Published in final edited form as:

J Alzheimers Dis. 2013 January 1; 36(3): 433–443. doi:10.3233/JAD-130113.

Screening by Telephone in the Alzheimer's Disease Anti-inflammatory Prevention Trial

Gila Z. Reckess, Ph.D.¹, Jason Brandt, Ph.D.^{1,2,3}, Cheryl A. Luis, Ph.D.⁴, Peter Zandi, Ph.D.⁵, Barbara Martin, Ph.D.⁶, John C. S. Breitner, M.D., Ph.D.^{7,8}, and the ADAPT Research Group

¹ Department of Psychiatry & Behavioral Sciences, The Johns Hopkins University School of Medicine Baltimore, MD

² Department of Neurology, The Johns Hopkins University School of Medicine Baltimore, MD

³ The Copper Ridge Institute Sykesville, MD

⁴ Roskamp Institute Sarasota, FL

⁵ Department of Mental Health The Johns Hopkins Bloomberg School of Public Health Baltimore, MD

⁶ Lancaster Heart & Stroke Foundation and Clinical Research Center Lancaster, PA

⁷ Douglas Mental Health University Institute Montreal, Canada

⁸ Department of Psychiatry, McGill University Faculty of Medicine Montreal, Canada

Abstract

Compared with in-person assessment methods, telephone screening for dementia and other cognitive syndromes may improve efficiency of large population studies or prevention trials. We used data from the Alzheimer's Disease Anti-Inflammatory Prevention Trial (ADAPT) to compare performance of a four-test Telephone Assessment Battery (TAB) that included the Telephone Interview for Cognitive Status (TICS) to that of a traditional in-person Cognitive Assessment Battery (CAB). Among 1,548 elderly participants with valid telephone and in-person screening results obtained within 90 days of each other, 225 persons were referred for a full cognitive diagnostic evaluation (DE) that was completed within six months of screening. Drawing on results from this panel of 225 individuals, we used the Capture-Recapture method to estimate population numbers of cognitively impaired participants. The latter estimates enabled us to compare the performance characteristics of the two screening batteries at specified cut-offs for detection of dementia and milder forms of impairment. Although our results provide relatively imprecise estimates of the performance characteristics of the two batteries, a comparison of their relative performance suggests that, at selected cut-off points, the TAB produces results broadly comparable to in-person screening and may be slightly more sensitive in detecting mild impairment. TAB performance characteristics also appeared slightly better than those of the TICS alone. Given its benefits in time and cost when screening for cognitive disorders, telephone screening should be considered for large samples.

Keywords

Telephone; neuropsychology; dementia; Alzheimer disease; mild cognitive impairment; prodromal period; memory disorders; geriatric assessment

Introduction

Brief, valid, and cost-effective methods of screening for cognitive disorders are essential for large, community-based studies. The demand for such methods will grow as interventions are developed to halt or slow the progression of pre-clinical Alzheimer's disease (AD) to dementia [1]. In-person cognitive screening is especially costly for large-scale research studies and for those with subjects in remote locations. These issues have prompted interest in telephone screening as an alternative [1].

Telephone assessment has obvious limitations. First, its test items are necessarily administered by voice, thus prohibiting (for example) the assessment of visuospatial cognition. The telephone examiner also has limited control over environmental factors that may adversely affect performance, and is typically unable to ensure adherence to instructions (e.g., not writing down test items or getting help from others nearby). Additional concerns, such as hearing impairment, may be especially problematic with elderly populations. Finally, it may be more difficult to build rapport over the telephone. Although these problems may be reduced as web-based video “telephoning” becomes more popular, they will likely persist to some degree.

Despite these limitations, several telephone-based cognitive screening tests have been developed [2-4], and they are generally well tolerated by older adults [5, 6]. Thus, telephone screening has been used successfully in several large-scale studies for detection of dementia [7-10]. Nonetheless, some drawbacks remain apparent. For example, the Telephone Interview for Cognitive Status (TICS), one of the most widely used telephone-based dementia screening tools [11, 12], has excellent sensitivity and specificity for dementia [12-14] but may lack adequate sensitivity to detect mild cognitive impairment (MCI) [14, 15].

The Alzheimer's Disease Anti-Inflammatory Prevention Trial (ADAPT) offered an opportunity to compare the results of telephone and in-person cognitive screening in a large cohort of older adults at risk for cognitive disorders. Anticipating a need for future cost-saving measures, ADAPT investigators developed a Telephone Assessment Battery (TAB) that built upon the TICS for use in the trial's final wave of cognitive screening. To assess the performance of the TAB, ADAPT personnel administered it in close temporal proximity to a standard in-person Cognitive Assessment Battery (CAB) that had been administered annually in the trial for several years. (See Table 1a for specific test items of both batteries.) The resulting data enabled us to compare the performance characteristics of the two batteries for the detection of dementia and of milder cognitive syndromes.

Materials and Methods

Participants

ADAPT was a multicenter, randomized, placebo-controlled trial of naproxen or celecoxib for the primary prevention of AD dementia. Its enrollees were older adults (age ≥ 70 at entry) who reported a family history of AD dementia but who were cognitively normal at entry. They were also required to have an available collateral respondent to provide third-

party information. The trial was funded by the National Institute on Aging between 2000 and 2007 [16].

ADAPT enrolled 2,528 volunteers at six dementia clinics in the U.S. (Baltimore, Maryland; Boston, Massachusetts; Rochester, New York; Seattle, Washington; Sun City, Arizona; and Tampa, Florida). All pertinent institutional review boards approved the study protocol, and written informed consent was obtained from all subjects and collateral respondents. All enrollees performed above specified cut-off scores on a brief battery of neuropsychological tests [16] and were found to have no active neurological or psychiatric illness or medical contraindication to use of non-steroidal anti-inflammatory drugs. Participants were also required to be proficient in English. The ADAPT protocol and its inclusion and exclusion criteria are available at <http://www.jhucct.com/adapt/manall43.pdf>.

After enrollment, participants were screened annually with the in-person CAB. Although the trial's treatments were discontinued in 2004 [17], the cohort was followed for an additional two years to assess extended efficacy and safety effects. At the final wave of cognitive screening, 1,860 individuals were administered both the CAB and the TAB. Of these, 312 were excluded from the analyses reported here because their paired screening tests were administered more than 90 days apart ($N = 164$) or because at least one of the tests was rated by examiners to be of questionable validity ($N = 148$). Causes for the latter judgment included hearing problems, interruptions during testing, or physical or emotional illness that interfered with completion of testing. The remaining 1,548 participants constituted the analysis pool for the present work (Figure 1). Using previously selected cut-off "trigger" scores (Table 1a), those who had screened positive on either test battery were asked to return for a more comprehensive dementia evaluation (DE) that including a standardized neuropsychological test battery (Table 1b). DE results were then reviewed by consensus panels for assignment to diagnostic groups as described below.

Test batteries and interpretation

All neuropsychological tests were administered by study personnel who had completed a detailed training and certification process. Training was overseen by a senior clinical and research neuropsychologist (J Brandt). Psychometrists converted raw to standard scores using age-, sex- and education-adjusted norms from the Cache County Memory Study [10]. At the completion of each testing session, psychometrists rated the validity of each participant's performance on a scale of 0 to 5; test sessions with ratings less than 4 were excluded.

Trigger criteria for the CAB and TAB are summarized in Table 1a. The CAB has been described previously [16]. The TAB included the well-known TICS [11, 12] as its core screening measure. The TICS includes 11 items that assess orientation, praxis, attention, calculation, language, and memory, and yields a total possible score of 41. It correlates highly with the Mini-Mental State Examination [11, 12, 18] and other neuropsychological tests [9, 19], has been translated and cross-validated in several languages [15, 20-25], and has also been adapted for computer-assisted administration [26]. The TICS has high test-retest reliability [9, 12, 13] and split-half reliability [19], and excellent sensitivity and specificity for detection of dementia [12-14]. However, less is known about its utility as a screening test for milder cognitive syndromes.

In an effort to maximize the TAB's sensitivity for mild cognitive impairment (MCI) [27] and similar syndromes, and also to ensure some direct comparability between the CAB and TAB, the latter included two other tasks: a test of generative verbal fluency (reciting supermarket items in one minute); and immediate and delayed story recall from the Rivermead Behavioral Memory Test (RBMT [28]). To obtain structured observations about

participants' cognitive and functional capacities, the TAB also included the 16-item version of the Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE [29]), completed by collateral respondents. To improve the validity of telephone testing, participants' spouses or other co-residents (often their collateral respondents) were asked to remove any visual memory aids (calendars, newspapers, paper and pens) and to turn off audible distracters (e.g., radio, television).

Participants were referred for a full DE if they met trigger criteria on either the CAB or the TAB. The consensus panel at each site that reviewed DE results included the site director, study physician, study nurse, and site neuropsychologist, at a minimum. Based on all available data, often including laboratory tests and neuroimaging, these clinicians assigned each participant to a diagnostic group. Participants were first designated as having either "normal cognition" or "impaired cognition." The latter group was then specified as having "dementia" or "cognitive impairment - no dementia" (CIND). The CIND group was then further specified by first judging whether participants met criteria for amnesic mild cognitive impairment (aMCI) as defined by Petersen et al. [27]. Participants were assigned to a category of "prodromal AD" (pAD) if they did not meet formal criteria for aMCI but were believed to have a clinical syndrome consistent with earliest stage AD (e.g., declining memory and borderline language or visuospatial skills without reported functional decline) [30]. For this analysis, we pooled the MCI and pAD groups. The remaining participants with CIND were classified as having "other CIND." These included participants with cognitive impairment with another neurological syndrome, such as stroke or Parkinson's disease, as well as those without severe memory impairment and those lacking a history of functional decline. This last group often displayed signs suggestive of vascular cognitive impairment. Thus, this analysis considered three diagnostic groups: (1) dementia (all causes); (2) MCI/pAD; and (3) other CIND.

Analyses

Analyses considered screening test performance for the first-level grouping (impaired versus normal cognition) and for the three diagnostic groups.

An important limitation of these analyses was the unknown status of the large number of individuals who did not trigger either on the CAB or the TAB, and who were therefore not referred for a DE. Thus, for each diagnostic group we could count only the numbers of *true positives* and *false positives* resulting from the two screening batteries. This meant that the only performance characteristic we could calculate directly was positive predictive value (PPV; i.e., the proportion of individuals who triggered a DE who were subsequently assigned to the specified diagnostic category). To estimate other performance characteristics (viz., sensitivity, specificity, negative predictive value), we needed a method for estimation of the "true" number of cases who would meet criteria for each diagnostic group if all participants had, in fact, completed a DE. Such estimates can be made using the capture-recapture method [31, 32].

Capture-recapture is often used in game and wildlife husbandry, where the technique can be used to estimate the total number of individuals (e.g., fish) within a bounded set (e.g., a lake). The general concept is that one can estimate the number of fish in a lake by collecting two separate samples. The first time the net is cast, a sample of fish is collected, counted, and tagged. The fish are then returned to the lake, and a second sample is collected. The resulting data – the number of fish caught by both castings of the net (i.e., the number of tagged fish in the second sample) and the number unique to each sample – are then used to estimate the total number of fish in the lake. In the current study, administration of the CAB was the first "casting of the net," and those individuals who triggered a DE that resulted in a diagnosis of cognitive impairment constituted the "first capture" (i.e., the tagged fish). The

TAB was the second “casting of the net,” and those whose TAB results then triggered a DE revealing cognitive impairment constituted the “second capture.” These values were then used to estimate the number of individuals in the original sample who were impaired but who did not meet trigger criteria on either the CAB or TAB (i.e., false negatives), which, in turn, was used to estimate the total number of individuals with impaired cognition in the sample. Table 2 demonstrates the method applied to the impaired cognition group.

The capture-recapture method has been applied in many epidemiological studies where it was not possible or practical to measure the size of a population directly. This method is clearly inferior to a true population count. Other than the positive predictive values, all performance characteristics reported here can only be viewed as rough estimates. Nonetheless, these estimates are informative when used to compare the two batteries, since each battery is equally subject to the limitations of the estimating procedures.

The capture-recapture method was used to estimate the number of cases in each diagnostic grouping. Parallel analyses used similar methods to evaluate the performance characteristics of the TICS alone. We also used paired t-tests to compare scores from similar tests within the CAB and TAB (RBMT and verbal fluency), probing differences between administration formats.

Results

Of the 1,548 participants, 333 were referred for follow-up dementia evaluation based on pre-determined trigger criteria. Of these, 278 participants (83.48%) completed their DE. To minimize the possibility of significant cognitive decline during the interval between screening and dementia evaluation, we excluded participants who, for any reason, had a DE completed more than six months after their relevant screening assessment. After exclusion of 53 such individuals, the proportion of CAB triggers who completed a DE ($76/101 = .75$) was similar to that of TAB completers ($198/294 = .67$), suggesting that failure to complete a DE in a timely manner was not due to factors unique to either of the batteries. The remaining sample consisted of 225 individuals who (a) had valid CAB and TAB screens within 90 days of each other, (b) triggered a DE on one or both of these screening assessments; and (c) completed a DE within six months of the triggering battery (see Figure 1).

The average time required to complete the TAB ($M = 15.87$ min.; $SD = 3.06$) was less than half that for the CAB ($M = 34.61$ min.; $SD = 5.59$). There were no statistically significant differences between scores on the two component tests that overlapped (supermarket fluency: $t(224) = -0.83$, $p = .41$, $r = .06$; Rivermead delayed recall: $t(244) = 1.85$, $p = .07$, $r = .12$), suggesting that the administration method did not significantly affect performance. Nonetheless, the TAB resulted in many more DE referrals than did the CAB. Of the 225 individuals studied here, 149 (66%) were sent for a DE because they fell below cut-offs on the TAB only; 27 (12%) had DEs triggered by the CAB only; and 49 (22%) triggered on both tests.

Overall, 87 participants (38.67%) had some form of impaired cognition on the DE (Figure 2). Only 11 (4.89%) of these were assigned to the dementia group, while 51 (22.67%) were assigned to the MCI/pAD group and 25 (11.11%) to the Other CIND group. Demographic characteristics of each of these diagnostics group are summarized in Table 3.

Table 4 summarizes the performance characteristics of each screening battery. Raw numbers for all analyses were adjusted by the overall rate of completion ($225/333$) to estimate values for the original sample of screen-positive individuals. The TAB captured more cases of

impaired cognition than the CAB, resulting in higher overall sensitivity. However, the TAB also yielded more false positives, resulting in lower specificity and PPV. Both batteries had excellent specificity and NPV, with the CAB yielding slightly higher values.

Regarding diagnostic subgroups, the TAB detected many more participants with MCI/pAD and Other CIND than did the CAB. However, the TAB failed to identify 2 of the 11 participants who ultimately were diagnosed with dementia, whereas the CAB successfully identified all 11. This pattern is evident in the sensitivity estimates; the TAB had higher sensitivity for MCI/pAD and Other CIND, whereas the CAB had higher sensitivity for dementia. Regarding specificity, both the TAB and CAB performed well, although the TAB's specificity (range: 0.82-0.85) was somewhat lower than the CAB's (range: 0.94-0.96).

The PPV of each screening battery was low (0.37) for all subgroups, with the highest values being for MCI/pAD. The TAB had a slightly lower PPV than the CAB for dementia and MCI/pAD despite detecting a larger number of individuals in the latter group. With respect to NPV, the CAB and TAB each performed well (NPV 0.88). NPV for the CAB was marginally higher for each group.

Table 4 also shows performance characteristics of the TICS alone. A total of 89 people met trigger criteria on the TICS, 62 of whom completed a DE. TICS sensitivity was slightly lower than the TAB for dementia and meaningfully lower for MCI/pAD and Other CIND. TICS specificity and NPV were slightly better than the TAB for all groups, and roughly equivalent to the CAB. TICS PPV was equal to or slightly better than the TAB. Overall, these results suggest that supplementing the TICS with the RBMT, Fluency, and IQCODE improved the detection of cases somewhat, but also increased the number of false positives.

Discussion

Using the capture-recapture approach to estimate the number of true cases in each diagnostic group, we compared the performance characteristics of the TAB, the CAB, and the TICS as applied in 1,548 enrollees from the final screening wave in the ADAPT. This was accomplished using data from the 225 individuals who underwent a complete dementia evaluation within six months of screening results that indicated the need for such evaluation. Performance characteristics of each battery were examined for 3 outcomes: dementia, MCI/pAD, and other forms of cognitive impairment ("Other CIND").

Our most important finding is that the TAB outperformed the CAB in detecting MCI/pAD but not in detecting dementia. For dementia, the CAB showed a very slight advantage, interpretation of which is unclear (it identified all 11 participants with dementia, whereas the TAB identified only 9). Not surprisingly, the TAB's higher sensitivity for impaired cognition in general was accompanied by reduced specificity for all three diagnostic groups. Nonetheless, specificity estimates for the TAB were quite acceptable (0.82), suggesting that the vast majority of individuals without impairment do pass the screening battery.

The proportion of false-positives to true-positives is suggested by the PPV, which reflects the proportion of those who test positive (i.e., triggered a dementia evaluation) who have the condition (i.e., were ultimately assigned to a clinical group after their DE). The PPV is important because it yields an index of apparently wasted effort pursued in the evaluation of screen-positive individuals. However, PPV is sensitive not only to the performance of the test but also to the prevalence (or "base rate") of the condition under investigation. In keeping with its specificity, the PPV for the TAB was less than 25% for all diagnostic groups. This shows that a considerably larger number of individuals who failed the screening battery were, in fact, unimpaired. Indeed, the CAB's PPVs were also low (9% to

37%), although slightly better than the TAB's for both dementia and MCI/pAD. Both batteries had excellent estimated NPVs for all diagnostic categories, suggesting that a large majority of individuals who pass the screeners are indeed cognitively normal.

We were curious to compare the performance of the TAB with that of the original TICS that formed its core. We found that the TICS alone demonstrated good sensitivity (0.73) and specificity (0.95) for dementia. Although it also showed excellent specificity (0.96) for MCI/pAD, its sensitivity (0.29) for the latter was very low (at least with the specified cut-off points) and, in our opinion, unacceptable. We note that somewhat better values have been reported by others [14, 15], probably because of methodological differences (e.g. different cut-points). Overall, however, our observed pattern of good sensitivity for the TICS in detection of dementia but reduced sensitivity for MCI/pAD is consistent with the many reports that show the TICS is most useful in detecting unambiguous impairment. Our findings also show that addition of 3 additional test measures (RBMT Delayed Recall, Fluency, and IQCODE) in the TAB improve the battery's performance for detection of milder syndromes.

Limitations

Our study capitalized on an opportunity to use data from the large and well characterized ADAPT cohort. However, several important limitations should be mentioned. First, and probably most importantly, we needed an estimation procedure to calculate the sensitivity, specificity and NPV of our screening measures. This need was driven by the fact that individuals received a dementia evaluation only if they “failed” one or both of the screening batteries. We employed the capture-recapture method of estimation, which is regularly used in epidemiological research but is less well known in the psychology literature. Estimation is clearly less ideal than direct measurement, but we believe that such procedures are nonetheless worthwhile. This is especially true for studies such as this one, where the primary goal is to *compare* two screening batteries. Although the precise performance characteristics reported here are only estimates, we can compare the *relative* merits of each battery because the same estimation method was applied to both batteries. It should also be noted that the precision of this estimation procedure is also reduced by small sample sizes in some of the diagnostic sub-groups. Most conspicuously, there were only 11 new dementia onsets discovered in this round of ADAPT screening, none of which were identified by the TAB alone. As a result, there was one cell with a value of zero when applying the capture-recapture method to the estimation of dementia cases in the original sample. While this may reduce the precision of the estimate, we do not believe it invalidates the result.

Another important consequence of the estimation procedure was that we were unable to evaluate the effect of varying cut-off points on the screening batteries using Receiver Operating Characteristics analysis. In some circumstances, at least, it is possible that different cut-offs might have yielded better results.

The pragmatics of a large-scale clinical trial also resulted in some suboptimal methods. First, the data collection process resulted in a large number of missing or excluded exams. Second, it would probably have been ideal to randomize or counterbalance the order of administration of the CAB and TAB, but the ADAPT group decided that this approach would have added burden and probable error. The group therefore opted to administer the TAB shortly after the final CAB (all participants having taken the CAB on two or more occasions previously). Third, to accommodate participant and staff availability, there was considerable variability (3 – 84 days) in the interval between administration of the CAB and of the TAB. The proportion of participants who completed their prescribed DE within a 90-day window of screening was not as high as we would have liked.

A different sort of limitation is the relative homogeneity of the ADAPT sample (Table 3). Despite strenuous efforts to recruit ethnic minorities, this cohort was a highly educated group of predominantly Caucasian, English-speaking volunteers who were at least 70 years old. It also had the unusual characteristic that all participants had at least one first-degree relative with AD. Appropriate caution is warranted when considering whether the current findings would generalize to other demographic groups.

Conclusions

In sum, our results suggest that telephone screening with the TAB can be useful for economical screening of older populations for cognitive disorders, and that the slightly longer TAB is an improvement over the TICS when detection of milder disorders is important. In addition to being far more efficient, telephone-based screening is more convenient for participants and avoids considerations such as time and means of conveyance, parking, and other burdensome aspects of in-person screening. These advantages may well outweigh whatever improvement in performance can be achieved by in-person examination. Thus, our work adds to a growing literature supporting telephone-based screening as the first step in two-tier methods for identification of cognitive impairment.

Acknowledgments

Dr. Brandt receives royalties from Psychological Assessment Resources, Inc. on sales of the Telephone Interview for Cognitive Status and the Hopkins Verbal Learning Test-Revised. This relationship is managed by the Johns Hopkins University according to its usual conflict of interest policies. Dr. Brandt received no royalties from the use of these tests in ADAPT.

ADAPT Research Group: This work was supported by the U.S. Department of Veterans Affairs Puget Sound Health Care System, Seattle, Washington, by the National Institutes of Health grants U01-AG-15477, R01-AG-24010, and P50-AG-05136, and by the N. Bud Grossman Center for Memory Research and Care, Minneapolis, MN, and McGill University. Celecoxib and matching placebo were provided by Pfizer, Inc. Naproxen sodium and its placebo were provided by Bayer Consumer Healthcare. No funding or supporting agency had any role in the design or conduct of the study; collection, management, or interpretation of the data; or preparation of the manuscript. Courtesy copies of the manuscript have been provided to the funding agency and to Pfizer, Inc., and Bayer Consumer Healthcare. The authors are deeply indebted to the outstanding staff of the ADAPT field sites and Coordinating Center and, especially, to the ADAPT participant cohort for their persistent commitment and generosity of time and effort.

List of Acronyms

AD	Alzheimer's disease
ADAPT	Alzheimer's Disease Anti-Inflammatory Prevention Trial
CAB	Cognitive Assessment Battery
CIND	Cognitive impairment, not dementia
DE	Dementia Evaluation
MCI	Mild cognitive impairment
TAB	Telephone Assessment Battery
TICS	Telephone Interview for Cognitive Status

References

1. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR Jr, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH. Toward defining the preclinical stages of

Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011; 7:280–292. [PubMed: 21514248]

2. Wolfson C, Kirkland SA, Raina PS, Uniat J, Roberts K, Bergman H, Furlini L, Pelletier A, Stroppe G, Angus CL, Keshavarz H, Szala-Meneok K. Telephone-administered cognitive tests as tools for the identification of eligible study participants for population-based research in aging. *Canadian Journal on Aging-Revue Canadienne Du Vieillissement*. 2009; 28:251–259.
3. Lewis B, Mills C, Mohs R, Hill J, Fillit H. Improving early recognition of Alzheimer's disease: A review of telephonic screening tools. *Journal of Clinical Outcomes Management*. 2001; 8:41–45.
4. Smith MM, Tremont G, Ott BR. A review of telephone-administered screening tests for dementia diagnosis. *Am J Alzheimers Dis Other Dement*. 2009; 24:58–69.
5. Nesselroade JR, Pedersen NL, McClearn GE, Plomin R, Bergeman CS. Factorial and criterion validities of telephone-assessed cognitive-ability measures - age and gender comparisons in adult twins. *Res Aging*. 1988; 10:220–234. [PubMed: 3175323]
6. Carpenter B, Strauss M, Ball A. Telephone assessment of memory in the elderly. *Journal of Clinical Gerontology*. 1995; 1:107–117.
7. Gallo J, Breitner J. Alzheimer's disease in the NAS-NRC Registry of Aging Twin Veterans, IV: Performance characteristics of a two-stage telephone screening procedure for Alzheimer's dementia. *Psychol Med*. 1995; 25:1211–1219. [PubMed: 8637951]
8. Lines C, McCarroll K, Lipton R, Block G. Telephone screening for amnesic mild cognitive impairment. *Neurology*. 2003; 60
9. Kang J, Weuve J, Grodstein F. Postmenopausal hormone therapy and risk of cognitive decline in community-dwelling aging women. *Neurology*. 2004; 63:101–107. [PubMed: 15249618]
10. Tschanz JT, Welsh-Bohmer KA, Skoog I, West N, Norton MC, Wyse BW, Nickles R, Breitner JCS. Dementia diagnoses from clinical and neuropsychological data compared - The Cache County study. *Neurology*. 2000; 54:1290–1296. [PubMed: 10746600]
11. Brandt J, Folstein M. Telephone Interview for Cognitive Status: Professional Manual. Psychological Assessment Resources, Inc.; Lutz, FL: 2003.
12. Brandt J, Spencer M, Folstein M. The Telephone Interview for Cognitive Status. *Neuropsychiatry Neuropsychol Behav Neurol*. 1988; 1:111–117.
13. Desmond DW, Tatemichi TK, Hanzawa L. The Telephone Interview for Cognitive Status (TICS) - Reliability and validity in a stroke sample. *Int J Geriatr Psychiatry*. 1994; 9:803–807.
14. 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. *Arch Neurol*. 2011; 68:607–614. [PubMed: 21555635]
15. Seo EH, Lee DY, Kim SG, Kim KW, Kim DH, Kim BJ, Kim MD, Kim SY, Kim YH, Kim JL, Kim JW, Moon SW, Park JH, Ryu SH, Yoon JC, Lee NJ, Lee CU, Jhoo JH, Choo IH, Woo JI. Validity of the telephone interview for cognitive status (TICS) and modified TICS (TICS-m) for mild cognitive impairment (MCI) and dementia screening. *Arch Gerontol Geriatr*. 2011; 52:E26–E30. [PubMed: 20471701]
16. ADAPT Research Group MC, McCaffrey LD, Breitner JC. Alzheimer's Disease Anti-Inflammatory Prevention Trial: Design, methods, and baseline results. *Alzheimer's & Dementia*. 2009; 5:93–104.
17. Group AR. Cognitive function over time in the Alzheimer's Disease Anti-inflammatory Prevention Trial (ADAPT). *Arch Neurol*. 2008; 65:896–905. [PubMed: 18474729]
18. Folstein MF, Folstein SE, Mchugh PR. Mini-Mental State - practical method for grading cognitive state of patients for clinician. *J Psychiatr Res*. 1975; 12:189–198. [PubMed: 1202204]
19. Black BS, Kasper J, Brandt J, Shore AD, German P, Burton L, Gallo JJ, Lyketsos C, Rabins PV. Identifying dementia in high-risk community samples: The memory and medical care study. *Alzheimer Dis Assoc Disord*. 2003; 17:9–18. [PubMed: 12621315]
20. Beeri M, Werner P, Davidson M, Schmidler J, Silverman J. Validation of the modified Telephone Interview for Cognitive Status (TICS-m) in Hebrew. *Int J Geriatr Psychiatry*. 2003; 18:381–386. [PubMed: 12766912]

21. Dal Forno G, Chiovenda P, Bressi F, Ferreri F, Grossi E, Brandt J, Rossini P, Pasqualetti P. Use of an Italian version of the Telephone Interview for Cognitive Status in Alzheimer's disease. *Int J Geriatr Psychiatry*. 2006; 21:126–133. [PubMed: 16416467]
22. Ferrucci L, Del Lungo I, Guralnik J, Bandinelli S, Benvenuti E, Salani B, Lamponi M, Ubezio C, Benvenuti F, Baroni A. Is the Telephone Interview for Cognitive Status a valid alternative in persons who cannot be evaluated by the Mini Mental State Examination? *Aging (Milano)*. 1998; 10:332–338. [PubMed: 9825025]
23. Konagaya Y, Washimi Y, Hattori H, Takeda A, Watanabe T, Ohta T. Validation of the Telephone Interview for Cognitive Status (TICS) in Japanese. *Int J Geriatr Psychiatry*. 2007; 22:695–700. [PubMed: 17431929]
24. van den Berg E, Ruis C, Jan Biessels G, Jaap Kappelle L, van Zandvoort M. The Telephone Interview for Cognitive Status (Modified): Relation with a comprehensive neuropsychological assessment. *J Clin Exp Neuropsychol*. 2012; 34:598–605. [PubMed: 22384819]
25. Vercambre MN, Cuvelier H, Gayon YA, Hardy-Leger I, Berr C, Trivaille C, Boutron-Ruault MC, Clavel-Chapelon F. Validation study of a French version of the modified telephone interview for cognitive status (F-TICS-m) in elderly women. *Int J Geriatr Psychiatry*. 2010; 25:1142–1149. [PubMed: 20054838]
26. Buckwalter JG, Crooks VC, Petitti DB. A preliminary psychometric analysis of a computer-assisted administration of the telephone interview of cognitive status-modified. *J Clin Exp Neuropsychol*. 2002; 24:168–175. [PubMed: 11992199]
27. Petersen. Mild cognitive impairment: Clinical characterization and outcome (vol 56, pg 303, 1999). *Arch Neurol*. 1999; 56:760–760. [PubMed: 10369319]
28. Wilson B, Cockburn J, Baddeley A, Hiorns R. The development and validation of a test battery for detecting and monitoring everyday memory problems. *J Clin Exp Neuropsychol*. 1989; 11:855–870. [PubMed: 2592527]
29. Jorm AF, Scott R, Cullen JS, Mackinnon AJ. Performance of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) as a screening-test for dementia. *Psychol Med*. 1991; 21:785–790. [PubMed: 1946866]
30. Mayeux R, Reitz C, Brickman AM, Haan MN, Manly JJ, Glymour MM, Weiss CC, Yaffe K, Middleton L, Hendrie HC, Warren LH, Hayden KM, Welsh-Bohmer KA, Breitner JC, Morris JC. Operationalizing diagnostic criteria for Alzheimer's disease and other age-related cognitive impairment-Part 1. *Alzheimers Dement*. 7:15–34. [PubMed: 21255741]
31. Laska EM. The use of capture-recapture methods in public health. *Bull World Health Organ*. 2002; 80:845–845. [PubMed: 12481204]
32. Whitfield K, Kelly H. Using the two-source capture-recapture method to estimate the incidence of acute flaccid paralysis in Victoria, Australia. *Bull World Health Organ*. 2002; 80:846–851. [PubMed: 12481205]
33. Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. *J Clin Psychiatry*. 1987; 48:314–318. [PubMed: 3611032]
34. Brandt, J.; Benedict, R. Hopkins Verbal Learning Test - Revised, professional manual. Psychological Assessment Resources; Odessa, FL: 2001.
35. Benedict RHB, Schretlen D, Groninger L, Dobraski M, Shpritz B. Revision of the brief visuospatial memory test: Studies of normal performance, reliability, and validity. *Psychol Assess*. 1996; 8:145–153.
36. Wechsler, D. Wechsler Adult Intelligence Scale - Revised. Manual. Psychological Corporation; New York: 1981.
37. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression screening scale - a preliminary-report. *J Psychiatr Res*. 1983; 17:37–49. [PubMed: 7183759]
38. Squire LR, Wetzel CD, Slater PC. Memory complaint after electroconvulsive-therapy - assessment with a new self-rating instrument. *Biol Psychiatry*. 1979; 14:791–801. [PubMed: 497304]
39. Clark CM, Ewbank DC. Performance of the dementia severity rating scale: A caregiver questionnaire for rating severity in Alzheimer disease. *Alzheimer Dis Assoc Disord*. 1996; 10:31–39. [PubMed: 8919494]

40. Morris JC, Heyman A, Mohs RC, Hughes JP, Vanbelle G, Fillenbaum G, Mellits ED, Clark C. The Consortium to Establish a Registry for Alzheimers-Disease (CERAD) .1. Clinical and neuropsychological assessment of Alzheimers- disease. *Neurology*. 1989; 39:1159–1165. [PubMed: 2771064]
41. Benton, A. Revised visual retention test. 4th Edition. Psychological Corporation; New York: 1974.
42. Benton, A. Multilingual aphasia examination. University of Iowa Press; Iowa City: 1988.
43. Smith, A. Symbol digit modalities test - Manual. Western Psychological Services; Los Angeles: 1982.
44. Zachary, R. Shipley institute of living scale - revised. Western Psychological Services; Los Angeles: 1991.
45. Cummings J, Mega M, Gray K, Rosenberg-Thompson S, Carusi D, Gornbein J. The neuropsychiatric inventory: Comprehensive assessment of psychopathology in dementia. *Neurology*. 1994; 44:2308–2314. [PubMed: 7991117]
46. Morris J. The clinical dementia rating (CDR): Current version and scoring rules. *Neurology*. 1993; 43:2412–2414. [PubMed: 8232972]

Figure 1.

Data selection flow-chart. White boxes depict each stage of data selection and the number of participants who met each respective selection criteria; black boxes show the number of participants excluded at each stage.

Figure 2.

Number of participants assigned to each diagnostic group. The Venn diagrams depict the number of subjects whose DE was triggered by the CAB only (black), TAB only (light gray), or both (dark gray). Within diagnostic categories, the size of each circle is proportionate to the size of the corresponding circle, but circle sizes are not proportionate across diagnostic groups.

Table 1a
Neuropsychological tests and Cutoff Scores for the Cognitive Assessment Battery (CAB) and Telephone Assessment Battery (TAB)

Test	Battery		Trigger threshold	
	CAB	TAB	Measure	Cutoff score
Modified Mini Mental State Examination (3MS-E) [33]	x		Total	82
Hopkins Verbal Learning Test - Revised (HVLT-R) [34]	x		Trial 4	4
Brief Visuospatial Memory Test - Revised (BVM-T-R) [35]	x		Trial 4	4 (ages 70-79)
			Trial 4	1 (ages 80+)
River-mead Behavioral Memory Test (RBMT) [28]	x	x	Delay	3
Generative fluency (Supermarket items)	x	x	Total	15
WAIS-R Digit Span [36]	x		Fwd (total)	3
			Bwd (total)	2
Telephone Interview for Cognitive Status (TICS) [12]		x	Total	30
Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) [29]		x	Average	> 3.31
Geriatric Depression Scale, 30-item version (GDS-30) [37]	x	x	N/A	N/A
Self-rating of memory [38]	x		N/A	N/A
Dementia Severity Rating Scale (DSRS) [39]	x		N/A	N/A

Note: To trigger a DE referral on the CAB, a participant had to perform below cutoff on the 3MS-E or on at least two other tests, at least one of which had to be a test of delayed memory (RBMT, HVLT-R or BVM-T-R). A DE referral was triggered on the TAB by performance below cutoff on any of the battery's four tests. Approximately halfway through the screening process, however, an obvious problem with excessive numbers of false positives prompted the investigators to alter the TAB cut-off criteria. After this date, a DE could be triggered by the TICS or the RBMT, or by any combination of two tests.

Table 1b

Neuropsychological tests in the Dementia Evaluation (DE)

Mini-Mental State Examination (MMSE) [18]
CERAD [40] Boston Naming Test - 15 item (BNT-15)
CERAD [40] Word list
CERAD [40] constructions
WMS-R [41] Logical Memory I & II
Benton Revised Visual Retention Test (RVRT) [41]
Controlled Oral Word Test (COWAT) [42]
Symbol Digit Modalities Test (SDMT) [43]
Shipley Vocabulary [44]
Neuropsychiatric Inventory (NPI) [45]
Clinical Dementia Rating Scale (CDR) [46]
Dementia Severity Rating Scale (DSRS) [39]

Table 2

Sample of the capture-recapture method

		TAB (recapture)		
		Yes	No	TOTAL
CAB (capture)	Yes	36	10	46
	No	41	x	?
TOTAL		77	?	N

Note. This example uses data for individuals ultimately assigned to the Impaired Cognition group. One unknown value (“x”) represents the number of participants who did not trigger a DE on either the CAB or TAB but who, had they completed a DE, would have been classified as cases. This value can be estimated by multiplying the number of unique cases detected by each battery (41×10), and dividing that value by the number of cases detected by both methods (36). The resulting value ($x = 11.39$) can be used to estimate the total number of individuals in the entire study sample who would meet criteria for the Impaired Cognition classification ($N = 36 + 10 + 41 + 11.39 = 98.39$).

Table 3

Demographic characteristics by final diagnosis.

	Normal Cognition			Dementia			MCI/pAD			Other CIND		
	CAB	TAB	TOTAL	CAB	TAB	TOTAL	CAB	TAB	TOTAL	CAB	TAB	TOTAL
N	179	148	189	11	9	11	28	46	51	7	22	25
Age (M(SD))	78.59(4.02)	78.74(4.03)	78.46(3.97)	79.64(4.32)	80.33(4.44)	79.64(4.32)	78.57(4.48)	78.15(4.29)	77.96(4.17)	76.43(1.72)	77.59(3.28)	77.40(3.12)
Sex (% Men)	54.20%	52.00%	53.40%	72.70%	88.90%	72.70%	60.70%	63.00%	62.70%	42.90%	50.00%	48.00%
Race (% Caucasian)	97.20%	96.60%	97.40%	90.90%	88.90%	90.90%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Education (M(SD))	14.69(2.65)	14.64(2.61)	14.68(2.64)	15.27(2.45)	15.67(2.45)	15.27(2.45)	14.68(3.14)	14.89(2.96)	15.04(2.91)	13.29(2.36)	14.05(2.79)	13.80(2.69)
Marital status (% Married)	67.60%	70.30%	67.70%	90.90%	100.00%	90.90%	78.60%	69.60%	70.60%	85.70%	63.60%	64.00%

† “CAB” and “TAB” subheadings denote participants who triggered a DE on each screening battery and were subsequently assigned to each respective diagnostic group.

* “TOTAL” refers to participants assigned to each respective diagnostic group, regardless of whether the DE was triggered by the CAB, TAB, or both.

Table 4

Performance characteristics of the CAB, TAB, and TICS

	Sensitivity			Specificity			PPV			NPV		
	CAB	TAB	TICS	CAB	TAB	TICS	CAB	TAB	TICS	CAB	TAB	TICS
Impaired Cognition (N=87)	0.47	0.78	0.33	0.98	0.87	0.97	0.61	0.39	0.52	0.94	0.91	0.97
Dementia (N=11)	1.00	0.82	0.73	0.94	0.82	0.95	0.14	0.05	0.13	0.97	0.89	0.95
MCI/pAD (N=51)	0.50	0.82	0.29	0.96	0.85	0.96	0.37	0.23	0.26	0.95	0.90	0.96
Other CIND (N=25)	0.18	0.57	0.21	0.94	0.82	0.95	0.09	0.11	0.13	0.94	0.88	0.95