Trends in the Risk of Cognitive Impairment in the United States, 1996–2014

Jo Mhairi Hale,[a,b] *Daniel C. Schneider,*[b] *Jutta Gampe,*[b] *Neil K. Mehta,*[c] *and Mikko Myrskylä*[b,d,e]

Background: Accumulating evidence suggests risk of cognitive impairment is declining in high-income countries. Much of this research uses longitudinal surveys in which learning over repeated tests may bias results. We analyze trends in cognitive impairment in the United States, accounting for prior test experience and selective mortality.
Methods: We use the Health and Retirement Study, a population-based, nationally representative panel dataset and include individuals ages 50 years and older in 1996–2014 (n = 32,784). We measure cognitive impairment and dementia using standard cutpoints of the modified Telephone Interview for Cognitive Status. We estimate logistic regression models for any impairment and dementia over time, adjusting for age, sex, and race/ethnicity, comparing models with and without adjustment for practice effects and education. We examine heterogeneity in trends by age, sex, race/ethnicity, and education.
Results: Models not controlling for test experience suggest that risk of cognitive impairment and dementia decreased over the study period. Controlling for test experience reverses the trend. In our primary models, prevalence of any cognitive impairment increased for women from 18.7% to 21.2% (annual change 0.7%, 95% confidence interval [CI], 0.1%, 1.3%) and for men from 17.6% to 21.0% (annual change 1.0%, CI, 0.5%, 1.4%). For dementia, women's annual increase was 1.7% (CI, 0.8%, 2.6%) and men's 2.0% (CI, 1.0%, 2.9%). If not for education, the increase would have been stronger. Increased risk was particularly rapid for Latinas, the least educated, and older ages.

Conclusions: Risk of cognitive impairment increased from 1996 to 2014. Uncovering determinants of increasing cognitive impairment risk should become a research priority. See video abstract: http://links.lww.com/EDE/B702.

Keywords: Cognitive impairment; Dementia; Learning; Practice effects; Trends

(*Epidemiology* 2020;31: 745–754)

Increasing life expectancy and population aging have inspired a surge of research on aging-related cognitive impairment, including dementia. Approximately 5.8 million people in the United States have a diagnosis of Alzheimer's disease (AD), comprising 60%–80% of dementia cases.[1] Deaths attributable to AD rose by two-thirds from 2000 to 2013, pushing it into the sixth leading cause of death in the United States.[2] Trends in cognitive impairment and dementia are, therefore, critically important for population health and health policy.

Recent research, however, suggests that risk of later-life cognitive impairment may be declining in many high-income countries, including the United States.[3–12] Three of the oft-proposed mechanisms for such trends are, first, more recent cohorts have higher educational attainment, building protective cognitive/brain reserve.[13–16] Second, recent cohorts had more health-promoting early-life environments, such as better nutrition.[7,17–19] Third, medical developments for treatment of vascular diseases, e.g., the use of statins, have decreased the incidence of risk factors for cognitive impairment.[3,8,20,21]

An important part of the evidence on declining risk of cognitive impairment comes from nationally representative longitudinal studies. We argue that these studies are likely to be biased because the longitudinal nature is not appropriately accounted for in the study design.[22] In particular, the Health and Retirement Study (HRS), a large national, longitudinal survey, has been used repeatedly to examine dementia trends. Results from the HRS and other longitudinal studies consistently suggest that trends in cognitive impairment in the United States are improving.[5,23–25] However, most of these studies (with the exception of Choi et al.[26]) do not account for one potentially important factor—individuals in the HRS are surveyed repeatedly using the same (counting backwards, serial 7s) or similar (word recall) metrics. An extensive body of literature, especially in medical fields, suggests that repeated

exposure to analogous tests creates a "practice effect" that may mask cognitive decline and bias trends.[22,27–32]

Practice effects are a specific type of panel conditioning, whereby respondents' responses are impacted by previous participation.[33–38] In the HRS, the average number of tests taken among the individuals in the 1996–1998 waves is 1.4; in the 2000–2002 waves, it is 3.0; and in the 2012–2014 waves, it is 6.0. As, over time, individuals who repeatedly respond to the survey accumulate test-taking experience, test scores may increase even though cognitive ability did not.[28,29,32,39] Thus, an increase in test scores over time may reflect either true increase or greater test-taking experience. The effect size associated with panel conditioning varies across contexts, and evidence suggests that the magnitude of practice effect for cognitive function testing may be large enough to mask several years of aging.[40] Therefore, when analyzing time trends in cognition, adjusting for the changing distribution of number of tests ("practice effects") is important. Adjusting for practice effects can be done in a multivariate regression model in the same way that one adjusts for other variables, such as racial/ethnic or educational composition.

We estimate age-specific time trends in cognitive impairment in the United States ages 50 and older population using a large-scale, population-based, nationally representative, longitudinal survey. Our key contributions are that we estimate the population-level trends after controlling for prior test experience, providing less-biased estimates than prior literature; we model multiple specifications of practice effects, testing the sensitivity of estimated trends to measurement; we also analyze to what extent the estimated trends would be biased if test experience is ignored. Most prior studies do not account for test experience, and to our knowledge, no population-based study using the HRS has analyzed the magnitude of the bias when prior test experience is ignored. We additionally examine heterogeneity in time trends by age, sex, race/ethnicity, and educational attainment. Our results show that without adjustment average cognitive function scores seem to be increasing over time; but, once practice effects are adjusted, the time trends are flat or negative.

METHODS

Study Population

The Health and Retirement Study (HRS) is a nationally representative, biennial panel survey of US residents age 50 and older and their spouses. The University of Michigan Institutional Review Board granted ethical approval for the Health and Retirement Study. It includes information on health, demographic factors, educational attainment, and a version of the Telephone Interview for Cognitive Status (TICS-M), specifically modified to be sensitive to pathological cognitive decline and minimize ceiling effects.[39,41] The University of Michigan conducts the HRS,[42,43] which is sponsored by the National Institute on Aging (grant number NIA U01AG009740). We use RAND Version P of the HRS, selecting respondents ages 50

and older and all the waves in which the TICS-M measures we utilize were administered consistently (1996–2014). We use the University of Michigan Survey Research Center's imputed TICS-M values.[44]

The raw HRS data set contains 37,495 individuals with 226,564 respondent or proxy interviews. The observation count reduces because of restrictions to age 50 and over and waves 1996–2014 (38,493 observations), exclusion of zero-weighted observations (7,503 observations), missing values in the regression variables (1,199 observations), and data anomalies, such as death being recorded before valid interview records (133 observations). The resulting 179,236 observations correspond to 32,784 subjects in the sample for our logistic regression models. In the eAppendix (http://links.lww.com/EDE/B677), we report results that use a continuous cognitive score as the outcome; in these results proxy responses that are categorical were excluded. As a result the sample size was 165,926 observations, corresponding to 31,696 subjects (13,310 fewer observations).

Procedures

Cognitive impairment is defined based on total score on TICS-M measures that reflect neurophysiological health:[45] immediate (0–10 points) and delayed recall (0–10 points), serial 7s (0–5 points), and backward counting from 20 (0–2 points). The range is 0–27; higher scores indicate better cognitive function. Although score on the TICS-M is not indicative of a clinical mild cognitive impairment or dementia diagnosis, we use standard cutpoints—no cognitive impairment (NCI) 12–27, cognitive impairment no dementia (CIND) 7–11, and dementia 0–6—that were validated against the Aging, Demographics, and Memory Study (ADAMS).[46] The ADAMS selects a subset of respondents from the HRS sample to undergo extensive clinical evaluation for cognitive impairment, which has subsequently been used by Langa, Crimmins, and others to establish cutpoints in the HRS for CIND and dementia.[46,47] For simplicity, we adhere to the norm of referring to the categories as CIND and dementia, despite that the lack of clinical diagnosis means it is only probable CIND or dementia. We combine CIND and dementia to create a category "any impairment" and analyze both any impairment and dementia.

We follow Langa and colleagues[5] in using a measure constructed from proxy interviews to indicate categories of cognitive impairment when respondents do not complete the TICS-M. Before 2000, this includes the proxy's assessment of the respondent's memory (0–4) and number of difficulties with instrumental activities of daily living (0–5) (NCI = 0–2, CIND = 3–4, and dementia = 5–9 points). From 2000 to 2014, the evaluation of the interviewer is included (0 = no cognitive limitation, 1 = some cognitive limitation, 2 = cognitive limitation prevents completion of the interview). For these years, 0–2 is NCI, 3–5 is CIND, and 6–11 is dementia. Our results using categorical coding of time or waves 2000–2002 as the reference time (Figure 1) show that our conclusions are robust

to this shift in definition. We also include a control for interview mode (face-to-face, telephone, or proxy).

In our primary regressions, we include time as a continuous variable (exact date of interview). We also analyze time trends using categorical values of time, corresponding to paired HRS waves (1995–1998, 2000–2002, 2004–2006, 2008–2010, 2012–2014). We include practice effects as a categorical variable for test number (first, second, third to fourth, fifth–seventh, and eighth test or higher), and, in alternative specifications, analyze test number as a continuous variable, square root of continuous test number, and broader categories (first, second, third, fourth, or higher). Proxy interviews receive their own category and do not increment the test number count, as they do not accumulate practice effects.

We control for sex (self-reported as binary, woman/man), race/ethnicity (nonHispanic White, Black [Hispanic or nonHispanic], nonBlack Hispanic, and nonHispanic "Other"; henceforth, White, Black, Latinx, Other), education (less than high school/general educational development, high-school diploma, associate degree or higher), and age at interview (age and age squared or age group 50–64, 65–74, 85+). We use RAND's complex survey weights that include residents of nursing homes and that were designed for longitudinal analysis (variable RwWTCRNH).[48]

Statistical Analyses

Using a sequence of logistic regression models, we regress a binary indicator of any impairment or dementia on time and selected covariates. Model 1 includes age, age squared, race/ethnicity, and an interaction between sex and interview mode (women with proxies are more than twice as likely to be reported as cognitively impaired than men with proxies). This model is expected to replicate the results on positive trends in cognitive function that others have found. Model 2 introduces controls for test number. This model is our primary model and is expected to produce less-biased estimates of time trends than model 1, as it removes the biasing effect of practice. Model 3 includes education as a potential mechanism driving trends in cognitive impairment. To analyze heterogeneity in trends across subpopulations, we estimate model 2, adding interactions between time and: sex, age, race/ethnicity, and education.

Furthermore, we conduct a series of robustness checks that respond to methodologic concerns, including analyzing the robustness of our results to time-on-study and multicollinearity among age, period, and practice effects. We restrict analyses to individuals who have the same test number and vary how we specify time, age, and practice effects. Additional robustness checks, presented in the eAppendix; http://links.lww.com/EDE/B677, include the following: (1) we replicate models 1–3 using a continuous measure of cognitive function; (2) we restrict the sample to respondents ages 65 and over to allow comparisons with other findings; (3) we test various cutpoints for categorizations of any impairment and dementia to examine sensitivity to cutpoint definition; (4) we estimate a joint model that includes both mortality and measurement of cognitive function to investigate whether selective mortality influences our results; and (5) to determine whether
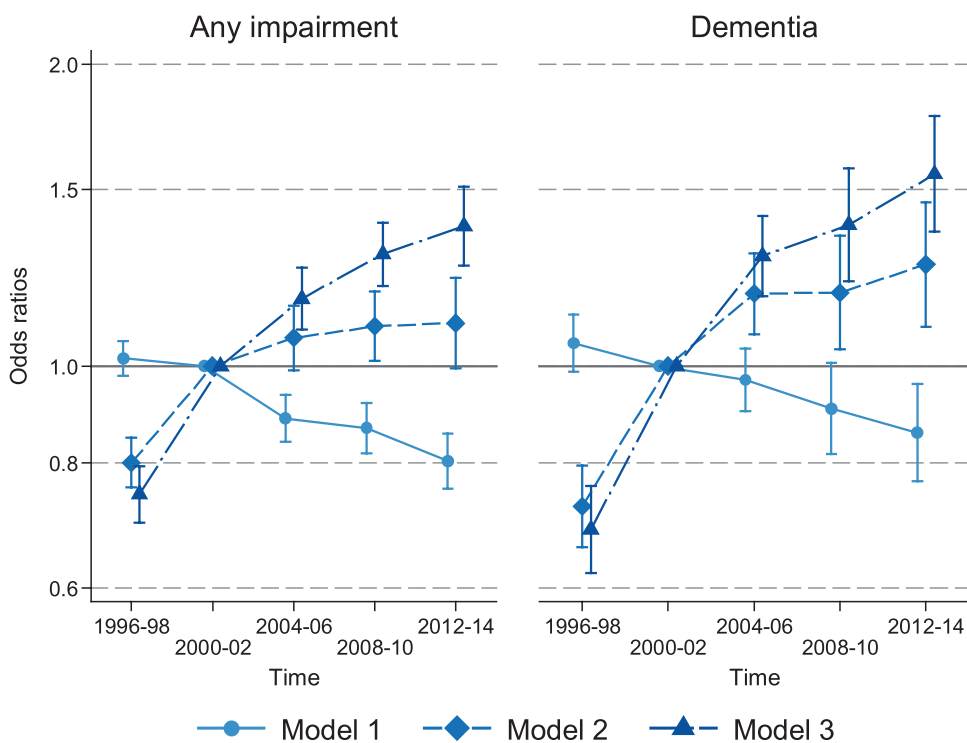

FIGURE 1. Odds ratios and 95% confidence intervals for time from logistic regression models of any cognitive impairment or dementia (Model 1 includes pairs of survey waves, age, age squared, race/ethnicity, and the interaction between gender and interview mode. Model 2 adds test number. Model 3 adds educational attainment).

multicollinearity biases results, we conduct a simulation in which we study whether the parameters can be reliably estimated from a dataset whose correlation structure mimics that of the actual data.

RESULTS

Descriptive results shown in Table 1 document declines in percent of any cognitive impairment and dementia, although mean cognitive function score remains stable. Mean test number increases across time. There is a strong decline across time in percent proxy interviews. The sample becomes more ethnically diverse and more highly educated across time.

Table 2 shows the regression results for any cognitive impairment and dementia. Model 1 shows that if one does not account for prior tests, the time trends in cognitive impairment appear positive, with a substantial decline in both any cognitive impairment and dementia. Over a 10-year period, odds of any impairment decrease by a factor of 0.85 (95% CI = 0.82, 0.89). For dementia, the odds decrease by 0.87 (95% CI = 0.82, 0.93). Model 2 includes test number as a covariate. This changes the estimated time trend such that odds of any impairment increase by a factor of 1.14 (95% CI = 1.06, 1.22) over a 10-year period. For dementia, the corresponding increase is 1.29 (95% CI = 1.16, 1.44). Model 3 introduces educational attainment as a covariate and shows that the increases would have been larger if educational attainment had not increased, with odds of any impairment increasing by a factor of 1.37 (95% CI = 1,29, 1.46) and dementia

by 1.54 (95% CI = 1.39, 1.70). Figure 1 illustrates the results using survey waves instead of continuous time and confirms that once practice effects are controlled (model 2), the odds of cognitive impairment are lower in 1996–1998 compared with 2000–2002 and higher in all later waves, suggesting increased trends in prevalence of cognitive impairment over time. Although the association is not strictly linear, across the period 1996–2014, there is no evidence of improvement.

Table 3 shows predicted prevalence of cognitive impairment based on models 1 and 2. Not controlling for practice effects (model 1), the results consistently suggest declining trends for any cognitive impairment and dementia for women and men. Controlling for test experience (model 2) reverses the trend: prevalence of any cognitive impairment increased for women from 18.7% to 21.2% (annual change 0.7%, CI = 0.1%, 1.3%) and for men from 17.6% to 21.0% (annual change 1.0%, CI = 0.5%, 1.4%). For dementia, the change among women was from 5.1% to 7.0% (annual change 1.7%, CI = 0.8%, 2.6%) and among men from 3.8% to 5.4% (annual change 2.0%, CI = 1.0%, 2.9%).

Table 4 shows the odds ratios of dementia for a 10-year change in time for models 1 and 2 for subpopulations. Most groups experience increasing dementia once we control for practice effects. We find no evidence for a gender difference (women OR = 1.31, CI = 1.13, 1.52; men = 1.28, CI = 1.13, 1.44). Analyses by age show that the pattern is strongest among the oldest (age 85+) compared with younger groups. Odds of dementia have increased across time among Latinas (OR = 1.79,

TABLE 1. Descriptive Characteristics of the Health and Retirement Study Analytical Sample Over Time (1996–2014), Number of Persons 32,784

	1996–1998	2000–2002	2004–2006	2008–2010	2012–2014	Total
Age (years, mean)	66.4	67.8	65.9	66.4	67.0	66.7
Women %	55.2	55.8	54.6	54.5	54.2	54.8
Any cognitive impairment %	20.0	22.1	19.5	19.5	18.6	19.8
Dementia %	5.0	6.8	5.6	5.2	4.7	5.4
Cognitive function score (mean; range 0–27)	15.7	15.6	15.7	15.6	15.7	15.6
Number of tests (mean)	1.4	3.0	4.0	5.0	6.0	4.2
Proxy interviews %	6.0	10.2	7.1	5.5	4.4	6.5
Race/ethnicity %						
White	83.3	82.4	80.6	79.2	77.8	80.3
Black	9.0	9.3	9.5	9.9	10.2	9.7
Latinx	5.7	6.1	7.2	7.9	8.4	7.2
Other	2.0	2.2	2.7	3.1	3.5	2.8
Educational attainment %						
Less than high school	26.2	23.7	18.7	15.9	13.6	18.8
HS/GED/some college[a]	52.2	52.6	52.7	52.1	51.5	52.2
Associate degree or higher	21.6	23.7	28.6	32.0	34.9	29.0
Person-Waves	33,646	35,722	36,654	36,971	36,243	179,236

All measures, including any cognitive impairment and dementia, include information from proxy informants, except continuous cognitive function score, which is available for self-respondents only, and number of tests that are calculated based on self-responses only.

[a]HS indicates high school. GED is general educational development, a set of tests that provides certification considered equivalent to a US high school diploma.

TABLE 2. OR and 95% CI from Logistic Regression Models for Any Cognitive Impairment and Dementia

	Any Cognitive Impairment			Dementia		
	M1	M2	M3	M1	M2	M3
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Time (unit: 10 years)	0.85 (0.82, 0.89)	1.14 (1.06, 1.22)	1.37 (1.29, 1.46)	0.87 (0.82, 0.93)	1.29 (1.16, 1.44)	1.54 (1.39, 1.70)
Race/ethnicity (ref. White)						
Black	4.8 (4.4, 5.3)	4.6 (4.2, 5.1)	3.4 (3.2, 3.7)	4.4 (3.9, 5.1)	4.3 (3.8, 4.9)	3.1 (2.8, 3.4)
Latinx	4.2 (3.6, 4.8)	4.0 (3.5, 4.6)	2.1 (1.8, 2.4)	3.3 (2.8, 3.9)	3.2 (2.7, 3.7)	1.8 (1.5, 2.1)
Other	2.4 (1.6, 3.5)	2.3 (1.6, 3.4)	2.1 (1.6, 2.8)	2.0 (1.2, 3.4)	1.9 (1.1, 3.4)	1.7 (1.1, 2.8)
Man (ref. woman)	1.1 (1.1, 1.2)	1.1 (1.0, 1.2)	1.1 (1.1, 1.2)	1.1 (1.0, 1.2)	1.0 (0.92, 1.1)	1.0 (0.93, 1.1)
Age (unit: 10 years)	1.4 (1.3, 1.6)	1.9 (1.7, 2.1)	1.4 (1.3, 1.6)	2.4 (2.1, 2.9)	3.2 (2.7, 3.7)	2.5 (2.1, 3.0)
Age squared (unit: 10 years)	1.1 (1.1, 1.1)	1.1 (1.0, 1.1)	1.1 (1.1, 1.1)	1.0 (1.0, 1.1)	1.0 (1.0, 1.0)	1.0 (1.0, 1.0)
Interview mode (ref. face-to-face)						
Telephone	0.80 (0.76, 0.83)	0.79 (0.75, 0.83)	0.81 (0.77, 0.85)	0.82 (0.75, 0.89)	0.78 (0.72, 0.86)	0.82 (0.75, 0.89)
Proxy	6.4 (5.7, 7.3)	4.4 (3.9, 5.0)	4.5 (3.9, 5.2)	19 (17, 22)	11 (9.4, 12)	11 (9.5, 13)
Telephone × man	1.1 (1.0, 1.2)	1.1 (1.0, 1.2)	1.2 (1.1, 1.3)	0.93 (0.82, 1.1)	0.95 (0.84, 1.1)	1.0 (0.89, 1.1)
Proxy × man	0.42 (0.36, 0.49)	0.46 (0.40, 0.54)	0.40 (0.34, 0.47)	0.33 (0.29, 0.39)	0.37 (0.32, 0.43)	0.33 (0.28, 0.39)
Test number (ref. 1st)						
2nd		0.94 (0.88, 1.00)	0.94 (0.88, 1.00)		0.91 (0.83, 1.00)	0.91 (0.83, 1.00)
3–4		0.77 (0.72, 0.82)	0.78 (0.73, 0.83)		0.66 (0.59, 0.74)	0.66 (0.59, 0.74)
5–7		0.58 (0.52, 0.64)	0.60 (0.54, 0.66)		0.46 (0.40, 0.53)	0.48 (0.42, 0.55)
8+		0.46 (0.40, 0.53)	0.48 (0.43, 0.54)		0.33 (0.27, 0.40)	0.34 (0.28, 0.42)
Educational attainment (ref. less than high school)						
HS/GED[a]			0.31 (0.29, 0.33)			0.35 (0.32, 0.38)
Associate +			0.13 (0.12, 0.14)			0.18 (0.15, 0.21)
Constant[b]	0.054 (0.048, 0.061)	0.048 (0.043, 0.055)	0.17 (0.15, 0.20)	0.0039 (0.0031, 0.0049)	0.0039 (0.0031, 0.0048)	0.012 (0.0092, 0.015)
Person-waves	179,236	179,236	179,236	179,236	179,236	179,236

M1 includes covariates time, race/ethnicity, age, age squared, and the interaction between sex and interview mode. M2 adds test number. M3 adds educational attainment.

[a]HS indicates high school diploma. GED is general educational development, a set of tests that provides certification considered equivalent to a US high school diploma.

[b]The constant represents the baseline odds (not odds ratio) when categorical variables are at their reference values and age and time at values to which these were centered (age 50, year 2000).

CI = 1.39, 2.29), White men (OR = 1.31, CI = 1.13, 1.51), White women (OR = 1.29, CI = 1.13, 1.48), and Black men (OR = 1.24, CI = 1.06, 1.45). Within educational groups, the odds ratios for the time trend do not indicate improvement even before controlling for practice effects. When controlling for practice effects, each education group shows higher odds ratios for time, and this is strongest among the least educated.

Robustness

The eAppendix; http://links.lww.com/EDE/B677 presents results and detailed discussion of our extensive robustness analyses. In summary, analyses in which we use a continuous score (eTable 1; http://links.lww.com/EDE/B677) restrict the sample to ages 65 and older (eTables 2, 3; http://links.lww.com/EDE/B677), and test for sensitivity in the cutpoints for cognitive impairment (eTable 4; http://links.lww.com/EDE/B677) all suggest that our findings of increasing odds in cognitive impairment are robust. Joint models in which we simultaneously model cognitive impairment and death do not indicate that our results are biased by selective mortality

(eTable 5; http://links.lww.com/EDE/B677). We also use three approaches to test whether collinearity among age, time, and practice effects biases results.

First, we use a subsampling strategy to purge estimation results from the effects of prior test experience and time-in-study. We create four subsamples restricted to first, second, fourth, or sixth test, and estimate for each subsample the association between cognitive impairment and time. Since within subsample the number of tests taken is constant, we do not need to control for number of tests. Only age, age squared, race/ethnicity, sex, interview mode, and interaction between sex and interview mode are controlled. Figure 2 shows the odds ratios over time. Within each subsample, the results suggest increasing or flat odds, but no decrease over time. These findings provide further evidence that there is no declining trend in odds of any cognitive impairment or dementia over the time period from 1996 to 2014.

Second, models in which we use combinations of alternative specifications for age (quadratic, categorical), time (continuous, two different categorizations), and practice effects (continuous, square root, two different categorizations)

TABLE 3. Predicted Prevalence and Annual Percent Change in Any Cognitive Impairment and Dementia for Ages 50 and Over from 1996 and 2014

	Any Cognitive Impairment (%)			Dementia (%)		
	1996	**2014**	**Annual % Change[a] (95% CI)**	**1996**	**2014**	**Annual % Change[a] (95% CI)**
Women						
Model 1	22.6	18.4	−1.1 (−1.5, −0.8)	6.7	5.6	−1.0 (−1.5, −0.8)
Model 2	18.7	21.2	0.7 (0.1, 1.3)	5.1	7.0	1.7 (0.8, 2.6)
Men						
Model 1	21.4	18.3	−0.9 (−1.2, −0.5)	5.1	4.3	−0.9 (−1.6, −0.2)
Model 2	17.6	21.0	1.0 (0.5, 1.4)	3.8	5.4	2.0 (1.0, 2.9)

Model 1 includes covariates time, race/ethnicity, age, age squared, and the interaction between gender and interview mode. Model 2 adds test number.
[a]Annual percent change is calculated as $100 \times [(P2014/P1996)^{(1/18)} - 1]$, where P2014 and P1996 are the predicted probabilities in years 2014 and 1996, respectively.

TABLE 4. Subpopulation Analysis: OR and 95% CI for the 10-year Time Trend from Regression Models for Dementia, Estimated Separately for Subpopulations

	Women		Men	
	Model 1	**Model 2**	**Model 1**	**Model 2**
	OR (95% CI)	**OR (95% CI)**	**OR (95% CI)**	**OR (95% CI)**
Sex-stratified	0.86 (0.79, 0.93)	1.31 (1.13, 1.52)	0.89 (0.82, 0.98)	1.28 (1.13, 1.44)
Age group				
50–64	1.07 (0.81, 1.43)	1.18 (0.89, 1.57)	1.01 (0.83, 1.24)	1.04 (0.85, 1.27)
65–74	0.78 (0.67, 0.91)	1.13 (0.95, 1.34)	0.73 (0.61, 0.88)	1.03 (0.85, 1.25)
75–84	0.78 (0.70, 0.86)	1.16 (0.99, 1.35)	0.91 (0.80, 1.03)	1.44 (1.23, 1.68)
85+	0.94 (0.82, 1.08)	1.38 (1.18, 1.62)	1.01 (0.85, 1.21)	1.65 (1.32, 2.07)
Race/ethnicity				
White	0.83 (0.77, 0.90)	1.29 (1.13, 1.48)	0.87 (0.78, 0.98)	1.31 (1.13, 1.51)
Black	0.77 (0.67, 0.88)	1.16 (0.96, 1.40)	0.93 (0.80, 1.08)	1.24 (1.06, 1.45)
Latinx	1.28 (0.98, 1.68)	1.79 (1.39, 2.29)	0.95 (0.68, 1.32)	1.18 (0.87, 1.60)
Other	0.69 (0.45, 1.07)	0.95 (0.62, 1.47)	0.98 (0.55, 1.75)	1.28 (0.73, 2.23)
Education				
<HS	1.06 (0.94, 1.19)	1.67 (1.42, 1.96)	1.15 (0.97, 1.36)	1.58 (1.32, 1.90)
HS/GED[a]	1.03 (0.93, 1.14)	1.55 (1.32, 1.81)	1.04 (0.92, 1.18)	1.42 (1.22, 1.66)
Associate +	0.89 (0.68, 1.17)	1.27 (0.97, 1.66)	1.09 (0.84, 1.41)	1.52 (1.17, 1.99)

Model 1 does not adjust for number of tests taken. Model 2 includes number of tests as a covariate. The models are estimated separately for men and women. Time is in 10-year units. Results by age group are obtained by including an interaction term between age group and time. Results by race/ethnicity are obtained by including an interaction term between race/ethnicity and time. Results by education are obtained by including an interaction term between education and time. Models denoted by "model 1" include as covariates age, age squared, race/ethnicity, and interview mode, except the model estimating age group–specific results, which includes categorical age instead of continuous and quadratic age. Models denoted by "model 2" additionally include the number of tests taken as a covariate.
[a]GED is general educational development, a set of tests that provides certification considered equivalent to a US high school diploma.
HS indicates high school diploma.

predominantly indicate increasing odds ratios over time of any cognitive impairment and dementia prevalence (eTable 6; http://links.lww.com/EDE/B677). In the large number of coefficients produced through these alternative specifications (252), a small number of results suggest weakly negative trends (e.g., for dementia, there are three negative coefficients out of 84). The exceptional coefficients that do not suggest flat or increasing trends are based on the crudest categorizations in which both age and

practice effects are coded in such coarse categories that they are likely to only partially control for the underlying concepts.

Third, we use a simulation experiment to test whether the survey circumstances of sample refreshment (reducing the collinearity between age and time) and nonresponse (reducing collinearity between number of tests taken, age, and time) alleviate collinearity enough for robust coefficient estimation. In short (eAppendix V; http://links.lww.com/EDE/B677 for details), after generating simulated cognitive

A

Any cognitive impairment
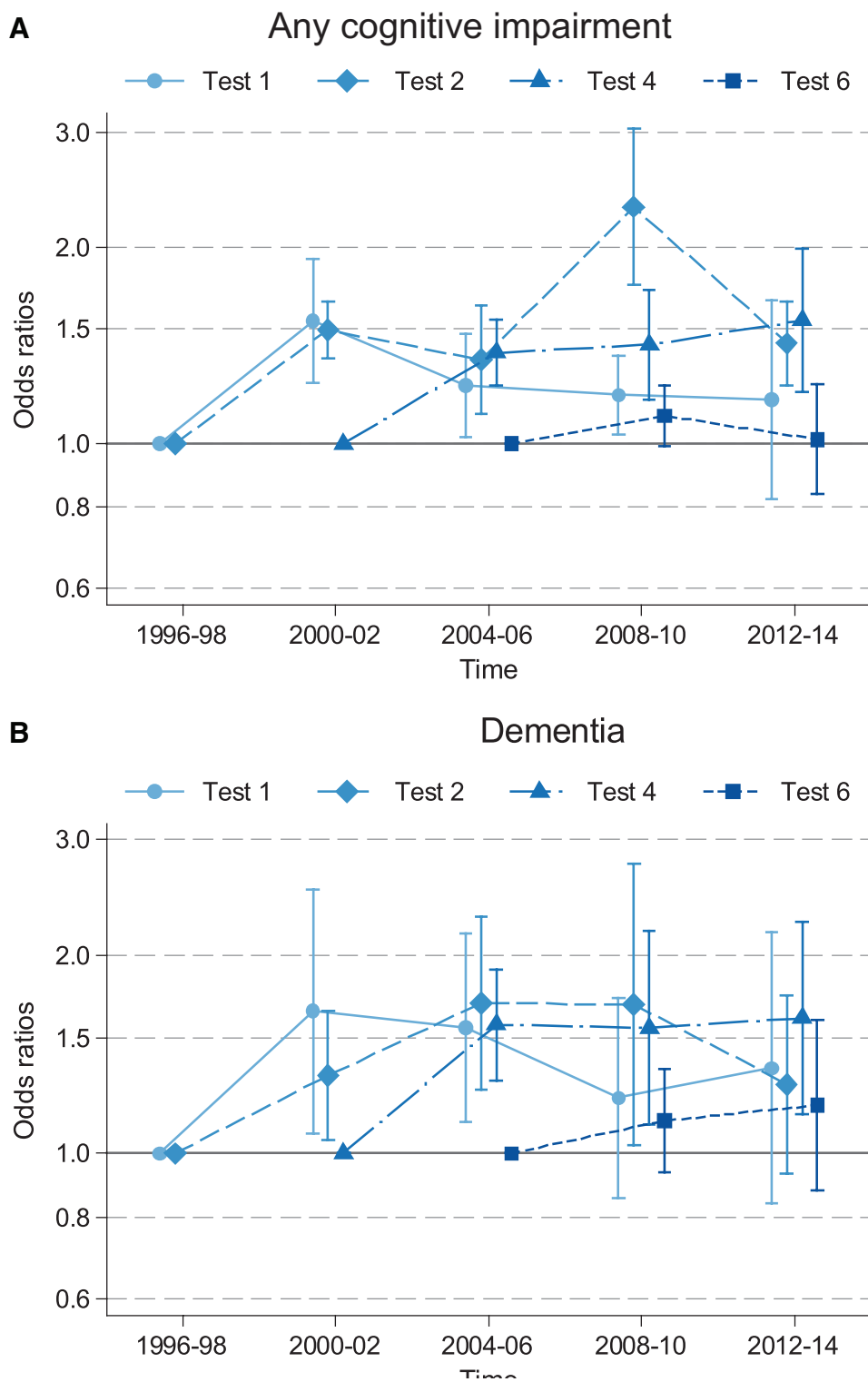

B

Dementia

FIGURE 2. Odds ratios and 95% confidence intervals for time from logistic regression models for any impairment (A) and dementia (B), estimated from subsamples restricted to first, second, fourth, or sixth test (The reference category is 1996–1998 for tests 1 and 2, 2002–2004 for test 4, and 2004–2006 for test 6).

function trajectories for HRS respondents, we reestimate the simulation model from the simulated data and determine that we can estimate the coefficients accurately despite collinearity (eTable 7; http://links.lww.com/EDE/B677). In addition, this exercise demonstrates that if practice effects are present, their omission can introduce a serious bias to the time trend.

DISCUSSION

Using the HRS, a nationally representative, population-based panel survey, we examine trends in the prevalence of both any cognitive impairment and dementia across almost 20 years, accounting for prior test experience. We find consistency across our results that once prior test experience is

considered, trends in cognitive impairment increase. This holds for all subpopulations and is true for prevalence of any impairment, cognitive impairment without dementia, and dementia, as well as when measuring cognitive function continuously. Latinas, the least educated, and the oldest age group experienced the largest increases in prevalence of impairment. Especially for the least educated and oldest, this may be driven by lengthening life expectancies, in general, as well as people living longer with dementia.

From a transdisciplinary perspective, the discrepancy between our results and other findings of declining risk of cognitive impairment is not surprising.[3,5,7,49,50] Medical researchers warn that when assessments of cognitive decline are based on repeated, analogous tests, results will be biased if practice effects are not considered; respondents' test-taking ability improves over repeated exposure, inflating cognitive function scores.[27–32] Indeed, community-based or clinical studies that use a medical evaluation for diagnosis are less likely to find declining trends in dementia prevalence or incidence than population-based survey data.[8,51–56]

Some studies acknowledge possible concerns with practice effects, but dismiss them as unlikely to be important.[49,57] However, more recent studies that explicitly model practice effects in contexts other than population-based research on cognitive impairment find effect sizes large enough to bias results.[12,26,28,29,32,39,58,59] Calamia and colleagues[40] review the practice effects literature and suggest that the effect size may be up to a quarter of a standard deviation. They conclude that practice effects cannot be ignored, a position echoed by the American Academy of Clinical Neuropsychology, as well as other cognitive scientists.[27,28,30,60]

Methodologic Considerations

While there is general agreement among cognitive scientists that we should account for practice effects when modeling cognitive decline, a range of details makes practice effects challenging to model. Practice effects are design-specific, and therefore likely vary strongly across studies. How practice effects accumulate is uncertain:[40,60,61] the largest practice effect may be between the first and second test[39,62]; there may be diminishing returns[61]; or effects may last only up until a certain number of tests.[12,58] We use several categorizations for practice effects to test the robustness of findings to functional form. That practice effects, time, and aging are collinear could also pose problems. We test for various specifications of all three variables and find our results remarkably consistent, except for the crudest categorizations wherein age and practice effects only partially control for the concepts they are supposed to measure (eTable 6; http://links.lww.com/EDE/B677).

Additionally, practice effects may depend on a variety of factors, including age, cohort, genetics (e.g., apolipoprotein E ε4 status), sex, race/ethnicity, health, level of education, baseline cognition, and time to death. In brief, the research on the association between practice effects and age, cohort,

education, and baseline cognitive function are inconclusive; there is less support for an interaction between practice effects and sex, race/ethnicity, apolipoprotein E ε4, or cardiovascular health.[12,59,63,64] Evidence does suggest that education and baseline cognitive function may be positively associated with practice effects[29,39]; although, others find no pattern.[59,63–66] In this analysis, however, the emphasis is not on the specification of the practice effects, but on the importance of accounting for them when analyzing trends in prevalence from panel data. Our results show the importance of accounting for practice effects when analyzing prevalence of cognitive impairment. Further research should analyze whether practice effects have a similar impact on incidence estimates.

Other survey design factors are also important. The number of tests taken, at the individual level, is correlated with the number of years that a person has stayed in the survey, which also captures differential attrition by cognitive status. It is possible that respondents who have poor cognitive function are more likely to leave the study, due to failing function, comorbidities, or even stigma associated with inability to complete the survey.[26] Because of the high correlation between number of tests and time-on-study, interpretation of the coefficient for prior tests is difficult. This coefficient can capture the effects of both prior tests and selective attrition. However, our robustness checks examining individuals at different waves of participation who have the same number of prior tests (Figure 2) and our simulation exercise (eAppendix V; http://links.lww.com/EDE/B677) should mitigate these concerns. Furthermore, our contribution is not focused on unbiased estimation of the practice effect, but on accurately estimating time trends in cognitive impairment. The latter is robust to the interpretational vagueness that comes with collinearity between number of tests taken and number of years in the survey.

Another survey design feature is that the HRS shifted to more face-to-face interviews in 2006; since then, the use of proxies has declined from approximately 10% to only 5.6% in 2014.[5,10] Proxy-use is positively associated with cognitive impairment,[67] so excluding proxies may introduce negative bias—a larger share of the population may appear cognitively impaired in more recent years when fewer proxies were used. Though this may be the case in modeling the continuous measure for cognitive function (eTable 1; http://links.lww.com/EDE/B677), our logistic regression analyses include proxy interviewees. We also include an interaction between interview mode and sex because women with proxies are more than twice as likely to be assessed as demented than proxied men. All results pointing to the same conclusion lends confidence that changes in interview mode are not driving our results.

Of course, despite all efforts to minimize the impact of nonrandom loss to follow-up, it is possible that such missingness still influences our estimated trends. However, it seems unlikely that such missingness would differentially influence

our results that control versus do not control for practice effects. Thus, even in the presence of biasing missingness, it seems unlikely that this missingness would influence our conclusions regarding the importance of controlling for prior test experience.

As we do, Choi and colleagues[26] attempt to control for various survey design features in analyzing the HRS, including prior testing. They also find no downward trend in cognitive impairment. Our analyses, however, diverge in several dimensions. Most important, although Choi et al.[26] note that their estimate of the time trend (no improvement) differs from much of the literature (that finds improvement), they write that analyzing the reasons for this difference is beyond the scope of their article. They speculate that the reasons might be related to, among other factors, how health conditions are modeled. However, practice effects are not mentioned among the potential candidates that could be responsible for the differences in time trends between the Choi et al.[26] article and the majority of the literature. Our article directly addresses this issue and shows how whether one controls or does not control for practice effects changes the estimated time trend in cognitive impairment. Additional important differences between Choi et al.[26] and our analysis include that Choi et al.[26] provide a rich analysis of socioeconomic disparities in cognitive impairment, but do not analyze dementia, and they focus on a relatively young population (age 55–69).

In sum, once we take prior testing experience into consideration, we find no evidence for improving trends in prevalence of any cognitive impairment or dementia from 1996 to 2014. These results are remarkably robust to alternative modeling specifications. Although there are distinct challenges in modeling practice effects, researchers estimating trends in cognitive impairment based on panel data should account for prior experience. Otherwise, trends will be downwardly biased, which could be misinterpreted as a decline in prevalence. Our results showing increasing time trends in cognitive impairment and dementia indicate that the population-level burden of cognitive impairment may be underestimated.

REFERENCES

1. Alzheimer's Association. 2019 Alzheimer's disease facts & figures. *Alzheimers Dement*. 2019;15:321–387.
2. McGinnis JM. Mortality trends and signs of health progress in the United States: improving understanding and action. *JAMA*. 2015;314:1699–1700.
3. Manton KC, Gu XL, Ukraintseva SV. Declining prevalence of dementia in the U.S. elderly population. *Adv Gerontol*. 2005;16:30–37.
4. Langa KM, Larson EB, Karlawish JH, et al. Trends in the prevalence and mortality of cognitive impairment in the United States: is there evidence of a compression of cognitive morbidity? *Alzheimers Dement*. 2008;4:134–144.
5. Langa KM, Larson EB, Crimmins EM, et al. A comparison of the prevalence of dementia in the United States in 2000 and 2012. *JAMA Intern Med*. 2017;177:51–58.
6. Larson EB, Langa KM. What's the "Take Home" from research on dementia trends? *PLoS Med*. 2017;14:e1002236.
7. Matthews FE, Arthur A, Barnes LE, et al.; Medical Research Council Cognitive Function and Ageing Collaboration. A two-decade comparison of prevalence of dementia in individuals aged 65 years and older from three geographical areas of England: results of the Cognitive Function and Ageing Study I and II. *Lancet*. 2013;382:1405–1412.
8. Wu YT, Beiser AS, Breteler MMB, et al. The changing prevalence and incidence of dementia over time - current evidence. *Nat Rev Neurol*. 2017;13:327–339.
9. Qiu C, von Strauss E, Bäckman L, Winblad B, Fratiglioni L. Twenty-year changes in dementia occurrence suggest decreasing incidence in central Stockholm, Sweden. *Neurology*. 2013;80:1888–1894.
10. Hudomiet P, Hurd MD, Rohwedder S. Dementia prevalence in the United States in 2000 and 2012: estimates based on a nationally representative study. *J Gerontol Ser B*. 2018;73(suppl_1):S10–S19.
11. Satizabal CL, Beiser AS, Chouraki V, Chêne G, Dufouil C, Seshadri S. Incidence of dementia over three decades in the Framingham Heart Study. *N Engl J Med*. 2016;374:523–532.
12. Dodge HH, Zhu J, Hughes TF, et al. Cohort effects in verbal memory function and practice effects: a population-based study. *Int Psychogeriatr*. 2017;29:137–148.
13. Stern Y. The concept of cognitive reserve: a catalyst for research. *J Clin Exp Neuropsychol*. 2003;25:589–593.
14. Cummings JL, Vinters HV, Cole GM, Khachaturian ZS. Alzheimer's disease: etiologies, pathophysiology, cognitive reserve, and treatment opportunities. *Neurology*. 1998;51(1 suppl 1):S2–S17.
15. Jones RN, Manly J, Glymour MM, Rentz DM, Jefferson AL, Stern Y. Conceptual and measurement challenges in research on cognitive reserve. *J Int Neuropsychol Soc*. 2011;17:593–601.
16. Richards M, Deary IJ. A life course approach to cognitive reserve: a model for cognitive aging and development? *Ann Neurol*. 2005;58:617–622.
17. Lindeboom M, Portrait F, van den Berg GJ. Long-run effects of a nutritional shock early in life: the Dutch Potato famine of 1846-1847. *J Health Econ*. 2010;29:617–629.
18. Roseboom T, de Rooij S, Painter R. The Dutch famine and its long-term consequences for adult health. *Early Hum Dev*. 2006;82:485–491.
19. Hale JM. Cognitive disparities: the impact of the great depression and cumulative inequality on later-life cognitive function. *Demography*. 2017;54:2125–2158.
20. Suthers K, Kim JK, Crimmins E. Life expectancy with cognitive impairment in the older population of the United States. *J Gerontol B Psychol Sci Soc Sci*. 2003;58:S179–S186.
21. Cramer C, Haan MN, Galea S, Langa KM, Kalbfleisch JD. Use of statins and incidence of dementia and cognitive impairment without dementia in a cohort study. *Neurology*. 2008;71:344–350.
22. Weuve J, Proust-Lima C, Power MC, et al.; MELODEM Initiative. Guidelines for reporting methodological challenges and evaluating potential bias in dementia research. *Alzheimers Dement*. 2015;11:1098–1109.
23. Crimmins EM, Saito Y, Ki J, Kim JK. Change in cognitively healthy and cognitively impaired life expectancy in the United States: 2000–2010. *SSM Popul Heal*. 2016;2:793–797.
24. Crimmins EM, Saito Y. Trends in healthy life expectancy in the United States, 1970-1990: gender, racial, and educational differences. *Soc Sci Med*. 2001;52:1629–1641.
25. Crimmins EM, Saito Y, Kim JK, Zhang YS, Sasson I, Hayward MD. Educational differences in the prevalence of dementia and life expectancy with dementia: changes from 2000 to 2010. *J Gerontol Psychol Sci Soc Sci*. 2018;73:S20–S28.
26. Choi H, Schoeni RF, Martin LG, Langa KM. Trends in the prevalence and disparity in cognitive limitations of Americans 55-69 years old. *J Gerontol B Psychol Sci Soc Sci*. 2018;73(suppl_1):S29–S37.
27. Rabbitt P, Diggle P, Smith D, Holland F, Mc Innes L. Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. *Neuropsychologia*. 2001;39:532–543.
28. Rabbitt P, Diggle P, Holland F, McInnes L. Practice and drop-out effects during a 17-year longitudinal study of cognitive aging. *J Gerontol B Psychol Sci Soc Sci*. 2004;59:P84–P97.
29. Goldberg TE, Harvey PD, Wesnes KA, Snyder PJ, Schneider LS. Practice effects due to serial cognitive assessment: implications for preclinical Alzheimer's disease randomized controlled trials. *Alzheimers Dement (Amst)*. 2015;1:103–111.
30. Heilbronner RL, Sweet JJ, Attix DK, Krull KR, Henry GK, Hart RP. Official position of the American Academy of Clinical Neuropsychology on serial neuropsychological assessments: the utility and challenges

of repeat test administrations in clinical and forensic contexts. *Clin Neuropsychol*. 2010;24:1267–1278.

31. Duff K, Callister C, Dennett K, Tometich D. Practice effects: a unique cognitive variable. *Clin Neuropsychol*. 2012;26:1117–1127.

32. Wesnes K, Pincock C. Practice effects on cognitive tasks: a major problem? *Lancet Neurol*. 2002;1:473.

33. Shih, RA, Lee, J, Das, L. Harmonization of cross-national studies of aging to the health and retirement study cognition. Working Papers WR-861/7, RAND Corporation; 2011:1–101.

34. Wooden M, Li N. Panel conditioning and subjective well-being. *Soc Indic Res*. 2014;117:235–255.

35. Warren JR, Halpern-Manners A. Panel conditioning in longitudinal social science surveys. *Sociol Methods Res*. 2012;41:491–534.

36. Das M, Toepoel V, van Soest A. Nonparametric tests of panel conditioning and attrition bias in panel surveys. *Sociol Methods Res*. 2011;40:32–56.

37. Lazarsfeld PF. "Panel" studies. *Am Assoc Public Opin Res*. 1940;4:122–128.

38. French DP, Sutton S. Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it? *Br J Health Psychol*. 2010;15(pt 3):453–468.

39. Karlamangla AS, Miller-Martinez D, Aneshensel CS, Seeman TE, Wight RG, Chodosh J. Trajectories of cognitive function in late life in the United States: demographic and socioeconomic predictors. *Am J Epidemiol*. 2009;170:331–342.

40. Calamia M, Markon K, Tranel D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. *Clin Neuropsychol*. 2012;26:543–570.

41. Fong TG, Fearing MA, Jones RN, et al. Telephone interview for cognitive status: Creating a crosswalk with the Mini-Mental State Examination. *Alzheimers Dement*. 2009;5:492–497.

42. University of Michigan. *Health and Retirement Study Public Use Dataset*. 2017.

43. RAND Center for the Study of Aging. *RAND HRS Data, Version P*. 2017.

44. Fisher GG, Hassan H, Faul JD, Rodgers WL, Weir DR. *Health and Retirement Study Imputation of Cognitive Functioning Measures: 1992–2014*. Ann Arbor, MI; 2017.

45. Ghisletta P, Rabbitt P, Lunn M, Lindenberger U. Two thirds of the age-based changes in fluid and crystallized intelligence, perceptual speed, and memory in adulthood are shared. *Intelligence*. 2012;40:260–268.

46. Crimmins EM, Kim JK, Langa KM, Weir DR. Assessment of cognition using surveys and neuropsychological assessment: the Health and Retirement Study and the Aging, Demographics, and Memory Study. *J Gerontol B Psychol Sci Soc Sci*. 2011;66(suppl 1):i162–i171.

47. Langa KM, Kabeto MU, Weir D. Report on race and cognitive impairment using HRS. *2010 Alzheimer's Disease Facts and Fgures*. 2009.

48. Bugliari D, Campbell N, Chan C, et al. *RAND HRS Data Documentation, Version P*. 2016;1580.

49. Freedman VA, Aykan H, Martin LG. Aggregate changes in severe cognitive impairment among older Americans: 1993 and 1998. *J Gerontol Psychol Sci Soc Sci*. 2001;56:S100–S111.

50. Plassman BL, Langa KM, Fisher GG, et al. Prevalence of cognitive impairment without dementia in the United States. *Ann Intern Med*. 2008;148:427–434.

51. Prince M, Ali GC, Guerchet M, Prina AM, Albanese E, Wu YT. Recent global trends in the prevalence and incidence of dementia, and survival with dementia. *Alzheimers Res Ther*. 2016;8:23.

52. Wu YT, Fratiglioni L, Matthews FE, et al. Dementia in western Europe: epidemiological evidence and implications for policy making. *Lancet Neurol*. 2015;4422:1–9.

53. Rocca WA, Petersen RC, Knopman DS, et al. Trends in the incidence and prevalence of Alzheimer's disease, dementia, and cognitive impairment in the United States. *Alzheimers Dement*. 2011;7:80–93.

54. Wiberg P, Waern M, Billstedt E, Östling S, Skoog I. Secular trends in the prevalence of dementia and depression in Swedish septuagenarians 1976–2006. *Psychol Med*. 2013;43:2627–2634.

55. Pérès K, Brayne C, Matharan F, et al. Trends in prevalence of dementia in French farmers from two epidemiological cohorts. *J Am Geriatr Soc*. 2017;65:415–420.

56. Parker MG, Ahacic K, Thorslund M. Health changes among Swedish oldest old: prevalence rates from 1992 and 2002 show increasing health problems. *J Gerontol A Biol Sci Med Sci*. 2005;60:1351–1355.

57. Freedman VA, Aykan H, Martin LG. Another look at aggregate changes in severe cognitive impairment: further investigation into the cumulative effects of three survey design issues. *J Gerontol B Psychol Sci Soc Sci*. 2002;57:S126–S131.

58. Dodge HH, Wang CN, Chang CC, Ganguli M. Terminal decline and practice effects in older adults without dementia: the MoVIES project. *Neurology*. 2011;77:722–730.

59. Gross AL, Benitez A, Shih R, et al. Predictors of retest effects in a longitudinal study of cognitive aging in a diverse community-based sample. *J Int Neuropsychol Soc*. 2015;21:506–518.

60. Salthouse TA. Influence of age on practice effects in longitudinal neurocognitive change. *Neuropsychology*. 2010;24:563–572.

61. Vivot A, Power MC, Glymour MM, et al. Jump, hop, or skip: modeling practice effects in studies of determinants of cognitive change in older adults. *Am J Epidemiol*. 2016;183:302–314.

62. Machulda MM, Pankratz VS, Christianson TJ, et al. Practice effects and longitudinal cognitive change in normal aging vs. incident mild cognitive impairment and dementia in the Mayo Clinic Study of Aging. *Clin Neuropsychol*. 2013;27:1247–1264.

63. Bartels C, Wegrzyn M, Wiedl A, Ackermann V, Ehrenreich H. Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing. *BMC Neurosci*. 2010;11:118.

64. Sánchez-Benavides G, Gispert JD, Fauria K, Molinuevo JL, Gramunt N. Modeling practice effects in healthy middle-aged participants of the Alzheimer and Families parent cohort. *Alzheimers Dement (Amst)*. 2016;4:149–158.

65. Lamar M, Resnick SM, Zonderman AB. Longitudinal changes in verbal memory in older adults: distinguishing the effects of age from repeat testing. *Neurology*. 2003;60:82–86.

66. Duff K, Chelune G, Dennett K. Within-session practice effects in patients referred for suspected dementia. *Dement Geriatr Cogn Disord*. 2012;33:245–249.

67. Weir DR, Faul JD, Langa KM. Proxy interviews and bias in the distribution of cognitive abilities due to non-response in longitudinal studies: a comparison of HRS and ELSA. *Longit Life Course Stud*. 2014;2:170–184.