

# Predictors of dementia misclassification when using brief cognitive assessments

Janice M. Ranson, MSc, Elżbieta Kuźma, PhD, William Hamilton, MD, Graciela Muniz-Terrera, PhD, Kenneth M. Langa, MD, PhD, and David J. Llewellyn, PhD

## Correspondence

Dr. Llewellyn  
david.llewellyn@exeter.ac.uk

*Neurology: Clinical Practice* April 2019 vol. 9 no. 2 109-117 doi:10.1212/CPJ.0000000000000566

## Abstract

### Background

Brief cognitive assessments can result in false-positive and false-negative dementia misclassification. We aimed to identify predictors of misclassification by 3 brief cognitive assessments; the Mini-Mental State Examination (MMSE), Memory Impairment Screen (MIS) and animal naming (AN).

### Methods

Participants were 824 older adults in the population-based US Aging, Demographics and Memory Study with adjudicated dementia diagnosis (DSM-III-R and DSM-IV criteria) as the reference standard. Predictors of false-negative, false-positive and overall misclassification by the MMSE (cut-point <24), MIS (cut-point <5) and AN (cut-point <9) were analysed separately in multivariate bootstrapped fractional polynomial regression models. Twenty-two candidate predictors included sociodemographics, dementia risk factors and potential sources of test bias.

### Results

Misclassification by at least one assessment occurred in 301 (35.7%) participants, whereas only 14 (1.7%) were misclassified by all 3 assessments. There were different patterns of predictors for misclassification by each assessment. Years of education predicts higher false-negatives (odds ratio [OR] 1.23, 95% confidence interval [95% CI] 1.07–1.40) and lower false-positives (OR 0.77, 95% CI 0.70–0.83) by the MMSE. Nursing home residency predicts lower false-negatives (OR 0.15, 95% CI 0.03–0.63) and higher false-positives (OR 4.85, 95% CI 1.27–18.45) by AN. Across the assessments, false-negatives were most consistently predicted by absence of informant-rated poor memory. False-positives were most consistently predicted by age, nursing home residency and non-Caucasian ethnicity (all  $p < 0.05$  in at least 2 models). The only consistent predictor of overall misclassification across all assessments was absence of informant-rated poor memory.

### Conclusions

Dementia is often misclassified when using brief cognitive assessments, largely due to test specific biases.



## MORE ONLINE

### Podcast

[NPub.org/NCP/podcast9-2b](http://NPub.org/NCP/podcast9-2b)

Clinical guidelines for dementia recommend a standardised cognitive assessment for suspected dementia in primary care, although it is uncertain which assessments are most appropriate.<sup>1</sup>

From the University of Exeter Medical School (J.M.R., E.K., W.H., D.J.L.); Centre for Dementia Prevention, University of Edinburgh (G.M.-T.), United Kingdom; The Alan Turing Institute (D.J.L.), London, United Kingdom; and Medical School (K.M.L.), Institute for Social Research, Institute for Healthcare Policy and Innovation, Veterans Affairs Ann Arbor Healthcare System, University of Michigan.

Funding information and disclosures are provided at the end of the article. Full disclosure form information provided by the authors is available with the full text of this article at [Neurology.org/cp](http://Neurology.org/cp).

Diagnostic accuracy studies of brief cognitive assessments suitable for use in primary care frequently recommend the Mini-mental State Examination (MMSE)<sup>2–5</sup> and memory impairment screen (MIS)<sup>6–10</sup> although diagnostic accuracy estimates vary considerably.<sup>9</sup> One alternative is animal naming (AN) which may have utility as a very brief assessment.<sup>11,12</sup>

Brief cognitive assessments can result in false-negative and false-positive misclassification of dementia compared to a gold-standard diagnosis.<sup>13</sup> Potential consequences of misclassification depend on the clinical circumstance and the context in which they are used, for example as a screening tool, for case-finding in clinical practice or in response to presentation of signs and symptoms of dementia. False-negatives may prevent or delay diagnosis, meaning missed opportunities for planning and timely access to treatment and services. False-positives may cause unnecessary referrals and investigations, affecting patients, families and health services.

Predictors of dementia misclassification by brief cognitive assessments have not been studied in-depth. The MMSE is the most studied assessment, with biases including age, education and ethnicity.<sup>14,15</sup> An analysis of highly educated participants found 16 or more years of education is associated with false-negatives by the MMSE, and an education-adjusted cut-off may improve diagnostic accuracy.<sup>16</sup> An investigation of factors contributing to false-positives found associations with old age, manual social class and visual impairment.<sup>17</sup> Additionally, nurse assessments of MMSE performance suggest sensory impairment, low education and stroke may be associated with false-positives.<sup>18</sup> There is a lack of evidence for factors predicting misclassification across a variety of assessments. This knowledge may assist clinical decision-making and inform dementia identification strategies. The objective of this study was to identify predictors of dementia misclassification when using brief cognitive assessments to detect possible dementia in primary care.

## Methods

### Data and study population

We used data from the Aging, Demographics and Memory Study (ADAMS) described in detail elsewhere.<sup>19</sup> The ADAMS provides a comprehensive neuropsychological assessment and dementia diagnosis for a subsample of the Health and Retirement Study (HRS), a population-based longitudinal study of aging in the US.<sup>20</sup> Eight hundred fifty-six participants aged 70–110 completed the ADAMS baseline assessment between August 2001 and January 2004. A clinical research nurse and neuropsychology technician conducted a 3- to 4-hour assessment in the participant's home. This included a standardised neurologic examination, blood pressure reading using a blood pressure cuff, pulse taken from the wrist, inspection of medications, collection of buccal DNA samples for APOE genotyping, self-report depression measure using the Composite International Diagnostic Interview,<sup>21</sup> comprehensive neuropsychological

test battery and a 7-minute video recording of cognitive and neurologic assessments. An informant was interviewed regarding the participant's medical history, current medications, lifestyle, family history and functional and cognitive abilities. Missing data on ADAMS variables were supplemented by equivalent data collected in the HRS assessment carried out closest to the ADAMS assessment. We used a cleaned and harmonised version of the HRS data provided by the RAND Center for the Study of Aging.<sup>22</sup>

### Standard protocol approvals, registrations, and patient consents

The ADAMS was approved by the Institutional Review Boards of both the University of Michigan and Duke University. All participants provided written informed consent.

### Dementia assessment

Dementia diagnoses were adjudicated by an expert panel of geropsychiatrists, neurologists, internists and cognitive neuroscientists. Diagnoses were assigned using *Diagnostic and Statistical Manual of Mental Disorders* (DSM) III-R<sup>23</sup> and DSM-IV<sup>24</sup> criteria based on all information obtained during the assessment and revised based on available medical record history if appropriate. Specific diagnoses were then grouped into 3 diagnostic categories of “all-cause dementia,” “cognitive impairment, not dementia” (CIND) and “normal cognitive function.” The reference standard in the present analysis combines the CIND and “normal cognitive function” categories, resulting in a binary diagnosis of “all-cause dementia” or “no dementia.”

### Brief cognitive assessments

Assessments included the MMSE,<sup>25</sup> MIS<sup>26</sup> and AN.<sup>27</sup> The MMSE consists of 22 items measuring orientation to time (5 points), orientation to place (5 points), registration of 3 words (3 points), attention and calculation (5 points), recall of 3 words (3 points), language (8 points) and visual construction (1 point).<sup>28</sup> Scores range from 0 to 30, with lower scores representing poorer cognitive function. The most commonly used cut-off scores for dementia are <24 and <25.<sup>13</sup> We used <24 which provided greater overall diagnostic accuracy in the analytic sample. The MIS consists of free and cued 4-word recall. Scores range from 0 to 8 and a score of <5 indicates probable dementia. AN is a semantic verbal fluency test requiring as many animals as possible to be named within 60 seconds, with no maximum score. In the ADAMS total scores ranged from 0 to 33. There is no agreed clinically defined cut-off for AN, though those used by previous research studies include <9,<sup>29</sup> <12<sup>30</sup> and <14.<sup>12,31</sup> We used <9, which provided the greatest overall diagnostic accuracy in the analytic sample. For all 3 assessments we regarded items not completed due to cognitive impairment as zero. Where a test was discontinued due to cognitive impairment the completed items were scored. Items not completed for any reason other than cognitive impairment were regarded as missing responses.

## Outcome variables

For each of the 3 assessments, binary outcome variables were false-negative (vs true-positive), false-positive (vs true-negative), and overall misclassification (false-negatives and false-positives vs true-negatives and true-positives).

## Predictor variables

### Sociodemographics

Sex (male/female), education (years of formal education), ethnicity (Caucasian/African American/Hispanic) and nursing home status (nursing home/living independently) were based on self- or proxy-reports in the HRS wave from which the participant was recruited. Age (years) was recorded at the time of the baseline ADAMS assessment.

### Subjective cognition

For both self-rated and informant-rated poor present memory, ratings of “poor” memory were contrasted to a combined category of “fair,” “good,” “very good” or “excellent” memory. Self-rated memory decline was defined by a memory rating of “worse” or “much worse” vs “same” “better” or “much better” compared to 2 years ago. Informant-rated memory decline was defined by an affirmative report of the participant having more difficulty with memory or thinking than in the past.

### Sources of systematic measurement error/bias

Motor impairment (including paresis or a tremor in the dominant hand) and illiteracy were based on the ADAMS neuropsychological examination post-visit report. Visual impairment and auditory impairment were based on self- and informant-reported poor vision and hearing (excluding adequately corrected impairment).

### Dementia risk factors

Risk factors were based on Deckers et al.<sup>32</sup> systematic review, including both modifiable and non-modifiable factors identified by population-based prospective studies. We selected variables investigated in at least 5 studies, with a statistically significant association with dementia ( $p < 0.05$ ) in 50% or more of the studies. This resulted in 18 potential risk factors. Of these, 10 were available in the current dataset. These were depression, hyperlipidemia, hypertension, heart problems, diabetes, stroke, APOE E4, smoking status, physical activity and self-reported cognitive impairment (included in subjective cognition above). History of depression was defined by an informant-reported 2-week period of feeling sad, blue or depressed at any time. Hyperlipidemia (high cholesterol or high triglycerides) was based on informant-reported diagnosis at any time. History of hypertension, heart problems (including heart attack, heart disease, coronary thrombosis, other heart problems or current prescription for heart medication), diabetes and stroke (excluding TIA) were defined by self- or informant-reported diagnosis at any time. Hypertension, heart problems and diabetes were also defined if the participant was in receipt of prescription medication for the condition, and hypertension was additionally defined by systolic blood pressure  $\geq 140$  mm Hg or diastolic blood

pressure  $\geq 90$  mm Hg<sup>33</sup> on 2 consecutive measurements during the examination. APOE E4 status was based on a buccal sample identifying at least one APOE  $\epsilon 4$  allele. Current smoking was defined by either self- or informant-reported current smoking of cigarettes or cigars. Self-reported low physical activity was defined by participation in vigorous physical activity on average less than 3 times per week<sup>34</sup> during the previous 12 months.

## Statistical analysis

### Classification accuracy

Receiver operating characteristic (ROC) curves and areas under the curve (AUC) were used to compare the discrimination performance of the assessments. We calculated the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of each assessment at the specified cut-points using a 13.9% population prevalence of dementia in US adults aged 70 and older.<sup>35</sup>

### Identification of predictors

To identify independent predictors of false-negative, false-positive and overall misclassification on each assessment, we followed the multivariable fractional polynomial approach recommended by Royston and Sauerbrei.<sup>36</sup> Logistic regression using backwards elimination (BE) with a rejection criterion of  $p \geq 0.05$  was conducted on 1000 random bootstrap replication samples with replacement. In each replication, BE was combined with a closed test function selection procedure (FSP). FSP is a systematic search for non-linearity that investigates the most appropriate functional form for each continuous predictor, considering first- and second-order fractional polynomial terms to 4 degrees of freedom. The incorporation of this procedure when selecting variables ensures that variables are not erroneously disregarded due to inappropriate modelling of non-linear associations. The association between a continuous predictor and the outcome is modelled as linear unless a more complex term provides a significantly better fit ( $\alpha = 0.05$ ).

### Misclassification of dementia

Potential predictors of false-negatives, false-positives and overall misclassification on each assessment were identified based on the proportion of times they were selected by the BE regression model in the bootstrapped samples. For each of the analyses, variables with a bootstrap inclusion fraction (BIF) of 50% or more were entered in to a final multivariate logistic regression model. The functional form of continuous predictors most frequently selected by the FSP was used in the final model. Odds ratios, 95% confidence intervals and  $p$ -values were calculated based on the final regression models. All analyses were performed using STATA, version 14.2.<sup>37</sup>

## Sensitivity analyses

### Missing data

Primary analyses were restricted to participants with complete data on all outcome and predictor variables. In the first sensitivity analysis, the final regression models were re-

**Table 1** AUC, sensitivity, specificity, PPV and NPV of the MMSE, MIS and AN with a dementia prevalence of 13.9% (N = 824)

	AUC	Sensitivity	Specificity	PPV	NPV
<b>MMSE (&lt;24)</b>	0.94	93.1 (89.6–95.8)	71.3 (67.2–75.1)	34.4 (31.3–37.5)	98.5 (97.7–99.0)
<b>MIS (&lt;5)</b>	0.91	82.1 (77.2–86.4)	85.4 (82.1–88.3)	47.5 (42.3–52.8)	96.7 (95.8–97.4)
<b>AN (&lt;9)</b>	0.92	78.4 (73.2–82.9)	89.5 (86.6–92.0)	54.6 (48.3–60.8)	96.2 (95.4–97.0)

Abbreviations: AN = animal naming; AUC = area under the curve; MIS = memory impairment screen; MMSE = mini-mental state examination; NPV = negative predictive value; PPV = positive predictive value.

estimated using multiply imputed data for all of the predictors (analyses were still restricted to those with complete data on all 3 brief cognitive assessments). Multivariate imputation by chained equations was conducted according to guidance by White et al.<sup>38</sup> Predictive mean matching to 3 nearest neighbours was used for continuous variables, and logistic regression for binary variables. Variables were imputed iteratively, from the one with the least missing data to the most. The imputation model contained a wider range of auxiliary variables than the predictors used in the present analyses (table e-1, [links.lww.com/CPJ/A71](https://links.lww.com/CPJ/A71)). Twenty data sets were imputed, this number being greater than the percentage of cases with incomplete data (15.5%).

### Lower dementia prevalence

Individuals with poor cognitive function were oversampled in the ADAMS cohort, thus inflating the prevalence of dementia. In the second sensitivity analysis we estimated the final regression models using survey weights to adjust the

proportion of dementia cases to a population-representative prevalence of 13.9%.

### Data availability

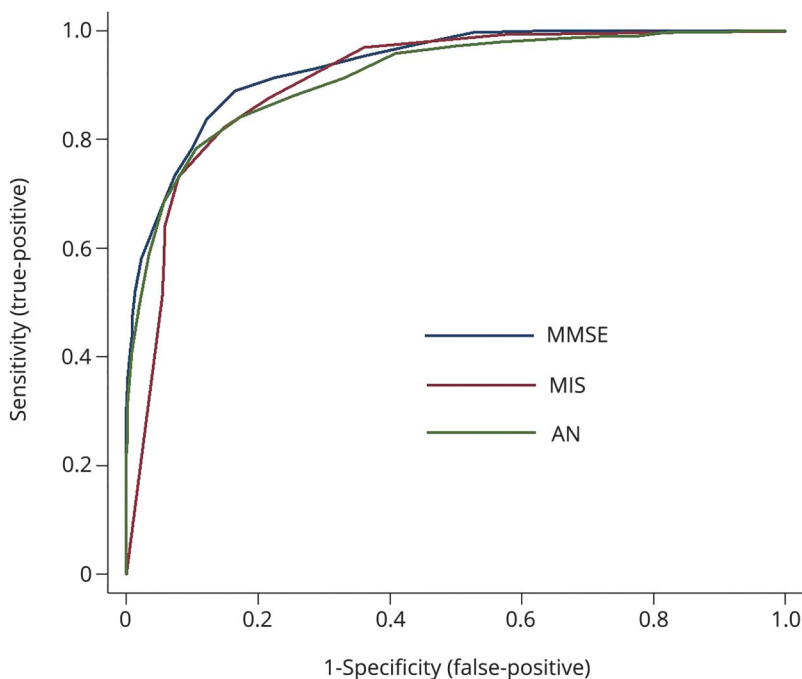
All HRS and RAND HRS public data are available online from the HRS to registered users without restriction. Sensitive health data in the ADAMS are additionally available to researchers who qualify for access. More information can be found at ([hrs.isr.umich.edu/data-products](https://hrs.isr.umich.edu/data-products)).

## Results

### Sample characteristics

Of the 856 participants who completed the ADAMS baseline assessment, 824 had complete data on all outcome and predictor variables. Detailed characteristics of the study participants as classified by each cognitive assessment are presented in table e-2, [links.lww.com/CPJ/A71](https://links.lww.com/CPJ/A71). The

**Figure** Discrimination accuracy of the MMSE, MIS and AN for all-cause dementia (N = 824)



## There were no consistent predictors of either false-negatives or false-positives across all assessments.

prevalence of dementia in this sample was 35.3% (N = 291) and of those without dementia, 231 (43.3%) met the criteria for CIND. The sample had a mean age of 81.62 (SD 7.11), 10.14 mean years of education (SD = 4.29), 479 (58.1%) females and 77 (9.3%) resident in a nursing home. The ethnicity of the sample consisted of 595 (72.2%) Caucasians, 148 (18.0%) African Americans and 81 (9.8%) Hispanics.

### Classification accuracy of the MMSE, MIS and AN

Table 1 reports the AUC, sensitivity, specificity, PPV and NPV for each assessment. The AUC for the MMSE is greater than for the MIS ( $p = 0.02$ ). The AUC for AN did not differ from either the MMSE or the MIS (both  $p > 0.05$ ). The ROC curves are shown in the figure

### Rates of misclassification

Misclassification by at least one assessment occurred in 301 (35.7%) cases, 113 (13.4%) were misclassified by 2 or more, and 14 (1.7%) were misclassified by all 3. Overall dementia misclassification rates for the MMSE, MIS and AN, were 21%, 16% and 14% respectively. Rates of false-negatives and false-positives for each assessment are shown in table 2. For the MMSE, MIS and AN, the number of participants with false-positives that met the criteria for CIND were 114 (74.5%), 64 (82.1%) and 46 (82.1%) respectively. The number with true-negatives that met the criteria for CIND were lower; 117 (30.8%), 167 (36.7) and 185 (38.8) for each assessment respectively.

### Selection of predictors for inclusion in the final models

Of the 22 candidate predictors, 6 did not have a BIF  $\geq 50\%$  in any of the models and so were not analysed any further.

These were diabetes, sex, hearing impairment, motor impairment, self-rated memory decline and stroke. See table e-3, [links.lww.com/CPJ/A71](https://links.lww.com/CPJ/A71) for the BIF of each predictor.

### Predictors of false-negative and false-positive misclassification

Predictors of false-negative and false-positive misclassification in the 6 final logistic regression models are shown in table 3. All functional forms of continuous variables were linear, with the exception of age, for which a second-order fractional polynomial transformation (3, 3) was used in model 2.

Seven variables predicted false-negatives on only one assessment: Education for the MMSE; age, APOE E4 non-carrier, depression, and absence of informant-rated memory decline for the MIS; not residing in a nursing home and physical activity for AN. Absence of self- and informant-rated poor memory both predicted false-negatives on the MIS and AN. There were no consistent predictors of false-negatives across all assessments.

Five variables predicted false-positives on only one assessment: Lower education, illiteracy, visual impairment, and APOE E4 for the MMSE; and Hispanic ethnicity for AN. No predictors were specific to false-positives on the MIS. Three predictors were associated with false-positives on 2 assessments: Age for the MIS and AN; and nursing home residency and African-American ethnicity for the MMSE and AN. There were no consistent predictors of false-positives across all assessments.

### Predictors of overall misclassification

Predictors of overall misclassification in the final logistic regression models for each cognitive assessment are shown in table 4. All functional forms of continuous variables were linear. Seven variables predicted overall misclassification on only one assessment: African-American ethnicity for the MMSE; age, visual impairment, APOE E4 non-carrier and depression for the MIS; No hyperlipidemia and informant-rated memory decline on AN. Lower education and heart problems predicted overall misclassification on both the MMSE and AN. Absence of informant-rated

**Table 2** Classification of dementia by the MMSE, MIS and AN (N = 824)

Reference diagnosis	Classification	MMSE <24		MIS <5		AN <9	
		N	%	N	%	N	%
<b>Dementia</b>	True-positive	271	32.9	239	29.0	228	27.7
	False-negative	20	2.4	52	6.3	63	7.7
<b>No dementia</b>	True-negative	380	46.1	455	55.2	477	57.9
	False-positive	153	18.6	78	9.5	56	6.8

Abbreviations: AN = animal naming; MIS = memory impairment screen; MMSE = mini-mental state examination.

## Absence of informant-rated poor memory predicted overall misclassification on all assessments.

poor memory predicted overall misclassification on all assessments.

### Sensitivity analyses

#### Missing data

Similar results were obtained by estimating the final regression models using multiply imputed data (tables e-4 and e-5, links. [lww.com/CPJ/A71](http://lww.com/CPJ/A71)). All associations remained consistent with the complete case analyses, with the exception of 2 slight

differences relating to the prediction of false-negatives on the MIS: Smoking emerged as a significant predictor, and informant-rated memory decline became non-significant.

#### Lower dementia prevalence

Using survey weights to reflect a lower dementia prevalence produced a similar pattern of results (tables e-6 and e-7, links. [lww.com/CPJ/A71](http://lww.com/CPJ/A71)). Some slight differences emerged, although all associations were in the same direction as in the primary analyses. For false-negative misclassification, 3 variables became significant: Non-nursing home residency and absence of informant-rated poor memory for MMSE, and smoking for the MIS. Two variables became non-significant: Education for the MMSE and age for the MIS. For false-positive misclassification, 2 variables became non-significant: Illiteracy for the MMSE, and Hispanic ethnicity for AN. For overall misclassification, 4 predictors became non-significant:

**Table 3** Predictors of false-negative and false-positive misclassification of dementia by the MMSE, MIS and AN brief cognitive assessments at the chosen thresholds in logistic regression models

Predictors	Predictors of false-negative misclassification (N=291)			Predictors of false-positive misclassification (N=533)		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	MMSE	MIS	AN	MMSE	MIS	AN
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age (p1)		1.24 <sup>a</sup> ,* (1.03–1.49)			1.12 <sup>***</sup> (1.08–1.16)	1.10 <sup>***</sup> (1.05–1.15)
Age (p2)		0.92 <sup>a</sup> ,* (0.85–0.99)				
Nursing home resident	0.16 (0.02–1.30)		0.15 <sup>**</sup> (0.03–0.63)	8.59 <sup>**</sup> (1.98–37.35)		4.85* (1.27–18.45)
Years of education	1.23 <sup>**</sup> (1.07–1.40)			0.77 <sup>***</sup> (0.70–0.83)		
African-American ethnicity				3.61 <sup>***</sup> (1.90–6.86)		4.50 <sup>***</sup> (2.27–8.93)
Hispanic ethnicity				1.12 (0.48–2.64)		3.86 <sup>**</sup> (1.71–8.72)
Illiteracy				4.87 <sup>**</sup> (1.48–15.97)		
Visual impairment				2.54 <sup>***</sup> (1.54–4.17)		
APOE E4		0.21 <sup>***</sup> (0.09–0.50)		2.02* (1.15–3.56)		
Smoking		3.64 (0.98–13.51)				
Hyperlipidemia						0.58 (0.29–1.17)
Low physical activity			0.33* (0.15–0.77)			
Depression		4.19 <sup>***</sup> (2.01–8.73)				
Hypertension	4.12 (0.52–32.44)					
Self-rated poor memory		0.30* (0.10–0.85)	0.32* (0.11–0.87)			
Informant-rated poor memory	0.23 (0.05–1.11)	0.19 <sup>**</sup> (0.07–0.50)	0.29 <sup>**</sup> (0.12–0.67)			
Informant-rated memory decline		0.09 <sup>**</sup> (0.02–0.50)				

Abbreviations: AN = animal naming; CI = confidence interval; MIS = memory impairment screen; MMSE = mini-mental state examination; OR = odds ratio. The ORs are for a 1-year increase in age or years of education.

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

<sup>a</sup>A fractional polynomial transformation (3, 3) was used for age.

**Table 4** Predictors of overall misclassification of dementia by the MMSE, MIS and AN brief cognitive assessments at the chosen thresholds in logistic regression models (N = 824)

Predictors	Model 7	Model 8	Model 9
	MMSE	MIS	AN
	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age		1.06*** (1.03–1.09)	
Years of education	0.86*** (0.81–0.90)		0.91*** (0.87–0.95)
African-American ethnicity	2.49*** (1.55–4.00)		
Visual impairment		1.66* (1.11–2.48)	
APOE E4		0.50** (0.30–0.46)	
Heart problems	1.48* (1.01–2.18)		1.92** (1.23–3.00)
Hyperlipidemia			0.45** (0.28–0.75)
Depression		2.00** (1.33–3.00)	
Informant-rated poor memory	0.10*** (0.04–0.29)	0.19*** (0.08–0.46)	0.34** (0.16–0.73)
Informant-rated memory decline			1.71* (1.08–2.72)

Abbreviations: AN = animal naming; CI = confidence intervals; MIS = memory impairment screen; MMSE = mini-mental state examination; OR = odds ratio. The ORs are for a 1-year increase in age or years of education.

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

heart problems for the MMSE, and APOE E4 non-carrier, depression and absence of informant-rated poor memory for the MIS.

## Discussion

This study investigated predictors of misclassification of dementia when using 3 brief cognitive assessments. Over a third of participants were misclassified overall by at least one assessment, whereas very few were misclassified on all 3 assessments. A wide range of predictors were identified, though these were generally specific to one assessment. Only one predictor was consistent across all assessments: Absence of informant-rated poor memory for overall misclassification.

Our finding that higher education is associated with false-negative misclassification, and lower education is associated with false-positive misclassification by the MMSE is consistent with previous research, and supports the suggestion of education-adjusted scoring.<sup>14–16,18</sup> Our findings also concur with previous studies suggesting non-Caucasian ethnicity<sup>14,15</sup> and visual impairment<sup>17,18</sup> are associated with false-positive misclassification by the MMSE. However, the present study did not find that age<sup>14,15,17</sup> or stroke<sup>18</sup> are predictors of misclassification by the MMSE when adjusting for the other factors. Older age was however predictive of false-positive misclassification by the MIS and AN. Stroke was not a predictor in any of the models.

We found nursing home residency to be associated with lower scores on the AN assessment (predicting lower false-negatives and greater false-positives). The reasons for this

require investigation and may include a tendency for reduced verbal communication in this setting. A lower threshold for AN scores may therefore reduce dementia misclassification in nursing home residents. This finding may also reflect a potential association between acute illness and false positives, due to nursing home placement following hospitalization. In this case, re-assessment following recovery may be worthwhile. Overall misclassification was associated with absence of informant-rated poor memory for all 3 assessments, suggesting the importance of considering informant-reported cognition in addition to brief cognitive assessment results. This can be done by incorporating informant-reported cognition into the test (e.g., the General Practitioner assessment of Cognition<sup>39</sup>) or by combining brief cognitive assessment results with a complementary informant-rating scale (e.g., the Informant Questionnaire on Cognitive Decline<sup>40</sup>).

This study has some limitations. The brief cognitive assessment results contributed to the reference standard dementia diagnosis made by the consensus panel in the analytic dataset. Consequently, there is a risk of incorporation bias with overestimation of diagnostic accuracy.<sup>41</sup> Therefore, the overall misclassification rates, which were not themselves the focus of this study, may be underestimated. However, this would not affect interpretation of the predictors of misclassification. Similarly, the predictors used in this study were available to the consensus panel. They may therefore have contributed to consideration of whether a participant met the criteria for dementia, although any potential incorporation bias would reduce, rather than increase the strength of detecting an association with misclassification. Eight candidate predictors

were not available in the analytic dataset: Mid-life obesity, renal dysfunction, high homocysteine, metabolic syndrome, inflammation, unsaturated fat intake, gait problems and frailty. These potential predictors therefore warrant investigation in future studies. Finally, it would also be useful to investigate alternative brief cognitive assessments, for example the Mini-Cog and GPCOG which have been recommended for dementia identification.<sup>42</sup>

Strengths of this study include identification of the predictors of dementia misclassification across a range of brief cognitive assessments, highlighting the differential prediction of false-negative and false-positive misclassification of dementia. The use of a large population-based cohort with adjudicated dementia diagnosis, and analysis of a wide evidence-based selection of candidate predictors, should provide a reliable estimate of the prevalence and predictors of misclassification by brief cognitive assessments. Furthermore, use of advanced modelling techniques such as bootstrapping and fractional polynomials improve the stability in the selection of predictors and selection of the appropriate functional form for continuous predictors.

Given that overall rates of misclassification by brief cognitive assessments are substantial (14%–21% in the current study), there is considerable scope for improvement. Knowledge of factors which predict misclassification and are readily available in clinical practice may improve clinical decision making by enhancing the selection and interpretation of assessments. If an assessment is known to produce biased results for a given patient group, an alternative and more appropriate assessment can be selected. For example, misclassification by the MMSE is related to African-American ethnicity. Alternatively, stratified cut-points could be provided to adjust for known biases, for example level of education on the MMSE and AN. Failing that, clinical judgement can be used to help interpret assessment results more appropriately when the biases affecting misclassification on a given assessment are known.

It is interesting to note that for all 3 assessments, the majority of participants without dementia who had false-positive assessment results met the criteria for CIND (75%–83% across the 3 assessments). These participants may be in the very early stages of conversion to dementia. In addition to potential biases contributing to misclassification, this should also be considered when deciding on the course of action for patients who perform poorly on a brief cognitive assessment and are not subsequently diagnosed with dementia. Of those without dementia who had true-negative results, a smaller but still considerable proportion met the criteria for CIND (31%–39%). Therefore, for those with low or borderline cognitive assessment results, re-assessment to detect further decline may be appropriate.

Dementia is often misclassified when using brief cognitive assessments, due to a wide range of test specific biases. The only consistent predictor of overall misclassification across all

assessments was absence of informant-rated poor memory. Brief cognitive assessments should therefore take into account both informant-rated cognition and multiple test specific biases in order to improve overall accuracy and clinical decision making.

## Author contributions

J.M. Ranson, drafting and revision of the manuscript, study design and concept, analysis and interpretation of data. E. Kuźma, critical revision of manuscript for intellectual content. W. Hamilton, critical revision of manuscript for intellectual content. K.M. Langa, critical revision of manuscript for intellectual content. G. M.-Terrera, critical revision of manuscript for intellectual content. D.J. Llewellyn, study design and concept, acquisition of data, critical revision of manuscript for intellectual content, study supervision, obtaining funding.

## Study funding

This work was supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1 (D.J. Llewellyn), the Halpin Trust (J.M. Ranson, D.J. Llewellyn and E. Kuźma), Mary Kinross Charitable Trust (D.J. Llewellyn and E. Kuźma), and National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) for the South West Peninsula (D.J. Llewellyn). The HRS is sponsored by the US National Institute on Aging (U01 AG009740), and is conducted by the Institute for Social Research, University of Michigan. K.M. Langa was supported by grants from the National Institute on Aging (P30 AG053760 and P30 AG024824). None of the funding sources had any role in the design of the study, in the analysis and interpretation of the data or in the preparation of the manuscript.

## Disclosure

The authors report no disclosures relevant to the manuscript. Full disclosure form information provided by the authors is available with the full text of this article at [Neurology.org/cp](http://Neurology.org/cp).

## TAKE-HOME POINTS

- Misclassification of dementia by brief cognitive assessments is common, with a different pattern of predictors for false-negative and false-positive misclassification.
- Predictors of misclassification depend on the assessments used, although informant-rated cognition is a consistent predictor of misclassification across all assessments.
- Knowledge of test-specific biases may enhance clinical decision-making by informing the choice of assessment and threshold for certain settings and patient groups.



## Publication history

Received by *Neurology: Clinical Practice* May 30, 2018. Accepted in final form September 17, 2018.

## References

1. Ngo J, Holroyd-Leduc JM. Systematic review of recent dementia practice guidelines. *Age Ageing* 2015;44:25–33.
2. Harvan JR, Cotter V. An evaluation of dementia screening in the primary care setting. *J Am Acad Nurse Pract* 2006;18:351–360.
3. Holsinger T, Plassman BL, Stechuchak KM, Burke JR, Coffman CJ, Williams JW Jr. Screening for cognitive impairment: comparing the performance of four instruments in primary care. *J Am Geriatr Soc* 2012;60:1027–1036.
4. Mitchell AJ, Malladi S. Screening and case finding tools for the detection of dementia. Part I: evidence-based meta-analysis of multidomain tests. *Am J Geriatr Psychiatry* 2010;18:759–782.
5. Mitchell AJ, Malladi S. Screening and case-finding tools for the detection of dementia. Part II: evidence-based meta-analysis of single-domain tests. *Am J Geriatr Psychiatry* 2010;18:783–800.
6. Brodaty H, Low LF, Gibson L, Burns K. What is the best dementia screening instrument for general practitioners to use? *Am J Geriatr Psychiatry* 2006;14:391–400.
7. Holsinger T, Deveau J, Boustani M, Williams JW Jr. Does this patient have dementia? *JAMA* 2007;297:2391–2404.
8. Ismail Z, Rajji TK, Shulman KI. Brief cognitive screening instruments: an update. *Int J Geriatr Psychiatry* 2010;25:111–120.
9. Lin JS, O'Connor E, Rossom RC, Perdue LA, Eckstrom E. Screening for cognitive impairment in older adults: a systematic review for the U.S. preventive services task force. *Ann Intern Med* 2013;159:601–612.
10. Milne A, Culverwell A, Guss R, Tuppen J, Whelton R. Screening for dementia in primary care: a review of the use, efficacy and quality of measures. *Int Psychogeriatr* 2008;20:911–926.
11. Sebaldt R, Dalziel W, Massoud F, et al. Detection of cognitive impairment and dementia using the animal fluency test: the DECIDE study. *Can J Neurol Sci* 2009;36:599–604.
12. Canning SJ, Leach L, Stuss D, Ngo L, Black SE. Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. *Neurology* 2004;62:556–562.
13. Tsoi KK, Chan JY, Hirai HW, Wong SY, Kwok TC. Cognitive tests to detect dementia: a systematic review and meta-analysis. *JAMA Intern Med* 2015;175:1450–1458.
14. Velayudhan L, Ryu S-H, Raczek M, et al. Review of brief cognitive tests for patients with suspected dementia. *Int Psychogeriatr* 2014;26:1247–1262.
15. Boustani M, Peterson B, Hanson L, Harris R, Lohr KN. Screening for dementia in primary care: a summary of the evidence for the U.S. preventive services task force. *Ann Intern Med* 2003;138:927–937.
16. Spering CC, Hobson V, Lucas JA, Menon CV, Hall JR, O'Bryant SE. Diagnostic accuracy of the MMSE in detecting probable and possible Alzheimer's disease in ethnically diverse highly educated individuals: an analysis of the NACC database. *J Gerontol A Biol Sci Med Sci* 2012;67:890–896.
17. Jagger C, Clarke M, Anderson J, Battcock T. Misclassification of dementia by the mini-mental state examination—are education and social class the only factors? *Age and Ageing* 1992;21:404–411.
18. Riih a J, Isoaho R, Ojanlatva A, Viramo P, Sulkava R, Kivel a SL. Poor performance in the mini-mental state examination due to causes other than dementia. *Scand J Prim Health Care* 2001;19:34–38.
19. Langa KM, Plassman BL, Wallace RB, et al. The aging, demographics, and memory study: study design and methods. *Neuroepidemiology* 2005;25:181–191.
20. Juster FT, Suzman R. An overview of the health and retirement study. *J Hum Resour* 1995;30:S7–S56.
21. Robins LN, Wing J, Wittchen HU, et al. The composite international diagnostic interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988;45:1069–1077.
22. RAND HRS Data, Version P. Santa Monica, CA: RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration; 2016.
23. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-III-R*. New York: Springer; 1987.
24. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*; 2000.
25. Folstein MF, Folstein SE, McHugh PR. Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.
26. Buschke H, Kuslansky G, Katz M, et al. Screening for dementia with the memory impairment screen. *Neurology* 1999;52:231–238.
27. Rosen WG. Verbal fluency in aging and dementia. *J Clin Neuropsychol* 1980;2:135–146.
28. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. *J Am Geriatr Soc* 1992;40:922–935.
29. Grober E, Hall C, McGinn M, et al. Neuropsychological strategies for detecting early dementia. *J Int Neuropsychol Soc* 2008;14:130–142.
30. Fuchs A, Wiese B, Altiner A, Wollny A, Pentzek M. Cued recall and other cognitive tasks to facilitate dementia recognition in primary care. *J Am Geriatr Soc* 2012;60:130–135.
31. Heun R, Papassotiropoulos A, Jennessen F. The validity of psychometric instruments for detection of dementia in the elderly general population. *Int J Geriatr Psychiatry* 1998;13:368–380.
32. Deckers K, van Boxtel MP, Schiepers OJ, et al. Target risk factors for dementia prevention: a systematic review and Delphi consensus study on the evidence from observational studies. *Int J Geriatr Psychiatry* 2015;30:234–246.
33. National Institute for Health and Care Excellence. Hypertension in adults: diagnosis and management [online]. Available at: [nice.org.uk/guidance/cg127](http://nice.org.uk/guidance/cg127). Accessed December 6, 2017.
34. World Health Organization. Global recommendations on physical activity for health [online]. Available at: [apps.who.int/iris/bitstream/10665/44399/1/9789241599979\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/44399/1/9789241599979_eng.pdf). Accessed December 6, 2017.
35. Plassman BL, Langa KM, Fisher GG, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology* 2007;29:125–132.
36. Royston P, Sauerbrei W. MFP: multivariable model-building with fractional polynomials. In: *Multivariable Model-building*: John Wiley & Sons, Ltd; 2008: 115–150.
37. *STATA Statistical Software* [computer program]. Version 14. College Station, TX: StataCorp LP; 2015.
38. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:377–399.
39. Brodaty H, Pond D, Kemp NM, et al. The GPCOG: a new screening test for dementia designed for general practice. *J Am Geriatr Soc* 2002;50:530–534.
40. Jorm AF, Korten AE. Assessment of cognitive decline in the elderly by informant interview. *Br J Psychiatry* 1988;152:209–213.
41. Roever L. Types of bias in studies of diagnostic test accuracy. *Evid Based Med Pract* 2016;2:e113.
42. Cordell CB, Borson S, Boustani M, et al. Alzheimer's association recommendations for operationalizing the detection of cognitive impairment during the medicare annual wellness visit in a primary care setting. *Alzheimers Dement* 2013;9:141–150.

---

## Practical Implications

*Neurology® Clinical Practice* is committed to providing clinical insights helpful to neurologists in everyday practice. Each Full Case includes a “Practical Implications” statement, a pearl of wisdom for the practicing clinician.

---