Subtle mistakes in self-report surveys predict future transition to dementia

Stefan Schneider[1,2], Doerte U. Junghaenel[1,2], Elizabeth M. Zelinski[2], Erik Meijer[3], Arthur A. Stone[1,4], Kenneth M. Langa[5], Arie Kapteyn[3]

[1]Dornsife Center for Self-Report Science & Center for Economic and Social Research, University of Southern California, CA, USA

[2]Leonard Davis School of Gerontology, University of Southern California, CA, USA

[3]Dornsife Center for Economic and Social Research, University of Southern California, CA, USA

[4]Deparment of Psychology, University of Southern California, CA, USA

[5]Department of Internal Medicine, Institute for Social Research and VA Center for Clinical Management Research, University of Michigan, Ann Arbor, MI, USA

Corresponding author:
Stefan Schneider, Ph.D.
Dornsife Center for Self-Report Science
and Center for Economic & Social Research
University of Southern California
635 Downey Way
Los Angeles, CA 90089-3332
Email: schneids@usc.edu

Conflicts of interest

All authors received support from National Institute on Aging grant R21AG061364 for the present manuscript. In the past 36 months, outside the submitted work, Dr. Schneider was supported by National Institute on Aging grants R01AG068190 and R21AG061364. Dr. Junghaenel was supported by National Institute of Arthritis and Musculoskeletal Diseases grant R21AR074020 and National Institute on Aging grant R21AG061364. Dr. Zelinski received consulting fees for NIA-sponsored research Blue Marble Health, received honoraria for participation in NIH study sections, and participated on two data safety and monitoring boards (both funded by NIA grants), one at the University of Florida and the other at Posit Science corporation. Dr. Meijer was supported by grants from the NIH and the Social Security Administration (SSA), and received royalties from Springer for a book he co-edited in 2008. Dr. Stone was supported by NIH grants R01AG051903, U01AG054580, K01AG050811, R37AG057685, R01AG062277, R21AG060447, R21AR074020, R01AG067624, R01AG068190, received consulting fees from the Gallup organization, IQVia, CiVi Pharmaceutical, and Adelphi Values, received honoraria from Stanford University and from the University of Copenhagen, and received support for travel from Adelphi Values for attending meetings. Dr. Langa was supported by NIH grants paid to the University of Michigan, received consulting fees from Harvard University, Johns Hopkins University, Michigan State University, University of Southern California, Peking University, and Dartmouth University, received payment for expert testimony from a legal firm regarding decisional capacity for an individual with cognitive impairment, and participated on an advisory board at Indiana University. Dr. Kapteyn was supported by NIH grants U01AG054580 and R01AG067624, and received payments for travel pertaining to the Survey of Health, Aging, and Retirement in Europe (SHARE).

Abstract

INTRODUCTION: We investigate whether indices of subtle reporting mistakes derived from responses in self-report surveys are associated with dementia risk.

METHODS: We examined 13,831 participants without dementia from the prospective, population-based Health and Retirement Study (mean age 69±10 years, 59% women). Participants' response patterns in 21 questionnaires were analyzed to identify implausible responses (multivariate outliers), incompatible responses (Guttman errors), acquiescent responses, random errors, and the proportion of skipped questions. Subsequent incident dementia was determined over up to 10 years of follow-up.

RESULTS: During follow-up, 2074 participants developed dementia and 3717 died. Each of the survey response indices was associated with future dementia risk controlling for confounders and accounting for death as a competing risk. Stronger associations were evident for participants who were younger and cognitively normal at baseline.

DISCUSSION: Mistakes in the completion of self-report surveys in longitudinal studies may be early indicators of dementia among middle-aged and older adults.

Keywords

Cognitive impairment; dementia; early detection; epidemiology; functional abilities; longitudinal; population-based; prospective; self-report surveys; survey response behaviors.

1. Background

Identifying preclinical markers that are predictive of future transition from healthy cognition to mild cognitive impairment and dementia is of paramount importance [1,2]. Earlier detection of cognitive decline could facilitate delays in dementia onset or progression once effective interventions are available, which could have a significant impact on incidence rates, quality of life, and health care costs. Next to a range of genetic and biological markers [3,4], decrements in everyday functional abilities are among the earliest and strongest signals that predict future dementia [4-9]. Subtle reductions in the efficiency, speed, and consistency of performing instrumental activities of daily living (IADLs) and other cognitively demanding tasks have been observed up to a decade before diagnosis [1,7].

Assessment of reductions in functional abilities in research on older populations has proven challenging. Most population-based longitudinal studies rely on subjective ratings of daily instrumental functioning. Although subjective performance ratings have proven useful in dementia research, it is widely acknowledged that the accuracy of these ratings can be impacted by memory and other biases and that they are not precise indicators of actual performance [10-12]. Conversely, available objective, behavior-based tools for assessing functional abilities (e.g., errors on standardized tests of goal-directed tasks such as using the telephone, making toast and coffee, etc.) are difficult to implement, burdensome for respondents, and costly due to specialist time and equipment involved [13,14].

In view of these challenges, recent research has recognized the enormous potential of developing objective yet cost-effective indicators of functional abilities that can be reliably inferred from commonly occurring behaviors. Most notably, accumulating evidence from studies that passively monitored computer use behavior via electronic data capture suggests that cognitively impaired older adults show less consistent engagement in computer use [15,16], less efficient mouse movements [17], more irregular keystroke behavior [18], and increased latency to complete online questionnaires [19], compared to adults with normal cognition.

Here, we examine the possibility that objective indicators of functioning that are sensitive to cognitive decline can be gleaned directly from the way people complete survey assessments. The rationale for these indicators is that completing a questionnaire or survey is in itself a complex and cognitively demanding task that requires attention, working memory, executive functioning, and short- and longer-term memory [20,21], comparable to the demands of other complex instrumental activities. Prior research suggests that individuals with cognitive deficits are more likely to display suboptimal response patterns with more subtle mistakes, including more skipped questions and inconsistent or implausible answer patterns [21-26]. Building on this prior research, we examine whether such survey response patterns can serve as early indicators of future dementia onset.

Admittedly, completing a survey is a small and uncommon slice of everyday functioning. However, self-report surveys represent a large component of population-based cohort studies that are a major resource for scientific knowledge about the epidemiology and etiology of dementia. Behavior-based functioning indicators that could be derived from existing survey data could capitalize on past and ongoing longitudinal studies, allowing predictions of incident dementia from objective functioning indicators collected many years earlier. Such indicators could contribute substantially to more comprehensive strategies for dementia detection from archival data [27-29]. In this study, we investigate whether older adults' survey response patterns predict subsequent incidence of dementia over a 10-year period in the Health and Retirement Study (HRS), accounting for death as a competing risk.

2. Methods

Study setting and population

The HRS is a longitudinal panel study of a US nationally representative sample of adults above 50 years of age that started in 1992 (http://hrsonline.isr.umich.edu/). Respondents are repeatedly interviewed every 2 years. The Psychosocial and Lifestyle Questionnaire (PLQ), a paper-and-pencil self-report survey, was introduced in the 2006 and 2008 waves (piloted in 2004). It was administered to a (mutually exclusive) random 50% of the HRS sample in each of the two waves, which served as baseline waves for the present analyses. Respondents were given the PLQ at the end of the HRS face-to-face interview to be returned in the mail, with response rates of 90% (in 2006) and 89% (in 2008) of those who completed the interview [30]. Our analyses included all respondents who completed the PLQ by themselves (excluded were 1.1% completed by proxy respondents). Out of 13,831 analyzed respondents, 13,448 (97.2%) completed the PLQ on paper and returned it by mail, and 383 (2.8%) completed it with an interviewer over the phone. Non-respondents and excluded participants were 11% more likely to have dementia at baseline, and 5% less likely to be female, but did not differ in age compared to those analyzed. Included participants were followed after PLQ completion until the onset of dementia, death, loss to follow-up, or the 2016 HRS interview. All participants provided informed consent as part of the HRS.

Indices derived from survey responses

Five different indices of participants' survey response patterns were derived from the PLQ. A common feature of all indices is that they focus on *how* individuals complete these surveys rather than the content being sought by the questions – that is, they reflect different types of response *behaviors*. The indices are described in Table 1 (see online supplement for statistical details). They included (a) skipping questions (item nonresponse), (b) inconsistent responses (random response errors), (c) implausible response patterns (multivariate outliers, i.e., unusual combinations of scores across PLQ items), (d) incompatible responses ("Guttman errors"), and (e) agreeing with statements regardless of content (acquiescence). Each of these response patterns have previously been associated with impaired cognitive functioning and suboptimal information processing [22-26].

We derived the indices from 102 questions included in 21 reliable and valid multi-item PLQ scales that were administered both in 2006 and 2008 (for psychometric information and internal consistency reliabilities of each scale, see [30]). We did not use PLQ portions that were modified across the two waves or were applicable only to respondent subgroups (i.e., questions about respondents' jobs, spouse or children were excluded). Included questionnaires comprised a range of constructs commonly assessed in psychosocial research (e.g., life satisfaction, anxiety, personality; see supplemental Table S1).

Dementia status

We ascertained dementia status using the criteria by Langa and Weir, which were developed for the HRS to classify respondents as either having normal cognition, cognitive impairment—no dementia (CIND), or dementia [31,32]. For self-respondents, the classification is based on cognitive tests of immediate and delayed free recall, serial seven subtractions, and backwards counting from 20, administered in the HRS cognitive battery [31], with respondents scoring 0-6 on a 28-point scale classified as having dementia, 7-11 as CIND, and 12-27 as normal. We also utilized information from proxy respondents to reduce sample attrition, where dementia categorization is based on proxy-reported respondent memory, proxy-reported IADL problems, and interviewer-assessed cognitive limitations; respondents scoring 0-2 on a 12-point summary limitations scale were classified as normal, 3-5 as CIND, and 6-11 as having dementia [31,32]. Missing scores on the subtests for dementia categorization were accommodated using imputations provided for the HRS cognitive tests [33] and proxy reports [34]. The classification cut-points have been identified using data from the Aging, Demographics, and Memory Study (ADAMS), an HRS sub-study in which clinical diagnoses were obtained by means of 3–4 hour in-home neuropsychological and clinical assessments together with expert clinician adjudication [31]. Using the ADAMS dementia diagnosis as gold-standard, the categorization correctly classifies 78% of HRS respondents (76% for self-respondents and 84% for proxy-respondents) [31,32].

Covariates and competing risk of death

The selection of covariates was based on potential confounders of the effects of functional abilities reflected in the survey response patterns [35]. We included the demographic variables age, sex, race (White, African American, other races), ethnicity, marital status, years of education, and wealth (quintiles); health variables, including smoking (smokes now, smoked in the past, never smoked), drinking (never drinks, <8 drinks per week, 8+ drinks per week), body mass index (BMI categories underweight, normal, overweight, obese), and exercise (less than once/month, 1-4 times/month, more than once/week); and physical conditions, including hypertension, diabetes, heart disease, and stroke. Measurement of these covariates took place at HRS interviews before participants were given the PLQ. Additionally, we statistically controlled for participants' scale scores on each of the 21 PLQ questionnaires from which the survey response patterns were derived.

To account for the competing risk of death, mortality data were coded from the HRS exit interview or spouse-reported year of death information. The month of death was recorded up until the end of year 2016, at which point the study was right-censored.

Self-reported functional limitations

Limitations in IADLs were identified by self-reports of difficulties using a telephone, taking medication, handling money, shopping, and preparing meals (score range =0-5). These measures from the baseline HRS interviews were included to juxtapose the prognostic accuracy of the survey response patterns against an established measure of functional limitations known to predict dementia risk [4-9].

Statistical analysis

In initial cross-sectional analyses, we compared the mean scores on each response pattern by current dementia status (i.e., respondents concurrently classified as having normal cognition, CIND, or dementia) using univariate ANOVAs.

Cox proportional hazards regression models were used to examine relationships between each of the survey response patterns (entered as independent variable in separate models) and subsequent incident dementia (dependent variable, considered an absorbing state after first being observed in a given wave); respondents with dementia at baseline were excluded from these models. Model 1 adjusted for age and sex, entered as covariates in the Cox regression models. Model 2 adjusted for age, sex, race, ethnicity, marital status, education, wealth, smoking, drinking, BMI, exercise, hypertension, diabetes, heart disease, stroke, and each of the 21 PLQ scale scores as covariates. Model 3 additionally accounted for death as a competing event using Fine and Gray's proportional subdistribution hazards regression model [36]. Inspecting the Schoenfeld residuals did not indicate violations of the proportional hazards assumption for predictors or covariates in the models. To control for Type I error inflation due to multiple comparisons, statistical significance was evaluated at a Bonferroni-corrected level of P <.003, adjusted for 5 parameters across 3 models (P =.05/15). The primary models tested linear associations of the survey response indices; potential curvilinearities were explored by adding quadratic terms.

In sensitivity analyses, we excluded respondents with more than 10% missing values on the PLQ from Model 3 (N =932; 6.7%). Further, we conducted analyses stratified by the year of PLQ administration (year 2006 vs 2008, to evaluate potential period effects [37]), and by respondents who completed the PLQ on paper versus on the phone (to evaluate mode of administration effects [38]); respective group differences in associations between response patterns and dementia incidence were evaluated using interaction terms.

Age, sex, and cognitive status (cognitively normal vs CIND) at baseline were evaluated as potential moderators by testing their interaction with the response patterns in Model 3. For moderated effects by age, we present age-stratified results (≤75 vs >75 years) and used age as a continuous moderator variable for significance testing. Statistical significance of moderated effects was evaluated using a Bonferroni-corrected level of P <.003, adjusted for 5 parameters across 3 moderators (P =.05/15).

To quantify the capacity of the response patterns to serve as prognostic markers of the risk of developing dementia, we estimated time-dependent receiver operating characteristic (ROC) curves [39]. For each response pattern, we examined time-dependent area under the ROC curve (AUC) statistics for increasingly longer epochs of follow-up time from baseline (PLQ assessment to 2 years, 4 years, and so on), to evaluate the evolution of their prospective discriminatory abilities over time. Corresponding AUC statistics based on IADL self-reports were also obtained for descriptive comparison. AUCs were estimated using the cumulative/dynamic definition by Heagerty et al. [39] and accounting for death as competing risk [40].

In all analyses, the survey response patterns were standardized to facilitate comparison. Missing values on covariates (median 1.2% missing [range 0%-3.9%]) were imputed using five multiple imputations. In post-hoc analyses, we additionally imputed scores on response patterns for HRS respondents who failed to complete the PLQ (using five imputations, based on observed covariates and dementia outcomes). Statistical analyses were performed using the PHREQ and MI/MIANALYZE procedures in SAS 9.4 (Cary, NC). The R package timeROC was used for ROC analyses [41].

3. Results

Descriptive sample characteristics are shown in Table 2. During up to 10 years of follow-up (median =8.0 years), 2074 (15.0%) individuals developed dementia, and 3717 (26.9%) died. The total observed follow-up time was 93,886 person-years, 89.4% of the possible total time without loss to follow-up [42].

Cross-sectional associations between survey response patterns and dementia categories

The 5 survey response patterns were positively correlated with each other, ranging from r =.11 to r =.88 (Table 3). Mean scores on each of the response patterns differed significantly by baseline cognitive status (P <.0001), with mean differences ranging from .25

to .63 *z*-scores when comparing participants categorized as CIND versus cognitively normal, from .17 to .28 *z*-scores comparing participants with dementia versus CIND, and from .42 to .88 *z*-scores comparing participants showing dementia versus normal cognition at baseline (Table 3).

<u>Associations between survey response patterns and subsequent dementia risk</u>

In proportional hazard regression models adjusting for age and sex (Model 1), higher values on each of the 5 response patterns were associated with a significantly greater risk of developing dementia (P <.0001), with hazard ratios (HRs) ranging from HR =1.22 (95% confidence interval [CI] =1.16-1.28) per standard deviation (SD) increase in acquiescent responses to HR =1.82 (95% CI =1.73-1.92) per SD increase in multivariate outlier responses (Table 4). The estimates were attenuated after adjusting for all covariates (Model 2) and when death as a competing risk was additionally accounted for (Model 3); however, the associations with incident dementia remained significant for each of the survey response patterns in these models (P <.0001, Table 4).

We found no significant curvilinear effects after fitting quadratic terms for item nonresponse (P =.17), random response errors (P =.61), multivariate outlier responses (P =.54), Guttman errors (P =.96), and acquiescence (P =.57). When analyses were restricted to respondents with <10% missing values on the PLQ, item non-response was no longer significantly associated with incident dementia after Bonferroni correction (P =.007); estimates for the remaining four response patterns were not meaningfully affected (supplemental Table S2). Results were similar to those in the primary analyses when scores on response patterns for respondents who failed to complete the PLQ were multiply imputed, and the estimates did not significantly differ between PLQ administration years (2006 vs 2008) and when comparing respondents completing the PLQ on paper versus via phone (supplemental Tables S3-S5).

Effect moderation by age, sex, and baseline cognitive status

Examining moderated effects by age yielded significant interactions between age and indices of item nonresponse, random errors, multivariate outliers, and Guttman errors (*P* <.0001). Higher values on these response patterns were consistently associated with increased dementia risk for respondents aged <75 years at baseline, whereas the associations were less pronounced at ages 75+ (Table 5). Baseline cognitive status similarly moderated the effects of these response patterns (*P* <.0001); associations with dementia risk were significant for cognitively normal respondents and non-significant for individuals with CIND at baseline (see Table 6). We found no significant interactions with sex (supplemental Table S6).

Time-dependent prognostic accuracies of dementia risk

AUCs quantifying the ability of the response patterns to predict the onset of dementia are shown in Figure 1. AUCs remained similar for increasingly longer epochs of follow-up time. The highest prognostic accuracy was evident for multivariate outlier responses, with AUCs attenuating slightly from 2 years (AUC =.70, 95% CI =.67-.72) to 10 years (AUC =.68, 95% CI =.66-.69), and it was overall lowest for acquiescent responses, with slight increases 2 years (AUC =.53, 95% CI =.51-.56) to 10 years (AUC =.56, 95% CI =.54-.58). For comparison, prognostic accuracies for IADL reports ranged between these values, with AUCs attenuating from 2 years (AUC =.62, 95% CI =.60-.64) to 10 years (AUC =.59, 95% CI =.58-.60).

4. Discussion

Self-report surveys are ubiquitous in longitudinal studies on aging. Our results indicate that subtle mistakes in self-report surveys are meaningfully associated with cognitive impairment and cognitive decline. Cross-sectionally, each of the investigated response patterns discriminated cognitively normal respondents from those with CIND (small to medium effect sizes) and those classified as having dementia (medium to large effect sizes) at baseline [43]. Prospectively, all response patterns predicted the risk of developing dementia with stable prognostic accuracies over up to 10 years of follow-up. The results are in line with prior research demonstrating that early signals of incident dementia can be discerned from characteristic features of individuals' responses to cognitively demanding questions, recorded many years prior [28].

Of the five response patterns examined, four (item nonresponses, random errors, multivariate outliers, Guttman errors) demonstrated very similar prognostic accuracies, comparable to or higher than those for self-reported IADL deficits. While loss of independence and major IADL limitations represent important disease milestones, the survey response patterns may result from functional limitations that predate disability and may develop gradually and early in the disease process [4-9]. This assumption is supported by our finding that IADL deficits had the strongest prognostic capability when they were assessed close to diagnosis, whereas the prognostic capability of the response patterns remained more consistent for increasingly long time-windows. The fifth response pattern, acquiescence, was overall a less accurate indicator of dementia risk. Arguably, acquiescence is to a larger extent driven by general tendencies in self-reporting (e.g., due to personality and cultural norms [44]) that are not inherently related to cognitive or functional abilities.

Alternative explanations for the observed relationships are also possible. Worse general biological trajectories may commonly underlie both suboptimal survey response behaviors and dementia risk. Many of the risk factors associated with dementia also predict an earlier death [45], but our analyses accounted for the fact that participants who might have had the most severe risks of developing dementia are likely to have died before any dementia diagnosis. Mistakes in survey responding have also been associated with mood disorders such as depression [24]; however, it is unlikely that the effects were driven by mental health problems given that multiple mental health measures from the PLQ were statistically controlled. We also cannot rule out that selection bias influenced our findings, although results remained very similar when data from those who did not complete the PLQ were multiply imputed.

Age and baseline cognitive status moderated the associations between each response pattern (except acquiescence) and dementia risk, with stronger associations for younger (vs older) and cognitively normal (vs CIND) respondents. Individuals with CIND may already face more obvious functioning deficits, such that subtle mistakes when performing cognitively demanding tasks may be less prognostically relevant at this stage. Respondents with dementia at baseline were excluded from analysis, and older participants who might have shown subtle functional limitations prior to diagnosis may have been less represented due to selective survival effects. We also cannot rule out that older respondents and those with early and mild cognitive impairment received assistance with completing the PLQ from others at home, which may have led to the obfuscation of associations between survey response patterns and future dementia in these respondents.

Our study has several limitations. Dementia status was derived from a limited set of cognitive tests and informant reports. Although validation studies have demonstrated 78% accuracy of dementia diagnoses based on these tests compared to detailed clinical evaluation in ADAMS [31], the results need to be replicated using clinically confirmed dementia diagnoses and extended to dementia subtypes. We also did not examine which specific cognitive abilities are being tapped by the different response patterns. We speculate that they may capture behavioral manifestations of multiple cognitive functions involved in goal-directed activities (e.g., “everyday cognition” [16,46]), including remembering the details of questionnaire instructions, consistently attending to the details of each question in deciding the best answer, flexibly adapting responses to changing answer formats, and sustaining effort to complete all questions. Furthermore, even though we did not find differential relationships by mode of survey administration, the investigated response patterns were for the largest part limited to paper-and-pencil assessments, and it is not clear whether the results generalize to other survey modalities (e.g., in-person interviews, online surveys) [38].

Although the present research focused on survey responses in longitudinal aging research, it may be possible to adapt the presented approach to survey responses in other settings. For example, in medical care settings, response patterns extracted from surveys routinely administered during check-in for appointments could potentially supplement information from standardized cognitive tests. In clinical trial research, response patterns extracted from health questionnaires might supplement functioning measures that serve as trial end-points, in line with FDA recommendations for early-stage Alzheimer's disease trials that encourage the development of novel approaches for the evaluation of early functional deficits [47]. These are avenues for future research.

In conclusion, our findings demonstrate that mistakes in the completion of self-report surveys in longitudinal studies may be early indicators of dementia among middle-aged and older adults. Work is underway to evaluate the prognostic ability of the survey response patterns in multiple national and international panel studies that administer self-report questionnaires across different study populations, languages, survey types, and administration modes.

Funding information

This work was supported by a grant from the National Institute on Aging (R01AG068190). The Health and Retirement Study is funded by the National Institute on Aging (U01 AG009740) and the Social Security Administration, and performed at the Institute for Social Research, University of Michigan.

Author contributions

Stefan Schneider and Doerte U. Junghaenel conceptualized the study aims and designed the study. Stefan Schneider analyzed the data. Elizabeth M. Zelinski, Arthur A. Stone, and Arie Kapteyn contributed to the conceptualization of the design, and critically revised the manuscript. Erik Meijer and Kenneth Langa provided technical and statistical expertise and revised the manuscript. All authors contributed to interpreting the data and critically revising it for important intellectual content. All authors read and approved the final manuscript.

Table 1: Definitions of the survey response patterns

Survey response pattern	Expected response behavior	Interpretation of the response pattern	Operationalization
Item non-response	Respondent should complete all survey items	Overt disengagement from response process [22]	Proportion of items skipped (missing values) by respondent
Random response errors	Answers should be internally consistent, whereby scores for items addressing the same concept are more similar than scores for items addressing different concepts	Random variability in attention or fluctuating cognitive performance [48]	Magnitude of random variance around the respondent's "true" scale scores, estimated from multilevel models
Multivariate outlier responses	Profile of a respondent's scores across all items should not overly deviate from the majority in the sample	Overall profile of responses is implausible (i.e., statistically unlikely), suggesting that some answers were made by mistake [24]	Mahalanobis distance of respondent's scores across all questionnaire items
Guttman errors	If a respondent endorses an item that expresses a strong opinion toward an object, items that express weaker opinions toward that same object should be endorsed at the same or higher levels	Incompatible responses to questions on the same scale (e.g., responding that one [a] is able to run a mile, and [b] cannot walk a short distance), suggesting incoherent processing of	Normed Guttman errors [49] calculated for each questionnaire

		the questions	
Acquiescent responses	Respondents are expected to engage in initial comprehension and subsequent reevaluation processes for each question [50], where comprehension involves tacit acceptance of the premise (akin to "yes, I understand"), and reevaluation involves deciding on the optimal answer	"Yea-saying" regardless of item content; suggesting tacit acceptance of a statement without cognitive efforts to reevaluate the response [26,50]	Two-factorial nominal response model separating acquiescent and substantive response factors in each questionnaire [23]

Table 2: Characteristics of the study sample at baseline

Characteristic (units)	Values	Sample size
Age in years (mean, *SD*)	69.2 (9.9)	13,831
Female	59.1%	13,831
Race		13,830
White	82.8%	
African American	13.1%	
Other race	4.1%	
Hispanic	8.0%	13,830
Married	63.1%	13,830
Years of education		13,813
0-11 years	21.5%	
12 years	34.7%	
13-15 years	21.8%	
<15 years	22.0%	
Wealth quartiles		13,831

< $52,100	23.6%	
$52,100 - $204,900	24.9%	
$205,000 - $547,000	25.7%	
>547,000	25.9%	
Smoking status		13,735
Smokes now	12.9%	
Smoked in the past	43.9%	
Never smoked	43.2%	
Drinking status		13,807
Heavy drinkers (8+ drinks/week)	9.6%	
Light drinkers (<8 drinks/week)	41.4%	
Never drinks	49.0%	
Body mass index categories		13,668
Underweight (<18.5 kg/m^2)	1.3%	
Normal weight (18.5 - 24.9 kg/m^2)	28.7%	
Overweight (25.0 - 29.9 kg/m^2)	38.1%	
Obese (BMI 30+ kg/m^2)	31.9%	
Exercise		13,821
Never exercises	62.0%	
Exercises 1-4 times/month	14.7%	
Exercises more than once/week	23.3%	
Hypertension	56.7%	13,724
Diabetes	19.9%	13,657
Heart disease	24.1%	13,664
Stroke	7.9%	13,638
IADL limitations (mean, *SD*)	0.21 (.68)	13,830

Table 3: Cross-sectional associations of survey response patterns with baseline dementia status

Survey response pattern	Intercorrelations: Item non-response	Random response errors	Multivariate outlier responses	Guttman errors	Mean (*SD*) by dementia category: Normal cognition (N=11071)	CIND (N=2253)	Dementia (N=507)	Mean differences (Tukey-corrected 95% CIs): CIND vs. normal cognition	Dementia vs. CIND	Dementia vs. normal cognition
Item non-response	--				-.11 (.83)	.40 (1.32)	.69 (1.71)	.52 (.46-.57)	.28 (.17-.40)	.80 (.70-.90)
Random response errors	.21	--			-.12 (.95)	.44 (1.08)	.61 (1.12)	.56 (.51-.61)	.17 (.05-.28)	.72 (.62-.83)
Multivariate outlier responses	.22	.88	--		-.13 (.96)	.49 (.97)	.75 (.93)	.63 (.57-.68)	.25 (.14-.36)	.88 (.78-.98)
Guttman errors	.17	.79	.81	--	-.11 (.94)	.41 (1.11)	.60 (1.09)	.53 (.47-.58)	.19 (.08-.30)	.72 (.61-.82)
Acquiescent responses	.11	.47	.32	.36	-.06 (.95)	.19 (1.14)	.37 (1.25)	.25 (.19-.30)	.18 (.06-.29)	.42 (.32-.53)

Note: Survey response patterns are expressed as z-scores (mean=0, SD=1 in the full sample) to compare means across indices with different units. CIND = cognitively impaired, not demented; SD = standard deviation; CI = confidence interval.

Table 4: Associations of survey response patterns with incident dementia

	Model 1	Model 2	Model 3
Survey response pattern	HR (95% CI)	HR (95% CI)	HR (95% CI)
Item non-response	1.54 (1.44-1.64)	1.17 (1.10-1.25)	1.16 (1.08-1.24)
Random response errors	1.63 (1.55-1.71)	1.24 (1.18-1.32)	1.18 (1.12-1.25)
Multivariate outlier responses	1.82 (1.73-1.92)	1.38 (1.30-1.48)	1.30 (1.23-1.39)
Guttman errors	1.58 (1.51-1.65)	1.24 (1.18-1.31)	1.18 (1.12-1.24)
Acquiescent responses	1.22 (1.16-1.28)	1.21 (1.15-1.28)	1.15 (1.09-1.21)

Note: Respondents with dementia at baseline were excluded, N=13,324. Model 1 is adjusted for age and sex. Model 2 is additionally adjusted for race, ethnicity, marital status, education, wealth, smoking, drinking, BMI, exercise, hypertension, diabetes, heart disease, stroke, and the 21 Psychosocial and Lifestyle Questionnaire scale scores. Model 3 additionally accounts for death as a competing event. Hazard ratios (HRs) were obtained with Cox regression models in Models 1 and 2, and with Fine and Gray's proportional subdistribution hazards regression models in Model 3. HRs above 1.00 denote that the hazards of dementia increase with a higher value of the survey response pattern. To compare HRs across indices of survey response patterns with different units, HRs are expressed per standard deviation difference in the survey response pattern. HR = hazard ratio; CI = confidence interval.

Table 5: Associations of survey response patterns with dementia risk moderated by baseline age and baseline cognitive status

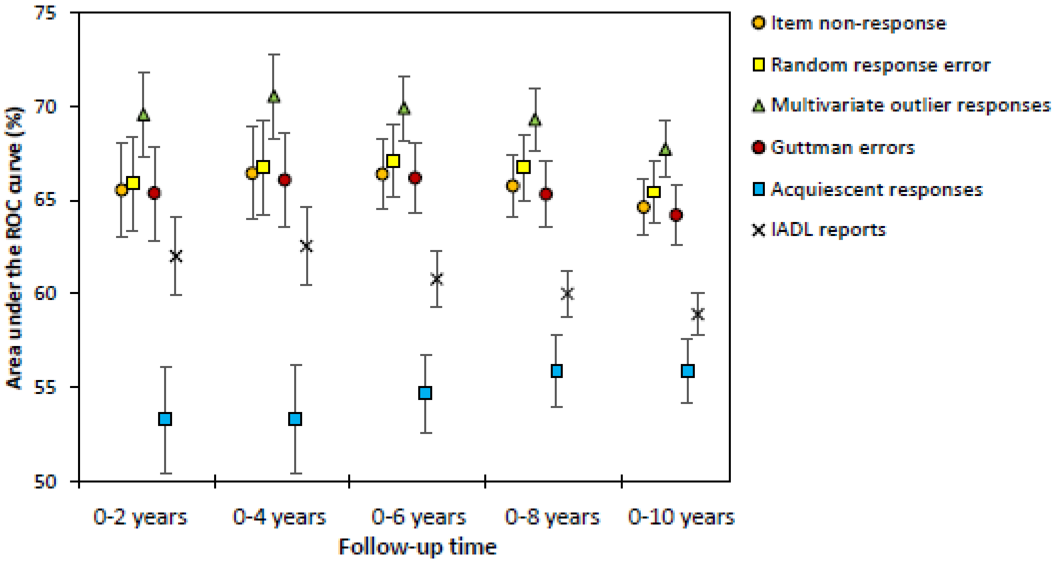
	Moderator: baseline age			Moderator: baseline cognitive status		
	Age <75 years	Age 75+ years	Interaction	Normal cognition	CIND	Interaction
Survey response	HR (95%	HR (95%	*P* value	HR (95%	HR (95%	*P* value

pattern	CI)	CI)		CI)	CI)	
Item non-response	1.28 (1.18-1.38)	1.09 (1.00-1.18)	<.0001	1.29 (1.20-1.40)	1.04 (0.96-1.12)	<.0001
Random response errors	1.32 (1.21-1.43)	1.04 (0.97-1.12)	<.0001	1.25 (1.15-1.37)	1.04 (0.96-1.12)	<.0001
Multivariate outlier responses	1.46 (1.32-1.60)	1.11 (1.02-1.21)	<.0001	1.37 (1.25-1.50)	1.09 (0.99-1.20)	<.0001
Guttman errors	1.29 (1.21-1.38)	1.03 (0.97-1.10)	<.0001	1.25 (1.16-1.34)	1.03 (0.97-1.10)	<.0001
Acquiescent responses	1.17 (1.09-1.25)	1.09 (1.02-1.17)	.08	1.13 (1.05-1.21)	1.05 (0.99-1.13)	.14

Note: Respondents with dementia at baseline were excluded, N=13,324. Hazard ratios (HRs) were obtained from Fine and Gray's proportional subdistribution hazards regression models, accounting for death as a competing event, and adjusted for continuous age at baseline, sex, race, ethnicity, marital status, education, wealth, smoking, drinking, BMI, exercise, hypertension, diabetes, heart disease, stroke, and the 21 Psychosocial and Lifestyle Questionnaire scale scores. We tested the significance of age interactions through modeling a product term of the unstandardized response patterns with continuous age. HRs are expressed per standard deviation difference in the survey response pattern. HR = hazard ratio; CI = confidence interval; CIND = cognitively impaired, not demented.

Figure Captions

Figure 1: Time-dependent area under the receiver operating characteristic (ROC) curves for survey response patterns as prognostic markers of dementia risk. *Note*: Area under the ROC curve (AUC) statistics are shown for increasingly long follow-up times within the study timeframe, using the cumulative/dynamic definition of time-dependent AUCs and accounting for death as competing risk. AUC values based on instrumental activities of daily living (IADL reports) are shown to place the predictive accuracy of the survey response patterns in the context of an established early marker of progression to dementia. Error bars represent 95% confidence intervals. ROC = receiver operating characteristic; IADL = instrumental activities of daily living.



References

[1] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer’s disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & dementia. 2011;7:280-92.

[2] Cummings J, Aisen PS, DuBois B, Frölich L, Jack CR, Jones RW, et al. Drug development in Alzheimer’s disease: the path to 2025. Alzheimer's research & therapy. 2016;8:39.

[3] Fiandaca MS, Mapstone ME, Cheema AK, Federoff HJ. The critical need for defining preclinical biomarkers in Alzheimer's disease. Alzheimer's & Dementia. 2014;10:S196-S212.

[4] Edmonds EC, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW. Subtle cognitive decline and biomarker staging in preclinical Alzheimer’s disease. Journal of Alzheimer's disease. 2015;47:231-42.

[5] Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C, Davies P, Goldberg TE, Initiative AsDN. Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's disease neuroimaging initiative. Archives of general psychiatry. 2011;68:961-9.

[6] Tarnanas I, Tsolaki A, Wiederhold M, Wiederhold B, Tsolaki M. Five-year biomarker progression variability for Alzheimer's disease dementia prediction: Can a complex instrumental activities of daily living marker fill in the gaps? Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2015;1:521-32.

[7] Verlinden VJ, van der Geest JN, de Bruijn RF, Hofman A, Koudstaal PJ, Ikram MA. Trajectories of decline in cognition and daily functioning in preclinical dementia. Alzheimer's & Dementia. 2016;12:144-53.

[8] Pérès K, Helmer C, Amieva H, Orgogozo JM, Rouch I, Dartigues JF, et al. Natural history of decline in instrumental activities of daily living performance over the 10 years preceding the clinical diagnosis of dementia: a prospective population-based study. Journal of the American Geriatrics Society. 2008;56:37-44.

[9] Fauth EB, Schwartz S, Tschanz JT, Østbye T, Corcoran C, Norton MC. Baseline disability in activities of daily living predicts dementia risk even after controlling for baseline global cognitive ability and depressive symptoms. International journal of geriatric psychiatry. 2013;28:597-606.

[10] Okonkwo OC, Wadley VG, Griffith HR, Belue K, Lanza S, Zamrini EY, et al. Awareness of deficits in financial abilities in patients with mild cognitive impairment: going beyond self-informant discrepancy. The American Journal of Geriatric Psychiatry. 2008;16:650-9.

[11] Debettignies BH, Mahurin RK, Pirozzolo FJ. Insight for impairment in independent living skills in Alzheimer's disease and multi-infarct dementia. Journal of clinical and experimental neuropsychology. 1990;12:355-63.

[12] Loewenstein DA, Argüelles S, Bravo M, Freeman RQ, Argüelles T, Acevedo A, et al. Caregivers' judgments of the functional abilities of the Alzheimer's disease patient: a comparison of proxy reports and objective measures. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2001;56:P78-P84.

[13] Jekel K, Damian M, Wattmo C, Hausner L, Bullock R, Connelly PJ, et al. Mild cognitive impairment and deficits in instrumental activities of daily living: a systematic review. Alzheimer's research & therapy. 2015;7:17.

[14] Lindbergh CA, Dishman RK, Miller LS. Functional disability in mild cognitive impairment: a systematic review and meta-analysis. Neuropsychology review. 2016;26:129-59.

[15] Kaye J, Mattek N, Dodge HH, Campbell I, Hayes T, Austin D, et al. Unobtrusive measurement of daily computer use to detect mild cognitive impairment. Alzheimer's & Dementia. 2014;10:10-7.

[16] Seelye A, Mattek N, Sharma N, Riley T, Austin J, Wild K, et al. Weekly observations of online survey metadata obtained through home computer use allow for detection of changes in everyday cognition before transition to mild cognitive impairment. Alzheimer's & Dementia. 2018;14:187-94.

[17] Seelye A, Hagler S, Mattek N, Howieson DB, Wild K, Dodge HH, et al. Computer mouse movement patterns: A potential marker of mild cognitive impairment. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2015;1:472-80.

[18] Stringer G, Couth S, Brown L, Montaldi D, Gledson A, Mellor J, et al. Can you detect early dementia from an email? A proof of principle study of daily computer use to detect cognitive and functional decline. International journal of geriatric psychiatry. 2018;33:867-74.

[19] Seelye A, Mattek N, Howieson DB, Austin D, Wild K, Dodge HH, et al. Embedded online questionnaire measures are sensitive to identifying mild cognitive impairment. Alzheimer disease and associated disorders. 2016;30:152.

[20] Tourangeau R, Rips LJ, Rasinski K. The psychology of survey response. Cambridge, UK: Cambridge University Press; 2000.

[21] Schwarz N, Knäuper B. Cognition, aging, and self-reports. In: Park D, Schwarz N, editors. Cognitive aging: a primer. Philadelphia, PA: Psychology Press; 2012. p. 233-52.

[22] Colsher PL, Wallace RB. Data quality and age: health and psychobehavioral correlates of item nonresponse and inconsistent responses. Journal of Gerontology. 1989;44:P45-P52.

[23] Schneider S. Extracting Response Style Bias From Measures of Positive and Negative Affect in Aging Research. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2018;73:64-74.

[24] Conijn JM, van der Ark LA, Spinhoven P. Satisficing in mental health care patients: The effect of cognitive symptoms on self-report data quality. Assessment. 2020;27:178-93.

[25] Kutschar P, Weichbold M, Osterbrink J. Effects of age and cognitive function on data quality of standardized surveys in nursing home populations. BMC geriatrics. 2019;19:1-10.

[26] Lechner CM, Rammstedt B. Cognitive ability, acquiescence, and the structure of personality in a sample of older adults. Psychological assessment. 2015;27:1301-11.

[27] Ienca M, Vayena E, Blasimme A. Big data and dementia: charting the route ahead for research, ethics, and policy. Frontiers in medicine. 2018;5:13.

[28] Snowdon DA, Kemper SJ, Mortimer JA, Greiner LH, Wekstein DR, Markesbery WR. Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the Nun Study. Jama. 1996;275:528-32.

[29] Whalley L, Starr J, Athawes R, Hunter D, Pattie A, Deary I. Childhood mental ability and dementia. Neurology. 2000;55:1455-9.

[30] Smith J, Fisher GG, Ryan L, Clarke P, House J, Weir D. Health and Retirement Study Psychosocial and Lifestyle Questionnaire 2006 - 2010: Documentation Report. Ann Arbor, MI: University of Michigan; 2013.

[31] 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. Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2011;66:i162-i71.

[32] Langa KM, Larson EB, Crimmins EM, Faul JD, Levine DA, Kabeto MU, et al. A comparison of the prevalence of dementia in the United States in 2000 and 2012. JAMA internal medicine. 2017;177:51-8.

[33] McCammon RJ, Fisher GG, Hassan H, Faul JD, Rogers W, Weir DR. Health and Retirement Study Imputation of Cognitive Functioning Measures: 1992-2016. Ann Arbor: University of Michigan; 2019.

[34] Langa K, Weir DR, Kabeto MU, Sonnega A. Langa-Weir Classification of Cognitive Function (1995 Onward). Anne Arbor: University of Michigan; 2020.

[35] VanderWeele TJ. Principles of confounder selection. European journal of epidemiology. 2019;34:211-9.

[36] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. Journal of the American statistical association. 1999;94:496-509.

[37] Schaie KW. A general model for the study of developmental problems. Psychological bulletin. 1965;64:92-107.

[38] Feveile H, Olsen O, Hogh A. A randomized trial of mailed questionnaires versus telephone interviews: response patterns in a survey. BMC medical research methodology. 2007;7:1-7.

[39] Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics. 2000;56:337-44.

[40] Saha P, Heagerty P. Time-dependent predictive accuracy in the presence of competing risks. Biometrics. 2010;66:999-1011.

[41] Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med. 2013;32:5381-97.

[42] Clark TG, Altman DG, De Stavola BL. Quantification of the completeness of follow-up. The Lancet. 2002;359:1309-10.

[43] Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

[44] Van Vaerenbergh Y, Thomas TD. Response styles in survey research: A literature review of antecedents, consequences, and remedies. International Journal of Public Opinion Research. 2013;25:195-217.

[45] Hayes-Larson E, Ackley SF, Zimmerman SC, Ospina-Romero M, Glymour MM, Graff RE, et al. The competing risk of death and selective survival cannot fully explain the inverse cancer-dementia association. Alzheimer's & Dementia. 2020;16:1696-703.

[46] Seligman SC, Giovannetti T, Sestito J, Libon DJ. A new approach to the characterization of subtle errors in everyday action: Implications for mild cognitive impairment. The Clinical Neuropsychologist. 2014;28:97-115.

[47] Kozauer N, Katz R. Regulatory innovation and drug development for early-stage Alzheimer's disease. N Engl J Med. 2013;368:1169-71.

[48] Gamaldo AA, An Y, Allaire JC, Kitner-Triolo MH, Zonderman AB. Variability in performance: identifying early signs of future cognitive impairment. Neuropsychology. 2012;26:534-40.

[49] Meijer RR. The number of Guttman errors as a simple and powerful person-fit statistic. Applied Psychological Measurement. 1994;18:311-4.

[50] Knowles ES, Condon CA. Why people say" yes": A dual-process theory of acquiescence. Journal of Personality and Social Psychology. 1999;77:379-86.

Research in Context

1. Systematic review: We reviewed the literature through traditional methods (e.g., PubMed) and based on references of relevant articles. Although studies have provided some evidence of relationships between the way people complete survey assessments and their cognitive functioning, none has investigated associations between participants' mistakes in the completion of self-report surveys and future dementia risk in longitudinal studies. Relevant studies are cited.
2. Interpretation: Our findings show that several indices of subtle reporting mistakes derived from response patterns in self-report surveys are associated with risk of developing dementia over 10 years of follow-up.
3. Future directions: The manuscript proposes a strategy for obtaining objective, behavior-based indicators of functioning deficits directly from survey response patterns in existing longitudinal studies. This approach may contribute to the identification and characterization of functional abilities that are predictive of transition from cognitively normal to dementia in older adults.