Assessment of Cognition Using Surveys and Neuropsychological Assessment: The Health and Retirement Study and the Aging, Demographics, and Memory Study

Eileen M. Crimmins,[1] Jung Ki Kim,[1] Kenneth M. Langa,[2,3] and David R. Weir[4]

[1]Andrus Gerontology Center, University of Southern California, Los Angeles.
[2]Division of General Medicine, Department of Medicine, University of Michigan, Ann Arbor.
[3]Ann Arbor Veterans Affairs Center for Clinical Management Research, Michigan.
[4]Institute for Social Research, University of Michigan, Ann Arbor.

Objectives. This study examines the similarity of cognitive assessments using 1 interview in a large population study, the Health and Retirement Study (HRS), and a subsample in which a detailed neuropsychiatric assessment has been performed (Aging, Demographics, and Memory Study [ADAMS]).

Methods. Respondents are diagnosed in ADAMS as demented, cognitively impaired without dementia (CIND), or as having normal cognitive function. Multinomial logistic analysis is used to predict diagnosis using a variety of cognitive and noncognitive measures from the HRS and additional measures and information from ADAMS.

Results. The cognitive tests in HRS predict the ADAMS diagnosis in 74% of the sample able to complete the HRS survey on their own. Proxy respondents answer for a large proportion of HRS respondents who are diagnosed as demented in ADAMS. Classification of proxy respondents with some cognitive impairment can be predicted in 86% of the sample. Adding a small number of additional tests from ADAMS can increase each of these percentages to 84% and 93%, respectively.

Discussion. Cognitive assessment appropriate for diagnosis of dementia and CIND in large population surveys could be improved with more targeted information from informants and additional cognitive tests targeting other areas of brain function.

Key Words: Cognitive assessment—Cognitive impairment—Dementia.

LIFE expectancy at older ages has been increasing, and the number of older persons has been growing rapidly. Understanding the demands for personal care and support as well as medical care posed by an aging population requires estimating demands related to declines in both the cognitive and physical health of older persons. Only in the past 20 years have large-scale nationally representative population surveys included assessment of cognitive functioning (Herzog & Wallace, 1997). These tests included in surveys are viewed as assessments of cognitive functioning rather than diagnostic of dementia; diagnosis of dementia is normally done by health professionals after an in-depth neuropsychological examination which is beyond the capability of large omnibus surveys. The Health and Retirement Study (HRS) has included extensive measures of cognitive functioning since its beginning in 1992. In an effort to make national estimates of dementia and cognitive impairment without dementia (CIND), the Aging, Demographics, and Memory Study (ADAMS), a substudy of the HRS, has performed a detailed neuropsychological and clinical assessment with a smaller subsample of the large national data set (Langa et al., 2005). There is a need to find ways to ascertain cognitive impairment and dementia that are less costly than neuropsychological examinations and more precise than current population surveys (Evans et al., 2011; Weir et al., 2011).

This article combines data from the two sources (HRS and ADAMS) to determine how to improve on the large-scale survey data available for many national populations to better identify those with cognitive impairment or dementia. Specifically, we will use the diagnostic data from ADAMS and additional information from the detailed neuropsychological examination in ADAMS to see how we can improve our ability to use information collected in HRS and other large population studies to assess dementia and milder forms of cognitive impairment. The aim of this analysis is, thus, to improve the ability of analysts to use the results from large population surveys to characterize cognitive functioning in demographically representative populations. Being able to do this will help us to track the overall burden of dementia in the population; being able to classify individual survey participants by cognitive status will allow us to relate cognitive status to extensive individual indicators of behavioral, social, economic, and psychological information to better understand the causes and effects of having dementia.

BACKGROUND

The HRS is a large nationally representative sample of persons aged 50 years and older interviewed biannually since 1992. From its beginning, the HRS included measures of cognitive functioning that could be used to determine cognitive decline and onset of cognitive impairment in a large population study that was carried out both in-person and on the telephone (Herzog & Wallace, 1997). Although this was novel at the time this study was begun, similar measures have now been included in a number of other national studies including the English Longitudinal Study of Ageing, the National Social Life, Health, and Aging Project, the Surveys of Health, Ageing, and Retirement in Europe, the Indonesia Family Life Surveys, and the Korean Longitudinal Study of Aging. Initial evaluation of the data from these cognitive tests indicated that they had reasonable response rates, psychometric properties, and construct validity (Herzog & Wallace, 1997). Since then, the HRS cognitive data have been used extensively to characterize differences and change in cognitive ability in the population (Alley, Suthers, & Crimmins, 2007; Freedman, Aykan, & Martin, 2001; Langa, Llewellyn, et al., 2009; Lièvre, Alley, & Crimmins, 2008; Masel & Peek, 2009; Zelinski, Crimmins, Reynolds, & Seeman, 1998; Zivin, Kabeto, Kales, & Langa, 2009) as well as to indicate the size of the cognitively impaired population (Freedman et al., 2001; Lièvre et al., 2008; Suthers, Kim, & Crimmins, 2003).

A central issue is whether cognitive impairments and cognitive decline related to age are separate processes or whether they reflect cognitive differences along a continuum. Not all agree on this. Some view dementia as cognitive change that is not a part of normal aging but rather a disease process, and some view cognitive decline to be part of normal aging (Baltes & Nesslerode, 1979; Sliwinski, Hofer, Hall, Buschke, & Lipton, 2003). Others have come to see the changes in cognition with age as occurring across the same continuum over substantial periods of time and with multiple long-term end points (Fisher, Plassman, Heeringa, & Langa, 2008).

Cognitive functioning in the HRS was assessed using an adapted version of the Telephone Interview for Cognitive Status (TICS; Brandt, Spencer, & Folstein, 1998). Similar tests have been used for assessing cognitive impairment and dementia (Plassman, Newman, Welsh, & Breitner, 1994; Welsh, Breitner, & Magruder-Habib, 1993). This relatively short battery has been administered by interviewers both over the phone and in-person to over 30,000 persons who have participated in the HRS.

The ADAMS is based on a stratified random subsample of HRS respondents selected for intensive study of cognitive functioning and to provide national estimates of the prevalence of dementia and CIND. A sample of 1,770 respondents, aged 70 years and older, both community dwelling and living in nursing homes, were selected on the basis of cognitive performance in the HRS on the wave before interview (2000 or 2002; Langa et al., 2005). Those with poor cognitive function were oversampled. From this selected group, 28% refused to participate in ADAMS, 13% were deceased, 3% could not be contacted, and 7% had other reasons for non-participation. This resulted in 856 completed assessments or a 56% participation rate among the non-deceased eligible for ADAMS (Heeringa et al., 2009). Extensive analysis demonstrated no relationship between nonresponse and cognitive status and sampling weights correct for both the selection rules and nonresponse. Because the ADAMS sample was drawn using a complex sample design, it can be weighted to represent the national population.

The ADAMS examination is an in-person structured assessment that lasts 3–4 hr. It was conducted by two professionals, a nurse and a neuropsychology technician in the person's home. In addition, the HRS cognitive battery was also applied in the ADAMS sample. Final diagnosis of dementia was made by a consensus panel.

HRS data have been used to estimate the size of the cognitively impaired population in the United States; Suthers and colleagues (2003) estimated that 9.5% of the population aged 70 years and older in the United States is cognitively impaired. The ADAMS data, which have been weighted to represent the same national population as the HRS, have also been used to estimate the national prevalence of dementia (13.7% of the age 71+ population) and CIND (22%; Plassman et al., 2007, 2008). Obviously, these two approaches provide quite different pictures of size of the problem of cognitive impairment. The question this article addresses is whether the data from the neuropsychological testing performed in ADAMS, when linked to HRS data, can be used to improve the classification of cognitive status for individuals who have data collected in population surveys like HRS and, further, to ask what additional testing in ADAMS but not in HRS would contribute most to improving classification beyond what is achievable with the current HRS cognitive battery.

DATA AND METHODS

Data

As indicated above, the ADAMS subsample of HRS respondents aged 70 years and older was chosen on the basis of cognitive scores at the HRS 2000 and 2002 interviews. This subsample received a neuropsychiatric assessment as well as detailed medical, functioning, cognitive, and physical exams; 856 people completed all the assessments in the ADAMS interview. We use data from the 2000 and 2002 HRS, and the first ADAMS interview to determine whether responses collected in one administration of the HRS can be used to predict subsequent neuropsychological assessed diagnosis in ADAMS. We require complete data be available on the variables used to define

cognition in the HRS resulting in an initial sample size of 819. We use data from the HRS file that includes imputations for missing data. The missing cases ($N = 37$) are all self-respondents to the survey with some missing data on cognitive measures. Among the missing cases, about half ($N = 18$) are diagnosed as demented in ADAMS, whereas 12 are classified as CIND and 7 as normal.

Measurement of Cognition

Tests used to assess cognitive functioning in the HRS included 10 word immediate and delayed recall tests of memory, a serial 7s subtraction test of working memory, counting backwards to assess attention and processing speed, an object naming test to assess language, and recall of the date and president and vice-president to assess orientation (TICS). Composite scores using all the items create a measure of cognitive functioning, which can range from 0 to 35.

There have been two approaches suggested for using the HRS to define cognitive impairment and dementia in the population. Herzog and Wallace (1997), in initial analyses of HRS, suggested that a score of 8 or less on the composite cognitive measure could be viewed as cognitive impairment. We use the Herzog–Wallace approach as one definition of cognitive impairment in the HRS. We continue to use the term "cognitive impairment" with this approach, as has been done in the literature, because this was not viewed as a diagnostic assessment. The suggested cutoff value of 8 has been used in multiple studies based on these data, although sometimes a higher cutoff is used for longitudinal studies to reflect the practice effect gained from repeated testing (Lièvre et al., 2008; Suthers et al., 2003). Although we use the suggested value of 8 in our analysis, we examine the effect of increasing the cutoff value.

After the ADAMS data became available, Langa and Weir (Langa, Kabeto, & Weir, 2009) developed an approach to defining dementia and CIND using the HRS data. They developed cut-points for the HRS cognitive measures that would produce the same population distribution of cognitive states estimated by ADAMS (i.e., "equipercentile equating"). In order to apply this approach to participants under 65 years, they could not make use of the orientation and naming items from the full HRS cognitive battery described above but included the same immediate and delayed recall items, the serial 7s, and backward counting so that the range of the composite measure is 0–27. They classify respondents who score from 0 to 6 as demented, 7 to 11 as CIND, and 12 to 27 as normal. We use the Langa–Weir approach as the second definition in the analysis that follows, and we use their cut-points to classify people as demented and CIND. CIND is defined in ADAMS as mild cognitive or functional impairment, reported by the participant or informant, or impaired test performance on the neuropsychological measures that does not reach the severity of dementia (Plassman et al., 2008).

Because of a need to understand health and functioning among the entire older population, persons who cannot answer for themselves due to either their physical or mental health may have a proxy respondent to the HRS. Proxies do not answer the cognitive battery on behalf of the respondents; however, proxies report on seven behavioral symptoms of the sample member designed to allow assessment of cognitive impairment. The seven symptoms are known as a Jorm scale and include hallucinations, getting lost, memory, judgment, organization, wandering, and ability to be left alone (Jorm, 1994; Jorm, Scott, & Jacomb, 1989). Having two or more of the seven Jorm symptoms has been used to define cognitive impairment in studies that have used the Herzog and Wallace cutoff (Lièvre et al., 2008; Suthers et al., 2003).

Langa–Weir have used different information provided from the proxy respondents to define cognitive status: a direct assessment of memory ranging from excellent to poor (Score 0–4), an assessment of limitations in five instrumental activities of daily living (IADLs; managing money, taking medication, preparing hot meals, using phones, and doing groceries; Score 0–5). They have also included the interviewer assessment of difficulty completing the interview because of cognitive limitation (Score 0–2 indicating, none, some, and prevents completion) in their cognitive score. Using this information from proxies and informants, high scores are classified as demented (6–11), with medium scores (3–5) as CIND.

In ADAMS, individuals are diagnosed as having normal cognitive function or being demented or having CIND using information taken from a detailed neuropsychological assessment and a consensus diagnosis. The cognitive testing in ADAMS included the Mini-Mental State Examination (MMSE), Boston naming test, digit span (combined forward and backward tasks), Symbol Digit Modality Test, animal fluency, word list three trial learning, construction praxis copying, Trail Making Test, Wechsler Memory Scale, Fuld Object Memory Test, Shipley vocabulary test, and the WRAT 3 blue reading test (Langa et al., 2005). In ADAMS, proxy respondents or persons close to the individual also report on behavioral patterns for all respondents but this is done using items in the Blessed Dementia Ratings (Blessed, Tomlinson, & Roth, 1968): ability to do household tasks, handle small amounts of money, remember short list, find way indoors, find way on a familiar street, grasp situations, recall recent events, dwell in past, eat, dress, and toilet.

Methods

First, we use the two approaches previously described with the HRS data to classify individual cognitive status and the ADAMS diagnosis to provide insights into which persons are being differentially classified in HRS and ADAMS. For self-respondents, the first approach (Herzog–Wallace)

uses the full range of cognitive functioning information in the HRS and defines a score of 8 or less as cognitively impaired; the second (Langa–Weir) uses a subset of the HRS cognitive functioning information and defines a score of 6 or less as demented and 7–11 as CIND.

We then perform a set of multinomial logistic analyses to determine how well we can reproduce the ADAMS diagnoses for individuals based on a set of variables from HRS and ADAMS. We fit multiple models where the outcome is the ADAMS diagnosis (i.e., normal, CIND, demented) for self-respondents in HRS including data first from HRS and then adding a few ADAMS scales. In the series of five equations, we first include only the HRS cognitive measures; then we add age, sex, education, activities of daily living and IADL functioning; two additional cognitive tests from ADAMS; and, finally, we add the AD-AMS proxy score from the Blessed Scale. Using the results of the equations, people are classified into one of the three outcomes, and we assess the fit of the models by the percent of cases correctly classified into the three ADAMS diagnoses. We also fit multiple models for proxy respondents to examine our ability to predict ADAMS classifications for people with proxy respondents using information from the HRS and ADAMS. Our goal is to see how close we can come to reproducing the "gold-standard" ADAMS classification with current HRS measures and some additional measures that could potentially be added to HRS. Finally, we use the ADAMS data itself to see how well the same equation does at prediction of ADAMS diagnoses and then how this is improved with the additional information in ADAMS.

RESULTS

Comparison of the Prevalence of Cognitive Impairment, CIND, and Dementia in HRS and ADAMS

To begin our descriptive analysis, we compare the age-specific prevalence of dementia in the ADAMS sample with the prevalence of cognitive impairment in the full HRS estimated using the older Herzog–Wallace criteria and the prevalence of likely dementia based on the more recent Langa–Weir criteria. As indicated above, the prevalence of cognitive impairment estimated using the Herzog–Wallace measure is lower than that estimated from ADAMS (Figure 1). Using the Langa–Weir approach and the whole HRS samples for 2000 and 2002, the prevalence is somewhat higher in the 71–79 and the 90+ age group and lower in between. The prevalence of CIND based on the Langa–Weir approach is somewhat higher than the prevalence in ADAMS at ages 71–79 and lower at ages 80 and older.

ADAMS Diagnosis and HRS Classification

We next examine the individual-level cross-classification of how individuals are classified using the criteria based on
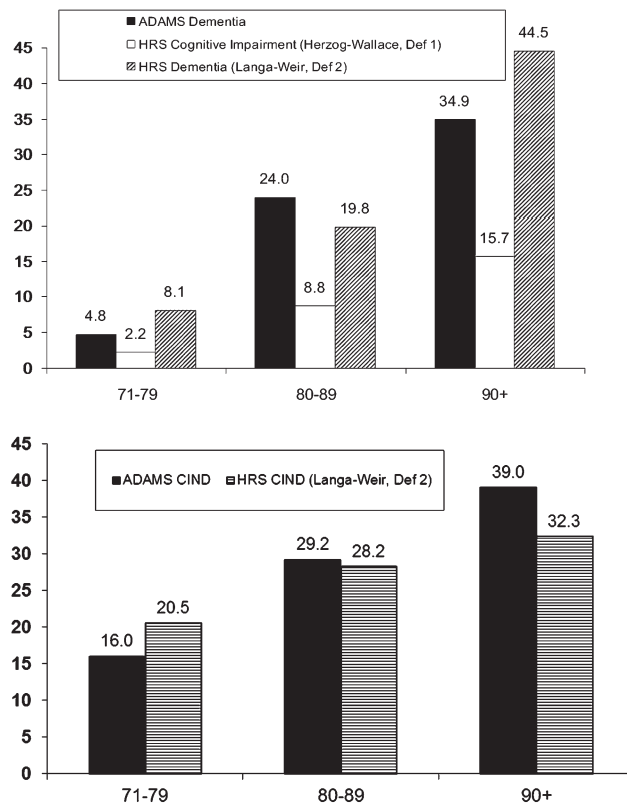

Figure 1. Prevalence of dementia in ADAMS and from two analyses of HRS; prevalence of CIND from ADAMS and from HRS. Source: ADAMS from Plassman and colleagues (2007, 2008); HRS—Definition 1: Herzog–Wallace from Suthers and colleagues (2003) based on HRS Wave 1 and Wave 3; HRS—Definition 2: Langa–Weir based on calculations from HRS 2000 and 2002 for the 71 + sample.

HRS data in the HRS wave prior to their ADAMS interview and how they are diagnosed in ADAMS. On average, there was a 13-month lag between the two interviews. Using the Herzog–Wallace definition of cognitive impairment, just over a third (36.63%) of those diagnosed as demented in ADAMS are classified as cognitively impaired in HRS (Table 1). (When the cutoff is ≤9, 38% of those diagnosed as demented in ADAMS are classified that way in HRS.) Only a small percentage of those diagnosed as normal or CIND in ADAMS are classified as impaired in HRS (0.89%; Table 1).

Table 1. Classification of Cognitive Functioning in HRS by ADAMS Diagnosis: HRS—Definition 1: Herzog–Wallace

	HRS		
ADAMS	Cognitively intact	Cognitively impaired	Total
Demented	63.37%	36.63%	100%
	n = 159	n = 131	
CIND or normal	99.11%	0.89%	100%
	n = 509	n = 20	

Notes: N = 819; % weighted. HRS cognitive impairment for self-respondents is 8 or less of 35; Jorm scale for proxy cases has a score of at least 2 of 7. ADAMS = Aging, Demographics, and Memory Study; CIND = cognitively impaired without dementia; HRS = Health and Retirement Study.

Table 2. Classification of Cognitive Functioning in HRS by ADAMS Diagnosis: HRS—Definition 2: Langa–Weir

	HRS			
ADAMS	Cognitively intact	CIND	Cognitively impaired (demented)	Total
Demented	17.17%	30.65%	52.18%	100%
	n = 33	n = 72	n = 185	
CIND	59.61%	27.58%	12.81%	100%
	n = 86	n = 83	n = 60	
Normal	87.17%	11.38%	1.44%	100%
	n = 235	n = 51	n = 14	

Notes: N = 819; % weighted. HRS dementia for self-respondents is score 6 or less, 7–11 is CIND, 12+ (of 27) is Normal; with proxy reports of mental status, instrumental activities of daily living difficulties, and interviewer assessment of cognition, for proxy cases, score 6+ (of 11) is demented, 3–5 is CIND, 2 or less is Normal. ADAMS = Aging, Demographics, and Memory Study; CIND = cognitively impaired without dementia; HRS = Health and Retirement Study.

Using the Langa–Weir approach, more people are correctly classified as demented (52.18%) among those diagnosed in ADAMS as demented (Table 2). Only about 17% of those whose diagnosis is dementia in ADAMS would be classified as cognitively intact using this approach. About a quarter of those with CIND diagnosed in ADAMS are classified as CIND using the HRS information (27.58%); the majority (59.61%) of the CIND-diagnosed group is classified as intact in HRS. Most of those who are classified as "normal" in ADAMS are also classified that way using the Langa–Weir approach (87.17%).

ADAMS Diagnosis and HRS Classification for Self-Respondents and for Proxies

Cognitive impairment is a significant obstacle to completing a long and difficult interview like the HRS. To maintain participation of the impaired, the HRS accepts interviews with proxies when the respondent is unable or unwilling to do an interview on their own. In the subsample of HRS and ADAMS respondents we are working with, 76% (unweighted, 89% weighted) were self-respondents and 24% had proxy respondents in the HRS (unweighted, 11% weighted). However, among the people who would be classified as cognitively impaired using the Herzog–Wallace definition, 79% (unweighted, 87% weighted) had

proxy respondents, meaning that most of the cognitive impairment based on the definition is coming from proxy responses rather than the responses to the cognitive tests by the respondents themselves. Among those classified as demented using the ADAMS data, 54% (unweighted, 49% weighted) had proxy respondents in HRS. This means that the assessment of dementia is highly dependent on proxy reporting. We examine the cross-classifications shown above separating the sample into those who were self-respondents and those who have a proxy response in HRS.

Using the Herzog–Wallace approach, self-respondents in HRS who are diagnosed as demented in ADAMS are rarely classified as cognitively impaired (6.48%; Table 3). Among those with proxy respondents, this percentage was 68.14%. Self-respondents who are diagnosed as CIND or normal in ADAMS are almost never categorized as cognitively impaired in HRS (0.34%). People with proxy respondents who are diagnosed as CIND or normal in ADAMS are more likely than self-respondents diagnosed as CIND or normal in ADAMS to be classified as cognitively impaired using the Herzog–Wallace approach (10.12%).

Using this Langa–Weir definition, among the self-respondents in HRS who were diagnosed as demented in ADAMS about a quarter (23.93%) were classified as demented in HRS (Table 4). About a quarter of self-respondents with a diagnosis of CIND in ADAMS would be classified that way in HRS (24.56%). Most of those diagnosed as normal in ADAMS would be classified that way using the Langa–Weir approach (87.67%).

Among the demented in ADAMS with a proxy respondent, most would be correctly classified (81.70%) using the Langa–Weir definition. Among those with a proxy in HRS and diagnosed as CIND, about half (51.84%) would be classified as CIND using HRS data. Most of those judged to have normal cognitive functioning in ADAMS who had a proxy in HRS would be assessed as normal using the Langa–Weir method (74.80%). All the others would be classified as CIND.

Multivariate Classification Models

We can also use multiple classification models and information from HRS, ADAMS, and from a combination of the two surveys to compare how well sets of variables predict the ADAMS classification of cognitive functioning among

Table 3. Classification of Cognitive Functioning in HRS by Proxy Status and ADAMS Diagnosis: HRS—Definition 1: Herzog–Wallace

	HRS					
	Self (n = 620)			Proxy (n = 199)		
ADAMS	Cognitively intact	Cognitively impaired	Total	Cognitively intact	Cognitively impaired	Total
Demented	93.52%	6.48%	100%	31.86%	68.14%	100%
	n = 116	n = 18		n = 43	n = 113	
CIND or Normal	99.66%	0.34%	100%	89.88%	10.12%	100%
	n = 472	n = 14		n = 37	n = 6	

Notes: N = 819; % weighted. HRS cognitive impairment for self-respondents is 8 or less of 35; Jorm scale for proxy cases is at least 2 of 7. ADAMS = Aging, Demographics, and Memory Study; CIND = cognitively impaired without dementia; HRS = Health and Retirement Study.

Table 4. Classification of Cognitive Functioning in HRS by Proxy Status and ADAMS Diagnosis: HRS—Definition 2: Langa–Weir

| | HRS | | | | | | | |
| | Self ($n = 620$) | | | | Proxy ($n = 199$) | | | |
ADAMS	Cognitively intact	CIND	Cognitively impaired (demented)	Total	Cognitively intact	CIND	Cognitively impaired (demented)	Total
Demented	30.34%	45.73%	23.93%	100%	3.41%	14.89%	81.70%	100%
	$n = 28$	$n = 57$	$n = 49$		$n = 5$	$n = 15$	$n = 136$	
CIND	63.88%	24.56%	11.56%	100%	25.31%	51.84%	22.85%	100%
	$n = 81$	$n = 68$	$n = 49$		$n = 5$	$n = 15$	$n = 11$	
Normal	87.67%	10.83%	1.50%	100%	74.80%	25.20%	0.00%	100%
	$n = 227$	$n = 47$	$n = 14$		$n = 8$	$n = 4$	$n = 0$	

Notes: $N = 819$; % weighted. HRS dementia for self-respondents is score 6 or less, 7–11 is CIND, 12+ (of 27) is Normal 27; with proxy reports of mental status, instrumental activities of daily living difficulties, and interviewer assessment of cognition, for proxy cases, score 6+ (of 11) is demented, 3–5 is CIND, 2 or less is Normal. ADAMS = Aging, Demographics, and Memory Study; CIND = cognitively impaired without dementia; HRS = Health and Retirement Study.

the HRS self-respondents (Table 5). We examine a series of multinomial logistic equations that predict the ADAMS diagnosis (demented, CIND, or normal), and then, using the results, each case is classified into one of the categories to determine the percent correctly classified, assuming that ADAMS represents correct classification. The resulting equations also indicate how individual variables predict dementia and CIND relative to normal cognitive functioning. In the first equation, we use the cognitive assessment measures in HRS; in subsequent models, we add demographic, functioning, and informant measures from HRS. Finally, we add the results of three neuropsychological tests from ADAMS to see how we can improve prediction for those who are self-respondents in HRS. The three tests we add from ADAMS are the informant-provided Blessed Scale, the digit span test, and the animal fluency test. The latter two tests are indicators of types of cognitive functioning that are not well covered in the HRS battery of cognition used in 2000 and 2002 but that could be added to population surveys with additional time. The digit span task is a task that assesses working memory; animal fluency provides an assessment of retrieval fluency. These two tests are chosen from the available ADAMS tests because they tap cognitive abilities similar to those assessed with two new tests added to the revised HRS cognitive battery in 2010 —number series and the animal fluency test. With only the HRS cognitive battery as it stood prior to 2010, 74% of the self-respondents will be correctly classified into their ADAMS diagnosis (Table 5, Model 1). The coefficients of the equations indicate that higher scores on most of the subsets of cognitive measures, immediate and delayed recall, serial 7's, and the TICS in HRS are significantly associated with a lower likelihood of being demented or CIND; however, immediate recall is not significantly associated with dementia diagnosis and delayed recall is not significantly associated with CIND diagnosis. Using the Langa–Weir grouping into categories based on the 27-item scale results in lower prediction of the ADAMS diagnosis (69.2%) (Model 2).

Correct prediction is increased by 2.7% with the incorporation of demographic, education, and functioning informa-

tion from HRS (Model 1 vs. Model 3), so that more than three quarters of self-respondents are correctly classified (76.9%). Once controls for the cognitive scores are included in Model 3, education is not related to diagnosis of dementia or CIND. Difficulty with money management is particularly highly related to being diagnosed with dementia.

In an attempt to determine how much we could improve prediction of the ADAMS diagnosis by adding additional tests to HRS, we add the animal fluency and the digit span tests and increase prediction by about 2% (Model 4). Higher scores on both tests reduce the likelihood of being classified as CIND and demented. In the 5th equation, we add informant reports from the Blessed Test in ADAMS and increase the correct classification by 5% to 84.1%. This suggests that information from an informed observer is potentially more valuable than the additional cognitive tests.

Using the results from the 5th equation in Table 5, we examine the predicted classification to see where the misclassified cases fall relative to diagnosis (Table 6). Among those diagnosed as demented in ADAMS, 79% (76 of 96) are correctly predicted. Most incorrectly predicted cases fall in the CIND category. Among those diagnosed as normal, the correct prediction is about the same as that for demented, 80%. The rest are mostly predicted as CIND. Among those diagnosed with CIND, only 68% are correctly predicted. False prediction is in both directions with a somewhat greater number being predicted to be demented (29) and normal (22).

Next, we examine how well we can predict proxy respondents using the Jorm scale, the proxy assessment of memory, IADL functioning, and the interviewer assessment of cognitive ability from HRS and then include the Blessed Scale from ADAMS. In this case, we use a logistic approach to predicting dementia relative to CIND. We exclude those diagnosed as normal as so few proxy respondents are actually classified as having normal functioning in ADAMS. About 84% of people with proxy respondents are correctly classified into CIND or dementia using the Jorm scale (Table 7). Compared with the Jorm scale, the HRS interviewer assessment is somewhat less successful in predicting whether people are CIND or demented (Model 2 in Table 7,

Table 5. Odds Ratios From Multinomial Logistic Regression Predicting ADAMS Diagnosis of Demented or CIND (normal as reference) Among HRS Sample Who Self-Respond (weighted)

	Model 1 (n = 620)		Model 2 (n = 620)		Model 3 (n = 620)		Model 4 (n = 585)		Model 5 (n = 578)	
	Demented in ADAMS	CIND in ADAMS	Demented in ADAMS	CIND in ADAMS	Demented in ADAMS	CIND in ADAMS	Demented in ADAMS	CIND in ADAMS	Demented in ADAMS	CIND in ADAMS
Age					1.20 (1.14–1.28)	1.11 (1.07–1.15)	1.14 (1.06–1.22)	1.09 (1.04–1.13)	1.16 (1.05–1.27)	1.09 (1.04–1.14)
Female					1.02 (0.47–2.22)	0.92 (0.60–1.42)	0.97 (0.39–2.39)	0.95 (0.59–1.51)	0.81 (0.25–2.66)	0.92 (0.54–1.57)
Low education (0–5 years)					0.36 (0.07–1.89)	1.30 (0.41–4.14)	0.08 (0.01–0.56)	0.69 (0.20–2.40)	0.11 (0.01–1.34)	0.55 (0.14–2.18)
Mid-education (6–11 years)					0.45 (0.20–1.02)	0.76 (0.47–1.21)	0.40 (0.16–1.02)	0.59 (0.35–0.97)	0.90 (0.27–3.03)	0.60 (0.33–1.07)
Immediate recall	0.80 (0.61–1.04)	0.78 (0.66–0.93)			0.73 (0.54–1.00)	0.80 (0.66–0.97)	0.77 (0.54–1.09)	0.89 (0.72–1.09)	0.93 (0.59–1.47)	0.93 (0.73–1.19)
Delayed recall	0.62 (0.50–0.75)	0.91 (0.792–1.03)			0.65 (0.52–0.81)	0.95 (0.82–1.09)	0.67 (0.51–0.88)	0.95 (0.81–1.11)	0.55 (0.38–0.80)	0.92 (0.77–1.10)
Serial 7s	0.78 (0.65–0.94)	0.86 (0.76–0.97)			0.68 (0.54–0.86)	0.83 (0.72–0.96)	0.78 (0.59–1.02)	0.88 (0.76–1.03)	0.87 (0.61–1.24)	0.92 (0.77–1.10)
TICS	0.59 (0.45–0.78)	0.67 (0.53–0.84)			0.60 (0.43–0.82)	0.67 (0.52–0.86)	0.66 (0.45–0.95)	0.73 (0.56–0.94)	0.68 (0.42–1.11)	0.68 (0.50–0.92)
HRS cognitive score Definition 2 (Score 12+ is Ref.)										
7–11			9.83 (5.67–17.05)	4.06 (2.59–6.36)						
0–6			28.36 (13.91–57.78)	9.80 (5.14–18.70)						
Dressing difficulty					0.33 (0.08–1.39)	2.76 (1.41–5.40)	0.37 (0.06–2.12)	3.42 (1.70–6.85)	0.47 (0.06–3.64)	3.70 (1.68–8.16)
Bathing difficulty					1.30 (0.42–4.02)	0.96 (0.45–2.04)	1.38 (0.37–5.16)	0.90 (0.41–1.99)	2.49 (0.45–13.64)	0.98 (0.40–2.37)
Eating difficulty					4.34 (0.96–19.68)	0.94 (0.31–2.86)	6.01 (1.18–30.74)	0.86 (0.28–2.66)	6.84 (0.72–65.51)	0.97 (0.27–3.50)
Managing money difficulty					9.72 (2.51–37.65)	3.90 (1.33–11.46)	9.77 (2.06–46.31)	4.45 (1.40–14.09)	8.02 (1.13–56.79)	6.21 (1.63–23.62)
Making phone difficulty					2.38 (0.59–9.62)	0.61 (0.17–2.16)	1.93 (0.33–11.44)	0.62 (0.15–2.54)	1.01 (0.10–10.17)	0.34 (0.07–1.64)
Digit span—ADAMS							0.81 (0.70–0.93)	0.88 (0.81–0.95)	0.81 (0.67–0.98)	0.85 (0.77–0.93)
Animal fluency—ADAMS							0.72 (0.65–0.80)	0.85 (0.81–0.90)	0.65 (0.55–0.76)	0.85 (0.79–0.91)
Blessed Scale—ADAMS									11.18 (6.80–18.36)	4.66 (3.27–6.64)
% Correctly classified										
Unweighted, %	56.4		55.5		63.9		67.7		76.7	
Weighted, %	74.2		69.2		76.9		79.1		84.1	

Notes: ADAMS = Aging, Demographics, and Memory Study; CIND = cognitively impaired without dementia; HRS = Health and Retirement Study; TICS = Telephone Interview for Cognitive Status.

Table 6. Model Predicted Classification into Demented, CIND, and Normal (using Table 5, Model 5 with age, gender, education, HRS cognitive scores, ADL/IADL difficulties, digit span, animal fluency, ADAMS Blessed) by ADAMS Diagnosis; Self-Respondents in HRS (not weighted)

		ADAMS diagnosis		
		Demented	CIND	Normal
Predicted	Demented	13.2%	5.0%	0.9%
		n = 76	n = 29	n = 5
	CIND	3.3%	18.9%	10.2%
		n = 19	n = 109	n = 59
	Normal	0.2%	3.8%	44.6%
		n = 1	n = 22	n = 258

Notes: ADAMS = Aging, Demographics, and Memory Study; ADL = activities of daily living; CIND = cognitively impaired without dementia; HRS = Health and Retirement Study; IADL = instrumental activities of daily living.

75.2%). Adding the memory assessment and the IADL difficulty score to the interviewer assessment marginally increases the prediction power (Model 3). The Blessed Scale alone results in the prediction power of 92.3% (Model 4). Correct assignment for this group improves marginally when all the other information is added (93.1%, Model 5). The Blessed instrument thus outperforms all the available information from proxies currently in the HRS.

Prediction from ADAMS scores

Our assessment of how well we can predict ADAMS diagnoses to this point has primarily relied on information from HRS to see how we can predict ADAMS diagnoses. There is a time lag between HRS and ADAMS that is likely to result in changes in the variables we use and reduce our prediction of the ADAMS diagnosis. We now turn to using the information within ADAMS, first to see whether the prediction is improved with contemporaneous measures and second to add additional test results from ADAMS to see how much we can continue to improve prediction. Some of these additional tests might be candidates to be added to surveys at a cost of more time and some of them would just not be practical for survey inclusion. We begin by fitting a model with data from ADAMS that is similar to Model 5

from Table 5 by replacing the cognitive measures assessed in HRS with the same measures as they were asked in ADAMS. This results in the correct classification 81.6% or 84.9% when weighted (Table 8). Next, we examine the effect of adding each of 11 additional ADAMS tests individually to the first model in Table 8. Of these 11 models, 7 of the added tests do not increase the explanatory power (MMSE, Boston naming test, constructional praxis test, Shipley vocabulary test, Symbol Digit Modalities Test, Fuld Object Memory evaluation, and Controlled Oral Word Association Test), and we do not consider these tests further. Each of four other tests adds from 0.95% to 3.52% to the percent correctly classified when added to the first model in Table 8. We add the four tests (Benton Visual Retention Test, Wechsler Memory Scale, the Dementia Severity Rating Scale, and the Trail Making Test) together to the model and increase our overall prediction by 6.1% (weighted) to 91.0% (Model 2). Because the N has been reduced by the addition of these tests, we rerun the Model 1 with the same cases in Model 3 and we find that our addition to the variance explained in these cases has only been 2%. Finally, for comparison, we include a model that includes only age, gender, and education and MMSE as this is a commonly used simple screener for dementia in clinical settings (Model 4). Correct classification occurs in 76% of the cases, lower than with the HRS measures in Model 1.

We should note that these measures are not all appropriate for administration in a large survey, and most of them require in-person administration, so it is not just a matter of adding more indicators and more time and being able to produce more accurate diagnostic results. But the issue of time is important. The HRS battery of cognitive tests takes 5.4 min on average. Adding the digit span, animal fluency, and the Blessed score would add 10 min (estimated from Brenda Plassman). The extra four tests added in the second equation of Table 8 would increase time by about 28 min.

DISCUSSIONS AND CONCLUSIONS

We have shown that in this large national sample of older people, results from the neuropsychological diagnostic

Table 7. Odds Ratios From Logistic Regression Predicting ADAMS Diagnosis of Demented (CIND as reference; normal excluded) Among HRS Sample With Proxy Respondents

	Model 1 (N = 187)	Model 2 (N = 187)	Model 3 (N = 187)	Model 4 (N = 187)	Model 5 (N = 184)	Model 6 (N = 184)
Jorm score (0–7)	4.63* (1.91–11.23)			3.82* (1.46–10.01)		1.90 (0.39–9.18)
Memory (0–4)			2.39* (1.25–4.55)	1.11 (0.51–2.40)		0.76 (0.23–2.51)
IADL difficulty			1.45* (1.01–2.10)	1.28 (0.84–1.95)		0.54 (0.20–1.45)
Interviewer's assessment of having cognitive limitation		4.98* (1.49–16.63)	1.37 (0.31–6.14)	1.19 (0.24–5.79)		0.80 (0.11–5.74)
Blessed Scale—ADAMS					4.02* (1.71–9.44)	5.54* (1.60–19.24)
% Correctly classified						
Unweighted, %	80.8	83.4	82.9	82.4	90.2	90.2
Weighted, %	84.4	75.2	78.0	86.1	92.3	93.1

Notes: ADAMS = Aging, Demographics, and Memory Study; CIND = cognitively impaired without dementia; HRS = Health and Retirement Study; IADL = instrumental activities of daily living.

*p < .05.

Table 8. Percent Correctly Classified into Demented, CIND, and Normal (reference) Among HRS Using Additional ADAMS Measures

% Correctly classified		
Unweighted	Weighted	Variables included in model
81.6	84.9	Model 1—age, gender, education, HRS cognitive measures administered in ADAMS, ADAMS ADL/IADL difficulties (dressing, eating, bathing, controlling bladder and bowel), digit span, animal fluency, ADAMS Blessed score (*N* = 667)
87.4	91.0	Model 2—Age, gender, education, HRS cognitive measures administered in ADAMS, ADAMS ADL/IADL difficulties (dressing, eating, bathing, controlling bladder and bowel), digit span, animal fluency, ADAMS Blessed score, Dementia Severity Rating Scale, Trail Making Test, Benton Visual Retention Test, Wechsler Memory Scale (immediate trial, delayed trial; *N* = 435)
85.1	89.1	Model 3—age, gender, education, HRS cognitive measures administered in ADAMS, ADAMS ADL/IADL difficulties (dressing, eating, bathing, controlling bladder and bowel), digit span, animal fluency, ADAMS Blessed score (*N* = 435)
69.5	75.9	Model 4—age, gender, education, MMSE (*N* = 814)

Notes: ADAMS = Aging, Demographics, and Memory Study; ADL = activities of daily living; CIND = cognitively impaired without dementia; HRS = Health and Retirement Study; IADL = instrumental activities of daily living; MMSE = Mini-Mental State Examination.

approach on a limited subsample can be used to develop methods to classify people as having dementia based on the cognitive assessments available in large population studies. The overall levels of dementia and of CIND estimated using this approach can be made similar to those directly estimated from the neuropsychological study. How well the estimates match the neuropsychological study depends on the definitions used and the information included.

For many social science users of large population surveys, it is very desirable to be able to classify each individual as cognitively impaired or not. Indications of cognitive state are useful as both dependent and independent variables. Using the measures on cognitive functioning available for self-respondents in the HRS, we can classify about 74% of this group into the three-category diagnosis provided from the ADAMS neuropsychological exam. This will be increased by about 5% with the addition of two new tests added to the HRS cognitive battery in 2010.

Our results show that informant information is very useful in improving categorization of self-respondents as well as that of proxy respondents. The fact that most people identified as cognitively impaired in HRS are classified in this category because of reports of behavioral problems by proxies rather than by assessment on cognitive testing, indicates the importance of both including proxy respondents in such studies and the importance of refining scales for use with proxies to identify those with dementia and cognitive loss. The Blessed Scale that was reported by an informant for all participants in the ADAMS study appeared to be a good predictor of cognitive status in the HRS sample. Most population surveys do not include informant information, and HRS only does so with proxies designated by respondents unwilling or unable to do the interview themselves, but it is an extension that would be valuable for identifying levels of cognitive impairment.

Certainly with more survey time and more attention to information from informants, the ability to classify the cognitive functioning of people in large population surveys could increase. However, the results from within the ADAMs data themselves show that it would require very significant increase in time to raise the ability much higher and that the marginal gain from adding most of the individual tests examined is quite small.

Our ability to correctly identify cognitive status with the HRS cognitive tests is not the same across cognitive ability. We are better able to identify those who have normal cognitive functioning and those who have dementia than those in the middle category of CIND. This category that is similar, but not identical, to mild cognitive impairment is inherently going to be more difficult to classify. Like dementia, there are a number of causes and subtypes of CIND (Plassman et al., 2008). The CIND group is likely to experience relatively rapid changes in cognitive ability over time and may have differential loss across cognitive abilities; both of which may make them more difficult to identify. However, the measures predictive of dementia and those predictive of CIND have been shown to be quite similar in our analysis.

Other approaches have been suggested for using the extensive survey data to estimate cognitive status. Fisher and colleagues (2008) have proposed using the data to produce a probability of dementia for each respondent. Such an approach could be compared with those presented here and could use some of the information obtained in our study to guide the models for estimating probabilities. Another approach would be to use extensive longitudinal data to better model the outcomes than our approach with cross-sectional data. Although this is a very data intensive approach, and not likely to be adopted by many analysts, it would be in concert with the fact that the cognitive battery now included in surveys was primarily designed to assess cognitive change in individuals rather than to diagnose dementia or severe cognitive impairment. There has been a longitudinal component to ADAMS, which has followed a subset of people diagnosed as mildly cognitively impaired. Results from analysis of change in cognitive functioning in this sample may be useful in further refining ability to estimate this group.

CORRESPONDENCE

Correspondence should be addressed to Eileen M. Crimmins, PhD, Andrus Gerontology Center, University of Southern California, Los Angeles, CA 90089-0191. E-mail: crimmin@usc.edu.

REFERENCES

Alley, D., Suthers, K., & Crimmins, E. M. (2007). Education and cognitive decline in older Americans: Results from the AHEAD sample. *Research on Aging*, 29, 73–94. doi:10.1177/0164027506294245

Baltes, P., & Nesselrode, J. (1979). History and rationale of longitudinal research. In P. Baltes & J. Nesselrode (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1–39). San Diego, CA: Academic Press.

Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *British Journal of Psychiatry*, 114, 797–811. doi:10.1192/bjp.114.512.797

Brandt, J., Spencer, M., & Folstein, M. (1988). The telephone interview for cognitive status. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, 1, 111–117.

Evans, D., Grodstein, F., Loewenstein, D., Kaye, J., Weintraub, S. (2011). Reducing case ascertainment costs in U.S. population studies of Alzheimer's disease, dementia, and cognitive impairment—Part 2. *Alzheimer's and Dementia*, 7, 110–123. doi:10.1016/j.jalz.2010.11.008

Fisher, G. G., Plassman, B. L., Heeringa, S. G., & Langa, K. M. (2008). Assessing the relationship of cognitive aging and processes of dementia. In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging interdisciplinary perspectives* (pp. 340–350). Los Angeles, CA: Sage.

Freedman, V. A., Aykan, H., & Martin, L. G. (2001). Aggregate changes in severe cognitive impairment among older Americans: 1993 and 1998. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 56, 100–111.

Heeringa, S., Fisher, G., Hurd, M., Langa, K., Ofstedal, M., Plassman, B., et al. (2009). *Aging, Demographics and Memory Study (ADAMS). Sample design, weighting and analysis for ADAMS*. Retrieved from http://hrsonline.isr.umich.edu/sitedocs/userg/ADAMSSampleWeights_Jun2009.pdf on July 12, 2010.

Herzog, A. R., & Wallace, R. B. (1997). Measures of cognitive functioning in the AHEAD study. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 52, 37–48. doi:10.1093/geronb/52B.Special_Issue.37

Jorm, A. F. (1994). Disability in dementia: assessment, prevention, and rehabilitation. *Disability and Rehabilitation*, 16, 98–109.

Jorm, A. F., Scott, R., & Jacomb, P. (1989). Assessment of cognitive decline in dementia by informant questionnaire. *International Journal of Geriatric Psychiatry*, 4, 35–39. doi:10.1002/gps.930040109

Langa, K. M., Kabeto, M., & Weir, D. (2009). *Report on race and cognitive impairment using HRS in, 2010 Alzheimer's disease facts and figures*. Retrieved from http://www.alz.org/documents_custom/report_alzfactsfigures2010.pdf on July 12, 2010.

Langa, K. M., Llewellyn, D. J., Lang, I. A., Weir, D. R., Wallace, R. B., Kabeto, M. U., . . . Huppert, F. (2009). Cognitive health among older adults in the United States and in England. *BMC Geriatrics*, 9, 23. doi:10.1186/1471-2318-9-23

Langa, K. M., Plassman, B. L., Wallace, R. B., Herzog, A. R., Heeringa, S. G., Ofstedal, M. B., . . . Willis, R. J. (2005). The aging, demographics, and memory study: Study design and methods. *Neuroepidemiology*, 25, 181–191. doi:10.1159/000087448

Lièvre, A., Alley, D., & Crimmins, E. M. (2008). Educational differentials in life expectancy with cognitive impairment among the elderly in the United States. *Journal of Aging and Health*, 20, 456–477. doi:10.1177/0898264308315857

Masel, M. C., & Peek, M. K. (2009). Ethnic differences in cognitive function over time. *Annals of Epidemiology*, 19, 778–783. doi:10.1016/j.annepidem.2009.06.008

Plassman, B. L., Langa, K. M., Fisher, G. G., Heeringa, S. G., Weir, D. R., Ofstedal, M. B., . . . Wallace, R. B. (2007). Prevalence of dementia in the United States: the aging, demographics, and memory study. *Neuroepidemiology*, 29, 125–132. doi:10.1159/000109998

Plassman, B. L., Langa, K. M., Fisher, G. G., Heeringa, S. G., Weir, D. R., Ofstedal, M. B., . . . Wallace, R. B. (2008). Prevalence of cognitive impairment without dementia in the United States. *Annals of Internal Medicine*, 148, 427–434.

Plassman, B. L., Newman, T., Welsh, K., & Breitner, J. (1994). Properties of the telephone interview for cognitive status. Application in epidemiological and longitudinal studies. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, 7, 235–241.

Sliwinski, M., Hofer, S., Hall, C., Buschke, H., & Lipton, R. (2003). Modeling memory decline in older adults: The importance of preclinical dementia, attrition, and chronological age. *Psychology of Aging*, 18, 658–671. doi:10.1037/0882-7974.18.4.658

Suthers, K., Kim, J. K., & Crimmins, E. (2003). Life expectancy with cognitive impairment in the older population of the United States. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 58, 179–186. doi:10.1093/geronb/58.3.S179

Weir, D., Wallace, R., Langa, K., Plassman, B., Wilson, R., Bennett, D., . . . Sano, M. (2011). Reducing case ascertainment costs in U.S. population studies of Alzheimer's disease, dementia, and cognitive impairment-Part 1. *Alzheimer's and Dementia*, 7, 94–109. doi:10.1016/j.jalz.2010.11.004

Welsh, K., Breitner, J., & Magruder-Habib, K. (1993). Detection of dementia in the elderly using telephone screening of cognitive status. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, 6, 101–110.

Zelinski, E. M., Crimmins, E., Reynolds, S., & Seeman, T. (1998). Do medical conditions affect cognition in older adults? Results from two studies. *Health Psychology*, 17, 504–512. doi:10.1037/0278-6133.17.6.504

Zivin, K., Kabeto, M. U., Kales, H. C., & Langa, K. M. (2009). The effect of depression and cognitive impairment on enrollment in Medicare Part D. *Journal of the American Geriatric Society*, 57, 1433–1440. doi:10.1111/j.1532-5415.2009.02348.x