Published in final edited form as:

Alzheimer Dis Assoc Disord. 2013 ; 27(4): 356–362. doi:10.1097/WAD.0b013e3182769c05.

Developing Dementia Prevention Trials: Baseline Report of the Home-Based Assessment Study

Mary Sano, PhD^{*}, Susan Egelko, PhD^{*}, Michael Donohue, PhD^{}, Steven Ferris, PhD[†], Jeffrey Kaye, MD^{‡,§}, Tamara L. Hayes, PhD[‡], James C. Mundt, PhD^{†§}, C.K. Sun, MS^{**}, Silvia Paparello, PhD^{**}, Paul S. Aisen, MD^{**}, and the Alzheimer Disease Cooperative Study Investigators**

^{*}Mount Sinai School of Medicine, New York, NY and the James J Peters VAMC, Bronx, NY

^{**}UCSD, La Jolla, CA

[†]New York University Langone Medical Center, New York, NY

[‡]Oregon Health & Science University, Portland, OR

[§]Portland VAMC, Portland, OR

^{†§}Center for Psychological Consultation, Madison, WI & University of Arkansas for Medical Sciences, Little Rock, AR

Abstract

This report describes the baseline experience of the multi-center, Home Based Assessment (HBA) study, designed to develop methods for dementia prevention trials using novel technologies for test administration and data collection. Non-demented individuals 75 years old were recruited and evaluated in-person using established clinical trial outcomes of cognition and function, and randomized to one of 3 assessment methodologies: 1) mail-in questionnaire/live telephone interviews (MIP); 2) automated telephone with interactive voice recognition (IVR); and 3) internet-based computer Kiosk (KIO). Brief versions of cognitive and non-cognitive outcomes, were adapted to each methodology and administered at baseline and repeatedly over a 4-year period. “Efficiency” measures assessed the time from screening to baseline, and staff time required for each methodology. 713 individuals signed consent and were screened; 640 met eligibility and were randomized to one of 3 assessment arms and 581 completed baseline. Drop out, time from screening to baseline and total staff time were highest among those assigned to KIO. However efficiency measures were driven by non-recurring start-up activities suggesting that differences may be mitigated over a long trial. Performance among HBA instruments collected via different technologies will be compared to established outcomes over this 4 year study.

Keywords

Alzheimer’s disease; clinical trials; in-home assessment; prevention studies

INTRODUCTION

Prevention trials for dementia and cognitive loss of aging will require effective, efficient, and economical methods of assessment. Traditional in-person visits to clinical assessment

sites are time consuming and costly. Most importantly, they exclude from participation some of the very cohorts at greatest risk for decline, such as those with extreme age, medical illnesses and immobility. Furthermore, current trials for disease prevention and development of diagnostics have required informants and have recruited convenience samples of highly educated non-diverse populations, despite evidence that black, Latino and those living alone without an informant are most likely to experience costly decline¹. The Home Based Assessment (HBA) study was designed to overcome these barriers by assessing elders in their home using a range of technologies for test administration and data collection. Over the past 10 years the Alzheimer Disease Cooperative Study (ADCS) has identified domains that are critical in the transition from cognitive health to dementia². They include cognition³, function⁴, behavior⁵, global clinical status⁶, quality of life⁷, and resource use⁸. In addition, we have begun to develop a screen for self-reported cognitive change. While early work developed comprehensive inventories that could be used across the span of cognitive status, this current protocol used brief versions selecting items with high sensitivity to change in a normal or mildly cognitively impaired elderly population. Further, we developed a measure of medication adherence as a performance-based “activity of daily living.”

In a pilot study⁹ we demonstrated the feasibility of assessing individuals in their home using established and new technologies. Specifically, participant elders were randomly assigned to be assessed using 1) telephone and paper forms that were mailed back to the study site (MIP), 2) an interactive voice response (IVR) system that used computer-automated telephone assessment, or 3) an Internet-enabled computer kiosk (KIO) that were installed in the participant’s home. Results from this pilot study established that these technologies were feasible for use with community dwelling elders. Here we describe the initiation and baseline characteristics of the cohort participating in a national, randomized trial to evaluate these assessment methods over four years.

METHODS

Site Selection

Recruitment was conducted from 28 sites located in metropolitan areas of the United States and selected by the ADCS based on completion of a site survey. Selection required confirmation of site access to community dwelling elders, typically living in concentrated areas including organized independent living facilities and informal neighborhoods with high incidence of elderly residence. In addition, the sites had to identify an Internet service provider serving the local community, and demonstrate staff with expertise in clinical trial recruitment who were prepared to conduct the full set-up of the protocol in the home.

Participants

Community dwelling elders signed written informed consent in accordance with local IRB standards and were screened to meet the study criteria. Inclusion criteria were: age 75 years or older; Mini-Mental State Examination of 26; or greater¹⁰; and independently living. (A study partner was desirable but not required). Exclusion criteria were: dementia; use of prescriptive cognitive-enhancing medication (specifically cholinesterase inhibitors and memantine) at screening; unwillingness to use the study provided multi-vitamin (MVI); other major medical conditions which cause specific cognitive impairment (neurological conditions including stroke, Parkinson’s disease, or active major psychiatric illness); and life expectancy of less than 5 years. At each study site, one in five enrollees was required to be a member of the minority community; accession was monitored and enrollment at a site was suspended until this ratio was achieved.

Procedures/Timeline

An in-person *screening visit* was conducted to determine eligibility. It consisted of a medical examination, a neurological exam with specific questions about memory complaint, and a neuropsychological battery taken from the Uniform Data Set (UDS) of the National Alzheimer Coordinating Center (NACC)¹¹. The tests in the neuropsychological battery included: Logical Memory, Immediate and Delayed; Digit Span: Forward and Backward; Category Fluency: Animal and Vegetable; Trail Making Test: Parts A and B; Digit Symbol Substitution; and Boston Naming Test. In addition, a 24-item ADCS ADL-MCI was administered¹². The clinician used this assessment battery to exclude those with dementia and categorized eligible participants as normal or Mild Cognitive Impairment (MCI) based on evidence of memory impairment from interview and available neuropsychological evaluation. An algorithmic categorization of MCI (vs. normal) was made centrally, based on education-adjusted Logical Memory delayed recall scores¹³. Blood was also collected for DNA extraction and apolipoprotein E genotyping.

Randomization to HBA Arms—Participants were randomized to one of three HBA arms. Two frequencies of assessment were nested within each arm. One of the two frequencies common to all arms was quarterly assessment. The second assessment frequency was set at annual visits for IVR and MIP, an interval commonly used in prevention trials. The second frequency for KIO was set at monthly visits, an interval compatible with the automated technology and possibly capable of capturing change in cognitive status at the earliest stage. (1) Mail-in/phone (MIP): Cognitive assessments were conducted by a trained evaluator during in-person telephone calls with the participants. Non-cognitive assessment and the experimental medication adherence procedures were conducted by mail-back paper forms. The telephone interactions were initiated by the evaluator contacting the participants at pre-scheduled times. The mail-in procedures were initiated by site mailings to the participants, who were instructed to provide return responses using pre-addressed mailers. Participants in the MIP arm were randomized to be assessed annually or quarterly during the study follow-up period. (2) Interactive Voice Response (IVR): Assessments in this arm were completed using a computer-automated telephone interface¹⁴, requiring no live staff time. A standard, large-key telephone was installed in the participant's home; the toll-free telephone number to access the HBA IVR assessment system and a unique participant identification number were programmed into telephone memory system. All cognitive, non-cognitive, and medication adherence assessments were administered through a speech-enabled, automated telephone interface. Participant responses were obtained and scored using automated speech recognition technology and/or touch-tone keypad entry. Visits were initiated by the participant calling in to the toll-free number at prescheduled times. Study staff were instructed to prompt participants to call if they missed a scheduled calling time. IVR participants were randomized to be assessed annually or quarterly. (3) Kiosk (KIO). A web-based computerized assessment, consisting of a computer kiosk (a touch-screen sensitive flat-panel monitor) with an attached telephone handset for recording verbal responses, was installed in the home and connected to the Internet via broadband. This typically required a staff member to attend the installation. All KIO participants required installation of internet access and the grant covered this expense, as well as the expense of ongoing access for the period of the grant. Cognitive and non-cognitive assessments were collected via the Internet requiring no live staff time. The visit was announced several days in advance on the KIO screen and then on the day of assessment initiated by a flashing screen telling the participant to begin the assessment. Participants were guided through the assessment by an intelligent on-screen pre-recorded video assistant. Medication adherence was measured via a separate MedTracker device¹⁵, with compartments for a week's supply of vitamins, which the participant was trained to fill weekly. The device recorded date and time of all compartment openings. The KIO

participants were assigned to either quarterly or monthly assessments. If the assessment was not completed on time, study staff called the participant as a reminder.

Training visit—The training visit took place in the participant's home for the IVR and KIO arms and either by phone or at the participant's home for the MIP arm. The training visit consisted of a review of the assessment procedures and a mock demonstration of test-taking by the participant. The baseline experimental assessment was scheduled within the next week.

Baseline visit—For those in the MIP arm, the cognitive battery was conducted by phone with a live tester; the tester reminded the participant to complete the paper-and-pencil non-cognitive instruments and return them as described above. For the other two arms, appointments were scheduled via their respective automated technologies. If the visit was not completed as scheduled, staff contacted the participant. No staff time was required for either the cognitive or non-cognitive assessment battery for IVR and KIO arms. However, the staff effort to reschedule or provide any assistance to complete the visit was captured on the Efficiency Form, as described below.

Follow-up visits—Each of the follow-up visits included a repeat of the baseline visit assessment tools, with the addition of medication adherence measurement. The results of these ongoing follow-up visits will be analyzed and reported following study completion.

Experimental Assessments – Cognitive

The cognitive portion of the in-home evaluation included the following tests: 1) Immediate Word List Recall (CERAD); 2) East Boston Memory Test- Immediate Recall; 3) Abbreviated TICS (Modified Telephone Interview for Cognitive Status; 4) Backward Digit Span (WAIS-R); 5) Adapted Trail Making Test: A and B; Adverse Event Checklist (filler test); 6) Delayed Word Recall; 7) Category Fluency (Animals); and 8) East Boston Memory Test-Delayed Recall. The scoring details have been previously published⁸. The order of the tests was designed to provide a delay of 15 to 20 minutes between the immediate learning and delayed recall for the Word List and East Boston Memory tasks and to avoid interference from Category Fluency. The expected overall time to administer the experimental cognitive performance battery was 20 to 25 minutes for all 3 arms. Only the MIP arm required staff time to collect data and this was recorded on the Efficiency Form. Each of the cognitive tests was preserved as much as possible to allow delivery by different technologies (i.e., live telephone interactions, IVR automated telephone interactions, and KIO computer-based audiovisual interactions) yet maintain enough features to enable performance comparisons to the standard face-to-face measures and across arms. The adaptations of Trail Making Test A and B to each of the in-home formats preserved the elements of following a sequence mentally and dealing with more than one stimulus or thought at a time; however, the original feature of visuomotor tracking could not be preserved, particularly in the IVR and MIP conditions. In the KIO arm, the Trails tests were scored based on the participant connecting the appropriate numbers and letters as traced with their finger on the touch screen. In light of these fundamental differences in adaptation, no inferential statistics were applied to compare Adapted Trail Making Test across groups.

Experimental Assessments -Non-cognitive

The non-cognitive portion of the home-based experimental assessment was completed by mail for the MIP arm, by automated telephone assessment (keypad and orally) for the IVR arm and by automated computer kiosk (orally and touch screen) interaction for the KIO arm. The expected time to complete the non-cognitive assessments was 20 minutes. The non-cognitive measures⁸ included: 1) Brief Cognitive Function Screening Instrument (BCFSI);

2) Quality of Life (QOL); 3) Behavioral Scale; 4) Instrumental Activities of Daily Living (IADL); 5) Clinical Global Impression of Change (CGIC); 6) Resource Use Inventory (RUI); 7) Participant Status; and 8) Performance-Based Medication Adherence¹⁴. Participants were instructed to take a study-provided MVI in the morning and evening each day. Adherence was measured according to the format of the arm and will be described in other reports.

As with the cognitive tests there were several differences in presentation of non-cognitive items across the three arms. In particular for the IADL and the BCFSI the IVR provides oral presentation only while MIP and KIO display all response choices in writing. Also the order of response option was adjusted to for each methodology.

Efficiency Measures

Two measures of efficiency were used to evaluate the methodologies: (1) the number of days from screening to baseline and (2) the amount of staff contact required to complete the experimental assessments, as systematically coded on The Efficiency Form. The frequency of contact and length of time spent with the participants in person or by phone was recorded along with the reason for the contact. These were categorized as follows: Training time (the time to familiarize a participant with tools and teach how to use the equipment and forms); Preparation time (the time in-home to set up the modality and answer questions), Time at baseline (all other non-scheduled contacts, including assurance to participants about the protocol or to remedy equipment problems, and address missing forms), and Testing time (administering the cognitive battery, for the MIP arm only). The Total Time (in minutes) was the sum of these times.

Trigger for In-person Evaluation During Longitudinal Follow-up

At each follow-up visit, three experimental test scores are evaluated to determine if sufficient worsening had occurred to warrant an in-person evaluation which would determine if the participant had indeed progressed to MCI or dementia. These “trigger” tests are Delayed Word Recall, BCFSI and IADL scale, the latter test score converted to a percentage, with 100% indicating best performance. The algorithm for trigger requires worsening in one or more of these tests and is specific to the initial participant categorization as either MCI or normal, and assigned frequency of assessment.

Statistical Analyses

The key demographic, clinical, and neuropsychological variables that were collected on the full cohort of participants in person at baseline were summarized by descriptive statistics. The descriptors for variables that were measured on a continuous scale were: means, standard deviations, and ranges of scores. Categorical variables (i.e., frequency counts) were summarized by percentages.

Statistical comparisons at baseline were conducted across arms on a number of continuous measures using one-way ANOVA's. These measures were: the three trigger experimental variables (Word List: Delayed recall, BCFSI, and IADL) and all efficiency variables: time from screening to baseline (in days), and the time that staff spent with the participant up to and including the first (baseline) experimental “visit.” If Bartlett's test of homogeneity of variances was non-significant ($P > .05$) for a given variable and the overall F test was significant, a Tukey post-hoc analysis was conducted. If Bartlett's test was significant ($P < .05$), Welch's t-tests¹⁶ were conducted to accommodate unequal variances, with Hochberg¹⁷ adjustments for multiple comparisons (i.e., pairwise comparisons = MIP:IVR, MIP:KIO, IVR:KIO). The number of times each subject was contacted by site personnel at baseline

was modeled as Poisson. All tests were 2-tailed with α set at .05. Statistical analyses were conducted with the R (version 2.14)¹⁸

RESULTS

A total of 713 community-dwelling seniors were screened (see Figure 1), of which 73 were not eligible. There were no significant differences in age, gender, race, or ethnicity between the group of 640 seniors who met eligibility requirements compared to the group who did not. However, the 640 who were eligible were better educated ($t[1,708] = 2.99$; $P = .004$), with a mean number of years of education of 15.7 (SD = 2.8), compared with a mean of 14.6 (SD = 2.9) for the 73 persons who failed screening.

Among the 28 participating sites a mean of 22.9 (SD = 8.6) subjects were randomized per site (range 5 to 40). Of the 640 subjects who were randomized to one of the 3 study arms, 59 subjects (9.2%) chose to discontinue prior to the baseline evaluation. There were no significant differences in age, education, gender, minority status, ApoE4 status, history of cardiovascular disease and hypertension, ADL-MCI scores, presence of MCI, or screening MMSE scores between those who did and did not continue participation. Likewise, within each of the three arms there were no significant differences in any of these variables between those who discontinued prior to baseline and those who completed baseline evaluation. However, there was a significant difference across arms in the dropout rate prior to baseline evaluation. As noted in Figure 1, the rate of dropout was 17% for KIO, 8% for IVR, and 2% for MIP (Fisher Exact test $P < .001$). In post-hoc pairwise Fisher tests the dropout rate for those participants assigned to the KIO arm was significantly higher than for the MIP ($P < .001$) and IVR ($P = .009$); the dropout for the IVR arm was significantly higher than the MIP arm ($P = .003$). Each of the three arms had two frequency conditions and higher frequency was associated with greater dropout for both high-technology arms even before baseline. IVR participants assigned to annual evaluation had a dropout rate of 7% compared with 10% for quarterly; KIO participants assigned to quarterly evaluation had a dropout rate of 15% compared with 20% for monthly; MIP, on the other hand, had a dropout rate of 0% for quarterly and a 4% for annual evaluation. (Fisher's Exact Test $P < .001$). Reasons for dropout were not categorized but recorded comments indicated inconvenience of equipment (e.g., lack of apartment space for the KIO) and too much of a time commitment to participate.

Table 1 provides demographic clinical and neuropsychological characteristics at baseline. The mean age at baseline was 81.0 years, ranging up to 98 years old; 74% had cardiovascular disease, 19% had MCI at baseline, 25% had one or more APOE 4 allele, and 28% admitted to a current memory problem by self-report.

The baseline scores on "trigger" tests are included in Table 2, with statistical comparisons across arms. There was no significant difference across arms on the Word List: Delayed recall, and the distribution was similar for both mean and standard deviation values in each of the arms and scores spanned the whole range (0-10). The mean BCFSI scores were different among groups [$F=13.4$, $df=2,576$, $P < .001$] with lower scores (i.e., fewer self-reported functional symptoms) in the MIP than both KIO and IVR groups. The variance of the IADL exhibited significant heterogeneity across groups, so pairwise comparisons were analyzed by Welch t -tests with Hochberg adjustment for multiple comparisons. The IADL score was the highest (i.e., least impaired) in the MIP arm and significantly above the KIO and IVR arms; the KIO scores were significantly higher than the IVR scores.

Table 3 presents efficiency measures. The number of days from screening to baseline was greater in the KIO arm (55.7 days $\pm$ 42.3), than in IVR 39.2 ($\pm$ 25.8) and in MIP 33.5 ($\pm$

25.5), which were not different from each other. The Total time (Table 3, last row) was longer in the KIO arm (280 minutes $\pm$ 314.5), compared with the MIP (78.3 $\pm$ 36.0) and the IVR (72.4 $\pm$ 38.0). The mean Testing time (33.8) was imputed for 27 MIP subjects with missing Testing times. The number of contacts ranged from 1 to 41 for KIO, 0 to 6 for IVR, and 1 to 11 for MIP. The distributions were highly skewed toward 0 with a median of 2 for all three arms. The difference in the rate of contact across the three arms was significant by the Poisson model likelihood ratio test ($P < .001$). The Poisson model estimated rate of contact was 2.01 ($P < .001$) in KIO, 1.22 ($P = .002$) in IVR, and 1.40 ($P < .001$) in MIP. Only the difference between IVR and MIP ($P = .112$) was not significant at the .001 level by the pairwise test with Hochberg adjustment.

DISCUSSION

This study examines home-based methods that might be used in dementia prevention trials. While those who failed the screening process had less education than those who were eligible, the HBA study successfully enrolled a diverse cohort of elders with more than 20% minority participants and a mean age in the eighth decade. The majority of the sample was female. Cardiovascular disease and hypertension were common, as expected in a group of this age. These clinical and demographic characteristics are associated with the risk of dementia, supporting the notion that we captured the population who would likely be enrolled in prevention studies.

Randomized assignment to the different assessment methods resulted in similar demographic and clinical features among the groups, suggesting that acceptability of these modalities in this age group is not biased by health or cultural variables. However, a higher dropout rate was associated with higher technology and more frequent assessments. This is an important observation because while more frequent assessment may yield more stable measures, this may come at a cost of greater attrition.

We selected both cognitive and functional outcomes to capture change that would mark the transition from “no dementia” to cognitive impairment or dementia. The cognitive outcome (Delayed recall) was comparable across arms despite very different methods of administration and scoring. Unexpectedly, the functional measure scores, captured by the questionnaires, were different across groups. Though the differences were small and all measures showed no evidence of floor effect, the level of impairment in scores for both IADL and BCFSI was least in the MIP group with KIO and IVR demonstrating more impairment. Of note, this pattern was also seen in the pilot study⁹ and may have been due to differences in format across the arms (paper-and-pencil for MIP vs. interactive questions by the automated examiner in KIO and IVR). These differences will be assessed further in the longitudinal phase of this study, along with the empirical question of differential sensitivity to worsening across arms.

The efficiency measures used in the study are a novel approach to assessing the feasibility of use of technology in clinical trials. We found longer times to study initiation (days to baseline) and greater staff time associated with the KIO. Staff expenditure was not quantified for dropouts prior to baseline; if these data had been available and included in the efficiency measures (i.e., efficiency as a function both of time to recruit and to assess the cohort) the KIO arm would have fared even worse. Going forward staff time will continue to be required for cognitive assessment in MIP for all follow-up study visits while staff time for KIO and IVR will not be required. It is possible that this advantage of automated assessment could mitigate the startup efficiency differences or even change the order of relative efficiency over the course of this 4-year trial, as was suggested in our HBA pilot

study which included a one-month follow-up². However, this is a matter of empirical investigation for a later report.

Several of the observations made here may be very specific to this study. For example, the computerized version of our evaluation (KIO) is a self-administered procedure, quite different from tester-assisted computerized testing; the time to train and set up the procedures may not reflect the experience of in-clinic computer testing. Nevertheless, the higher dropout for those assigned to the KIO arm even before the experimental procedures were in place suggests that computer interaction may not be welcome for this age cohort. Reasons for discontinuation once the baseline procedures had begun, particularly from the KIO and IVR, often reflected challenges of technology. It may also be expected that in the future many of these technological challenges may be lessened by improvements, such as smaller footprints using tablet computers as well as increased familiarity with technology for future cohorts of elders. We also acknowledge that while in-home assessment may allow us to expand the range of elders who might participate in research, it may be less desirable to some who would prefer the experience of a clinic visit with face-to-face staff interaction.

In summary, this work demonstrates the feasibility of recruiting and evaluating a community cohort and of conducting the assessments at home using a range of technology. This study also quantifies novel measures of efficiency. The longitudinal aspect of the study will allow us to assess the sensitivity of home-based assessment methods to capture the earliest transitions to cognitive impairment and dementia.

Acknowledgments

This work was supported by the following NIA grants: U01AG10483, P50AG005138, P30AG008051, and P30AG024978. Development of the Kiosk and MedTracker was supported in part by grants from NIA (P30-AG024978; P30-AG08017) and Intel Corporation.

The authors are grateful to the investigators of the participating sites of the Alzheimer's Disease Cooperative Study: Jeffrey Kaye at Oregon Health & Science University; Mary Pay at University of California, San Diego; Bruno Giordani at University of Michigan, Ann Arbor; Daniel Marson at University of Alabama, Birmingham; Jane Martin at Mount Sinai School of Medicine; Raj Shah at Rush University Medical Center; Ranjan Duara at Wien Center for Clinical Research; Constantine Lyketsos at Johns Hopkins University; Amanda Smith at University of South Florida, Tampa; Steven Ferris at New York University Medical Center; Jason Karlawish at University of Pennsylvania; Allison Caban-Holt at University of Kentucky; Ruth Mulnard at University of California, Irvine; Neill Graff-Radford at Mayo Clinic, Jacksonville; Martin Farlow at Indiana University; Christopher van Dyck at Yale University School of Medicine; Diana Kerwin at Northwestern University; Brigid Reynolds at Georgetown University; Marwan Sabbagh at Banner Sun Health Research Institute, Robert Stern at Boston University; Kathleen Smyth at Case Western Reserve University; John Olichney at University of California, Davis; Smita Kittur at Neurological Care of CNY; Douglas Scharre at Ohio State University; Kaycee Sink at Wake Forest University Health Sciences; J. Wesson Ashford at Stanford University; Richard King at University of Utah Center for Alzheimer's Care, Imaging and Research; and Charles Bernick at Cleveland Clinic Lou Ruvo Center for Brain Health.

Sincere thanks as well to Devon Gessert at UCSD for technical support on this manuscript. The authors would also like to acknowledge the contributions of Tracy Reyes and Ben Barth at Healthcare Technology Systems; Jacques H. de Villiers, Rachel Coulston, Esther Klabbers, John Paul Hosom, and Thomas Riley at the OHSU Center for Spoken Language and Understanding; Jessica Payne-Murphy at the OHSU Oregon Center for Aging & Technology; and Karen Stokes and Danielle Whitehair at UCSD.

References

1. Yaffe K, Fox P, Newcomer R, et al. Patient and caregiver characteristics and nursing home placement in patients with dementia. *JAMA*. 2002; 287(16):2090–7. [PubMed: 11966383]
2. Ferris SH, Aisen PS, Cummings J, et al. Alzheimer's Disease Cooperative Study Group. ADCS Prevention Instrument Project: overview and initial results. *Alzheimer Dis Assoc Disord*. 2006; 20(4 Suppl 3):S109–23. [PubMed: 17135805]

3. Salmon DP, Cummings JL, Jin S, et al. ADCS Prevention Instrument Project: Development of a brief verbal memory test for primary prevention clinical trials. *Alzheimer Dis Assoc Disord.* 2006; 20(4 Suppl 3):S139–46. [PubMed: 17135807]
4. Galasko D, Bennett DA, Sano M, et al. ADCS Prevention Instrument Project: assessment of instrumental activities of daily living for community-dwelling elderly individuals in dementia prevention clinical trials. *Alzheimer Dis Assoc Disord.* 2006; 20(4 Suppl 3):S152–69. [PubMed: 17135809]
5. Cummings JL, Raman R, Ernstrom K, et al. Alzheimer's Disease Cooperative Study Group. ADCS Prevention Instrument Project: behavioral measures in primary prevention trials. *Alzheimer Dis Assoc Disord.* 2006; 20:S147–51. [PubMed: 17135808]
6. Schneider LS, Clark CM, Doody R, et al. ADCS Prevention Instrument Project: ADCS-clinicians' global impression of change scales (ADCS-CGIC), self-rated and study partner-rated versions. *Alzheimer Dis Assoc Disord.* 2006; 20(4 Suppl 3):S124–38. [PubMed: 17135806]
7. Patterson MB, Whitehouse PJ, Edland SD, et al. ADCS Prevention Instrument Project: quality of life assessment (QOL). *Alzheimer Dis Assoc Disord.* 2006; 20(4 Suppl 3):S179–90. [PubMed: 17135811]
8. Sano M, Zhu CW, Whitehouse PJ, et al. Alzheimer Disease Cooperative Study Group. ADCS Prevention Instrument Project: pharmacoeconomics: assessing health-related resource use among healthy elderly. *Alzheimer Dis Assoc Disord.* 2006; 20(4 Suppl 3):S191–202. [PubMed: 17135812]
9. Sano MS, Egelko S, Ferris S, et al. Pilot Study to Demonstrate the Feasibility of a Multi-Center Trial of Home-Based Assessment of People Over 75 Years Old. *Alzheimer Dis Assoc Disord.* 2010; 24(3):256–63. [PubMed: 20592583]
10. Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975; 12:189–198. [PubMed: 1202204]
11. Morris JC, Weintraub S, Chui HC, et al. The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. *Alzheimer Dis Assoc Disord.* 2006; 20:210–216. [PubMed: 17132964]
12. Pedrosa H, De Sa A, Guerreiro M, et al. Functional evaluation distinguishes MCI patients from healthy elderly people--the ADCS/MCI/ADL scale. *J Nutr Health Aging.* 2010; 14:703–9. [PubMed: 20922349]
13. Mueller SG, Weiner MW, Thal LJ, et al. Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement.* 2005 Jul; 1(1): 55–66. [PubMed: 17476317]
14. Mundt JC, Geralt DS, Moore HK. Dial "T" for Testing: Technological Flexibility in Neuropsychological Assessment. *Telemed J E Health.* 2006; 12(3):317–323. [PubMed: 16796499]
15. Hayes, TL.; Hunt, JM.; Adami, AM., et al. An electronic pillbox for continuous monitoring of medication adherence. *Conference Proceedings : Engineering in Medicine and Biology Society;* 2006. p. 6400-3.
16. Welch BL. The generalization of 'student's' problem when several different population variances are involved. *Biometrika.* 1947; 34:28–35. [PubMed: 20287819]
17. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika.* 1988; 75:800–803.
18. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: URL 2011. <http://www.R-project.org/>

Figure 1.
Flow of Participants from Screening to Baseline

Table 1

Demographic, Clinical, and Neuropsychological Characteristics at Baseline In-Person Assessment: Full Cohort with Three Arms Combined

Characteristics at In-Person Baseline Assessment	Value* Full Sample N =581	Range
Age, in years	81.0 ± 4.4	75-98
Education, in years	15.6 ± 2.9	4-20
MMSE [†]	28.8 ± 1.2	26-30
Logical Memory: Immediate recall, max=25	12.4 ± 4.3	1-23
Logical Memory: Delayed recall, max=25	11.0 ± 4.7	0-25
Digit Span Backward: Length, max=7	4.8 ± 1.2	2-7
Category Naming: Animals	18.0 ± 5.0	6-39
Trail Making Test: Part A Time to Complete in sec. cutoff time=150"	42.3 ± 17.1	13-150
Trail Making Test: Part B Time to Complete in sec. cutoff time=300"	108.2 ± 53.9	30-300
Digit Symbol, max=93	40.0 ± 9.8	8-77
Boston Naming Test, max=30	26.4 ± 3.4	8-30
% Female	67	
% Racial/ethnic minority	22	
% Married	42	
% History of hypertension	59	
% Cardiovascular disease	74	
% Self-report current memory problem	28	
% Self-report current memory problem is a change from before	22	
% Self-report current memory problem worse than age peers	3	
% MCI, by assessment	19	
% APOE abnormality N = 471	24	

* Mean ± SD.

[†]The study population is defined by MMSE ≥ 26.

Table 2

Experimental home-based test performance: cognitive and functional measures used as “triggers” for full evaluation

“Trigger” Variables + Completion Score	MIP (N =207)	IVR (N =196)	KIO (N =178)	Overall <i>P</i> Post-hoc comparison <i>P</i> 's
Word List: Delayed recall, mean ± SD (max=10)	5.53 ± 2.4	5.48 ± 2.2	5.19 ± 2.5	=0.32 ^a (<i>ns</i>)
BCFSI, mean ± SD (max=8) ^b	1.13 ± 1.4	1.72 ± 1.6	1.89 ± 1.5	<.001 ^a KIO:IVR=.503 (<i>ns</i>) MIP:IVR<.001 MIP:KIO<.001
% IADL, mean ± SD (max=100%)	95.07 ± 8.3	89.78± 10.2	92.73 ± 8.4	<.001 ^c KIO:IVR=.007 MIP:IVR<.001 MIP:KIO<.011

BCFSI: Brief Cognitive Function Self-Inventory

IADL: Instrumental Activities of Daily Living

^aBased on 1-way ANOVA and Tukey multiple comparisons post-hoc analyses.

^bHigher scores indicate greater impairment.

^cBartlett test for homogeneity of variance was significant ($P < .001$) in the 1-way ANOVA. Therefore, *P* values were based on Welch *t*-tests with unequal variances and Hochberg adjusted pairwise comparisons.

Table 3

Efficiency Measures Across Three In-Home Assessment Arms

Efficiency Measures	MIP (N =207)	IVR (N =196)	KIO (N =178)	Overall <i>P</i> /Post-hoc comparison <i>P</i> 's *
Days to Baseline, mean $\pm$ SD	33.5 $\pm$ 25.5	39.2 $\pm$ 25.8	55.7 $\pm$ 42.3	<.001 KIO:MIP<.001 KIO:IVR <.001 IVR:MIP=.075 (<i>ns</i>)
Training Time (min), mean $\pm$ SD	25.6 $\pm$ 15.2	39.3 $\pm$ 20.5	76.7 $\pm$ 60.1	<.001 KIO:MIP<.001 KIO:IVR<.001 IVR:MIP<.001
Preparation Time (min), mean $\pm$ SD	16.0 $\pm$ 17.1	21.8 $\pm$ 20.8	130.3 $\pm$ 140.4	< .001 KIO:MIP<.001 KIO:IVR<.001 IVR:MIP<.519 (<i>ns</i>)
Time at Baseline, (other) (min), mean $\pm$ SD	3.2 $\pm$ 11.4	11.4 $\pm$ 17.2	73.5 $\pm$ 235.6	< .001 KIO:MIP<.001 KIO:IVR<.001 IVR:MIP<.534 (<i>ns</i>)
Testing Time (min), mean $\pm$ SD	33.4 $\pm$ 16.8	N/A	N/A	---
TOTAL TIME (#1-4 in min), mean $\pm$ SD	77.9 $\pm$ 35.9	72.4 $\pm$ 38.0	280.0 $\pm$ 314.5	<.001 KIO:MIP<.001 KIO:IVR<.001 IVR:MIP=.755(<i>ns</i>)

* Bartlett test for homogeneity of variance was significant ($P < .001$) for all time efficiency variables in 1-way ANOVA's. Therefore, *P* values were based instead on Welch *t*-tests with unequal variances and Hochberg adjustment for multiple comparisons.