



Published in final edited form as:

*Neurobiol Aging*. 2011 July ; 32(7): 1207–1218. doi:10.1016/j.neurobiolaging.2009.07.002.

## Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI

Susan M. Landau<sup>a</sup>, Danielle Harvey<sup>b</sup>, Cindee M Madison<sup>a</sup>, Robert A. Koeppe<sup>c</sup>, Eric M. Reiman<sup>d</sup>, Norman L. Foster<sup>e</sup>, Michael W. Weiner<sup>f</sup>, and William J. Jagust<sup>a,g,\*</sup> on behalf of and the Alzheimer's Disease Neuroimaging Initiative

<sup>a</sup>Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720-3190

<sup>b</sup>School of Medicine, Room 152, MS 1-C, One Shields Ave, University of California, Davis, CA 95616

<sup>c</sup>University of Michigan Medical School, 1500 E. Medical Center Drive, Ann Arbor, MI 48109

<sup>d</sup>Banner Alzheimer's Institute, 901 East Willetta Street, Phoenix, AZ 85006

<sup>e</sup>Department of Neurology, University of Utah, 650 Kommas Drive, #106-A, Salt Lake City UT 84108

<sup>f</sup>San Francisco Veteran's Administration Hospital, 4150 Clement St., San Francisco, CA 94121

<sup>g</sup>Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720

### Abstract

The Functional Activities Questionnaire (FAQ) and Alzheimer's Disease Assessment Scale – cognitive subscale (ADAS-cog) are frequently-used indices of cognitive decline in Alzheimer's disease (AD). The goal of this study was to compare FDG-PET and clinical measurements in a large sample of elderly subjects with memory disturbance. We examined relationships between glucose metabolism in FDG-PET regions of interest (FDG-ROIs), and ADAS-cog and FAQ scores in AD and mild cognitive impairment (MCI) patients enrolled in the Alzheimer's Disease Neuroimaging Initiative (ADNI). Low glucose metabolism at baseline predicted subsequent ADAS-cog and FAQ decline. In addition, longitudinal glucose metabolism decline was associated with concurrent ADAS-cog and FAQ decline. Additionally, a power analysis revealed that FDG-ROI values have greater statistical power than ADAS-cog to detect attenuation of cognitive decline in AD and MCI patients. Glucose metabolism is a sensitive measure of change in cognition and functional ability in AD and MCI, and has value in predicting future cognitive decline.

---

\*Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([www.loni.ucla.edu/ADNI](http://www.loni.ucla.edu/ADNI)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. ADNI investigators include (complete listing available at [www.loni.ucla.edu/ADNI/Collaboration/ADNI\\_Manuscript\\_Citations.pdf](http://www.loni.ucla.edu/ADNI/Collaboration/ADNI_Manuscript_Citations.pdf)).

© 2009 Elsevier Inc. All rights reserved.

Correspondence to: Susan M. Landau, [slandau@berkeley.edu](mailto:slandau@berkeley.edu), Tel: (510) 643-6616, Fax: (510) 642- 3192, 118 Barker Hall MC #3190, UC Berkeley, Berkeley, CA 94720-3190.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosure statement

There are no potential or actual conflicts of interest.

## Keywords

FDG-PET; Alzheimer's disease; Mild Cognitive Impairment

---

## 1. Introduction

Although cognitive tests are used frequently as outcome measures in clinical trials, there are a number of limitations associated with their use (Visser, 2006). The Alzheimer's Disease Assessment Scale (ADAS-cog) is the standard for measuring decline in clinical trials for mild to moderate AD, but several factors limit the utility of this test in a clinical setting. First, the symptomatic significance of improvement or decline on clinical tests has not been well established, making it difficult to set a standard for what is meant by meaningful improvement in order to evaluate potential disease treatments. For example, there is not strong evidence that ADAS-cog performance correlates with measures that are clinically meaningful for patients, such as performance of everyday tasks and social activities (Winblad et al., 2001). Second, scores are highly variable when measured longitudinally (Doraiswamy et al., 2001), perhaps due to the influence of factors like test administrator biases, practice effects, and time of day of testing. Finally, the neurobiological mechanisms that underlie test performance are not well understood, and this complicates the selection of a clinical test that is aligned with biological indicators of disease state.

An optimal outcome measure, then, would reflect clinically-significant patient function, provide reliable measurements with minimal variability, and track a physiologically relevant disease process. FDG-PET is a candidate measure, in that cerebral glucose metabolism is largely a measure of synaptic activity (Sokoloff, 1981) and loss of synapses is an early feature of AD that explains the mechanism of progressive cognitive decline (Terry et al., 1991). Patients with Alzheimer's disease (AD) and mild cognitive impairment (MCI) show well-documented patterns of reduced [ $^{18}\text{F}$ ]fluorodeoxyglucose uptake (FDG-PET) at rest in a network of parietal, posterior cingulate, temporal, and frontal regions (Herholz et al., 2002). While there are few existing longitudinal FDG-PET studies in AD and MCI (Alexander et al., 2002; Drzezga et al., 2003), there is some evidence that FDG-PET accurately predicts subsequent decline (Anchisi et al., 2005; Minoshima et al., 1997) and conversion to AD (Chetelat et al., 2003; Drzezga et al., 2003). However, these studies have relatively small sample sizes and have not established strong evidence for longitudinal associations between existing cognitive measures and FDG-PET.

The goal of this study was to examine the potential for use of FDG-PET as a biomarker in clinical trials of putative therapeutic treatments. Validation of FDG-PET for this purpose would require (1) evidence that longitudinal measurements are feasible in a multicenter clinical trial setting, (2) that FDG-PET accurately tracks AD progression, and (3) that FDG-PET provides adequate statistical power (e.g. required number of subjects per treatment arm).

Our FDG-PET measure was mean glucose metabolism uptake in a set of regions of interest (FDG-ROIs) developed *a priori* and chosen because they have been frequently cited as demonstrating hypometabolism in AD in comparable studies. Our clinical measurements included the ADAS-cog (Rosen et al., 1984) and the Functional Activities Questionnaire (FAQ). We chose the FAQ because it is more closely tied to functionally relevant abilities, such as accomplishing everyday tasks required to live independently (Pfeffer et al., 1982), than the ADAS-cog. The statistical approach employed mixed effects models, which are used frequently to examine factors predicting longitudinal decline in AD by accounting for differences in individual starting points, missing data, and different numbers of visits across

participants (Mungas et al., 2005; Pavlik et al., 2006). Here, we used these models to determine whether baseline and longitudinal FDG-PET measurements were associated with decline in ADAS-cog and FAQ. In addition, because of its functional relevance, changes in FAQ scores over successive assessments served as an outcome variable with which to compare the FDG-PET and ADAS-cog to one another as independent predictors. Finally, we compared the statistical power of FDG-ROIs to ADAS-cog and FAQ as potential outcome measurements in a clinical trial of a putative treatment for AD symptoms. We carried out all analyses for an MCI group, as well as an AD group, in order to examine the relationship between FDG-PET and clinical measures within a population that is more diverse and less impaired than AD subjects. Furthermore, since MCI is considered a transitional phase into AD, our analysis for the MCI group allowed us to determine whether FDG-PET is associated with cognitive changes that precede AD diagnosis. For all analyses, we used subject data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), an ongoing multisite imaging study with a large elderly participant population with a range of cognitive impairment.

## 2. Methods

### 2.1. The Alzheimer's Disease Neuroimaging Initiative (ADNI)

ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and non-profit organizations as a \$60 million, 5-year public-private partnership. The primary goal of ADNI is to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians in the development of new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The principal investigator of this initiative is Michael W. Weiner, M.D., VA Medical Center and University of California, San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. ADNI participants include approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years. Participants are evaluated at baseline, 6, 12, 18 (for MCI only), 24, and 36 months (although AD participants do not have a 36 month evaluation). For additional information see [www.adni-info.org](http://www.adni-info.org).

### 2.2. Subjects

ADNI is an ongoing study and enrollment was staggered, so not all participants had the same number of follow-up visits. The data we present here is from a subset of AD, MCI, and cognitively normal ADNI participants who had completed at least two visits at the time of this study. The numbers of subjects with available data up to the 2 year follow-up visit (baseline, 6, 12, 18, 24 months) are listed in Table 1A. (Note that according to the ADNI protocol only MCI subjects participate in an 18 month visit.)

For full inclusion/exclusion criteria see [www.adni-info.org](http://www.adni-info.org). Briefly, all subjects were between ages 55 and 90, had completed at least 6 years of education, and were fluent in Spanish or English. AD subjects were recruited with the intent to identify individuals with early stages of disease. They had a CDR of 0.5 or 1.0, a MMSE of 20 – 26 (inclusive), and

met the criteria set by the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer’s Disease and Related Disorders Association (NINCDS/ADRDA) (McKhann et al., 1984) for probable AD. MCI subjects were classified as single- or multi-domain amnesic MCI according to the criteria of Petersen et al. (Petersen, 2003). These criteria included a CDR of 0.5, MMSE scores between 24–30 (inclusive), a memory complaint verified by an informant, objective evidence of memory loss as measured by education-adjusted scores on the Wechsler Memory Scale Revised – Logical Memory II, absence of significant levels of impairment in other cognitive domains, and preserved activities of daily living. Cognitively normal subjects had MMSE scores between 24–30, a CDR of 0, no evidence of depression, and no memory complaints. All subjects were free of any other significant neurological disease besides suspected incipient or clinically diagnosed mild to moderate AD. All subjects gave written, informed consent prior to participation through the local IRBs at participating institutions.

### 2.3. Clinical measurements

From the battery of clinical tests acquired for ADNI participants, we selected two for analyses in conjunction with our FDG-PET measurements: the Alzheimer’s Disease Assessment Scale – cognitive subscale (ADAS-cog) and the Functional Activities Questionnaire (FAQ). The ADAS-cog is administered by a certified individual at each study site. It is based upon written and verbal responses of subjects that are related to fundamental cognitive functions (language, memory, praxis, comprehension) and relevant to AD. The total score is reported as a composite score of 11 items and ranges from 0 to 70, with a higher score indicating poorer cognitive function (Rosen et al., 1984). Different forms of the test were administered for each visit to reduce practice effects. The FAQ is a measure of the ability to perform 10 high-level skills used in daily tasks (shopping, preparing meals, handling finances, and understanding current events), each rated by a knowledgeable informant. Each test item is scored on a 4-point scale of increasing caregiver dependence such that a score of zero indicates that the patient does not need assistance with the task, and a score of five indicates that the patient is dependent on a caregiver to perform the task. The total score ranges from 0 to 50, again with a higher score indicating poorer functional performance. PET scans were performed within approximately 2 weeks of the clinical testing sessions (mean number of days between PET and clinical testing sessions: AD = 12.8 +/- 33.3; MCI = 10.4 +/- 22.5; cognitively normal = 10.2 +/- 16.8).

Table 1 summarizes baseline and annual rate of change measurements of demographics and clinical data for patient groups.

### 2.4. Data acquisition and preparation

Details of the ADNI PET data acquisition protocol are publicly available on the UCLA Laboratory of Neuroimaging (LONI) website ([http://www.loni.ucla.edu/ADNI/Data/ADNI\\_Data.shtml](http://www.loni.ucla.edu/ADNI/Data/ADNI_Data.shtml)). Briefly, PET images were acquired at a variety of scanners nationwide using either a 30-minute six frame scan acquired 30 – 60 minutes post-injection or a static 30-minute single frame scan acquired 30 – 60 minutes post-injection. Dynamic scans were coregistered to the first frame and averaged to create a single average image. Static or single-frame averaged images were then aligned along the AC-PC line to a standard 160×160×96 voxel image grid. A subject-specific intensity normalization mask was generated by scaling all images so the value of the voxels in each individual normalization mask summed to one. This was designed to account for intensity differences introduced by use of multiple scanners. The images were then filtered with a scanner specific filter function to produce images of uniform isotropic resolution of 8mm FWHM, the lowest resolution across all the scanners in this multi-center

study, and therefore the common denominator for spatial smoothing. This pre-processing, along with an image quality control analysis, was the starting point for our analysis.

## 2.5. FDG-ROI Generation

We developed a set of pre-defined regions of interest (FDG-ROIs) by identifying regions cited frequently in FDG-PET studies of AD and MCI patients. We conducted a meta-analysis in PubMed using all permutations of the following search terms: AD or Alzheimer's; MCI or Mild Cognitive Impairment; FDG-PET or FDG or glucose metabolism. Within the studies identified by these terms we isolated those that listed coordinates representing results of cross-sectional and/or longitudinal voxelwise analyses in which FDG uptake differed significantly between groups, changed in the same individuals over time, or correlated with cognitive performance. This resulted in a total of 292 MNI or Talairach coordinates and (if available) their accompanying Z-scores or T-values, of which 209 were from cross-sectional or correlational studies and 31 were coordinates from longitudinal studies. See Supplementary Table 1 for the list of studies used to generate the FDG-ROIs.

The following steps were carried out separately for (1) the set of coordinates from cross-sectional or correlational studies and (2) the set of coordinates from longitudinal studies. All coordinates were transformed into MNI space. Then intensity values were generated for coordinates that reflected a combination of the Z-score or t-value associated with the coordinate and the degree to which coordinates within the same region overlapped (indicating repeated citations of the same region across studies). All t-values were transformed to approximate Z scores. Then, overlapping Z scores, when they occurred, were added. The volumes were smoothed with a 14mm FWHM smoothing kernel. Finally, the volume was then intensity normalized using the maximum value, resulting in a map with values between 0 and 1. The cross-sectional coordinate map was then thresholded at 0.50, and this resulted in a set of four regions located in right and left angular gyri, bilateral posterior cingulate gyrus, and left middle/inferior temporal gyrus. Because the longitudinal map was composed of far fewer coordinates than the cross-sectional map and therefore had less regional consistency among coordinates, we thresholded the coordinate intensity values at a higher threshold (0.75), which resulted in a single ROI in right middle/inferior temporal gyrus. An additional longitudinal FDG-ROI in the prefrontal cortex was identified, but it did not meet our cluster size criterion (20 voxels) and signal to noise in this region was insufficient for analysis. All five FDG-ROIs were binarized prior to analysis and are illustrated in the Supplementary Figure.

The correlations between the five FDG-ROIs were statistically significant at baseline (all bivariate ROI correlations  $p < 0.001$ ; Pearson's  $R = 0.38$  to  $0.84$ ), so we generated a composite ROI by averaging across all five ROIs for each subject at each timepoint. Subsequent mixed effects models were calculated using this composite of FDG-ROIs. Baseline status and change for the five FDG-ROIs and the composite FDG-ROI measure are shown in Table 1 and Figure 1. (Note that although the right temporal gyrus FDG-ROI was generated from the longitudinal coordinate map, while the others were generated from the cross-sectional coordinate map, this region did not show greater longitudinal change than the other FDG-ROIs (Figure 1B), so it was included with the other 4 FDG-ROIs in the composite FDG-ROI.)

## 2.6. FDG-PET Image Processing

Spatial normalization of each individual's PET volumes to the SPM5  $^{15}\text{O-H}_2\text{O}$  PET template was conducted using SPM5 (Ashburner and Friston, 2005) (template voxel dimensions:  $91 \times 109 \times 91$ ; voxel size:  $2\text{mm} \times 2\text{mm} \times 2\text{mm}$ ).



To eliminate between-subject nuisance variability in tracer uptake, we used a reference region comprised of the cerebellar vermis, defined by the AAL region within the MNI atlas, and pons, which was manually traced on the MNI atlas. Individual PET volumes at each time point were then intensity normalized to this region. Finally, mean FDG uptake was extracted for each of the five ROIs for each subject at each timepoint.

## 2.7. Statistical analyses

Statistical analyses were carried out using SPSS 16.0. Summary baseline and change means and standard deviations were computed for FDG-ROIs and cognitive tests. Analysis of variance (ANOVA) and post-hoc two-sample t-tests were used to determine differences between groups (AD, MCI, cognitively normal) and were carried out at  $\alpha = 0.05$ , although post-hoc tests were still significant after accounting for multiple comparisons. For longitudinal change summary statistics shown in Table 1, annual change means and standard deviations were estimated by subtracting baseline measurements from 12 month measurements.

For our descriptive data summary (Table 1 and Figure 1), we used a simple baseline – 12m subtraction to show longitudinal change. However, for our regression models, we used mixed effects models to estimate change because we were interested in using a more sophisticated method of modeling the longitudinal data that also tolerated missing time points. Mixed effects regression models in longitudinal analyses make it possible to account for both within-subject variability and between-subject variability (Laird and Ware, 1982). Within-subject error coefficients represent variability in each individual's repeated measurements over time, while between-subjects error coefficients account for cross-subject variability in the effects of rate-of-change predictors on a time-varying dependent variable.

Here, the use of mixed effects models was advantageous in that it allowed us to model serial PET and behavioral measurements at baseline, 6, 12, 18, and 24 months while accounting for missing data and individual variability in between-scan intervals (Gould et al., 2001). First, annual rates of change for all longitudinal variables were calculated using mixed effects models with both fixed and random effects for time (slope) and a random intercept term, incorporating all available data for time points through 24 months. The coefficient for time represents the annual rate of change for each longitudinal variable (shown in Table 1C). Next, we carried out mixed effects models to evaluate relationships with two longitudinal dependent variables: ADAS-cog and FAQ. Separate models were conducted for AD and MCI subject groups (and for two combined subject groupings; AD+MCI, AD+MCI +Normal, see Supplementary Table 2). A fixed effect to account for group membership was included in models involving the combined groupings. Age, education, sex, and ApoE4 status (number of ApoE4 alleles) were included as covariates in all models described below. Each model also included a random slope to account for unexplained between-person variability in rate of change and a random intercept to account for variability in individual starting point, and statistical significance was set at  $\alpha < 0.05$ .

Independent variable data was prepared as follows. ADAS-cog and FDG-ROI measurements were each split into separate variables representing baseline level and a time-varying measure of change since baseline. Time in years since the initial visit was also computed for use in the models. All independent variables of interest (except for time since baseline) were standardized using the group mean and standard deviation, so that parameter estimates for independent variables of interest could be compared to one another and interpreted as units of change in the dependent variable.

We carried out two sets of mixed effects models: One examining baseline independent variables as predictors of change (Model 1), and one examining baseline level and

longitudinal change in independent variables as predictors of change (Model 2). Specifically, with Model 1, we examined the extent to which Baseline FDG-ROIs alone were associated with change in ADAS-cog and with change in FAQ. For this model, the interaction term for Baseline FDG-ROIs X exam date (in years since baseline) was our independent variable of interest in that it represents the degree to which Baseline FDG-ROIs were related to change in the dependent variable over time. With Model 2, the FDG-ROI change variable was added to the existing model to determine the extent to which FDG-ROI change was associated with concurrent change in the outcome variable (ADAS-cog or FAQ) when simultaneously accounting for baseline FDG-ROI levels.

Finally, using the same model types, we compared FDG-ROI and ADAS-cog variables as predictors of FAQ change. Baseline FDG-ROI and Baseline ADAS-cog scores were first entered simultaneously as independent variables (Model 1) to determine the extent to which these baseline scores were related to subsequent change in FAQ. Next, FDG-ROI change and ADAS-cog change were added to the existing baseline model (Model 2) to examine whether change in these tests was related to concurrent FAQ change.

## 2.8. Power calculations

We were interested in comparing the statistical power of FDG-ROIs with that of cognitive tests to detect attenuation of annual decline during a one-year clinical trial of a therapeutic treatment. The statistical power of a given measure depends on the observed rate of annual change and the variability of that rate of change. Separate analyses were carried out for AD and MCI groups (1) using estimates of annual rates of change based on all available data (up to 24 months post-baseline), and (2) using estimates of annual rates of change based on only baseline, 6 month, and 12 month time points, since 12 months may be a more realistic time period for a clinical trial.

We mean-centered FDG-ROI, FAQ, and ADAS-cog longitudinal measurements and fit each as a dependent measure in a mixed effects model with time as a fixed effect, and a random slope and random intercept. The parameter estimate of the time covariate was used as the mean rate of change estimate for a theoretical placebo group in a clinical trial. This placebo group mean was then used to calculate treatment group means demonstrating 25% and 33% attenuation in decline. Sample sizes required per equally-sized group to detect each treatment effect with power = 0.80 and  $\alpha = 0.05$ , assuming a two-sided test and linear rates of decline (Diggle et al., 2002), were computed. These calculations are based on the formula

$$\frac{2(z_{1-\alpha/2} + z_{power})^2 \hat{\sigma}_e^2}{j(t_j - \bar{t})^2 d^2}$$

where  $z_\alpha$  is the value from the standard normal distribution such that  $P[Z < z_\alpha] = \alpha$ ,  $d$  is the difference in annual change between the two arms,  $\hat{\sigma}_e^2$  is the residual standard error from the mixed effects model,  $t_j$  is the time (in years) of the  $j^{\text{th}}$  assessment and  $\bar{t}$  is the average (in years) of the scheduled visits (Diggle et al., 2002).

## 3. Results

### 3.1. Demographic, clinical, and neuroimaging data

Demographic, clinical, and neuroimaging summary data for each group is summarized in Table 1. Means (+/- SD) are shown for baseline clinical tests and FDG-ROI values (Table 1B) and for annual change in the same clinical tests and FDG-ROIs (Table 1C).

AD, MCI, and cognitively normal participant groups do not differ in age or gender ratios (Table 1B). However, AD patients had a lower education level than both MCI ( $t = 3.18$ ,  $p = 0.002$ ) and cognitively normal ( $t = 2.89$ ,  $p = 0.004$ ) groups. In addition, AD patients have

disproportionately higher ApoE 4 allele frequency compared with MCI and cognitively normal patients (Chi-Square = 38.04,  $p < 0.001$ ).

AD, MCI, and cognitively normal groups were compared for differences on baseline clinical measures. Each patient group pairwise comparison (AD and MCI, MCI and cognitively normal, AD and cognitively normal) resulted in group differences in the clinical measures (MMSE, ADAS-cog, FAQ) with MCI subjects scoring significantly higher than AD subjects (two-sample t-tests;  $p < 0.001$ ), and cognitively normal subjects scoring significantly higher than MCI subjects ( $p < 0.001$ ).

AD, MCI, and cognitively normal groups were also compared for differences on change measures. Note that for the clinical tests in our regression analyses (ADAS-cog and FAQ), a positive change represents greater impairment, whereas for the FDG-ROIs and MMSE, a negative change represents worsening (Table 1C). AD patients showed greater annual decline than MCI or cognitively normal subjects on all clinical tests (two-sample t-tests;  $p < 0.005$ ). MCI patients showed greater annual decline than cognitively normal subjects on ADAS-cog and FAQ ( $p < 0.001$ ), and MMSE ( $p < 0.05$ ).

Finally, baseline ADAS-cog and FAQ scores were correlated for AD ( $R = 0.44$ ,  $p < 0.001$ ) and MCI ( $R = 0.26$ ,  $p < 0.001$ ) groups, but not for cognitively normal subjects ( $p > 0.5$ ).

### 3.2. FDG-PET Regions of Interest

Group means are illustrated for the five separate FDG-ROIs (right and left angular gyri, bilateral posterior cingulate, right and left inferior temporal gyri) and the Composite FDG-ROI (Figure 1A). Each pairwise comparison demonstrated significant group differences for all FDG-ROIs (two-sample t-tests;  $p < 0.001$ ) such that mean metabolism was lowest for AD, moderate for MCI, and highest for cognitively normal subjects.

With respect to change (Figure 1B), AD patients show greater decline in all FDG-ROIs compared with MCI and normal groups ( $p < 0.001$ ). MCI patients showed greater annual decline than normal subjects for the Composite FDG-ROI ( $p < 0.001$ ), the left temporal FDG-ROI ( $p = 0.04$ ), and marginally for the posterior cingulate FDG-ROI ( $p = 0.08$ ), but not the other individual FDG-ROIs.

### 3.3. Mixed effects models: Predicting cognitive decline

Mixed effect model results are shown in Table 2. Parameter estimates are based on standardized values of the independent variables, and can therefore be compared across models and AD/MCI groups. Parameter estimates represent the number of points of change in the dependent variable expected with a one standard deviation increase in the independent variable. Note that parameter estimates typically have opposite signs since an increase in FDG-ROI measures represents an improvement but an increase in ADAS-cog measures represents worsening. All models controlled for age, education, sex, and ApoE4 status.

First, we assessed the degree to which measurements in the FDG-ROIs predict ADAS-cog (Table 2A) and FAQ decline (Table 2B) in AD and MCI groups. Low FDG-ROI values at baseline were associated with greater ADAS-cog decline for both subject groups (Table 2A, Model 1), and this relationship remained significant when FDG-ROI longitudinal change is added to the model (Model 2). Decreases in FDG-ROI measures over time were also strongly associated with concurrent ADAS-cog decline for both groups (Model 2). To visualize these relationships, we adjusted each variable for age, education, sex, and number of ApoE4 alleles, and plotted residuals (Figure 2). For both groups, lower baseline FDG-ROI values (Figure 2A) and decreases in FDG-ROIs over time (Figure 2B) were associated with increases in ADAS-cog scores. Because all time points used in the mixed models could



not be shown graphically, we used 12 month – baseline differences to estimate FDG-ROI and ADAS-cog change.

Low baseline FDG-ROI means also predict greater increases in FAQ over time, although the association is marginal for the AD group ( $p = 0.06$ ) (Table 2B, Model 1). When FDG-ROI longitudinal change is added to the model, the baseline FDG-ROI variable remains a significant predictor for the MCI group ( $p < 0.001$ ) but not the AD group ( $p = 0.12$ ). FDG-ROI decreases are associated with concurrent FAQ decline, although the association was marginal for the MCI group ( $p = 0.08$ ).

The second set of mixed effects models compare FDG-ROIs and ADAS-cog to one another as predictors of FAQ change in the same model (Table 2C). Baseline FDG-ROIs predict FAQ change for both groups, and baseline ADAS-cog also predicts FAQ change for the MCI group only (Table 2C, Model 1). When FDG-ROI and ADAS-cog longitudinal change variables are added to the model (Table 2C, Model 2), baseline ADAS-cog scores do not remain significant, while baseline FDG-ROIs remain marginally significant for both groups. The association between ADAS-cog decline and FAQ decline is significant for the MCI group only. Based on a comparison of the parameter estimates, FDG-ROI decline is a stronger predictor than ADAS-cog decline of concurrent FAQ decline for both groups.

In order to examine these relationships in a group with a broad range of cognitive ability, we also carried out the same analyses in combined subject groupings (AD + MCI, AD + MCI + Normal; Supplementary Table 2). In both combined groupings, low FDG-ROI means at baseline predicted both ADAS-cog and FAQ decline, and greater FDG-ROI declines were associated with concurrent ADAS-cog and FAQ decline. Finally, FDG-ROIs at baseline had a greater predictive value with FAQ decline than baseline ADAS-cog, although both were significant predictors; however, only ADAS-cog change (and not FDG-ROI change) was associated with concurrent FAQ change.

### 3.4. Power calculations

We performed power calculations in order to determine samples sizes per arm of AD and MCI subject groups that would be needed to detect 25% and 33% attenuation of decline in a clinical trial of a candidate treatment for symptoms of Alzheimer's disease. Power calculations were carried out using FDG-ROIs, ADAS-cog, and FAQ as potential outcome measures. Statistical power for each outcome measure depended on two parameters: rate of annual decline and residual standard deviation obtained from the mixed effects models. All power calculations assumed linear decline and equal numbers of subjects per treatment and placebo groups. Estimates for AD subjects based on all available data (Figure 3A) were very similar to estimates that included only data up to 12 months post-baseline (Figure 3B). Based on data up to 12 months post-baseline, FDG-ROIs required the lowest number of AD subjects per group to detect a 25% treatment effect (180 subjects per arm, compared with 312 for ADAS-cog and 300 for FAQ) and a 33% treatment effect (101 subjects, compared with 176 for ADAS-cog and 169 for FAQ). Sample sizes for MCI subjects were considerably higher (Figure 3C, 3D), and showed a different pattern, with FDG-ROIs requiring fewer subjects per arm (1271) than ADAS-cog (2175), but more than FAQ (452) for 33% treatment effect based on data up to 12 months post-baseline.

Post-hoc analyses revealed that use of the individual FDG-ROIs (e.g. posterior cingulate) did not improve the sample size estimates, since the individual ROIs had higher longitudinal variability than the composite region.

## 4. Discussion

The goal of this study was to examine the sensitivity of resting glucose metabolism (FDG-PET) to detect longitudinal change in both cognitive (ADAS-cog) and functional (FAQ) measurements within AD and MCI patient populations. We used a subset of participants from the ongoing ADNI study, which provided data from multiple timepoints up to 24 months post-baseline. Overall, we found strong evidence that lower baseline FDG-PET consistently predicts subsequent cognitive decline, and that longitudinal FDG-PET is associated with concurrent cognitive decline. These relationships were similar for functional outcomes, although associations were marginal in some cases. Importantly, an analysis of the statistical power of these measures to detect attenuation in decline for a putative AD treatment (Figure 3) revealed that use of FDG-ROIs would require fewer AD subjects to detect attenuation in decline (101 subjects per group for 33% treatment effect) than ADAS-cog (176 subjects) and FAQ (169 subjects). Sample sizes for the MCI group were considerably higher, although FDG-ROIs again required fewer subjects than ADAS-cog. However, for the MCI group, the FAQ had the lowest sample size estimate, perhaps because MCI subjects were close to ceiling on this test, leading to reduced variability and an artifactual increase in statistical power. This suggests that the FAQ may not be optimal for capturing subtle functional change in MCI. Overall, our findings suggest that FDG-ROIs are a reliable tool for detecting longitudinal change, and may exceed the power of standard clinical outcome measures.

Our finding that FDG-PET was more consistently associated with ADAS-cog than FAQ may be due to differences in test characteristics. The FAQ differs from the ADAS-cog in that it is not an index of cognitive function but instead a measure of the ability to carry out daily functions. Importantly, each test requires input from a person other than the study participant, and this may introduce subjectivity; the FAQ is completed by an informant, whereas the ADAS-cog is administered by a certified tester at the study site. Furthermore, as noted above, FAQ performance may be at ceiling in cognitively normal and MCI individuals, where there is subtle or no impairment and little change over time. 27% of MCI patients (compared with 1% of AD patients) had an FAQ score of 0, indicating little or no functional impairment.

Differing associations we observed for AD and MCI groups provide insight into the sensitivity of baseline and longitudinal FDG-PET in populations with varying levels of disease severity. Consistent with previous findings (Alexander et al., 2002), AD patients demonstrated lower FDG-PET uptake at baseline (Figure 1A) and greater longitudinal decline than MCI or cognitively normal participants across all cognitive tests and all ROIs of interest (Figure 1B). For both AD and MCI groups, lower baseline FDG-ROI measurements predicted greater subsequent decline on the ADAS-cog (Table 2A; Figure 2A) and the FAQ (Table 2B), although the latter association was marginal for the AD group. Greater longitudinal FDG-ROI decline was also associated with concurrent ADAS-cog (Figure 2B) and FAQ decline, although the latter was marginal for the MCI group (likely due to the ceiling effect discussed above). Finally, a comparison of FDG-ROIs and ADAS-cog as predictors of FAQ decline revealed that baseline and longitudinal FDG-ROI measures were marginally or significantly associated with FAQ change in all cases. Baseline and longitudinal ADAS-cog measures were associated with FAQ change for the MCI group but not the AD group.

Parameter estimates of FDG-ROI variables were generally higher in the AD group, likely reflecting greater decline in clinical measures for that group. The MCI group, on the other hand, experienced lower levels of decline, and was more variable, with some subjects experiencing decline and others remaining relatively stable. In addition, there was a closer

relationship between ADAS-cog decline and FAQ decline for the MCI group compared with the AD group (Table 2C), indicating strong consistency between these measures despite the reduced variability on the FAQ.

These data extend the findings of previous studies showing the value of FDG-PET for predicting subsequent decline in MCI patients (for example, Chetelat et al., 2003; Herholz et al., 1999; Minoshima et al., 1997) and normal older individuals (de Leon et al., 2001). Few studies, however, have examined longitudinal concurrent relationships between FDG-PET and cognitive measurements. Existing large multicenter FDG-PET studies have typically focused on cross-sectional analyses and diagnostic accuracy of FDG-PET for AD, rather than longitudinal decline (Herholz et al., 2002; Mosconi et al., 2008b). Nonetheless, our findings are consistent with the few existing longitudinal FDG-PET studies, which is not surprising since our ROIs were based in part on coordinates cited in these studies. In voxelwise analyses, declines in AD patients (Alexander et al., 2002) and in MCI patients who convert to AD (Drzezga et al., 2003; Fouquet et al., 2009) were reported in regions that overlapped with ours, as well as frontal regions. While we did identify a frontal ROI that survived our thresholding procedure during ROI generation, the region was eliminated because it was too small to give meaningful results.

Our results are also in agreement with other studies that have carried out power calculations using FDG-PET as an outcome measurement to detect clinical trial treatment effects based on data from normal individuals at genetic risk for AD (Reiman et al., 2001; Small et al., 2000) and in AD (Alexander et al., 2002). The latter study was based on a similar (12 month) followup period, and it reports sample sizes (ranging from 24 – 179, depending on brain region, for a 33% treatment effect) that are less than that required for the cognitive tests they examined. However, our method differs in that we used pre-defined ROIs as opposed to a voxelwise analysis where the results depend on the AD patients in the study. Nonetheless, both studies are in agreement in suggesting that FDG-PET may be a more reliable outcome measure than cognitive tests to detect attenuation of decline in clinical trials of AD patients. For MCI subjects, sample size estimates were considerably larger than the AD group, likely due to greater variability in disease severity. Additional analyses are currently being conducted to directly compare the power of different imaging modalities (i.e., FDG-PET and structural MRI) and different voxel-based, functionally and anatomically defined ROI and whole brain image analysis methods in terms of their estimated power to detect effects of putative AD-slowing treatments in randomized clinical trials.

There are several novel features of this study that improve on previous analyses. First, we used continuous measures of cognition as predictor and outcome variables, rather than binary conversion/nonconversion status as is used frequently in longitudinal studies (de Leon et al., 2001; Drzezga et al., 2003). The use of continuous outcome variables measuring cognition (for example, Chetelat et al., 2005; Herholz et al., 1999; Jagust et al., 2006; Mosconi et al., 2008a) may become increasingly important as clinical trials move to enroll milder patients and measure cognitive change, rather than conversion, as an outcome. Second, our study-independent ROIs differ from other studies that used standard atlas regions (e.g. Talairach, MNI) or regions that result from a voxelwise analysis. The motivation for this approach was that it allowed us to identify critical regions with more precision than is possible using anatomically-defined regions. When hypometabolism occurs in a subregion of a large atlas ROI such as the inferior temporal gyrus, this effect may be diluted when averaged across the entire atlas-based ROI. Furthermore, in studies using voxelwise analyses, the precise location and nature of the differences is dependent on the individuals in the study and the data processing methods used. Spatial normalization procedures are highly variable, and the success of implementing these procedures

successfully may introduce variability in the results. A limitation of our approach, however, is that the size and location of the most significant glucose metabolism decline for this group may not be adequately captured by the FDG-ROIs, whereas that would be optimized in a voxelwise analysis.

A final novel feature of this study was the use of the ADNI population, which made it possible to obtain serial cognitive and FDG-PET measurements acquired at a variety of sites and PET scanners up to 24 months post-baseline, which is a quantity of longitudinal data that has not been previously available. Current knowledge about cognitive and neural function in Alzheimer's disease has been pieced together from much smaller studies, since studies incorporating multiple study sites have been rarely conducted and they are not longitudinal. The results presented here show that it is possible to successfully replicate previous findings using multisite data and to examine models that have not been previously possible due to insufficient sample sizes or study length. In addition, a multisite study raises a number of methodological questions related to image processing and statistical analysis. For example, our method of collapsing across diagnostic groups (AD + MCI subjects; AD + MCI + cognitively normal subjects) was designed to treat disease progression as a continuum as opposed to discrete diagnostic states. For these groups, we found robust relationships between FDG-ROIs and clinical/functional change, perhaps because the sample sizes were largest and the use of continuous variables allowed us to detect subtle relationships at all levels of disease severity.

In summary, we found that baseline and longitudinal FDG-ROI measures are sensitive to change in both the ADAS-cog and a test of functional competence (FAQ), validating the cognitive and functional relevance of longitudinal changes in FDG-PET measurements. Our power analysis indicated that FDG-PET may be a reliable and clinically-useful measure of decline compared with ADAS-cog, particularly in AD patients. Strong associations observed between FDG-PET and ADAS-cog, in particular, indicate that FDG-PET could be useful in clinical trials for selecting subjects who likely to decline for clinical trials, or as an outcome measurement for monitoring clinically-relevant change over time. While the ADAS-cog is frequently used as an outcome measure in clinical trials, the clinical relevance of the small margins of change that are often cited as positive results (Rogers et al., 1998) is unclear, and it has substantial variability. The results we present here are part of an ongoing analysis of the extensive ADNI dataset that is not yet complete. Future analyses of ADNI data will address the question of role of the ApoE4 allele, which is known to play a role in FDG-PET decline (Reiman et al., 2001), CSF biomarkers such as A $\beta$ -42 and tau (Haense et al., 2008), and grey matter volume, which shows substantial reductions longitudinally (Jack et al., 1999; Mungas et al., 2005).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This study was supported by NIH grant U01 AG024904.

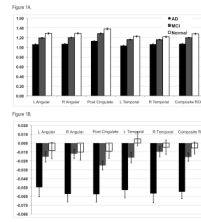
## References

- Alexander GE, Chen K, Pietrini P, Rapoport SI, Reiman EM. Longitudinal PET Evaluation of Cerebral Metabolic Decline in Dementia: A Potential Outcome Measure in Alzheimer's Disease Treatment Studies. *Am J Psychiatry*. 2002; 159:738–745. [PubMed: 11986126]
- Anchisi D, Borroni B, Franceschi M, Kerrouche N, Kalbe E, Beuthien-Beumann B, Cappa S, Lenz O, Ludecke S, Marcone A, Mielke R, Ortelli P, Padovani A, Pelati O, Pupi A, Scarpini E, Weisenbach

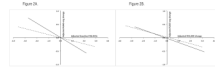
- S, Herholz K, Salmon E, Holthoff V, Sorbi S, Fazio F, Perani D. Heterogeneity of brain glucose metabolism in mild cognitive impairment and clinical progression to Alzheimer disease. *Arch Neurol*. 2005; 62:1728–1733. [PubMed: 16286547]
- Ashburner J, Friston KJ. Unified segmentation. *Neuroimage*. 2005; 26:839–851. [PubMed: 15955494]
- Chetelat G, Desgranges B, de la Sayette V, Viader F, Eustache F, Baron JC. Mild cognitive impairment: Can FDG-PET predict who is to rapidly convert to Alzheimer's disease? *Neurology*. 2003; 60:1374–1377. [PubMed: 12707450]
- Chetelat G, Eustache F, Viader F, De La Sayette V, Pelerin A, Mezenge F, Hannequin D, Dupuy B, Baron JC, Desgranges B. FDG-PET measurement is more accurate than neuropsychological assessments to predict global cognitive deterioration in patients with mild cognitive impairment. *Neurocase*. 2005; 11:14–25. [PubMed: 15804920]
- de Leon MJ, Convit A, Wolf OT, Tarshish CY, DeSanti S, Rusinek H, Tsui W, Kandil E, Scherer AJ, Roche A, Imossi A, Thorn E, Bobinski M, Caraos C, Lesbre P, Schlyer D, Poirier J, Reisberg B, Fowler J. Prediction of cognitive decline in normal elderly subjects with 2-[(18)F]fluoro-2-deoxy-D-glucose/positron-emission tomography (FDG/PET). *Proc Natl Acad Sci U S A*. 2001; 98:10966–10971. [PubMed: 11526211]
- Diggle, P.; Heagerty, P.; Liang, K-Y.; Zeger, S. 2 ed.. Oxford: Oxford University Press; 2002. Analysis of longitudinal data.
- Doraiswamy PM, Kaiser L, Bieber F, Garman RL. The Alzheimer's Disease Assessment Scale: evaluation of psychometric properties and patterns of cognitive decline in multicenter clinical trials of mild to moderate Alzheimer's disease. *Alzheimer Dis Assoc Disord*. 2001; 15:174–183. [PubMed: 11723368]
- Drzezga A, Lautenschlager N, Siebner H, Riemenschneider M, Willoch F, Minoshima S, Schwaiger M, Kurz A. Cerebral metabolic changes accompanying conversion of mild cognitive impairment into Alzheimer's disease: a PET follow-up study. *Eur J Nucl Med Mol Imaging*. 2003; 30:1104–1113. [PubMed: 12764551]
- Fouquet M, Desgranges B, Landeau B, Duchesnay E, Mezenge F, de la Sayette V, Viader F, Baron JC, Eustache F, Chetelat G. Longitudinal brain metabolic changes from amnesic mild cognitive impairment to Alzheimer's disease. *Brain*. 2009
- Gould R, Abramson I, Galasko D, Salmon D. Rate of cognitive change in Alzheimer's disease: methodological approaches using random effects models. *J Int Neuropsychol Soc*. 2001; 7:813–824. [PubMed: 11771624]
- Haense C, Buerger K, Kalbe E, Drzezga A, Teipel SJ, Markiewicz P, Herholz K, Heiss WD, Hampel H. CSF total and phosphorylated tau protein, regional glucose metabolism and dementia severity in Alzheimer's disease. *Eur J Neurol*. 2008; 15:1155–1162. [PubMed: 18803648]
- Herholz K, Nordberg A, Salmon E, Perani D, Kessler J, Mielke R, Halber M, Jelic V, Almkvist O, Collette F, Alberoni M, Kennedy A, Hasselbalch S, Fazio F, Heiss WD. Impairment of neocortical metabolism predicts progression in Alzheimer's disease. *Dement Geriatr Cogn Disord*. 1999; 10:494–504. [PubMed: 10559566]
- Herholz K, Salmon E, Perani D, Baron JC, Holthoff V, Frolich L, Schonknecht P, Ito K, Mielke R, Kalbe E, Zundorf G, Delbeuck X, Pelati O, Anchisi D, Fazio F, Kerrouche N, Desgranges B, Eustache F, Beuthien-Baumann B, Menzel C, Schroder J, Kato T, Arahata Y, Henze M, Heiss WD. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *Neuroimage*. 2002; 17:302–316. [PubMed: 12482085]
- Jack CR Jr, Petersen RC, Xu YC, O'Brien PC, Smith GE, Ivnik RJ, Boeve BF, Waring SC, Tangalos EG, Kokmen E. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology*. 1999; 52:1397–1403. [PubMed: 10227624]
- Jagust W, Gitcho A, Sun F, Kuczynski B, Mungas D, Haan M. Brain imaging evidence of preclinical Alzheimer's disease in normal aging. *Ann Neurol*. 2006; 59:673–681. [PubMed: 16470518]
- Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics*. 1982; 38:963–974. [PubMed: 7168798]
- McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of



- Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984; 34:939–944. [PubMed: 6610841]
- Minoshima S, Giordani B, Berent S, Frey KA, Foster NL, Kuhl DE. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann Neurol*. 1997; 42:85–94. [PubMed: 9225689]
- Mosconi L, De Santi S, Li J, Tsui WH, Li Y, Boppana M, Laska E, Rusinek H, de Leon MJ. Hippocampal hypometabolism predicts cognitive decline from normal aging. *Neurobiol Aging*. 2008a; 29:676–692. [PubMed: 17222480]
- Mosconi L, Tsui WH, Herholz K, Pupi A, Drzezga A, Lucignani G, Reiman EM, Holthoff V, Kalbe E, Sorbi S, Diehl-Schmid J, Pernecky R, Clerici F, Caselli R, Beuthien-Baumann B, Kurz A, Minoshima S, de Leon MJ. Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer's disease, and other dementias. *J Nucl Med*. 2008b; 49:390–398. [PubMed: 18287270]
- Mungas D, Harvey D, Reed BR, Jagust WJ, DeCarli C, Beckett L, Mack WJ, Kramer JH, Weiner MW, Schuff N, Chui HC. Longitudinal volumetric MRI change and rate of cognitive decline. *Neurology*. 2005; 65:565–571. [PubMed: 16116117]
- Pavlik VN, Doody RS, Massman PJ, Chan W. Influence of premorbid IQ and education on progression of Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2006; 22:367–377. [PubMed: 16954693]
- Petersen, RC. Conceptual overview. In: Petersen, RC., editor. *Mild Cognitive Impairment: Aging to Alzheimer's Disease*. New York: Oxford University Press; 2003. p. 1-14.
- Pfeffer RI, Kurosaki TT, Harrah CH Jr, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol*. 1982; 37:323–329. [PubMed: 7069156]
- Reiman EM, Caselli RJ, Chen K, Alexander GE, Bandy D, Frost J. Declining brain activity in cognitively normal apolipoprotein E epsilon 4 heterozygotes: A foundation for using positron emission tomography to efficiently test treatments to prevent Alzheimer's disease. *Proc Natl Acad Sci U S A*. 2001; 98:3334–3339. [PubMed: 11248079]
- Rogers SL, Farlow MR, Doody RS, Mohs R, Friedhoff LT. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group. *Neurology*. 1998; 50:136–145. [PubMed: 9443470]
- Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984; 141:1356–1364. [PubMed: 6496779]
- Small GW, Ercoli LM, Silverman DH, Huang SC, Komo S, Bookheimer SY, Lavretsky H, Miller K, Siddarth P, Rasgon NL, Mazziotta JC, Saxena S, Wu HM, Mega MS, Cummings JL, Saunders AM, Pericak-Vance MA, Roses AD, Barrio JR, Phelps ME. Cerebral metabolic and cognitive decline in persons at genetic risk for Alzheimer's disease. *Proc Natl Acad Sci U S A*. 2000; 97:6037–6042. [PubMed: 10811879]
- Sokoloff L. Relationships among local functional activity, energy metabolism, and blood flow in the central nervous system. *Fed Proc*. 1981; 40:2311–2316. [PubMed: 7238911]
- Terry RD, Masliah E, Salmon DP, Butters N, DeTeresa R, Hill R, Hansen LA, Katzman R. Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. *Ann Neurol*. 1991; 30:572–580. [PubMed: 1789684]
- Visser PJ. Role of cognitive testing in disease modifying AD trials. *J Nutr Health Aging*. 2006; 10:131–132. discussion 132–133. [PubMed: 16554947]
- Winblad B, Brodaty H, Gauthier S, Morris JC, Orgogozo JM, Rockwood K, Schneider L, Takeda M, Tariot P, Wilkinson D. Pharmacotherapy of Alzheimer's disease: is there a need to redefine treatment success? *Int J Geriatr Psychiatry*. 2001; 16:653–666. [PubMed: 11466744]

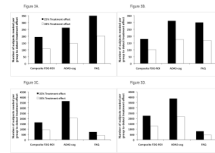


**Figure 1.** Histograms of FDG-ROIs baseline status (A) and longitudinal change (B) are shown for AD, MCI, and cognitively normal groups. Error bars represent standard error. Individual FDG regions of interest (FDG-ROIs) are defined based on a meta-analysis (see Methods). The Composite FDG-ROI is the average of the five individual FDG-ROIs and was used for all subsequent analyses. FDG-ROI means are listed in Table 1.



**Figure 2.**

Fitted regression lines are shown in order to visualize results of mixed effects models shown in Table 2A. As in the mixed effects models, we adjusted each variable for age, education, sex, number of ApoE4 alleles (and baseline FDG-ROI for the FDG-ROI change model), and plotted the standardized residuals. For the AD (solid line) and MCI (dashed line) groups, adjusted baseline FDG-ROI values (x-axis) are plotted against adjusted ADAS-cog change (12 month – baseline) (Figure 2A). Similarly, adjusted FDG-PET change (12 month – baseline) is plotted against adjusted ADAS-cog change for both groups (Figure 2B). The adjusted values can be interpreted as follows: a value of 1 on either axis indicates that the subject is 1 standard deviation higher for that measure than would be expected for an individual with that age, education, sex, and ApoE4 status.



**Figure 3.**

Histograms show sample sizes per group (treatment, placebo) that would be required to detect a 25% (black bars) or 33% (white bars) attenuation in annual decline of each outcome measure during a clinical trial of a candidate therapeutic treatment. Analyses were carried out for AD subjects with all available up to 24 months post-baseline (Figure 3A) and using only data up to 12 months post-baseline (Figure 3B), and similarly for MCI subjects with all available data (Figure 3C) and using data up to 12 months (Figure 3D). Power calculations assume linear rate of decline, equal treatment and placebo group sizes, and were carried out at power = 0.80,  $\alpha = 0.05$ , 2-tailed.

Table 1

Demographic, clinical, and FDG-PET data.

(A)	AD	MCI	Normal
N Baseline	95	206	102
N 6 Months	86	188	94
N 12 Months	74	176	85
N 18 Months	NA	76	NA
N 24 Months	26	66	45
(B)	AD	MCI	Normal
% Female	37%	32%	39%
Age (yrs)	77.2 +/- 7.5	76.6 +/- 7.2	77.8 +/- 4.6
Education (yrs)	14.7 +/- 3.2	15.8 +/- 2.9	15.9 +/- 3.2
Total num visits	3.0 +/- 0.9	3.8 +/- 1.2	3.3 +/- 0.9
ApoE4 allele freq	0.41	0.33	0.13
% Right-handed	90%	93%	96%
Baseline clinical data			
MMSE	23.4 +/- 2.1	27.2 +/- 1.7	29.0 +/- 1.1
ADAS-cog	19.2 +/- 6.6	10.8 +/- 4.2	6.8 +/- 3.0
FAQ	13.7 +/- 6.6	3.4 +/- 3.9	0.2 +/- 0.8
Baseline FDG-ROIs			
L Angular	1.06 +/- 0.18	1.19 +/- 0.15	1.29 +/- 0.15
R Angular	1.07 +/- 0.17	1.20 +/- 0.15	1.29 +/- 0.14
Post Cingulate	1.13 +/- 0.15	1.29 +/- 0.17	1.38 +/- 0.16
L Temporal	1.04 +/- 0.16	1.16 +/- 0.14	1.23 +/- 0.13
R Temporal	1.06 +/- 0.17	1.16 +/- 0.12	1.22 +/- 0.11
Composite ROI	1.07 +/- 0.13	1.20 +/- 0.13	1.28 +/- 0.12
(C)	AD	MCI	Normal
Annual change: clinical data			



(A)	AD	MCI	Normal
MMSE	-2.8 +/- 3.9	-0.4 +/- 2.2	0.1 +/- 1.4
ADAS-cog	3.8 +/- 6.2	1.0 +/- 4.2	-0.9 +/- 3.1
FAQ	4.0 +/- 5.1	1.6 +/- 3.7	0.1 +/- 1.5
<b>Annual change: FDG-ROIs</b>			
<b>L Angular</b>	-0.050 +/- 0.092	-0.015 +/- 0.079	-0.009 +/- 0.083
<b>R Angular</b>	-0.057 +/- 0.077	-0.012 +/- 0.073	-0.010 +/- 0.082
<b>Post Cingulate</b>	-0.058 +/- 0.076	-0.025 +/- 0.067	-0.009 +/- 0.074
<b>L Temporal</b>	-0.053 +/- 0.076	-0.016 +/- 0.080	0.005 +/- 0.071
<b>R Temporal</b>	-0.057 +/- 0.089	-0.010 +/- 0.075	-0.005 +/- 0.074
<b>Composite ROI</b>	-0.055 +/- 0.068	-0.016 +/- 0.063	-0.005 +/- 0.062

Sample sizes (A) of currently available data are listed for each time point. Demographic, clinical, and FDG-PET means +/- SD are listed for baseline status (B) and longitudinal change (C). Individual FDG regions of interest (FDG-ROIs) are defined based on a meta-analysis (see Methods). Composite FDG-ROI means were calculated for each subject at each timepoint by averaging across the five individual FDG-ROIs. FDG-ROI baseline status and longitudinal change are illustrated in Figure 1.

Table 2

Results of mixed effects models.

<b>(A) Dependent: ADAS-cog</b>						
			<b>AD</b>		<b>MCI</b>	
	$\beta$	P	$\beta$	P	$\beta$	P
<b>Model 1</b>						
FDG-ROIs						
Baseline	-1.96	+/-	0.65	0.003*	-0.54	+/- 0.21 0.012*
<b>Model 2</b>						
FDG-ROIs						
Baseline	-1.95	+/-	0.60	0.002*	-0.66	+/- 0.22 0.003*
FDG-ROIs						
Change	-3.25	+/-	0.71	<0.001*	-1.08	+/- 0.30 <0.001*
<b>(B) Dependent: FAQ</b>						
			<b>AD</b>		<b>MCI</b>	
	$\beta$	P	$\beta$	P	$\beta$	P
<b>Model 1</b>						
FDG-ROIs						
Baseline	-0.76	+/-	0.40	0.06	-0.89	+/- 0.20 <0.001*
<b>Model 2</b>						
FDG-ROIs						
Baseline	-0.64	+/-	0.41	0.12	-0.88	+/- 0.22 <0.001*
FDG-ROIs						
Change	-1.21	+/-	0.47	0.01*	-0.46	+/- 0.26 0.08

	(C) Dependent: FAQ		
	AD		MCI
	$\beta$	P	$\beta$
<b>Model 1</b>			
FDG-ROIs			
Baseline	-0.97 +/-	0.45 0.03*	-0.62 +/- 0.20 0.003*
ADAS-cog			
Baseline	-0.44 +/-	0.39 0.27	0.83 +/- 0.21 <0.001*
<b>Model 2</b>			
FDG-ROIs			
Baseline	-0.91 +/-	0.48 0.06	-0.39 +/- 0.21 0.06
FDG-ROIs			
Change	-1.20 +/-	0.61 0.05	-0.99 +/- 0.21 <0.001*
ADAS-cog			
Baseline	-0.53 +/-	0.41 0.20	-0.06 +/- 0.26 0.80
ADAS-cog			
Change	0.03 +/-	0.32 0.92	0.67 +/- 0.14 <0.001*

FDG-ROIs were examined as a predictor of (A) ADAS-cog change and (B) FAQ change. In (C), FDG-ROI and ADAS-cog measures were entered into the model simultaneously to determine their contributions to rate of FAQ change. Baseline variables were first entered as predictors (Model 1), and then change predictors were added to the model (Model 2). Standardized parameter estimates ( $\pm$ SEM) for association with change in the outcome ( $\beta$ ) and p-values are listed (asterisks indicate significant p-values). All models included covariates for age, education, sex, and number of ApoE4 alleles. See Methods for details.