American Journal of Epidemiology

Vol. 192, No. 4
https://doi.org/10.1093/aje/kwac197
Advance Access publication:
November 11, 2022

Original Contribution

Measurement of Prevalent Versus Incident Dementia Cases in Epidemiologic Studies

Emma Nichols*, Derek K. Ng, Bryan D. James, Jennifer A. Deal, and Alden L. Gross

* Correspondence to Emma Nichols, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 2024 E. Monument Street, Suite 2-700, Baltimore, MD 21205 (e-mail: enicho30@jh.edu).

Initially submitted February 24, 2022; accepted for publication November 1, 2022.

Because dementia is progressive, incident cases are on average milder than prevalent cases, affecting the performance of cognitive tests and questions on functional limitations (i.e., cognition/functional limitation items) used for dementia assessment. Longitudinal studies assess incident cases, while cross-sectional studies assess prevalent cases, but differences are not typically considered when researchers select items to include in studies. We used longitudinal data from the Religious Orders Study and Memory and Aging Project (ROSMAP) ($n = 3,446$) collected between 1994 and 2021 to characterize differences in associations between items (cognition: 35 items; functional limitations: 14 items) and incident or prevalent dementia using multinomial regression models with generalized estimating equations, controlling for ROSMAP cohort (Religious Orders Study or Memory and Aging Project), age, sex, race, and education. The association between a given item and incident dementia was significantly weaker than the association between the same item and prevalent dementia for 46 of 49 items. However, there was variability, with larger differences for some items, including naming a pencil (prevalence odds ratio = 0.02 (95% confidence interval: 0.02, 0.03); incidence odds ratio = 0.10 (95% confidence interval: 0.06, 0.17); P for difference < 0.001). Important differences exist in the performance of cognition/functional limitation items for measurement of incident versus prevalent dementia. Differences can inform the choice of items for cross-sectional studies of prevalent cases or longitudinal studies of incident cases, leading to reduced misclassification and increased statistical power.

aging; cognitive testing; dementia; functional limitations; measurement; study design

Abbreviations: ADL, activities of daily living; AUC, area under the receiver operating characteristic curve; CI, confidence interval; IADL, instrumental activities of daily living; MAP, Memory and Aging Project; OR, odds ratio; ROS, Religious Orders Study; ROSMAP, Religious Orders Study and Memory and Aging Project.

Editor's note: *An invited commentary on this article appears on page 535.*

Due to expected trends in population aging, the burden of dementia will continue to grow in importance (1, 2). Against this backdrop, continued investment in epidemiologic research on the causes and consequences of dementia is crucial. However, the inherent subjectivity in the definition of clinical dementia poses unique challenges in epidemiologic research (3). While physician-given or physician-adjudicated clinical diagnoses of dementia remain the gold standard, they are often too resource-intensive to administer in large-scale cohort studies. Ultimately, a lack of consensus on the best tools or criteria for dementia assessment results in wide heterogeneity, with 230 different ascertainment methods being documented across 237 scientific publications assessing the prevalence or incidence of dementia (4).

When studies do not rely on physician assessment, dementia is commonly ascertained using systematic assessments of cognition and functional limitations, the 2 core components of clinical diagnoses (3). Investigators then often use algorithmic approaches to identify dementia cases (5, 6). Algorithms are typically validated in cross-sectional samples, allowing researchers to assess algorithm perfor-

mance in distinguishing persons with prevalent dementia from those without dementia (7–9). However, despite the fact that dementia algorithms are often validated for detection of prevalent cases, these algorithms are commonly used to assess incident dementia cases in time-to-event analyses (10–12). This is problematic, because assessments and algorithms used to identify dementia may not necessarily be equally sensitive to new (i.e., incident) and prevalent cases alike.

Because dementia is a progressive condition, incident cases are, on average, milder than prevalent cases (13). However, tests of cognition (i.e., test items) used in the adjudication of dementia and in algorithmic definitions of dementia have differential precision for measurement at different severities of cognitive impairment (14). Relatedly, the difficulty of items assessing functional limitations is variable, and these items may also have differential precision and validity at different disease severities (15). A better understanding of the magnitude of potential differences in the performance of cognition/functional limitation items for the classification of incident dementia versus prevalent dementia could help characterize the impacts of suboptimal measurement in existing studies. Additionally, quantitative evidence on differences can be used to improve measurement in future studies; the most appropriate dementia assessment for a study trying to ascertain incident dementia cases may be different from the most appropriate assessment for a study trying to ascertain prevalent cases.

In this study, we aimed to estimate differences in the measurement of incident and prevalent dementia and assess which items on cognition and functional limitations are best suited to the measurement of prevalent cases of dementia in cross-sectional settings versus the measurement of incident dementia cases in longitudinal studies. We further assessed the impact of potential differences on the classification of prevalent versus incident dementia. Based on the results, we provide guidance for improving the measurement of dementia and the application of algorithmic approaches to dementia measurement in future epidemiologic studies.

METHODS

Sample

This study used harmonized data from the Religious Orders Study (ROS) and the Rush Memory and Aging Project (MAP), collectively known as ROSMAP. Both studies are ongoing prospective cohort studies with continuous recruitment and high retention, with only 7.8% of participants having withdrawn since study initiation (16). The ROS started in 1994 and enrolls priests and nuns from across the United States. The MAP began in 1997 and recruits volunteer participants from northeastern Illinois. We excluded individuals with missing covariate information ($n = 1$) and individuals with prevalent dementia at baseline ($n = 217$), yielding a final sample size of 3,446. Participants had on average 8.4 visits (range, 1–27) over an average follow-up period of 7.7 years (range, 0–30).

Measurement of cognitive functioning and functional limitations

A comprehensive neuropsychological battery was administered, including 32 items measuring orientation, memory, language, visuospatial functioning, and executive functioning (details are provided in Web Table 1 (available at https://doi.org/10.1093/aje/kwac197) and Table 1). In analyses, each continuous item was transformed by dividing scores by twice the standard deviation to facilitate comparisons between continuous and binary items (17). Items from the Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) scales were used to assess functional limitations (Table 1) (18, 19).

Covariates

Age at each visit, self-reported sex, years of educational attainment, self-reported ethnicity (Hispanic yes/no), and self-reported race were considered as covariates. Race was collapsed into a 3-level categorical variable (White, Black, and other) because of low numbers of participants in other racial categories.

Adjudication of dementia diagnoses

Dementia was adjudicated independently at each annual clinical visit using a 3-stage process (details are provided in Web Appendix 1); yearly adjudication allowed us to differentiate between incident and prevalent cases of dementia (20). An incident dementia visit was defined as the first visit with clinically adjudicated dementia; all visits after the visit with incident dementia were defined as prevalent dementia visits (i.e., dementia was an absorbing state). Therefore, incident dementia will coincide with more mild, early-stage disease, whereas prevalent dementia will encompass a wider range of cases including severe, late-stage disease. Because adjudicated dementia diagnoses in the ROSMAP sample were conducted independently at each assessment, 7.1% of participants had at least 1 visit after their initial visit with dementia in which they were ruled to not have clinical dementia (i.e., either mild cognitive impairment or normal). We considered these individuals to have prevalent dementia at these later assessments.

Statistical analyses

First, we assessed demographic characteristics and characterized the distributions of age, sex, and responses to items assessing cognition and functional limitations across person-visits (each unique combination of person and visit) by dementia status. Statistical models used the person-visit as the primary unit of analysis to maximize the data informing analytical models.

Primary analyses used multinomial logistic regression models with generalized estimating equations to evaluate associations between dementia status (no dementia, incident dementia, or prevalent dementia) and responses to each item on cognition or functional limitations. Models controlled

Table 1. Distributions of Demographic Characteristics and Responses to Items on Cognition and Functional Limitations Across All Study Visits in the Religious Orders Study and Memory and Aging Project (n = 3,446), 1994–2021

Characteristic	Standardized Item Difficulty[a]	Dementia Status					
		No Dementia (n = 25,258 Visits)		Incident Dementia (n = 919 Visits)		Prevalent Dementia (n = 2,461 Visits)	
		%	Mean (SD)	%	Mean (SD)	%	Mean (SD)
Age at visit, years			81.2 (7.6)		87.4 (6.6)		88.6 (6.5)
Female sex		73.4		73.6		80.6	
Cognitive functioning							
Orientation							
Year	−1.11	97.6		59.6		37.6	
Season	−0.36	91.2		67.6		47.8	
Date	−0.02	84.2		28.3		20.4	
Day of week	−1.05	97.7		61.0		42.0	
Month	−1.26	98.3		67.9		45.2	
State	−1.74	99.5		89.1		69.8	
County	−0.73	95.8		59.3		41.3	
City	−1.18	98.1		71.4		51.1	
Room	−0.71	95.6		67.3		50.6	
Address	−0.38	90.8		37.5		23.7	
Memory							
Three-word recall	0.31		2.6 (0.7)		1.2 (1.0)		0.8 (1.0)
Immediate word recall	1.10		19.3 (4.9)		11.1 (4.3)		9.9 (5.0)
Delayed word recall	1.09		6.2 (2.5)		1.6 (1.8)		1.3 (1.9)
Word recall recognition	−0.37		9.7 (1.0)		7.1 (2.8)		6.3 (3.2)
East Boston Memory Test—Immediate Recall	0.73		9.7 (1.9)		6.6 (2.7)		5.4 (3.3)
East Boston Memory Test—Delayed Recall	0.82		9.3 (2.3)		4.8 (3.4)		3.4 (3.6)
Logical memory immediate	1.13		13.1 (4.5)		5.7 (3.9)		4.6 (4.1)
Logical memory delayed	1.19		11.8 (5.0)		3.6 (3.4)		2.9 (3.7)
Language							
Name watch	−1.71	99.5		93.1		79.9	
Name pencil	−1.92	99.7		96.4		84.2	
Say phrase	0.01	85.0		61.0		48.5	
Read command	−0.98	97.3		84.3		69.5	
Follow 3-step command	0.13		2.7 (0.5)		2.2 (0.9)		1.8 (1.1)
Write sentence	−1.23	98.6		88.5		70.0	70.0
Category fluency	1.11		34.7 (9.9)		17.4 (7.6)		14.3 (8.5)
Boston Naming Test	0.13		14.0 (1.2)		12.0 (2.4)		10.7 (3.3)
Executive							
Spell backwards	−0.28		4.9 (0.5)		4.0 (1.5)		3.1 (2.1)
Symbol Digit Modalities Test	1.13		39.9 (11.1)		19.4 (10.9)		16.5 (11.6)
Digit span forward	0.93		8.2 (2.0)		6.8 (2.3)		6.3 (2.5)
Digit span backward	1.04		6.4 (2.1)		4.1 (2.0)		3.7 (2.1)
Number comparison	1.11		25.2 (7.4)		15.7 (7.6)		13.5 (8.2)
Digit ordering	1.11		7.4 (1.7)		5.2 (2.0)		4.8 (2.3)

Table continues

Table 1. Continued

Characteristic	Standardized Item Difficulty[a]	Dementia Status					
		No Dementia (n = 25,258 Visits)		Incident Dementia (n = 919 Visits)		Prevalent Dementia (n = 2,461 Visits)	
		%	Mean (SD)	%	Mean (SD)	%	Mean (SD)
Visuospatial							
Copy drawing	0.09	81.2		48.6		35.2	
Line orientation test	0.96		10.2 (3.1)		7.3 (3.4)		6.9 (3.8)
Progressive matrices	0.89		12.1 (2.7)		8.9 (2.7)		8.2 (3.1)
Functional limitations							
Difficulty using the telephone	−0.76	3.0		23.0		41.2	
Difficulty preparing a meal	0.19	11.8		50.8		64.9	
Difficulty performing light housework	0.34	12.1		46.3		58.0	
Difficulty performing heavy housekeeping	2.43	48.6		82.1		84.4	
Difficulty taking medications	−0.34	7.2		47.6		62.1	
Difficulty handling finances	0.37	11.4		57.2		69.3	
Difficulty shopping	0.60	16.5		62.4		73.0	
Difficulty traveling within the community	1.02	21.0		68.1		75.1	
Difficulty walking across a room	0.15	11.8		38.8		51.1	
Difficulty bathing	−0.03	10.0		45.0		56.2	
Difficulty dressing	−0.80	4.8		29.5		43.7	
Difficulty eating	−1.73	1.2		11.9		21.0	
Difficulty transferring from bed	−0.61	5.0		24.3		37.9	
Difficulty using the toilet	−0.82	4.1		22.8		37.0	

Abbreviation: SD, standard deviation.

[a] The standardized item difficulty is a comparable measure of item difficulty, computed as a weighted average of item thresholds from an item response theory model. Details are provided in Web Appendix 3.

for ROSMAP cohort (ROS vs. MAP), age at study visit, sex, race, ethnicity, and years of educational attainment, as these factors are associated with cognitive test performance and may be associated with disease severity due to either the biological progression of disease or associations with survival, study retention, and administrative censoring, which affect whether an individual would be captured as an incident or prevalent case (21, 22). In each regression model, we excluded records with missing data on the given item of interest (complete-case analysis) (Web Table 2). To adjust precision while ensuring that estimates would replicate findings from a hypothetical cross-sectional sample of individuals with no dementia, incident dementia, and prevalent dementia, we used an independent correlation structure and robust standard errors to account for correlation among visits within individuals. From models, we estimated adjusted odds ratios (ORs) describing the association between item responses and incident dementia as compared with no dementia, the association between item responses and prevalent dementia as compared with no dementia, and the ratio of ORs comparing the association for incident dementia with the association for prevalent dementia. Because dementia is not rare across strata of all items on cognition or functional limitations, ORs cannot be interpreted as relative risks, but instead were used to compare associations between incident and prevalent dementia. We also directly evaluated differences in classification performance by estimating the area under the receiver operating characteristic curve (AUC) for the classification of incident versus prevalent dementia and compared differences in AUC to estimated ORs. We further evaluated differences in classification performance for a larger set of items to obtain a more realistic estimate of potential classification differences (Web Appendix 2).

Next, we evaluated how the comparison between associations for incident and prevalent dementia varied with the

Table 2. Demographic Characteristics of Participants in the Religious Orders Study and Memory and Aging Project (*n* = 3,446) at Baseline, 1994–2021

Characteristic	Total (*n* = 3,446)		ROSMAP Cohort			
			Religious Orders Study (*n* = 1,376)		Memory and Aging Project (*n* = 2,070)	
	No.	%	No.	%	No.	%
Age at baseline, years[a]	78.0 (7.8)		75.4 (7.5)		79.7 (7.5)	
No. of follow-up visits[a]	8.4 (5.7)		10.5 (6.6)		7.0 (4.6)	
Duration of education, years[a]	16.3 (3.8)		18.5 (3.4)		14.9 (3.3)	
Female sex	2,525	73.3	982	71.4	1,543	74.5
Race						
White	3,201	92.9	1,269	92.2	1,932	93.3
Black	200	5.8	90	6.5	110	5.3
Other	45	1.3	17	1.2	28	1.4
Hispanic ethnicity	169	4.9	58	4.2	111	5.4

Abbreviation: ROSMAP, Religious Orders Study and Memory and Aging Project.
[a] Values are expressed as mean (standard deviation).

relative difficulty of items in the ROSMAP battery of cognitive and ADL/IADL indicators. To quantify item difficulty, we used item response theory methods (details are given in Web Appendix 3). To assess the relationship between item difficulty and ORs comparing the association for incident dementia with the association for prevalent dementia, we estimated correlation coefficients and examined scatterplots.

We conducted a supplementary analysis to evaluate the sensitivity of findings to our choice of model structure. We used Poisson regression with generalized estimating equations to quantify the association between years since dementia diagnosis and responses to items on cognition and functional limitations among persons with dementia. The outcome variable was the number of years since dementia diagnosis; each model evaluated the association between a single item and the number of years since dementia diagnosis. We used an independent correlation structure and controlled for the same variables as in the primary analyses. In these models, estimated exponentiated coefficients represent the expected multiplicative increase in years since dementia diagnosis given a 1-unit higher response to an item. We evaluated the similarity of results to primary analyses by calculating the correlation between coefficients and through visual inspection of scatterplots.

Finally, we investigated the sensitivity of results to bias from differential dropout. We used stabilized inverse probability weights for death and dropout (details are given in Web Appendix 4). We repeated our primary analysis with 3 sets of weights: combined weights, weights for dropout due to death, and weights for dropout due to other reasons. With the exception of item response theory models (conducted in Mplus, version 8 (Muthén and Muthén, Los Angeles, California)), statistical analyses were conducted in R, version 4.1.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

This study included 3,446 participants; about 60% were from the MAP, with the rest coming from the ROS. Participants from the ROS were younger at baseline (75.4 years vs. 79.7 years), with longer available follow-up time (10.5 visits vs. 7.0 visits) (Table 2).

The 3,446 participants included in the analysis attended 28,638 study visits over an average of 8.4 years. Of the study visits, 88% were among individuals without dementia at the time of the visit, 3% were incident dementia visits (i.e., the first visit at which an individual was diagnosed with dementia), and 9% were prevalent dementia visits (i.e., visits for individuals with dementia that were not "incident" visits) (Table 1). Across all items on cognitive functioning, scores were higher on average among persons with no dementia than among those with either incident or prevalent dementia, but were also higher on average among those with incident dementia as compared with prevalent dementia. The crude proportion of individuals who endorsed having functional limitations was highest among those with prevalent dementia, followed by those with incident dementia and then those without dementia.

Crude patterns indicating that associations between items on either cognition or functional limitations and dementia were stronger when considering prevalent dementia as compared with incident dementia held in multivariable multinomial regression models (Figures 1 and 2). For 46 out of 49 total items, there was a statistically significant difference in the strength of the association with prevalent

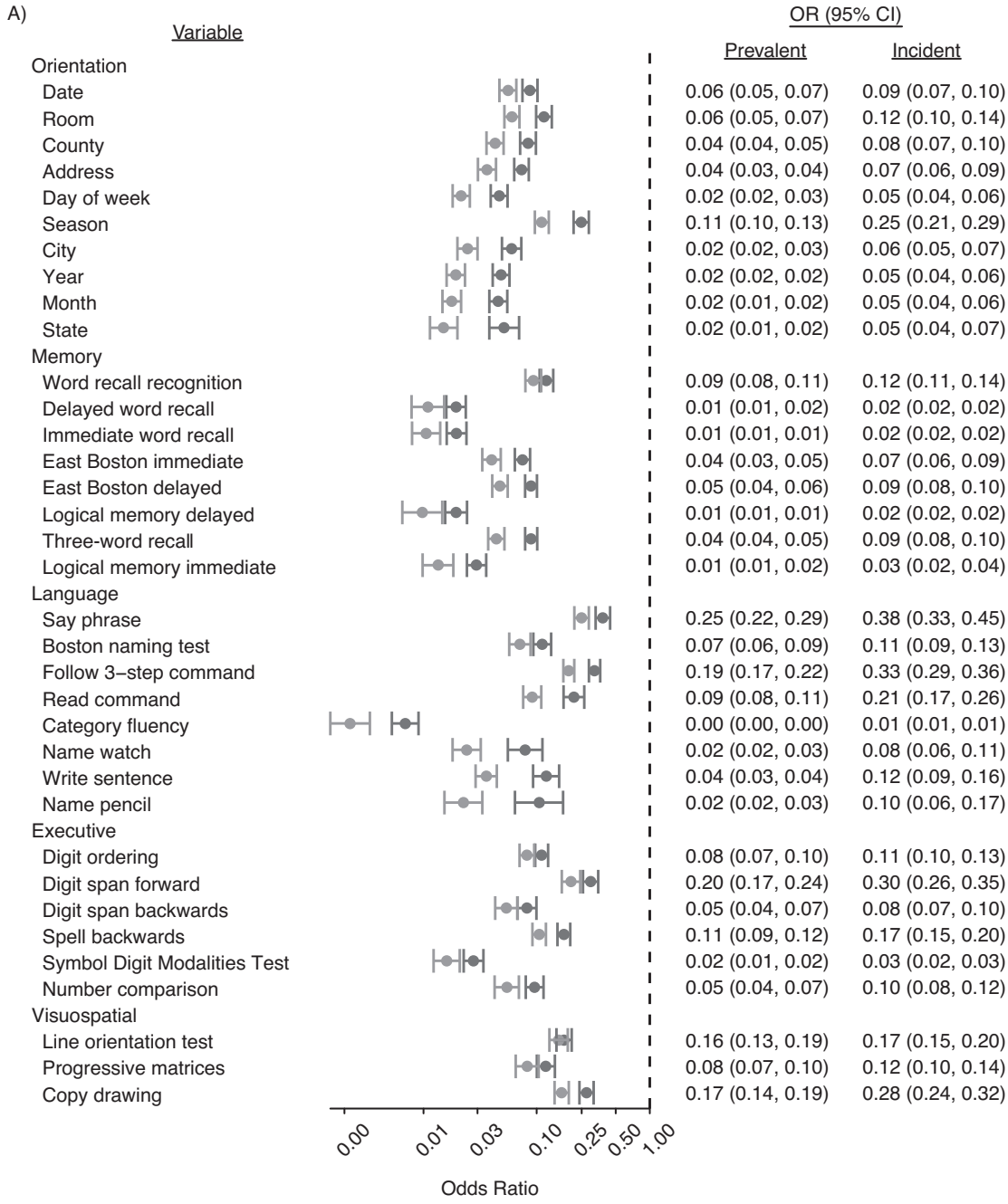

Figure 1 Continues

dementia versus the strength of the association with incident dementia.

Cognitive items

The orientation items with the largest discrepancies in their associations with prevalent versus incident dementia were naming the state and naming the month (Figure 1A). For example, for the item for naming the state, the OR for prevalent dementia was 0.02 (95% confidence interval (CI): 0.01, 0.02), whereas the OR for incident dementia was 0.05 (95% CI: 0.04, 0.07). Comparing these 2 estimates, the estimated ratio of ORs comparing the association for incident dementia with that for prevalent dementia in regards to naming the state was 3.43 (95% CI: 2.69, 4.37) (Figure 1B). Because better performance on cognitive tests is protective against dementia, ratios of ORs above 1 indicate that the item is better for discriminating prevalent cases than it is

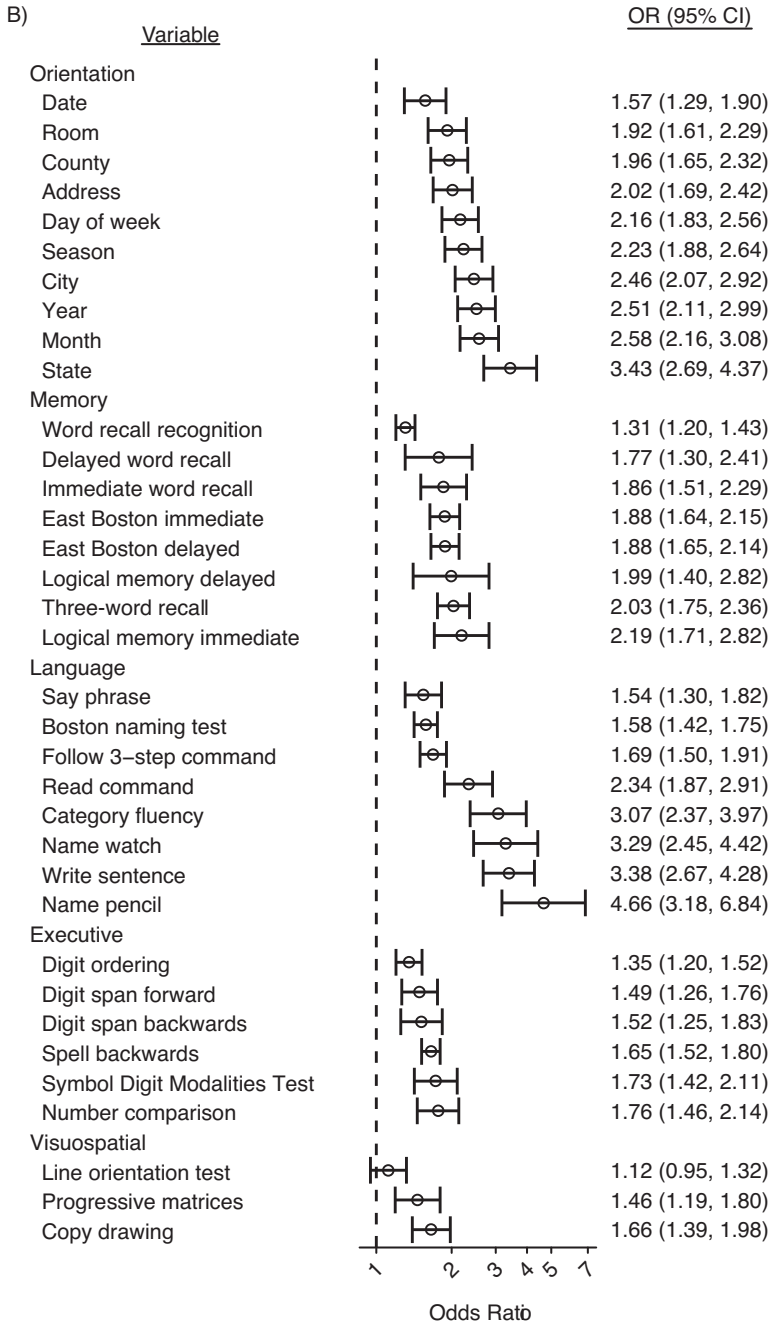

Figure 1. Associations between cognitive items and dementia according to dementia type (incident vs. prevalent) (A) and comparison of associations between cognitive items and incident dementia with associations between cognitive items and prevalent dementia (B), Religious Orders Study and Memory and Aging Project (n = 3,446), 1994–2021. Stratified odds ratios (ORs) (panel A) show the strength of the association between each item and prevalent dementia (light gray) or incident dementia (dark gray). ORs in panel B represent the OR comparing the effect for incident dementia with the effect for prevalent dementia. Larger ORs in panel B indicate that an item is more strongly related to prevalent dementia than to incident dementia. All models controlled for cohort (Religious Orders Study or Memory and Aging Project), age, sex, education, race, and ethnicity. Bars show the 95% confidence intervals (CIs).

for discriminating incident cases. Inverting this comparison, our estimate indicated that the strength of the association between the item on naming the state and incident dementia was 71% weaker (95% CI: 63, 77) than the association between the same item and prevalent dementia. The percent decrease in the strength of the estimate for incident dementia as compared with prevalent dementia was 61% (95% CI: 54, 68) for naming the month. However, despite

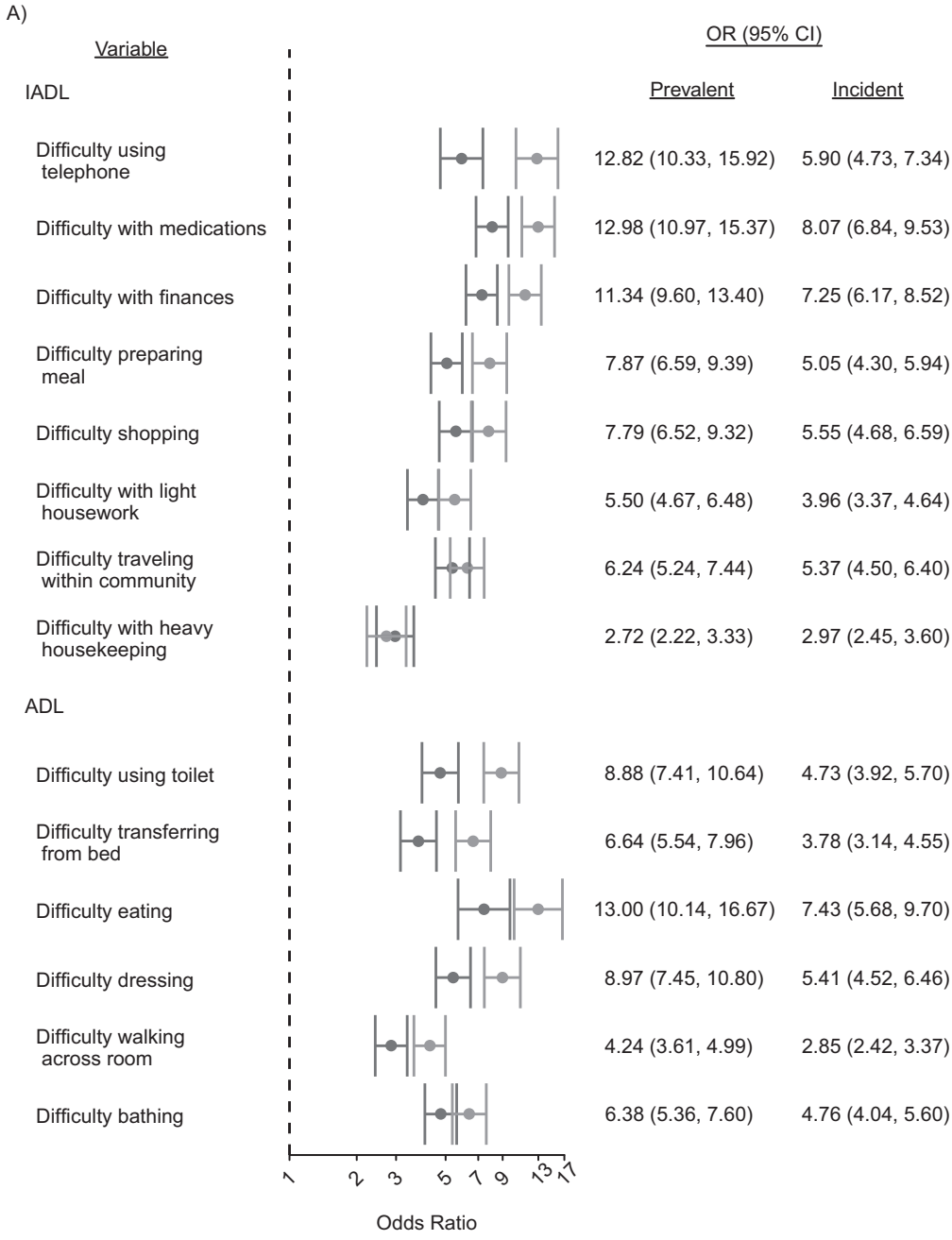

Figure 2 Continues

large observed differences in associations between incident and prevalent dementia for naming the state and naming the month, individual ORs describing the association between these items and either incident or prevalent dementia were stronger than the individual ORs for the association between either incident or prevalent dementia and the item on naming the season (Figure 1B).

Some items assessing language also showed large differences in their associations with incident versus prevalent dementia, including identifying a pencil, writing a sentence, and identifying a watch. While the difference between the associations of category fluency with incident dementia and prevalent dementia was large, the individual associations between this item and the 2 dementia categories were the strongest observed across all items, indicating that the item still performed well for the classification of incident dementia despite this difference. In general, there were smaller differences in the associations between items and either inci-

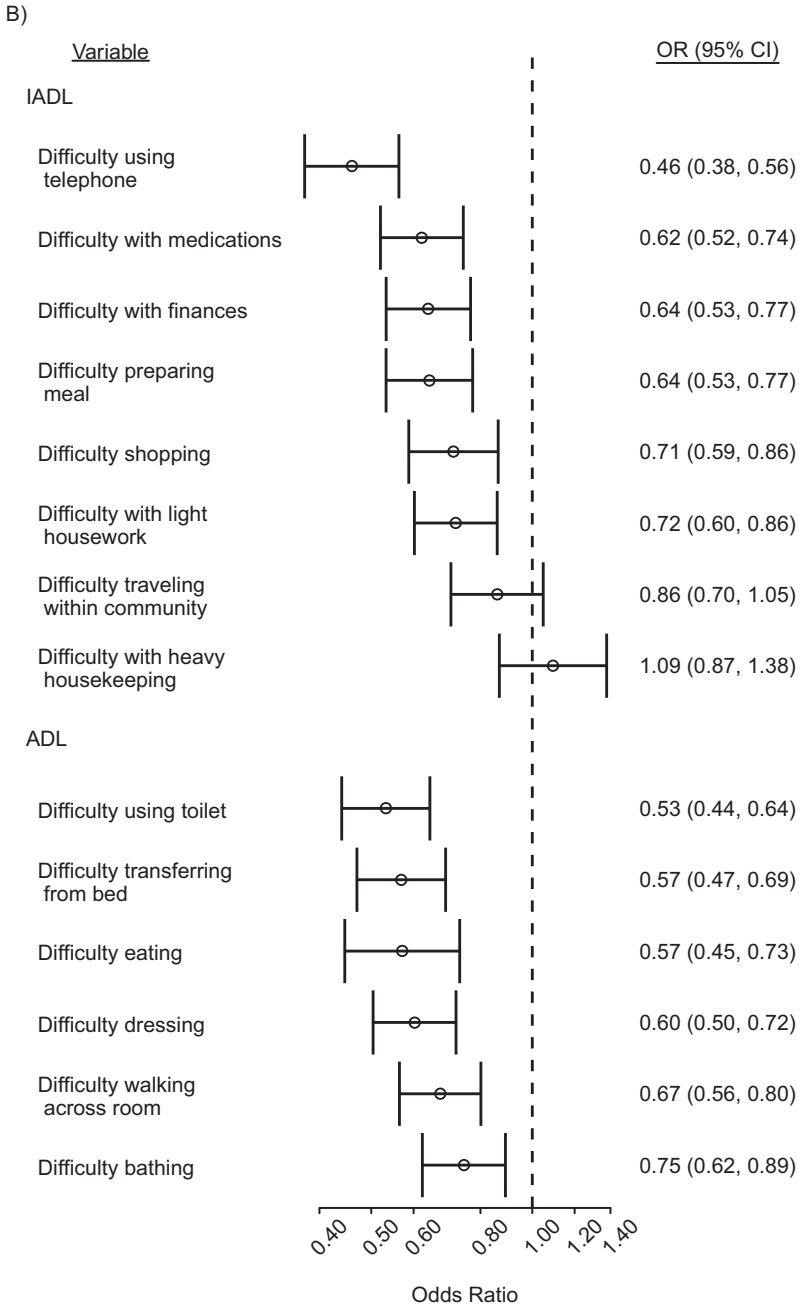

Figure 2. Associations between items on functional limitations and dementia according to dementia type (incident vs. prevalent) (A) and comparison of associations between items on functional limitations and incident dementia with associations between items on functional limitations and prevalent dementia (B), Religious Orders Study and Memory and Aging Project (n = 3,446), 1994–2021. Stratified odds ratios (ORs) (panel A) show the strength of the association between each item and prevalent dementia (light gray) or incident dementia (dark gray). ORs in panel B represent the OR comparing the effect for incident dementia with the effect for prevalent dementia. Smaller ORs in panel B (closer to 0) indicate that an item is more strongly related to prevalent dementia than to incident dementia. All models controlled for cohort, age, sex, education, race, and ethnicity. Bars show the 95% confidence intervals (CIs). ADL, activities of daily living; IADL, instrumental activities of daily living.

dent or prevalent dementia and items for memory, executive functioning, and visuospatial functioning (Figure 2).

The rank order of items in terms of the strength of their associations with prevalent dementia and the rank order of items in terms of their associations with incident dementia were fairly similar (Spearman correlation coefficient (r) = 0.95) (Table 3). However, even when ORs were similar, there were noticeable differences in the rank order of items,

which could affect the choice of items for use in shorter batteries. For example, the item on naming a state had the sixth strongest association with prevalent dementia but the 10th strongest association with incident dementia, while the item on naming a pencil had the 11th strongest association with prevalent dementia but the 21st strongest association with incident dementia.

Functional limitation items

The items on functional limitations with the largest differences in their associations with incident versus prevalent dementia were difficulty using the telephone, difficulty using the toilet, and difficulty eating (Figure 2A). For the item on difficulty using the telephone, the OR for the association with prevalent dementia was 12.82 (95% CI: 10.33, 15.92), whereas the OR for incident dementia was 5.90 (95% CI: 4.73, 7.34). The ratio of ORs describing the comparison between the OR for incident dementia and the OR for prevalent dementia was 0.46 (95% CI: 0.38, 0.56); in comparison, this estimate was 0.53 (95% CI: 0.44, 0.64) for difficulty toileting and 0.57 (95% CI: 0.47, 0.69) for difficulty transferring from bed (Figure 2B). These estimates indicate, for example, that the association between the item on difficulty using the telephone and incident dementia was 54% (95% CI: 44, 62) weaker than the association between the same item and prevalent dementia. Associations with incident versus prevalent dementia were not statistically different for 2 items: difficulty with heavy housekeeping and difficulty traveling within one's community. However, associations between these 2 items and either incident or prevalent dementia were still smaller compared with items with somewhat larger differences in their associations with incident and prevalent dementia, such as difficulty handling finances (Figure 2B).

Despite high correspondence in the rank order of items in terms of the strength of their associations with either prevalent or incident dementia ($r = 0.87$), we identified differences (Table 4). For example, the item with the sixth strongest association with prevalent dementia (difficulty toileting) had the 10th strongest association with incident dementia.

Classification performance

The AUC was greater for the classification of prevalent dementia than for the classification of incident dementia for all items. The difference in AUC was strongly correlated with estimated ratios of ORs (cognition: $r = 0.49$; functional limitations: $r = -0.92$) (Web Figure 1, Web Table 3). When items with the largest AUC differences (15 items (10 cognition, 5 functional limitations)) were used together in a single model for classification, the AUC for the classification of prevalent dementia was 0.95, as opposed to 0.90 for incident dementia.

Correlations with item difficulty

Differences in the associations with incident versus prevalent dementia were strongly correlated with item difficulty (Web Table 4, Web Figure 2). For both cognitive items ($r = -0.65$) and functional limitations items ($r = 0.88$), the difference in associations between incident and prevalent dementia was smaller for more difficult items. However, the direction of the correlations was reversed for the 2 sets of items because the endorsement of functional limitations was associated with increased dementia whereas better performance on cognitive items was associated with a lower probability of dementia.

Sensitivity analyses

Secondary models using a Poisson regression framework to assess the association between responses to cognition/functional limitation items and years of disease duration led to similar conclusions at the item level; coefficients of interest were highly correlated ($r = -0.97$) (Web Figures 3 and 4). Estimates of the difference in associations comparing associations with incident versus prevalent dementia were strengthened in sensitivity analyses using inverse probability of death and censoring weights (Web Figure 5).

DISCUSSION

There were large differences in the strength of associations between items measuring cognition and functional limitations and either incident or prevalent dementia. Results suggest that measures or algorithms evaluated and validated using cross-sectional samples of prevalent dementia cases may not have the same performance for the assessment of incident dementia in longitudinal studies. The magnitude of differences varied and was closely coupled with item difficulty, such that differences observed when comparing associations with incident dementia to associations with prevalent dementia were greater for easier items. Therefore, when cognitive assessments or algorithms rely on a large number of easier items, assessments may not be adequately suited for the identification of incident dementia, which tends to be more subtle in presentation. Use of easier items, such as those in the Mini-Mental State Examination, for the detection of incident dementia may result in algorithms or ascertainment procedures with reduced sensitivity. In studies where the primary goal is detection of incident dementia, the inclusion of more difficult assessments may improve the performance of algorithms or assessments.

Comparisons of the relative magnitude of differences in associations for incident dementia and prevalent dementia may also help researchers select items providing the most information across assessments of both incident and prevalent dementia. For example, in this study, the items for difficulty using the telephone and difficulty managing medications had similar ORs for their associations with prevalent dementia, but the difference in associations for incident dementia as compared with prevalent dementia was much larger for the item assessing difficulty using the telephone than for the item assessing difficulty managing medications. Therefore, while the items would be interchangeable if one were only interested in identifying prevalent dementia, if it were important to ascertain both incident and prevalent dementia the item on managing medications would be preferred.

Table 3. Rank Order of Items on Cognitive Functioning in Terms of the Strength of Their Associations With Either Prevalent Dementia or Incident Dementia, Religious Orders Study and Memory and Aging Project (n = 3,446), 1994–2021[a]

Rank	Item for Assessment			
	Prevalent Dementia	OR	Incident Dementia	OR
1	Category fluency	0.002	Category fluency	0.007
2	Logical memory delayed	0.010	Delayed word recall	0.019
3	Immediate word recall	0.011	Logical memory delayed	0.019
4	Delayed word recall	0.011	Immediate word recall	0.020
5	Logical memory immediate	0.013	Symbol Digit Modalities Test	0.028
6	State	0.015	Logical memory immediate	0.029
7	Symbol Digit Modalities Test	0.016	Month	0.046
8	Month	0.018	Day of week	0.047
9	Year	0.019	Year	0.048
10	Day of week	0.022	State	0.052
11	Name pencil	0.023	City	0.060
12	Name watch	0.024	Address	0.073
13	City	0.024	East Boston immediate	0.075
14	Write sentence	0.036	Name watch	0.079
15	Address	0.036	Digit span backward	0.082
16	East Boston immediate	0.040	County	0.084
17	County	0.043	Date	0.087
18	Three-word recall	0.044	East Boston delayed	0.089
19	East Boston delayed	0.047	Three-word recall	0.089
20	Digit span backwards	0.054	Number comparison	0.096
21	Number comparison	0.054	Name pencil	0.105
22	Date	0.056	Digit ordering	0.111
23	Room	0.060	Boston Naming Test	0.112
24	Boston Naming Test	0.071	Room	0.116
25	Digit ordering	0.082	Progressive matrices	0.121
26	Progressive matrices	0.082	Write sentence	0.121
27	Read command	0.091	Word recall recognition	0.122
28	Word recall recognition	0.093	Spell backwards	0.174
29	Spell backwards	0.105	Line orientation test	0.175
30	Season	0.111	Read command	0.213
31	Line orientation test	0.156	Season	0.248
32	Copy drawing	0.166	Copy drawing	0.276
33	Follow 3-step command	0.192	Digit span forward	0.300
34	Digit span forward	0.201	Follow 3-step command	0.325
35	Say phrase	0.250	Say phrase	0.385

Abbreviation: OR, odds ratio.

[a] Numbers depict the rankings, and ORs show the strength of the association between each item and either prevalent or incident dementia.

We used the term *incident dementia* to refer to individuals at the first visit at which dementia was identified and the term *prevalent dementia* to refer to individuals who had dementia at each subsequent visit. While categories of incident and prevalent dementia are important in the consideration of future epidemiologic studies of clinical dementia, the biological processes that underlie dementia start decades before clinical symptoms emerge (23). Therefore, our categorization of incident dementia does not align with the notion of a new disease process but rather aligns with a new detection of

Table 4. Rank Order of Items on Functional Limitations in Terms of the Strength of Their Associations With Either Prevalent Dementia or Incident Dementia, Religious Orders Study and Memory and Aging Project (n = 3,446), 1994–2021[a]

Rank	Item for Assessment			
	Prevalent Dementia	OR	Incident Dementia	OR
1	Difficulty eating	12.999	Difficulty taking medications	8.073
2	Difficulty taking medications	12.984	Difficulty eating	7.427
3	Difficulty using the telephone	12.823	Difficulty handling finances	7.251
4	Difficulty handling finances	11.341	Difficulty using the telephone	5.896
5	Difficulty dressing	8.973	Difficulty shopping	5.553
6	Difficulty using the toilet	8.881	Difficulty dressing	5.406
7	Difficulty preparing a meal	7.867	Difficulty traveling within the community	5.367
8	Difficulty shopping	7.795	Difficulty preparing a meal	5.052
9	Difficulty transferring from bed	6.642	Difficulty bathing	4.756
10	Difficulty bathing	6.384	Difficulty using the toilet	4.728
11	Difficulty traveling within community	6.245	Difficulty performing light housework	3.956
12	Difficulty performing light housework	5.499	Difficulty transferring from bed	3.779
13	Difficulty walking across a room	4.243	Difficulty performing heavy housekeeping	2.972
14	Difficulty performing heavy housekeeping	2.718	Difficulty walking across a room	2.854

Abbreviation: OR, odds ratio.
[a] Numbers depict the rankings, and ORs show the strength of the association between each item and either prevalent or incident dementia.

clinical symptoms, a common outcome in population-based dementia studies.

Because dementia is progressive, our groupings of incident and prevalent dementia are consistent with groupings of persons with mild disease as compared with moderate or severe dementia. Our results are thus aligned with both prior research indicating that assessments of cognition and function have better discrimination for any dementia compared with mild dementia and other findings that persons with mild dementia score more similarly on cognitive tests to those with no disease as compared with those with severe disease (24–27).

Across cognitive domains, language items had some of the largest differences in comparing associations between items and either incident dementia or prevalent dementia. In a prior study of Mini-Mental State Examination items, the items on writing a sentence and object-naming had low sensitivity to impairment as compared with longer, gold-standard assessments of writing and naming (28). Low sensitivity to mild impairment probably also affected associations between these items and incident dementia in this study, explaining larger observed differences in the associations with incident dementia compared with prevalent dementia. More difficult and therefore more sensitive assessments of language, such as the category fluency task, had stronger associations with both prevalent and incident dementia and therefore may be better suited for assessing incident and prevalent dementia.

It is important to understand whether observed differences in associations between items on cognition and functional limitations and either incident dementia or prevalent dementia would lead to differences in algorithm performance. Prior research showing that algorithms for mild cognitive impairment have worse performance than algorithms for dementia suggests that algorithmic performance depends on symptom severity (29, 30). When using an algorithm containing 15 items with the largest differences in classification performance, the AUC for prevalent dementia was 0.95 versus 0.90 for incident dementia, suggesting that observed differences can affect classification performance for incident dementia to some extent. If items with large differences in classification performance and lower classification performance overall (e.g., the item on difficulty using the telephone, naming a pencil, or naming a watch) are used in algorithms, algorithmic performance may be worse for the classification of incident dementia in longitudinal studies. However, if high-quality items with high precision across the range of cognitive abilities, such as category fluency or the delayed and immediate word recall tasks, are included in algorithms, performance for the classification of incident dementia may be adequate.

Suboptimal ascertainment of dementia due to poor choice of items for dementia measurement can lead to the loss of power in epidemiologic studies. Use of a set of items with strong associations with prevalent dementia but weaker associations with incident dementia in a longitudinal study that was focused on ascertaining incident dementia would lead to nondifferential misclassification. In a longitudinal study of a dementia risk factor, researchers would be more likely to miss mild, incident cases of disease and to incorrectly categorize them as noncases in both the exposed and unexposed groups. Assuming a true effect of increased

dementia risk, the exposed group would appear more similar to the unexposed group, and the study would have lower statistical power to detect true group differences. Accurate measurement of disease outcomes, such as either prevalent dementia or incident dementia, is a prerequisite for well-conducted epidemiologic research.

Strengths of this study include yearly clinical adjudication of dementia in a large cohort, which allowed us to differentiate between prevalent and incident dementia and reduce measurement error in the classification of incident dementia. Although some measurement error likely remains, 1-year intervals for evaluation are probably sufficient given the slow, progressive nature of clinical dementia. The availability and reliability of this measure allowed for the direct estimation of regression coefficients of interest, rather than requiring reliance on latent variable methods to estimate the constructs of interest. We used multiple model forms and statistical analyses to arrive at the same conclusions, showing the robustness of our findings to analytical decisions. Additionally, sensitivity analyses showed that primary analyses were conservative in comparison with models accounting for differential attrition due to death and dropout.

Some limitations should be considered. First, for each item considered, we conducted complete-case analysis, subsetting data to nonmissing records. However, one reason for missing data on cognitive tests may be that participants are unable to complete them, indicating potentially severe cognitive impairment (31). Given this missingness mechanism, our estimates are conservative; complete-case analysis probably underestimated differences between a given item and incident versus prevalent dementia. Second, while the availability of physician-adjudicated diagnoses of dementia is a strength of this study, we implicitly assumed that physicians' diagnoses do not have measurement error. Because adjudication was completed independently at each visit, 7.1% of participants diagnosed with dementia were later diagnosed as not having dementia. We assumed that participants had dementia at every visit after their first visit with a dementia diagnosis, but if the opposite were true (if individuals were only classified as having dementia after they were no longer ever classified as dementia-free) we would exclude less impaired individuals from both the incident dementia and prevalent dementia categories. Because the composition of both the incident and prevalent dementia groups would change similarly, the impact on comparisons between incident and prevalent dementia would likely be small. Additionally, while physicians conducting adjudication were not blinded to data on cognition and functional limitations, any circularity would apply equally across visits and therefore should not influence the assessment of differences in associations between items on cognition or functional limitations and either incident or prevalent dementia (20).

Third, measures of ADLs/IADLs were self-reported. If persons with severe disease are less likely to report difficulty with activities due to the effects of dementia, we would expect comparisons of associations for incident versus prevalent dementia to be conservatively biased. Because all ADL/IADL items were ascertained in a similar manner, comparisons between items would still be valid. Finally, potential limits to generalizability of the ROSMAP sample should be considered, given the high levels of educational attainment and the skewed racial distribution of the sample (>90% White). Lack of diversity in studies is a major challenge in dementia and Alzheimer disease research broadly, and the validation of measures or development of norms requires careful attention with respect to race and ethnicity (32–35). However, our primary objective was not validation but rather to compare the measurement of prevalent and incident dementia. While education and race influence cognitive testing, it is less likely that these characteristics would substantially alter comparisons made between associations with incident dementia and associations with prevalent dementia (36–38).

Overall, this study provided strong evidence showing differences in associations between items assessing cognition and functional limitations and either incident or prevalent dementia. Results can be used to help guide the selection of items for use in the planning of future studies seeking to ascertain dementia status. Findings demonstrate that the identification of incident (early-stage) or prevalent (late-stage) clinical dementia is a crucial consideration for the design of epidemiologic studies on dementia. We recommend that investigators think deeply about the goals of the measurement instruments they create to ascertain dementia.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, United States (Emma Nichols, Derek K. Ng, Jennifer A. Deal, Alden L. Gross); Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, United States (Bryan D. James); Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois, United States (Bryan D. James); and Cochlear Center for Hearing and Public Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, United States (Jennifer A. Deal).

This work was supported by the National Institutes of Health/National Institute on Aging (grants T32AG000247 and F31AG074599 to E.N., grants P30AG010161 and R01AG017917 to B.D.J., grant K01AG054693 to J.A.D., and grant K01AG050699 to A.L.G.) and the Cochlear Center for Hearing and Public Health, Johns Hopkins University.

We used deidentified data from the Religious Orders Study and Memory and Aging Project. Researchers can apply to access these data at https://www.radc.rush.edu/.

This work was presented at the 55th Annual Meeting of the Society for Epidemiologic Research, Chicago, Illinois, June 14–17.

The views expressed in this article are those of the authors and do not reflect those of the National Institutes of Health.

Conflict of interest: none declared.

REFERENCES

1. Alzheimer's Disease International. *World Alzheimer Report 2015. The Global Impact of Dementia: An Analysis of Prevalence, Incidence, Cost and Trends.* London, United Kingdom: Alzheimer's Disease International; 2015. https://www.alz.co.uk/research/world-report-2015. Accessed April 8, 2018.
2. Brookmeyer R, Johnson E, Ziegler-Graham K, et al. Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement.* 2007;3(3):186–191.
3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition: DSM-5*. Washington, DC: American Psychiatric Association; 2013.
4. Nichols E, Szoeke CEI, Vollset SE, et al. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* 2019;18(1):88–106.
5. Hong CH, Falvey C, Harris TB, et al. Anemia and risk of dementia in older adults: findings from the Health ABC Study. *Neurology.* 2013;81(6):528–533.
6. Sheehan OC, Haley WE, Howard VJ, et al. Stress, burden, and well-being in dementia and nondementia caregivers: insights from the Caregiving Transitions Study. *Gerontologist.* 2021;61(5):670–679.
7. Copeland JRM, Dewey ME, Griffiths-Jones HM. A computerized psychiatric diagnostic system and case nomenclature for elderly subjects: GMS and AGECAT. *Psychol Med.* 1986;16(1):89–99.
8. Prince MJ, de Rodriguez JL, Noriega L, et al. The 10/66 Dementia Research Group's fully operationalised DSM-IV dementia computerized diagnostic algorithm, compared with the 10/66 dementia algorithm and a clinician diagnosis: a population validation study. *BMC Public Health.* 2008;8(1):219.
9. Eramudugolla R, Mortby ME, Sachdev P, et al. Evaluation of a research diagnostic algorithm for DSM-5 neurocognitive disorders in a population-based cohort of older adults. *Alz Res Therapy.* 2017;9(1):15.
10. Sutin AR, Stephan Y, Terracciano A. Verbal fluency and risk of dementia. *Int J Geriatr Psychiatry.* 2019;34(6):863–867.
11. Liu H, Zhang Z, Choi S, et al. Marital status and dementia: evidence from the Health and Retirement Study. *J Gerontol B Psychol Sci Soc Sci.* 2020;75(8):1783–1795.
12. Gemmill A, Weiss J. The relationship between fertility history and incident dementia in the US Health and Retirement Study. *J Gerontol B Psychol Sci Soc Sci.* 2022;77(6):1118–1131.
13. Samtani MN, Raghavan N, Novak G, et al. Disease progression model for Clinical Dementia Rating–Sum of Boxes in mild cognitive impairment and Alzheimer's subjects from the Alzheimer's Disease Neuroimaging Initiative. *Neuropsychiatr Dis Treat.* 2014;10:929–952.
14. Proust-Lima C, Amieva H, Dartigues J-F, et al. Sensitivity of four psychometric tests to measure cognitive changes in brain aging-population–based studies. *Am J Epidemiol.* 2007;165(3):344–350.
15. Kempen GI, Suurmeijer TP. The development of a hierarchical polychotomous ADL-IADL scale for noninstitutionalized elders. *Gerontologist.* 1990;30(4):497–502.
16. Bennett DA, Buchman AS, Boyle PA, et al. Religious Orders Study and Rush Memory and Aging Project. *J Alzheimers Dis.* 2018;64(suppl 1):S161–S189.
17. Gelman A. Scaling regression inputs by dividing by two standard deviations. *Stat Med.* 2008;27(15):2865–2873.
18. Katz S, Ford AB, Moskowitz RW, et al. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *JAMA.* 1963;185(12):914–919.
19. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *J Am Geriatr Soc.* 1983;31(12):721–727.
20. Bennett DA, Schneider JA, Aggarwal NT, et al. Decision rules guiding the clinical diagnosis of Alzheimer's disease in two community-based cohort studies compared to standard practice in a clinic-based cohort study. *Neuroepidemiology.* 2006;27(3):169–176.
21. van Hooren SAH, Valentijn AM, Bosma H, et al. Cognitive functioning in healthy older adults aged 64–81: a cohort study into the effects of age, sex, and education. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn.* 2007;14(1):40–54.
22. Sloan FA, Wang J. Disparities among older adults in measures of cognitive function by race or ethnicity. *J Gerontol B Psychol Sci Soc Sci.* 2005;60(5):P242–P250.
23. Dubois B, Hampel H, Feldman HH, et al. Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. *Alzheimers Dement.* 2016;12(3):292–323.
24. Galvin JE, Roe CM, Powlishta KK, et al. The AD8: a brief informant interview to detect dementia. *Neurology.* 2005;65(4):559–564.
25. Vitaliano PP, Breen AR, Russo J, et al. The clinical utility of the Dementia Rating Scale for assessing Alzheimer patients. *J Chronic Dis.* 1984;37(9-10):743–753.
26. Paul R, Moser D, Cohen R, et al. Dementia severity and pattern of cognitive performance in vascular dementia. *Appl Neuropsychol.* 2001;8(4):211–217.
27. Zec RF, Landreth ES, Vicari SK, et al. Alzheimer Disease Assessment Scale: useful for both early detection and staging of dementia of the Alzheimer type. *Alzheimer Dis Assoc Disord.* 1992;6(2):89–102.
28. Feher EP, Mahurin RK, Doody RS, et al. Establishing the limits of the Mini-Mental State: examination of 'subtests'. *Arch Neurol.* 1992;49(1):87–92.
29. Yim D, Yeo TY, Park MH. Mild cognitive impairment, dementia, and cognitive dysfunction screening using machine learning. *J Int Med Res.* 2020;48(7):0300060520936881.
30. Zhu F, Li X, Tang H, et al. Machine learning for the preliminary diagnosis of dementia. *Sci Program.* 2020;2020:e5629090.
31. Hardy SE, Allore H, Studenski SA. Missing data: a special challenge in aging research. *J Am Geriatr Soc.* 2009;57(4):722–729.
32. Mooldijk SS, Licher S, Wolters FJ. Characterizing demographic, racial, and geographic diversity in dementia research: a systematic review. *JAMA Neurol.* 2021;78(10):1255–1261.
33. Gilmore-Bykovskyi AL, Jin Y, Gleason C, et al. Recruitment and retention of underrepresented populations in Alzheimer's disease research: a systematic review. *Alzheimers Dement (N Y).* 2019;5:751–770.
34. Manly JJ. Advantages and disadvantages of separate norms for African Americans. *Clin Neuropsychol.* 2005;19(2):270–275.
35. Brickman AM, Cabo R, Manly JJ. Ethical issues in cross-cultural neuropsychology. *Appl Neuropsychol.* 2006;13(2):91–100.

36. Manly JJ. Race, culture, education, and cognitive test performance among older adults. In: *Handbook of Cognitive Aging: Interdisciplinary Perspectives*. Thousand Oaks, CA: Sage Publications, Inc.; 2008:398–417.
37. Jones RN, Gallo JJ. Education and sex differences in the Mini-Mental State Examination: effects of differential item functioning. *J Gerontol B Psychol Sci Soc Sci.* 2002; 57(6):P548–P558.
38. Gibbons LE, Crane PK, Mehta KM, et al. Multiple, correlated covariates associated with differential item functioning (DIF): accounting for language DIF when education levels differ across languages. *Ageing Res.* 2011;2(1):19–25.