Published in final edited form as:

Int Psychogeriatr. 2013 August ; 25(8): 1325–1333. doi:10.1017/S1041610213000598.

Detection of Mild Cognitive Impairment and early stage dementia with an audio-recorded cognitive scale

Margaret C. Sewell¹, Xiaodong Luo¹, Judith Neugroschl¹, and Mary Sano^{1,2}

¹Mount Sinai School of Medicine, New York NY

²James J. Peters VAMC, Bronx, NY

Abstract

BACKGROUND—Physicians often miss a diagnosis of Mild Cognitive Impairment (MCI) or early dementia and screening measures can be insensitive to very mild impairments. Other cognitive assessments may take too much time or be frustrating to seniors. This study examined the ability of an audio-recorded scale, developed in Australia, to detect MCI or mild Alzheimer's disease and compared cognitive domain specific performance on the audio-recorded scale to in-person battery and common cognitive screens.

METHOD—Seventy-six subjects from the Mount Sinai Alzheimer's Disease Research Center were recruited. Subjects were 75 years or older, with clinical diagnosis of AD or MCI (n=51) or normal control (n=25). Participants underwent in-person neuropsychological testing followed by testing with the Audio-recorded Cognitive Screen (ARCS).

RESULTS—ARCS provided better discrimination between normal and impaired elders than either the Mini-Mental Status Exam (MMSE) or the clock drawing test. The in-person battery and ARCS analogous variables were significantly correlated, most in the .4 to .7 range, including verbal memory, executive function/attention, naming, and verbal fluency. The area under the curve generated from ROC curves indicated high and equivalent discrimination for ARCS and the in-person battery (0.972 vs. 0.988; p=0.23).

CONCLUSION—The ARCS demonstrated better discrimination between normal controls and those with mild deficits than typical screening measures. Performance on cognitive domains within the ARCS was well correlated with the in-person battery. Completion of the ARCS was accomplished despite mild difficulty hearing the instructions even in very elderly subjects, indicating that it may be a useful measure in primary care settings.

Keywords

Cognitive assessment; Alzheimer's disease; MCI; diagnostic accuracy

Corresponding Author: Margaret C. Sewell, Ph.D., The Alzheimer's Disease Research Center-Mount Sinai School of Medicine, 1 Gustave L. Levy Place, Box 1230, New York, NY 10029, Phone: 212-241-0188, Fax: 212-996-0987, margaret.sewell@mssm.edu. Alternate Corresponding Author: mary.sano@mssm.edu.

Conflict of Interest

None

Description of authors' roles:

M. Sewell designed the study, collected the data and wrote the paper. X. Luo analyzed the data and reviewed the paper. J. Neugroschl reviewed and edited the paper. M. Sano supervised study design, assisted in data analyses and edited and supervised the writing of the paper.

INTRODUCTION

Neuropsychological testing to determine the presence of memory and other cognitive deficits is critical in the diagnosis of the earliest stages of Alzheimer's disease (AD) and Mild Cognitive Impairment (MCI). Additionally, cognitive deficits that are captured in neuropsychological evaluations can be critical both to differential diagnosis and to identifying residual strengths and specific weaknesses useful in developing care management plans.

The need to assess cognitive function in the elderly is imperative since primary care physicians often do not detect early AD and MCI, with reports that more than half the cases are unrecognized (Valcour *et al.*, 2000; Kaduszkiewicz *et al.*, 2010). Reliance on a subjective memory complaint for early detection is not advisable because it is a common complaint that is neither sensitive nor specific to objective cognitive deficits (Purser *et al.*, 2006). Screening measures such as the Mini-Mental Status Examination (MMSE) are particularly insensitive to distinguishing MCI from normal controls or those with AD (Mitchell, 2009). Other tools such as the Montreal Cognitive Assessment (Nasreddine *et al.*, 2005), Test Your Memory (TYM) (Brown *et al.*, 2009) and telephone screening batteries (Manly *et al.*, 2011; Hill *et al.*, 2005) require a trained administrator or clinician and provide only a global score without information on specific profiles of deficits across different cognitive domains.

Other assessments require trained personnel such as the Brief Kingston Standardized Cognitive Assessment (Hopkins *et al.*, 2005) and the Philadelphia Brief Assessment on Cognition (Libon *et al.*, 2011), or additional technology such as computerized cognitive assessments (Inoue *et al.*, 2011; Doniger *et al.*, 2005). However, these assessments may not be convenient to a primary care office because they may use too much clinician time, be expensive, and may have poor acceptability to seniors because of frustration or unfamiliarity with technological demands (Sano *et al.*, 2010).

Schofield and colleagues (Schofield *et al.*, 2003) created the Audio-recorded Cognitive Screen (ARCS) which uses an automated audio delivery of test stimuli for cognitive tests through earphones, and is administered to unsupervised patients with written responses captured in a booklet. The ARCS assesses cognitive domains important in the assessment of elders with AD and other dementias: memory, verbal fluency, naming, visuospatial function and attention/executive function.

Validity and reliability have been described (Schofield *et al.*, 2010) in a large sample of Australian residents from community and clinic settings. High correlations were found between ARCS domain scores (memory, verbal fluency, naming, visuospatial function and attention/executive function) and scores from in-person evaluations. ARCS administration requires approximately 35 minutes, which is shorter than a standard neuropsychological test battery. Other advantages of a brief audio-recorded battery include low demand on the clinician or data collector and a technology with which older people are familiar. The ARCS has been described by Schofield and colleagues (Schofield *et al.*, 2010) as a "hybrid" instrument, one that encompasses the brevity of a cognitive screen and the breadth of a comprehensive battery.

In the present study we compared the ARCS to an in-person battery to determine 1) its ability to distinguish between impaired and non-impaired subjects 2) the correlation of performance within domains and 3) the ability to correctly detect subjects within diagnostic categories. We also evaluated the feasibility of use of the ARCS in an elderly cohort of individuals living in the metropolitan New York area by assessing the acceptability of the

administration and testing procedures. We hypothesize that the ARCS will be well-accepted in the elderly cohort and that the ARCS will provide a sensitive measure of diagnostic discrimination.

METHODS

Subjects

Subjects from the Mount Sinai Alzheimer's Disease Research Center (MSSM-ADRC) who previously consented to be approached for additional studies were invited to participate. Potential subjects were age 75 years or older, and had a most recent diagnostic evaluation of AD (non-AD dementias were excluded), MCI or normal control, with a Clinical Dementia Rating (CDR) of 0, 0.5, 1, or 2. Consistent with other studies (Weintraub *et al.*, 2009), the CDR was administered to our normal controls to confirm the absence of cognitive and functional impairment. We recruited 76 of 109 consecutively approached subjects. One additional subject signed consent but was not able to complete the study because of significant hearing loss. All participants had capacity to sign consent, and were given the option of being tested at the center or in their home. Participants were classified as either normal controls (n=25) or impaired subjects (n=51). The impaired subjects had an ADRC consensus diagnosis of Mild Cognitive Impairment or mild Alzheimer's disease (90% of the MCI and AD group had a global CDR of .5, 1, or 2 and 10% had a CDR of 0. The mean MMSE score was 23.5). Within the impaired subjects group of 51, a subgroup of 33 mildly impaired subjects was identified. This mildly impaired subgroup had a Global CDR $\leq$ 0.5 and included 21 MCI subjects and 12 AD subjects. Of the 33 mildly impaired subjects, 27 had a CDRSOB 0–2 and 6 had a CDRSOB 2.5–4.

Assessments/Instruments

Uniform Data Set (UDS) (Weintraub *et al.*, 2009)—This battery of neuropsychological tests, which takes approximately 60 minutes to complete not including scoring, is administered annually to all ADRC subjects by a trained research coordinator or neuropsychologist. It is collected under a separate Mount Sinai IRB approved protocol. This battery includes the Mini-Mental Status Exam, the Clinical Dementia Rating (CDR), the 15 item Geriatric Depression Scale (GDS), and standard tests of memory (Logical Memory), verbal fluency (animal naming) confrontation naming (Boston Naming Test, 30-item version), visuospatial function (Digit Symbol from the WAIS-III), and attention/executive function (Trail Making A and B, Digit Span from WAIS-III).

Clinical Assessment—The clinical evaluation was conducted by a geriatric psychiatrist or neurologist. A consensus diagnosis was reached with a physician (geriatric psychiatrist or neurologist) and a neuropsychologist, using the information from the clinical assessment, along with the UDS neuropsychological test data. No predetermined cut offs were used for the purposes of the consensus diagnosis.

Audio-Recorded Cognitive Screen (ARCS)—The ARCS (Schofield *et al.*, 2003) is an audio-recorded battery of cognitive tests that takes approximately 35 minutes to administer and less than five minutes to score. A high school volunteer or other non-professional provided instructions on the use of the compact disk device and headphones but the subjects were otherwise unsupervised as they listened to test instructions and wrote responses in a booklet. The 26-page booklet consists of lined pages for freehand responses, blank pages between subtests, naming pictures, a page with yes/no responses to be circled for recognition items. All instructions, including when to stop and turn to the next page, are on the audio-recording. The ARCS is comprised of the Newcastle Auditory Verbal Learning Test (NAVLT), a three-trial 12-word list learning test which includes a 10-minute delayed recall

and recognition trial; 60-second verbal fluency tests (letter, category and action fluency) where subjects write responses freehand, a 10-item naming test which uses pictures already printed in each response booklet, a clock drawing task, and the Hunter Attentional Task, an attention/executive function test which is similar to the Trail- Making Test. The first section of the Hunter Attentional Task-A (HAT-A) is a 40-item test with a 30 second trial in which capital letters are written next to lower case letters. The second section of the Hunter Attentional Task-B (HAT-B) consists of 40 randomly distributed circled and non-circled lower case letters that are pre-printed on the page. The subject has 30 seconds to write capital letters next to the non-circled letters, and lower case letters next to the circled letters. To assess speed of writing, the subject is instructed to write the word “table” as many times as possible in a 30 second interval on a blank page. The response booklets for the first 17 subjects were scored by the high school volunteer and the senior neuropsychologist to permit assessment of inter-rater reliability. The remainder was scored by the neuropsychologist.

To assess feasibility and acceptability of testing, the ARCS workbook includes questions (some answered yes/no, others answered on a Likert scale) regarding subjects’ perception of the testing and their performance including: “How do you think you performed on the test of memory and concentration?”, “Could you hear the instructions on the recording clearly?”, “Did you find the directions on the recording confusing or difficult to follow?”, “Did you give your best effort during this test?”, and finally, a question asking whether the subjects had any concerns that might have affected their test performance (e.g. nerves, fatigue, technical difficulties). The voice of the version used for this study was that of a US English speaker. All subjects were given the ARCS within six months of their most recent UDS testing date.

Data Collection and Analyses

In order to determine inter-rater reliability, we calculated the inter-class correlation coefficient (ICC) and the corresponding 95% confidence limits for the 14 ARCS variables that were analogous to the UDS variables. Demographic and global function scores were reported for the non-impaired and impaired groups. Feasibility (i.e. acceptability) items were summarized and compared across diagnostic groups. ARCS and UDS scores were compared across impairment groups and using t-tests. Correlations between UDS and ARCS domain scores were calculated using Spearman correlation coefficients to account for the potential nonlinear correlation between the scores. For the ARCS, an overall score for the five cognitive domains of memory (3 trials initial recall and 1 delayed recall of NAVLT), fluency (total of letter, category, action), naming, visuospatial (clock drawing) and executive/attention (HATA and HATB) were created by totaling the scores of the variables that comprise each of these tests in order to compare them to analogous domains on the UDS. Several ARCS variables (e.g. the recognition trial and three error indices) were not included in our analyses as there are no analogous variables in the UDS battery. We compared the Receiver Operating Characteristic (ROC) curves of these two test batteries in terms of their ability to correctly classify the non-impaired and impaired subjects. We also compared the ROC curves of the ARCS and the MMSE and Clock Drawing Test to determine if the ARCS was better able to classify impaired and non-impaired subjects than either the MMSE, the Clock Drawing, or the two combined. Individual scores from the ARCS and UDS were used in the regression analyses that generated the (ROC) curves. All the analyses were performed using SAS 9.1 for PC.

RESULTS

Inter-rater reliability

Two raters, the senior neuropsychologist and a high-school ADRC volunteer double-scored 17 of the 76 protocols, with the neuropsychologist scoring the remainder. Inter-class correlation coefficients (ICC) were calculated for each of the 14 ARCS variables. The ICC's ranged from 0.98–1.00.

Clinical and demographic comparisons

The study successfully recruited 76 elderly subjects with a mean age over 80 with no statistical difference between the two diagnostic groups on any demographic variable. Impaired and non-impaired subjects were well matched on age (impaired mean 81.9 ± 4.7 versus non-impaired 81.2 ± 4.0), education (better than high school: 51% versus 48%), gender (66% versus 48% male) and ethnicity (80% white for both groups). Mean MMSE scores, as expected, were significantly lower in the impaired group (23.5 ± 4.3) than the non-impaired group (28.1 ± 2.0), and CDR sum of boxes was higher in the impaired group (3.4 ± 3.2) than the non-impaired group (0.1 ± 0.2).

Mean scores for the impaired group on the Geriatric Depression Scale (3.0 ± 2.9) were significantly higher than the non-impaired group (1.7 ± 2.0), though both raw scores were low.

Feasibility Results

Using chi square analyses, no statistically significant differences were found between the impaired and non-impaired groups on any feasibility question. Results are summarized in Table 1.

To the question, “Could you hear the test instructions clearly?” a quarter of the sample ($n=19$) responded that they occasionally or frequently had difficulty hearing the test instructions (77% of the impaired subjects and 68% of the non-impaired subjects reported no difficulty hearing instructions). The mean age of those who reported difficulty hearing the instructions (83.4 ± 3.9) was significantly higher than those who did not report hearing problems (81.0 ± 4.5) ($t = -2.20$, $p < .03$). Age was not a statistically significant factor for any other feasibility question.

Performance on ARCS and UDS Variables

The results of Table 2 indicate that, as expected, the non-impaired group performed better than the impaired group on all measures with the exception of digit span forward. There was a significant difference ($p=0.01$) between the impaired (mean 9.7 words) and non-impaired groups (mean 11.6 words) on writing speed.

Correlations between ARCS and UDS domain scores are summarized in Table 3. The bolded entries are purported to assess the same domains of cognitive function. Some of the scores between domains (e.g. memory and executive function) are correlated, which is consistent with other neuropsychological test studies. The correlations within the expected domains are consistently modestly related. For example, the UDS and ARCS analogous variables were significantly correlated, most in the 0.4 to 0.7 range, including verbal memory, executive function, naming, verbal fluency, and attention. HAT-A/B and Trails-A/B scores are negatively correlated because HAT-A/B scores represent number of items correct and Trails A/B scores represent time. Information regarding how ARCS domain scores were generated, including factor analyses, may be found in Schofield's paper (Schofield *et al.*, 2010).

Ability to correctly classify subjects as impaired or non-impaired

In order to gauge our ability to classify subjects as impaired (MCI or early AD) or non-impaired, as defined by ADRC consensus conference, we analyzed the ROC curves generated from the individual UDS and ARCS variables. We first fitted two logistic regression models, one with the individual UDS variables and the other with the ARCS variables, both adjusted for age, gender, education and ethnicity. Then we generated the ROC curves based on these two models and compared the Area Under Curve (AUC) of these two ROC curves.

The AUC for the UDS variables was 0.988 and the AUC for the ARCS variables was 0.972. The difference was not statistically significant ($p=0.23$). The AUC results may be found in Figure 1. Examination of the graph suggests that, for the same amount of specificity, there is better sensitivity in some score ranges for the UDS battery than the ARCS. There is some circularity in the ROC analysis for the UDS but not for the ARCS, as diagnosis does include information from the ARCS. Because the AUC for the ARCS is not significantly worse than the AUC for the UDS, the existence of circularity makes our comparison more conservative.

Comparison of very mildly impaired and non-impaired subjects

We also compared the AUC of ROC curves for the ARCS in very mildly impaired ($N=33$ with a global $CDR \leq 0.5$ and a CDR_{SOB} of ≤ 2 in 27 subjects and a $CDR_{SOB} \geq 2.5$ in 6 subjects) and non-impaired subjects for the commonly used screening instruments MMSE and Clock Drawing Test, controlling for age, education and gender. The AUC for the ARCS full model was significantly better than for the MMSE alone (0.956. vs. 0.758; $p=0.018$), the clock drawing alone ($AUC=0.717$; $p=0.0002$) or the MMSE and Clock Drawing Test combined ($AUC=0.792$; $p=0.0035$). Based on this new classification, we regenerated ROC curves for individual UDS and ARCS variables. The AUC for the ARCS and UDS variables were not significantly different (ARCS $AUC=0.956$ vs. UDS $AUC=0.987$; $p=0.13$).

DISCUSSION

This study describes the use of the ARCS in an elderly cohort of individuals with a mean age above 80 with and without MCI and AD. With the exception of one subject (not included in the analyses) who had to discontinue due to inability to hear adequately the audio-recorded instructions, all other subjects were able to satisfactorily complete the ARCS. The ARCS data could be collected without the help of a professional in a variety of clinic and community settings. In general this supports the feasibility of assessment through the ARCS. A number of subjects reported that the instructions were confusing but most reported giving their best effort. A majority of the non-impaired subjects and half the impaired subjects perceived their test performance as satisfactory or better.

The ARCS took approximately 35 minutes to administer and 5 minutes to score. A high-school level research volunteer was adequately and reliably able to administer and score the ARCS. Though supervision is not needed for ARCS administration, in the current study, the volunteer or other staff member stayed with each subject while they completed the ARCS, initially to help adjust the volume on the Compact Disc player, and then to observe subjects' behavior. Since the subjects were not alone, we were not able to assess whether subjects "cheated" on the test (e.g. turn back to previous pages), a behavior that might be more likely of subjects knew they were not being observed and something that might be important for assessment in group settings. Schofield and colleagues (Schofield *et al.*, 2010) reported very little systematic cheating (fewer than 6 over 2,600 administrations). The current results suggest that the ARCS may be useful in settings that rely on personnel with little or no professional training in neuropsychology.

One challenge to the feasibility of using ARCS in this elderly sample was hearing loss. Approximately 26% of the sample reported either occasional or frequent difficulty hearing the audio-recorded instructions. Although there were no differences in reported hearing difficulty between the impaired and non-impaired groups, there were more reports of difficulty hearing the instructions among the older subjects. Though Schofield's studies do not specifically address hearing loss, our results may reflect the fact that the current sample (mean age 81.7) is considerably older than Schofield's original Australian sample (mean age was 59 years in one study (Schofield *et al.*, 2010) and 74 in another (Schofield *et al.*, 2003). The ARCS administration in the current study involved the use of a standard portable Compact Disc player and headphones. The ARCS may also be used with other devices including MP3, iPod, or computer audio. In future studies it may be useful to investigate the use of hearing assistive technologies (such as FM systems and infrared systems that provide better amplification or direct amplification into the listener's hearing aid).

Performance on tests of individual cognitive domains within the ARCS battery, which included measures of verbal memory, verbal fluency, attention/executive function, naming and visuospatial function, was significantly correlated with corresponding tests on the in-person battery. Though intercorrelations were evident among many tests, we were pleased that the UDS and ARCS analogous domains were significantly correlated. These results expand the findings of Schofield and colleagues (Schofield *et al.*, 2010) by observing a significant correlation between the ARCS and an in-person assessment in a significantly older sample.

Using the overall score of all cognitive domains, the ARCS' ability to distinguish between impaired (mild AD or MCI) and non-impaired diagnostic groups was found to be as good as the in-person test battery. These results were maintained when we examined a subset of mildly impaired subjects whose CDR was ≤ 0.5 , suggesting that the ARCS is useful in identifying mildly impaired individuals and distinguishing them from normal controls. The advantage of the ARCS over an in-person evaluation is that it reduces workload on the examiner and may be administered and scored by those with little training. The performance of our normal controls on the in-person battery was comparable to that reported by Weintraub and colleagues (Weintraub *et al.*, 2009) on the same tests in a sample of over 3,000 community-dwelling normal controls suggesting that our normal controls were relatively similar to this large heterogeneous sample. Our normal controls are comprised of those recruited through flyers and community lectures or may be spouses of demented subjects. It is possible that some normal controls came to the ADRC due to memory concerns, but a discussion of this is beyond the scope of our paper.

The ARCS was able to distinguish between diagnostic groups better than commonly used screening instruments, even in very mildly impaired subjects. These results are consistent with Schofield's description of the ARCS as a "hybrid" instrument, with the time benefits of a cognitive screen in conjunction with meaningful clinical information, such as domain scores, associated with longer neuropsychological test batteries.

To our knowledge this is the first use of the ARCS outside of Australia and the UK and the high correlation of audio and in-person assessment across critical cognitive domains suggests that common instruments can be developed and used across international settings. The ARCS may be particularly useful to medical generalists who are commonly required to make diagnostic decisions because of shortages of mental health specialists in low and middle income countries (Bruckner *et al.*, 2011) as well as in remote, rural areas of the US (Smalley *et al.*, 2010).

One limitation of the current study is that the sample was a self-selected group of research subjects who may have been predisposed to participate in additional studies. A limitation of any study with more than one time point of cognitive testing is possible practice effects. The fact that the ARCS and UDS administrations were spaced one to six months apart raises the possibility of practice effects on similar subtests for the shorter time frames or cognitive decline in the longer time frames. Furthermore, the sample size was small and was characterized in a tertiary medical setting. This is a highly educated, primarily Caucasian cohort that limits the generalizability of the findings. The normal controls were devoid of comorbid conditions such as movement disorders and depressive disorders that may be common in primary care practice. While this is a limitation regarding generalizability of the findings, it is often the healthiest elders who seek very early diagnosis of cognitive problems that may be particularly relevant to the primary care setting.

The current study uses ARCS raw scores as norms were not available during the data collection phase of our study. Since our data were initially collected, a website for the ARCS was created (Schofield, 2012). The website provides a scoring template/calculator for entering raw scores to generate demographically adjusted scores for individual tests, specific cognitive domains and for the global ARCS score. Norms are provided for age, education and gender, making interpretation of ARCS results more meaningful. The template/calculator can provide a quick indication of whether a patient is performing below expectation for their age on the domains most pertinent to dementia evaluations such as memory, verbal fluency and executive functioning.

In summary, the ARCS provided excellent discrimination between normal and mildly impaired elderly individuals, and was able to evaluate cognitive performance within specific key domains. It was well tolerated and the relative ease of administration suggests it may be useful in a wide range of settings.

Acknowledgments

The authors would like to express their gratitude to Talia Sandwick, B.A., for her work with subject recruitment and data collection.

REFERENCES

- Brown J, Pengas G, Dawson K, Brown LA, Clatworthy P. Self administered cognitive screening test (TYM) for detection of Alzheimer's disease: cross sectional study. *British Medical Journal*. 2009; 338:b2030.
- Bruckner TA, et al. The mental health workforce gap in low- and middle-income countries: a needs-based approach. *Bulletin of the World Health Organization*. 2011; 89:184–194. [PubMed: 21379414]
- Doniger GM, et al. Towards practical cognitive assessment for detection of early dementia: a 30-minute computerized battery discriminates as well as longer testing. *Current Alzheimer Research*. 2005; 2:117–124. [PubMed: 15974907]
- Hill J, et al. Validation of a brief screen for cognitive impairment (BSCI) administered by telephone for use in the medicare population. *Disease Management*. 2005; 8:223–234. [PubMed: 16117717]
- Hopkins RW, Kilik LA, Day DJ, Rows CP, Hamilton PF. The brief Kingston Standardized Cognitive Assessment--revised. *International Journal of Geriatric Psychiatry*. 2005; 20:227–231. [PubMed: 15717345]
- Inoue M, Jimbo D, Taniguchi M, Urakami K. Touch Panel-type Dementia Assessment Scale: a new computer-based rating scale for Alzheimer's disease. *Psychogeriatrics*. 2011; 11:28–33. [PubMed: 21447106]
- Kaduszkiewicz H, et al. Do general practitioners recognize mild cognitive impairment in their patients? *The Journal of Nutrition, Health and Aging*. 2010; 14:697–702.

- Libon DJ, et al. The Philadelphia Brief Assessment of Cognition (PBAC): A Validated Screening Measure for Dementia. *Clinical Neuropsychology*. 2011; 25:1314–1330.
- Manly JJ, Schupf N, Stern Y, Brickman AM, Tang MX, Mayeux R. Telephone-based identification of mild cognitive impairment and dementia in a multicultural cohort. *Archives of Neurology*. 2011; 68:607–614. [PubMed: 21555635]
- Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research*. 2009; 43:411–431. [PubMed: 18579155]
- Nasreddine ZS, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*. 2005; 53:695–699. [PubMed: 15817019]
- Purser JL, Fillenbaum GG, Wallace RB. Memory complaint is not necessary for diagnosis of mild cognitive impairment and does not predict 10-year trajectories of functional disability, word recall, or short portable mental status questionnaire limitations. *Journal of the American Geriatrics Society*. 2006; 54:335–338. [PubMed: 16460388]
- Sano M, et al. Pilot study to show the feasibility of a multicenter trial of home-based assessment of people over 75 years old. *Alzheimer Disease and Associated Disorders*. 2010; 24:256–263. [PubMed: 20592583]
- Schofield, P. ARCS: Audio-Recorded Cognitive Screen [Online]. Newcastle, Australia: 2012. Available: <http://www.cognitionhealth.com> [[Accessed September 15 2012]]
- Schofield PW, Lee SJ, Davies GC. Cognitive screening using a tape recorder: a pilot study. *Journal of the American Geriatrics Society*. 2003; 51:415–418. [PubMed: 12588588]
- Schofield PW, et al. The Audio Recorded Cognitive Screen (ARCS): a flexible hybrid cognitive test instrument. *Journal of Neurology, Neurosurgery and Psychiatry*. 2010; 81:602–607.
- Smalley KB, Yancey CT, Warren JC, Naufel K, Ryan R, Pugh JL. Rural mental health and psychological treatment: a review for practitioners. *Journal of Clinical Psychology*. 2010; 66:479–489. [PubMed: 20222125]
- Valcour VG, Masaki KH, Curb JD, Blanchette PL. The detection of dementia in the primary care setting. *Archives of Internal Medicine*. 2000; 160:2964–2968. [PubMed: 11041904]
- Weintraub S, et al. The Alzheimer's Disease Centers' Uniform Data Set (UDS): the neuropsychologic test battery. *Alzheimer Disease and Associated Disorders*. 2009; 23:91–101. [PubMed: 19474567]

Comparison of UDS and ARCS ROC curves based on full models

Test of H0: equal areas under curves -- $p=0.2259$

Figure 1. Receiver Operating Characteristic curves

The selection of a cut-off point of .80 results in a 14.3% false positive rate and a 30% false negative rate. This represents a significant improvement over the literature reporting that more than half of primary care physicians miss a dementia diagnosis (Valcour et al., 2000, Kaduszkiewicz et al., 2010).

Table 1Feasibility Question Results^a

Question	% Impaired Responding Yes	% Non-impaired Responding Yes	p-value
How do you think you performed on the test of memory and concentration (% satisfactory or better)	50	68	0.22
Did you find the directions on the recording confusing?	28	16	0.39
Did you give your best effort?	86	100	0.09
Do you believe "nerves" or anxiety affected your test performance?	36	32	0.80
Did you experience mental fatigue or exhaustion before starting this test?	16	8	0.48
Did you have technical difficulties (e.g. problems with the CD player or headphones)	8	4	0.66
Could you hear the instructions on the recording clearly? ^b	77	68	0.41

^a n=75 due to missing data

^b n=73 due to missing data

Table 2

Performance on ARCS and UDS Tests

Domain/Battery	Test	Non-impaired (N=25)	Impaired (N=51)	p-value
Writing speed				
ARCS	“Table” Writing	11.6 words	9.7 words	0.01
Memory				
UDS	Logical Memory initial	13.2(3.4)	6.7(4.3)	<0.0001
	Logical Memory delayed	12.8(3.8)	3.8(4.1)	<0.0001
ARCS	NAVLT initial	20.4(5.0)	13.0(5.7)	<0.0001
	NAVLT delayed	7.3(2.4)	2.7(2.6)	<0.0001
Verbal Fluency				
UDS	Animal Naming	17.9(4.0)	12.6(4.7)	<0.0001
ARCS	Action/Letter Fluency	34.6(8.8)	23.7(9.6)	<0.0001
Naming				
UDS	Boston Naming 30-item	26.3(2.7)	20.5(6.8)	<0.0001
ARCS	10-item Naming	8.2(1.2)	5.6(2.9)	<0.0001
Visuospatial				
UDS	WAIS Digit Symbol	39.6(7.5)	29.5(11.4)	<0.0001
ARCS	Clock Drawing	7.7(2.7)	4.3(3.7)	<0.0001
Attention/Executive				
UDS	Trails A	39.5(9.7)	59.2(33.5)	0.01
	Trails B	100.2(35.2)	191.4(94.0)	<0.0001
ARCS	HAT-A items correct	20.3(7.0)	15.8(11.0)	0.04
	HAT-B items correct	12.4(7.9)	8.0(7.5)	0.02

UDS=Uniform Data Set; ARCS=Audio-Recorded Cognitive Screen; NAVLT=Newcastle Auditory Verbal Learning Test; HAT=Hunter Attentional Task;

Table 3

Spearman Correlation Coefficients

	NAVLT initial ^c	NAVLT delayed	HAT A	HATB	Fluency Action/letter	Clock	ARCS Naming-10
Logical Memory Initial	0.56^a	0.61 ^a	0.37 ^a	0.52 ^a	0.50 ^a	0.60 ^a	0.58 ^a
Logical Memory Delayed	0.65 ^a	0.73^a	0.34 ^a	0.52 ^a	0.53 ^a	0.57 ^a	0.65 ^a
Digit Forward	0.25 ^b	0.06	0.24^b	0.19	0.46 ^a	0.28 ^b	0.16
Digit Backward	0.37 ^a	0.21	0.43^a	0.35 ^a	0.57 ^a	0.31 ^a	0.23 ^b
Animal Naming	0.59 ^a	0.53 ^a	0.51 ^a	0.48 ^a	0.61^a	0.60 ^a	0.65 ^a
Trails A	-0.32 ^a	-0.25 ^b	-0.45^a	-0.53 ^a	-0.53 ^a	-0.47 ^a	-0.29 ^b
Trails B	-0.52 ^a	-0.48 ^a	-0.64 ^a	-0.61^a	-0.62 ^a	-0.60^a	-0.42 ^a
Digit Symbol	0.41 ^a	0.40 ^a	0.68 ^a	0.66 ^a	0.65 ^a	0.52^a	0.42 ^a
Boston Naming-30	0.53 ^a	0.54 ^a	0.26 ^b	0.26 ^b	0.36 ^a	0.50 ^a	0.69^a

NAVLT=Newcastle Auditory Verbal Learning Test; HAT=Hunter Attentional Task; ARCS=Audio-recorded Cognitive Screen

^a p-value <=0.01;

^b p-value <=0.05;

^c n=75 due to missing data.