of dementia) z-scored: Expert algorithm, X=Loneliness z-score
(B.3.) M= Ln(odds of dementia) z-scored: LASSO algorithm, X=Loneliness z-score

LASSO, Least Absolute Shrinkage Selection Operator; M, mediator/moderator; TE, total effect; X, exposure