Twelve-month metabolic declines in probable Alzheimer's disease and amnestic mild cognitive impairment assessed using an empirically pre-defined statistical region-of-interest: Findings from the Alzheimer's Disease Neuroimaging Initiative

Kewei Chen [a,b,q,*], Jessica B.S. Langbaum [a,q], Adam S. Fleisher [a,p,q], Napatkamon Ayutyanont [a,q], Cole Reschke [a,q], Wendy Lee [a,q], Xiaofen Liu [a,q], Dan Bandy [a,q], Gene E. Alexander [c,q], Paul M. Thompson [e], Norman L. Foster [f], Danielle J. Harvey [g], Mony J. de Leon [h,i], Robert A. Koeppe [j], William J. Jagust [k], Michael W. Weiner [l,m,n], Eric M. Reiman [a,d,o,q]
and The Alzheimer's Disease Neuroimaging Initiative [1]

[a] Banner Alzheimer's Institute and Banner Good Samaritan PET Center, Phoenix, AZ, USA
[b] Department of Mathematics and Statistics, Arizona State University, Tempe, AZ, USA
[c] Department of Psychology and Evelyn F. McKnight Brain Institute, Tucson, AZ, USA
[d] Department of Psychiatry, University of Arizona, Tucson, AZ, USA
[e] Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Los Angeles, CA, USA
[f] Center for Alzheimer's Care, Imaging and Research and Department of Neurology, University of Utah, Salt Lake City, UT, USA
[g] Division of Biostatistics, Department of Public Health Sciences, University of California, Davis, USA
[h] Center for Brain Health, Department of Psychiatry, Langone Medical Center, New York University, USA
[i] The Nathan Kline Institute, Orangeburg NY, USA
[j] Division of Nuclear Medicine, Department of Radiology, University of Michigan, Ann Arbor, MI, USA
[k] School of Public Health and Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA, USA
[l] Department of Radiology, University of California San Francisco, San Francisco, CA, USA
[m] Department of Medicine, University of California San Francisco, San Francisco, CA, USA
[n] Department of Psychiatry, University of California San Francisco, San Francisco, CA, USA
[o] Division of Neurogenomics, Translational Genomics Research Institute, Phoenix, AZ, USA
[p] Department of Neurosciences, University of California, San Diego, USA
[q] Arizona Alzheimer's Consortium, Phoenix, AZ, USA

ARTICLE INFO

ABSTRACT

Alzheimer's disease (AD) is characterized by specific and progressive reductions in fluorodeoxyglucose positron emission tomography (FDG PET) measurements of the cerebral metabolic rate for glucose (CMRgl), some of which may precede the onset of symptoms. In this report, we describe twelve-month CMRgl declines in 69 probable AD patients, 154 amnestic mild cognitive impairment (MCI) patients, and 79 cognitively normal controls (NCs) from the AD Neuroimaging Initiative (ADNI) using statistical parametric mapping (SPM). We introduce the use of an empirically pre-defined statistical region-of-interest (sROI) to characterize CMRgl declines with optimal power and freedom from multiple comparisons, and we estimate the number of patients needed to characterize AD-slowing treatment effects in multi-center randomized clinical trials (RCTs). The AD and MCI groups each had significant twelve-month CMRgl declines bilaterally in posterior cingulate, medial and lateral parietal, medial and lateral temporal, frontal and occipital cortex, which were significantly greater than those in the NC group and correlated with measures of clinical decline. Using sROIs defined based on training sets of baseline and follow-up images to assess CMRgl declines in independent test sets from each patient group, we estimate the need for 66 AD patients or 217 MCI patients per treatment group to detect a 25% AD-slowing treatment effect in a twelve-month, multi-center RCT with 80% power and two-tailed alpha = 0.05,

* Corresponding author. Banner Alzheimer's Institute, PET Center, Banner Good Samaritan Medical Center, 1111 E. McDowell Road, Phoenix, AZ 85006, USA. Fax: +1 602 839 3498.
E-mail address: Kewei.Chen@bannerhealth.com (K. Chen).

roughly one-tenth the number of the patients needed to study MCI patients using clinical endpoints. Our findings support the use of FDG PET, brain-mapping algorithms and empirically pre-defined sROIs in RCTs of AD-slowing treatments.

Introduction

Alzheimer's disease (AD) is the most common form of disabling cognitive impairment in older adults (Evans et al., 1989). Given the extraordinary toll AD takes on patients and their families, the rapidly growing number of people in older age groups (Hebert et al., 2001), and the financial burden that it is projected to take on communities around the world by the time today's young adults become senior citizens (Brookmeyer et al., 1998), there is an urgent need to find demonstrably effective treatments for this disease. One of the major challenges to the development of AD treatments is finding the means to evaluate them in the most cost-effective, rapid and rigorous way (Reiman and Langbaum, 2009).

Due to the slow progression of cognitive decline and test–retest variability in the clinical endpoints used in randomized clinical trials (RCTs), it currently takes too many subjects, too much time and too much money to evaluate AD-slowing treatments, especially in the earliest symptomatic and pre-symptomatic stages of AD. For these and other reasons, researchers have sought to develop biomarker endpoints that reflect AD progression or pathology and that could help to assess putative AD-slowing treatments with better statistical power than clinical endpoints (Jones et al., 2009; Landau et al., 2009). To date, the best established biomarkers of AD progression are volumetric magnetic resonance imaging (MRI) measurements of brain shrinkage (Fox et al., 1999; Fox et al., 2000; Hua et al., 2009; Jack et al., 2008; Schuff et al., 2009) and fluorodeoxyglucose positron emission tomography (FDG PET) measurements of regional cerebral metabolic rate for glucose (CMRgl) decline (Alexander et al., 2002; Reiman and Langbaum, 2009; Reiman et al., 2001) and the best established biomarkers of AD pathology are fibrillar amyloid-β (Aβ) PET measurements using Pittsburgh Compound B (PiB) and other recently developed radioligands (Doraiswamy et al., 2009; Jack et al., 2009; Johnson et al., 2009; Klunk et al., 2004; Nyberg et al., 2009; Shoghi-Jadid et al., 2002; Small et al., 2006) and cerebrospinal fluid (CSF) Aβ, total tau and phospho-tau levels (Fagan et al., 2006; Fagan et al., 2007; Fagan et al., 2009; Hansson et al., 2006; Hansson et al., 2009; Li et al., 2007; Sunderland et al., 2004).

FDG PET studies find characteristic and progressive CMRgl reductions in posterior cingulate, precuneus, parietal, temporal and frontal regions in patients with AD and mild cognitive impairment (MCI) (Alexander et al., 2002; de Leon et al., 1983; Drzezga et al., 2003; Foster et al., 1983; Haxby et al., 1990; Herholz et al., 2002; Jagust et al., 1988; Langbaum et al., 2009; Minoshima et al., 1994; Mosconi et al., 2005; Mosconi et al., 2009) and in cognitively normal people at increased genetic risk for AD (Bookheimer et al., 2000; Reiman et al., 1996; Reiman et al., 2001; Reiman et al., 2004; Reiman et al., 2005; Small et al., 1995). A smaller number of studies have also found progressive CMRgl declines in medial temporal and occipital regions in these patients (Alexander et al., 2002; Mielke et al., 1994). We and our colleagues previously used FDG PET and statistical parametric mapping (SPM) to characterize twelve-month regional CMRgl declines in probable AD patients, and we estimated the number of patients per group needed to detect treatment effects in a twelve-month single-center RCT using the atlas coordinates associated with maximal CMRgl declines (two-tailed $p < 0.01$, uncorrected for multiple comparisons) (Alexander et al., 2002). The number of patients needed to detect significant treatment effects was about one-tenth of that needed using the Mini-Mental State Examination (MMSE) (Folstein et al., 1975), a measure of clinical progression, and roughly comparable to that reported in an earlier study using volumetric MRI measurements of whole-brain atrophy (Fox et al., 2000). In another study, we characterized 24-month regional CMRgl declines in cognitively normal, late-middle-aged carriers of the APOE ε4 allele, the major AD susceptibility gene and estimated the need for relatively fewer than 200 subjects per group to help test putative pre-symptomatic treatments in a 24-month RCT, thus suggesting the feasibility of using these biomarker endpoints to conduct prevention studies in a cost-effective way (Reiman et al., 2001).

Launched in 2004 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Foundation for the National Institutes of Health, private pharmaceutical companies, and nonprofit organizations, the AD Neuroimaging Initiative (ADNI) is an ongoing, multi-center, longitudinal study intended to assist in the early detection and tracking of AD and the design of multi-center clinical trials using brain-imaging and other endpoints. ADNI capitalizing on standardized data acquisition methods makes publicly available clinical, cognitive, volumetric MRI, FDG PET, Pittsburgh Compound B (PiB) PET, cerebral spinal fluid (CSF) and genetic data, and biological fluid samples in probable AD patients, MCI patients, and normal control (NC) subjects (http://www.loni.ucla.edu/ADNI/About/About_Funding.shtml). It is primarily intended to characterize and compare different measures of AD progression and pathology in the detection and tracking of AD, characterize the extent to which the biomarker measurements are correlated with measurements of cognitive and clinical decline, and compare these different measurements—including those using different imaging modalities and image-analysis techniques—in the estimated number of AD and MCI patients per group needed to detect effects of AD-slowing treatments in multi-center RCTs (Hua et al., 2009; Mueller et al., 2005a; Mueller et al., 2005b; Shaw et al., 2007).

As members of ADNI's PET Coordinating Center, we are responsible for conducting voxel-based analyses of FDG PET images using SPM. In this report, we characterize and compare twelve-month CMRgl declines in the probable AD patients, MCI patients, and NC subjects. We introduce the use of an empirically pre-defined spatially distributed statistical region-of-interest (sROI), comprised of the set of voxels consistently associated with longitudinal change, to optimize the power to characterize these declines and evaluate AD-slowing treatments in RCTs. This provides a single imaging endpoint that could be used in RCTs, free from issues associated with multiple comparisons. The sROI can also be customized to particular subject samples and study durations using voxel-based image-analysis techniques. sROIs were empirically pre-defined in AD and MCI patient training sets using SPM, and were then applied to independent AD and MCI patient test sets to demonstrate the reproducibility of these declines and their correlation with measures of cognitive decline, and estimate the number of patients needed per group to detect AD-slowing treatment effects in twelve-month multi-center RCTs. (For more information about ADNI, including the procedures and time-points not included in this report, please see http://adni-info.org/images/stories/adniproceduresmanual12.pdf).

Methods

Subjects

At the time of the analyses reported here, an ADNI database search indentified 69 mild AD patients, 154 amnestic MCI patients and 79

NCs from about 50 clinical sites who had undergone baseline and twelve-month follow-up FDG PET scans available for downloading from the ADNI Laboratory of Neuroimaging website (www.loni.ucla.edu/ADNI/). The AD patients met NINCDS-ADRDA criteria for probable AD (McKhann et al., 1984), had MMSE scores between 20 and 26, and had Clinical Diagnostic Rating scores (CDR) (Morris, 1993) of 0.5 or 1. The MCI patients had MMSE scores of at least 24 and met revised Petersen criteria for amnestic MCI (Petersen et al., 1999; Winblad et al., 2004), including informant-verified subjective memory complaint, objective memory loss measured by education-adjusted scores on the Logical Memory II subscale (delayed paragraph recall) of the Wechsler Memory Scale-Revised (Wechsler, 1981), a CDR of 0.5, absence of significant impairments in other cognitive domains, preserved activities of daily living (ADLs), and an absence of dementia. The NCs had an MMSE score of at least 24, a CDR of 0, and did not meet criteria for major depression, aMCI, or dementia. All ADNI subjects were 55–90 years old at the time of enrollment, provided their informed consent and were studied under guidelines approved by the human subjects committees at each respective institution. Additional information about ADNI selection criteria can be found at www.adni-info.org (Mueller et al., 2005a).

After the baseline and twelve-month follow-up images were acquired, the ADNI statistical core assigned subjects to independent training and test data sets in order to optimize image-analysis procedures in the training set and then characterize and compare different image modalities and specified image analyses in the independent test set. 40% of the subjects (31 probable AD patients, 65 MCI patients, and 32 NCs who had FDG PET scans) were assigned to the training set; 60% of subjects (38 probable AD patients, 89 MCI patients, and 47 NCs who had FDG PET scans) were assigned to the independent test set; with a comparable proportion of subjects in each subject group (probable AD, MCI, or NC), study arm (1.5 T MRI, 1.5 T MRI + PET, 1.5 T MRI + 3 T MRI), and age range (younger or older than age 76) in each of these data sets. The pre-specified training and test datasets are noted in the ADNI website (http://www.adni-info.org/images/stories/CrossVal/cross-validation.pdf).

Clinical ratings

Clinical ratings acquired at the time of each scan were used to help track the progression of cognitive impairment in each subject. They included the AD Assessment Scale-Cognitive Subscale (ADAS-Cog), CDR Sum of Boxes (CDR-SB) and CDR Global Scores, and MMSE. Other clinical ratings and neuropsychological tests are noted in the ADNI grant application at www.adni-info.org. The ADAS-Cog, a 70-point scale designed to measure the severity of cognitive impairment, is currently the most widely used cognitive measure in AD clinical trials (Rosen et al., 1984); eleven tasks are used to assess learning and memory, language production and comprehension, constructional and ideational praxis, and orientation. CDR-SB scores, ranging from 0 to 18, measure dementia severity by evaluating patients' performance in six domains: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. CDR global scores, ranging from 0 to 3, provide an overall assessment of clinical severity. The MMSE, ranging from 30 to 0, also measures the severity of cognitive impairment; tasks are used to assess orientation, registration, attention and calculation, recall, and language.

FDG PET acquisition and pre-processing

Baseline and twelve-month follow-up FDG PET scans were acquired according to a standardized protocol on 22 different PET scanners (7 GE models, 5 Philips models, and 10 Siemens models) from 45 imaging sites for ADNI participants included in the current study. A 30-min dynamic emission scan, consisting of six 5-min frames, was acquired starting 30 min after the intravenous injection

of 5 mCi of ^{18}F-FDG, as the subjects, who were instructed to fast for at least 4 h prior to the scan, lay quietly in a dimly lit room with their eyes open and with minimal sensory stimulation. Data were corrected for radiation-attenuation and scatter using transmission scans or X-ray CT, and reconstructed using reconstruction algorithms specified for each type of scanner as described at http://www.loni.ucla.edu/ADNI/Data/ADNI_Data.shtml.

Each dynamically acquired image was reviewed and pre-processed by ADNI PET Coordinating Center investigators at the University of Michigan using standardized procedures to identify artifacts and minimize scanner-dependent differences in FDG uptake. During the pre-processing stage, automated algorithms were used to register and average each subject's six 5-min emission frames, transform each subject's registered image into a $160 \times 160 \times 1.5$ mm voxel matrix with sections parallel to a horizontal section through the anterior and posterior commissures (without any adjustment for size or shape), normalize the images for individual variations in absolute image intensity, and apply a filter function previously customized for each scanner using a Hoffmann brain phantom scanned during the site qualification process to ensure an isotropic spatial resolution of 8 mm full-width-at-half-maximum (FWHM). The images were uploaded to the Laboratory of Neuroimaging (LONI) ADNI website at UCLA and ultimately downloaded from the LONI website in NIFTI format by investigators at the Banner Alzheimer's Institute for the analyses reported here.

SPM analyses

After downloading, images underwent further processing using the computer package, SPM (SPM5, http://www.fil.ion.ucl.ac.uk/spm/). The baseline and twelve-month follow-up image pair for each subject was first aligned to each other. They were then linearly and non-linearly transformed to the Montreal Neurological Imaging (MNI) Template, and smoothed using a Gaussian kernel. Regional PET count data was normalized for the variation in measurements from whole brain, an empirically characterized "spared ROI" or other candidate reference ROIs using proportional scaling as described below, and then used to compute statistical maps of twelve-month CMRgl declines. In order to characterize twelve-month CMRgl declines, data from both the training and test data sets were used in each patient group. The training dataset was subsequently used to optimize the settings used to empirically define the procedures associated with maximal twelve-month sROI CMRgl declines in each patient group; and the test dataset was ultimately used to estimate the number of patients needed to detect treatment effects in each patient group in a twelve-month RCT using the empirically pre-defined settings and sROI.

Twelve-month CMRgl declines in each patient group

To characterize twelve-month CMRgl decline in each patient group (Figs. 1A and B), the baseline and twelve-month follow-up images from both the training and test datasets were smoothed with a Gaussian kernel with 12 mm full-width-at-half-maximum (FWHM); whole-brain measurements were computed using the spm_global sub-routine and used to normalize the baseline and twelve-month follow-up scans for the individual variation in whole-brain measurements; and the general linear model (GLM) based simple *t*-test was used to examine the twelve-month CMRgl difference images in each subject group and the independent two-sample *t*-test was used to compare the CMRgl decline difference in each patient group to that in the NCs.

Optimizing sROI settings in each patient group's training dataset

In the training data set, we characterized sROI declines using the combination of a) different FWHM values of the smoothing Gaussian
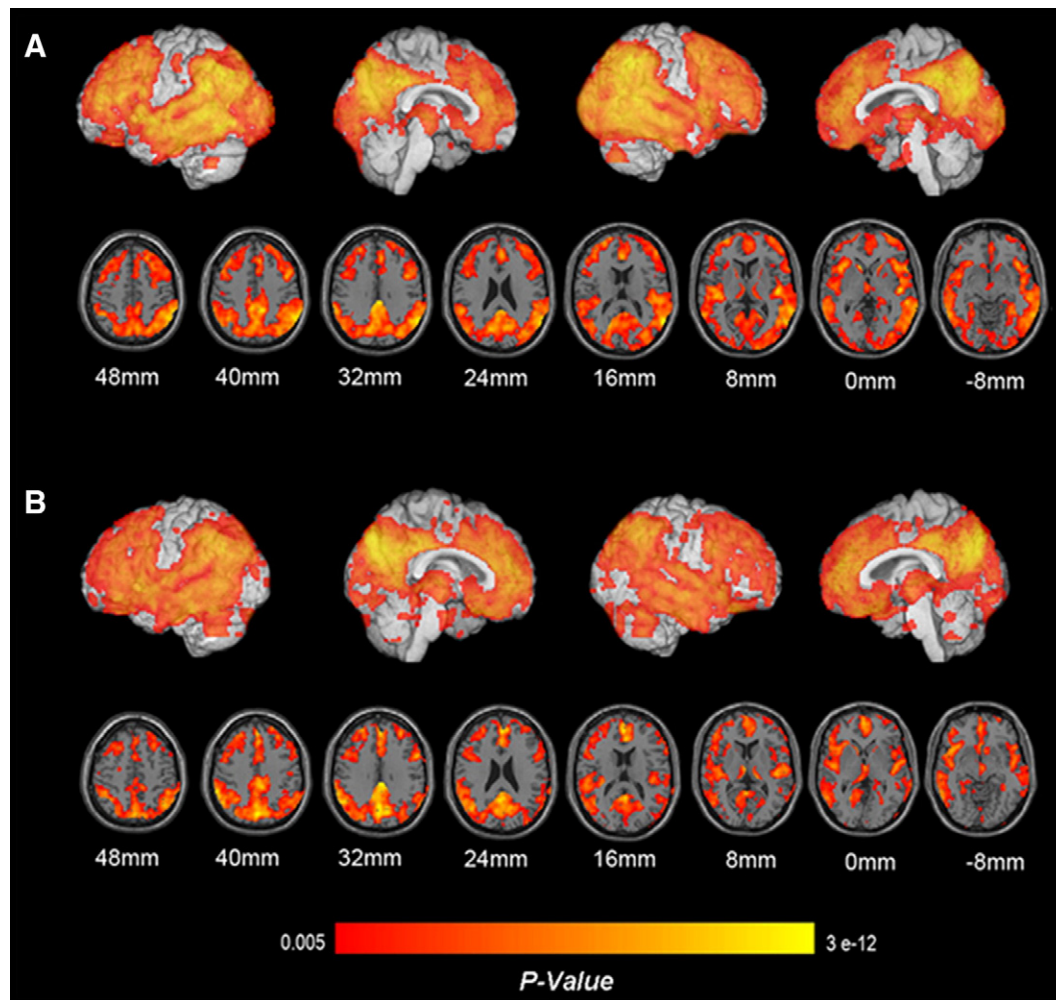
Fig. 1. Statistical brain maps of significant twelve-month CMRgl declines in A) probable AD patients and B) MCI patients ($p < 0.005$, uncorrected for multiple comparisons). The brain maps were generated using baseline and follow-up images acquired in all of the subjects in each patient group (except for those studied on the HRRT and HiRez scanners) and smoothing to a spatial resolution of 12 mm FWHM. Findings are projected onto the lateral and medial surfaces of the left and right cerebral hemispheres and are also shown on horizontal sections at the specified distances above or below a horizontal section between the anterior and posterior commissures; the right side of each section corresponds to the right hemisphere.

kernel; b) measurements from different candidate reference regions to normalize PET images for the individual variation in absolution PET counts; and c) different t-score thresholds to define the cluster of voxels in the candidate sROI. The FWHM values of the Gaussian smoothing kernel, candidate reference regions, and candidate sROI t-score thresholds are described below.

SPM5 voxel-based statistical procedure was repeatedly applied in batch mode to the baseline and twelve-month follow-up images from the probable AD patient and MCI patient training datasets, respectively, to identify the optimal combination of the candidate settings noted above and detailed below to empirically define an sROI consisting of the set of voxels associated with the most significant twelve-month sROI CMRgl declines in each patient group. The optimal FWHM of the smoothing kernel, reference region, and sROI would then be applied to the respective patient group's independent test dataset to estimate the number of patients needed to detect a 25% effect of an AD-slowing treatment (i.e., a 25% slowing of the CMRgl decline in the sROI) with 80% power, two-tailed $\alpha = 0.05$, and no need to correct for multiple regional comparisons in a twelve-month RCT. We introduced this strategy to provide greater statistical power than that using ROIs defined using anatomical landmarks, which may not correspond closely to the set of voxels with maximal CMRgl decline. In contrast to the use of the individual atlas coordinates associated with maximal CMRgl decline, this strategy would also provide a single

endpoint free from the inflated type-I error associated with multiple comparisons.

In the following paragraphs, we describe the procedures used to find the optimal combination of the smoothing Gaussian kernel FWHM, reference region, and sROI significance threshold to characterize twelve-month CMRgl declines using the probable AD and MCI training data sets.

Finding the best normalization factor, best smoothness and best sROI threshold for each patient group

The training data set was used to determine the best reference region to normalize baseline and twelve-month follow-up scans for the variation in FDG PET measurements in each patient group. The candidate regions included whole-brain, pontine, cerebellar, somatosensory, motor, or thalamic regions suggested to be relatively spared in AD patients, and relatively spared sROIs. Whole-brain measurements were automatically defined using the spm_global sub-routine. Pontine ROI measurements were characterized by Dr. Foster's laboratory using NeuroStat software as previously described (Minoshima et al., 1995) and downloaded from the LONI website. Cerebellar ROI measurements were made available using a cerebellum mask provided by Dr. de Leon's laboratory. Somatosensory, motor and thalamic ROI measurements were automatically characterized using the automatic anatomical labeling (AAL) routine in SPM5, along with AAL ROIs 57–58, 1–2, and

77–78, respectively (Tzourio-Mazoyer et al., 2002). Relatively spared sROIs were empirically defined from *t*-score maps as the set of voxels associated with twelve-month regional-to-whole-brain CMRgl increases (which we interpret as relative sparing in comparison with the rest of the brain) using the General Linear Model in SPM5, proportional scaling for the variation in whole-brain measurements, and several different thresholds (0.01, 0.005, 0.001, 0.0005, 0.0001, 0.00005, and 0.00001, uncorrected for multiple comparisons, being irrelevant to the aim of our study). In addition to the scanner specific smoothness already applied to the images downloaded from the LONI website, the patient group's training set was also used to evaluate the effects of applying Gaussian kernels of different FWHM sizes, i.e., 0 mm (no smoothing), 5 mm, 8 mm and 12 mm. Finally, different voxel-level statistical thresholds of $p = 0.01$, 0.005, 0.001, 0.0005, 0.0001, 0.00005, and 0.00001 (uncorrected for multiple comparisons) were applied to define the spatially distributed cluster of voxels associated with twelve-month CMRgl decline in each patient group, and were also evaluated using the training dataset.

SPM5 was repeatedly performed in batch mode using every combination of normalization factors, smoothness and sROI significance thresholds noted above to find the best normalization factor, smoothness and sROI to apply to each patient group's test set, characterize twelve-month sROI CMRgl declines and estimate the number of patients needed to detect AD-slowing treatment effects on these declines with no need to correct for multiple comparisons. Since the trajectory of CMRgl declines in different brain regions is unlikely to be linear during the progression of AD, the approach used in the training set permitted us to customize the selected procedures for the two different patient groups and a twelve-month interval between scans.

Estimating statistical power in each patient group's independent test dataset

The reference regional normalization factor, smoothness and the sROI determined from the training dataset were subsequently used to construct to characterize twelve-month CMRgl declines in the respective patient group's independent test dataset. This sROI was applied to the independent testing datasets to estimate the number of mildly affected probable AD and amnestic MCI patients, respectively, needed per group to detect a 25% AD-slowing treatment effect (i.e., reflecting 25% mean sROI/spared ROI CMRgl decline reduction observed in the treatment group as compared to no decline reduction in the placebo group) with 80% power and two-tailed $\alpha = 0.05$ in a twelve-month, multi-center, double-blind, parallel group, placebo-controlled RCT. Sample size estimates were computed using the standard procedure for two-independent samples (Rosner, 1990) as recommended by the ADNI biostatistical core. They also used the same procedure to characterize and compare sample size estimates using different imaging modalities, image-analysis procedures, and clinical ratings. The sample size estimates were based on twelve-month sROI/spared CMRgl decline ratios in the probable AD group and the MCI group, respectively. We also correlated the sROI CMRgl decline with clinical (ADAS-cog, CDR-SB and MMSE) declines and compared sROI power results to that based on clinical (ADAS-cog, CDR-SB, and MMSE) endpoints. The sROI defined for AD was also used in the NC group testing dataset, permitting us to compare mean sROI declines in each of the three groups using analysis of covariance (ANCOVA). Since an AD-slowing treatment may not be able to attenuate the brain changes associated with normal aging (Fox et al., 2000), we also performed power analyses providing accounting for the mean sROI/spared ROI CMRgl declines associated with "normal aging," using the mean decline in the NC group (i.e., reflecting 25% of the difference between the mean twelve-month sROI/spared ROI CMRgl decline anticipated in the placebo group and the mean decline in the NC group.) Please note that these estimates do not consider any potentially treatable declines in the NC group related to pre-symptomatic AD, and they do not yet consider the variance in the NC group's declines. Since the ADNI PET Core originally identified possible problems in the quality of images acquired on the Siemens HRRT and BioGraph HiRez scanners, the analysis was re-performed in 25 probable AD patients, 74 MCI patients and 36 NCs after excluding the baseline and follow-up images acquired on these scanners. Those findings are reported here, even though new image reconstruction algorithms have been suggested to correct the problem on the BioGraph HiRez scanner and on the HRRT scanner (www.loni.ucla.edu/twiki/bin/view/ADNI/ADNIPETCore).

Results

Subject characteristics

Each subject group's demographic characteristics, proportion of subjects with 0, 1, or 2 copies of the *APOE ε4* allele, baseline clinical ratings, twelve-month clinical declines, and mean (SD) interval between the baseline and approximately twelve-month follow-up scans are shown in Table 1. With the exception of a slightly lower mean educational level in the probable AD group ($p = 0.01$), the probable AD, MCI and NC groups did not differ significantly in their mean age, or sex distribution. As expected from their enrollment criteria, the three groups differed in their mean baseline MMSE, CDR-SB and ADAS-Cog ratings and in their twelve-month changes in these clinical ratings ($p < 0.001$). As expected, the probable AD and MCI groups each had a significantly larger proportion of *APOE ε4* carriers than the NC group ($p = 1.3e-5$). There was little variation and no significant group differences in the time interval between the baseline and approximately twelve-month follow-up scans. (For this reason, twelve-month CMRgl declines are virtually identical whether they are computed as between-session differences as reported here or as the slope in decline over each individual subject's between-scan interval).

Twelve-month CMRgl declines

Statistical brain maps of the twelve-month regional-to-whole-brain CMRgl declines in the probable AD and MCI patient groups are shown in Fig. 1. In contrast to the brain maps in Fig. 2, the maps in Fig. 1 were generated from all of the subjects in the patient group, using images from both the training and test data sets, and after

Table 1
Subject characteristics[a].

	AD ($n = 69$)	MCI ($n = 154$)	NC ($n = 79$)	*p*-value
Age (years)	75.3 (6.9)	74.9 (7.4)	76.0 (4.7)	0.53
Sex (% male/female)	61/39	66/34	66/34	0.77
Educational level (years)	14.7 (3.0)	16.0 (2.9)	15.9 (3.2)	0.01
Baseline MMSE	23.6 (1.9)	27.2 (1.7)	29.0 (1.2)	$2.8e^{-57}$
Baseline ADAS-Cog	18.4 (6.2)	10.9 (4.2)	6.6 (3.0)	$1.2e^{-41}$
Baseline CDR-SB	4.5 (1.5)	1.5 (0.8)	0 (0.1)	$8.8e^{-92}$
12-month MMSE change	−2.6 (3.7)	−0.5 (2.3)	0.2 (1.5)	$1.1e^{-10}$
12-month ADAS-Cog change	3.8 (6.0)	1.0 (4.3)	−0.8 (3.1)	$5.4e^{-9}$
12-month CDR-SB change	1.2 (2.0)	0.5 (1.2)	0 (1.1)	$4.1e^{-7}$
Between-scan interval (months)	12.0 (0.6)	12.12 (0.8)	12.00 (0.8)	0.45
APOE ε4 gene dose (% with 0/1/2 alleles)	30.4/52.2/ 17.4	48.7/38.3/ 13.0	72.2/25.3/ 2.5	$1.3e^{-5b,c}$

Group differences in sex and *APOE ε4* gene dose distributions were analyzed using Fisher's exact Chi-square (χ^2) tests. Scheffé-multiple comparison test was used to compare the differences between each pair of means.

Group differences in other baseline measurements and twelve-month changes were analyzed using ANOVA.

MMSE = Mini-Mental State Exam; ADAS-Cog = AD Assessment Scale-Cognitive Subscale; CDR-SB = Clinical Dementia Rating Sum of Boxes.

[a] Unless otherwise indicated, data are mean (SD).
[b] *APOE ε4* gene dose significantly differs between AD and NC.
[c] *APOE ε4* gene dose significantly differs between MCI and NC.

smoothing them to a spatial resolution of 12 mm FWHM. There was an extensive cluster of voxels associated with regional-to-whole-brain CMRgl decline in each patient group, including 38,181 voxels ($p = 1.5e$–89, uncorrected for multiple comparisons) in the probable AD group and 23,691 voxels ($p = 9.5e$–70) in the MCI patient group. The atlas coordinates, magnitude, spatial extent, and the associated significances in each patient group are shown in Table 2. The probable AD and MCI groups each had significant twelve-month CMRgl declines bilaterally in posterior cingulate, precuneus, medial parietal, lateral parietal, medial temporal, frontal and occipital cortex ($p < 0.001$, uncorrected for multiple comparisons). In each of the brain regions the decline in AD patients were significantly greater than those in the NC group ($p < 0.001$, uncorrected for multiple comparisons). Greater declines in MCI patients, when compared to the NC group, were also observed in precuneus, temporal regions ($p < 0.001$, uncorrected for multiple comparisons). (Twelve-month CMRgl declines in each subject group's APOE ε4 carrier and non-carrier sub-groups will be described in a separate report).

Settings associated with maximal sROI CMRgl decline in the training dataset and sample size estimates in the independent test dataset

Using the training dataset to determine the setting associated with optimal CMRgl decline, the following settings were identified as optimal for the AD group a) a spatial resolution of 8 mm FWHM (i.e.,

no additional smoothing), b) a spared region defined by a significance threshold of $p < 0.0005$, consisting of 6665 voxels associated with twelve-month regional-to-whole-brain CMRgl increases and located primarily in white matter and cerebellum, and c) a decline sROI defined by a significance threshold of $p < 0.0005$, consisting of 32,807 voxels associated with twelve-month regional-to-spared sROI decline and located in the same regions known to be preferentially affected in cross-sectional and longitudinal brain-mapping studies of AD. For the MCI group, the optimal settings were a) a spatial resolution of 8 mm FWHM (i.e., no additional smoothing), b) a spared sROI located primarily in white matter and somatosensory cortex and consisted of 2454 voxels associated with twelve-month regional-to-whole-brain CMRgl increases, as determined by a significance threshold of $p < 0.0001$, and c) a decline sROI defined by a significance threshold of $p < 0.000001$, consisting of 9212 voxels associated with twelve-month regional-to-spared sROI decline and located in the same regions known to be preferentially affected in cross-sectional and longitudinal brain-mapping studies of AD.

Statistical brain maps of the implicated spared region (in the blue color scale) and decline sROI (in the red-to-yellow color scale) in the probable AD and MCI patient groups are shown in Fig. 2. In contrast to the brain maps in Fig. 1, these maps were generated from only those subjects in the training group, using the extracted values from the spared region as the normalization factor and with no additional smoothing. The sROIs associated with twelve-month CMRgl decline in
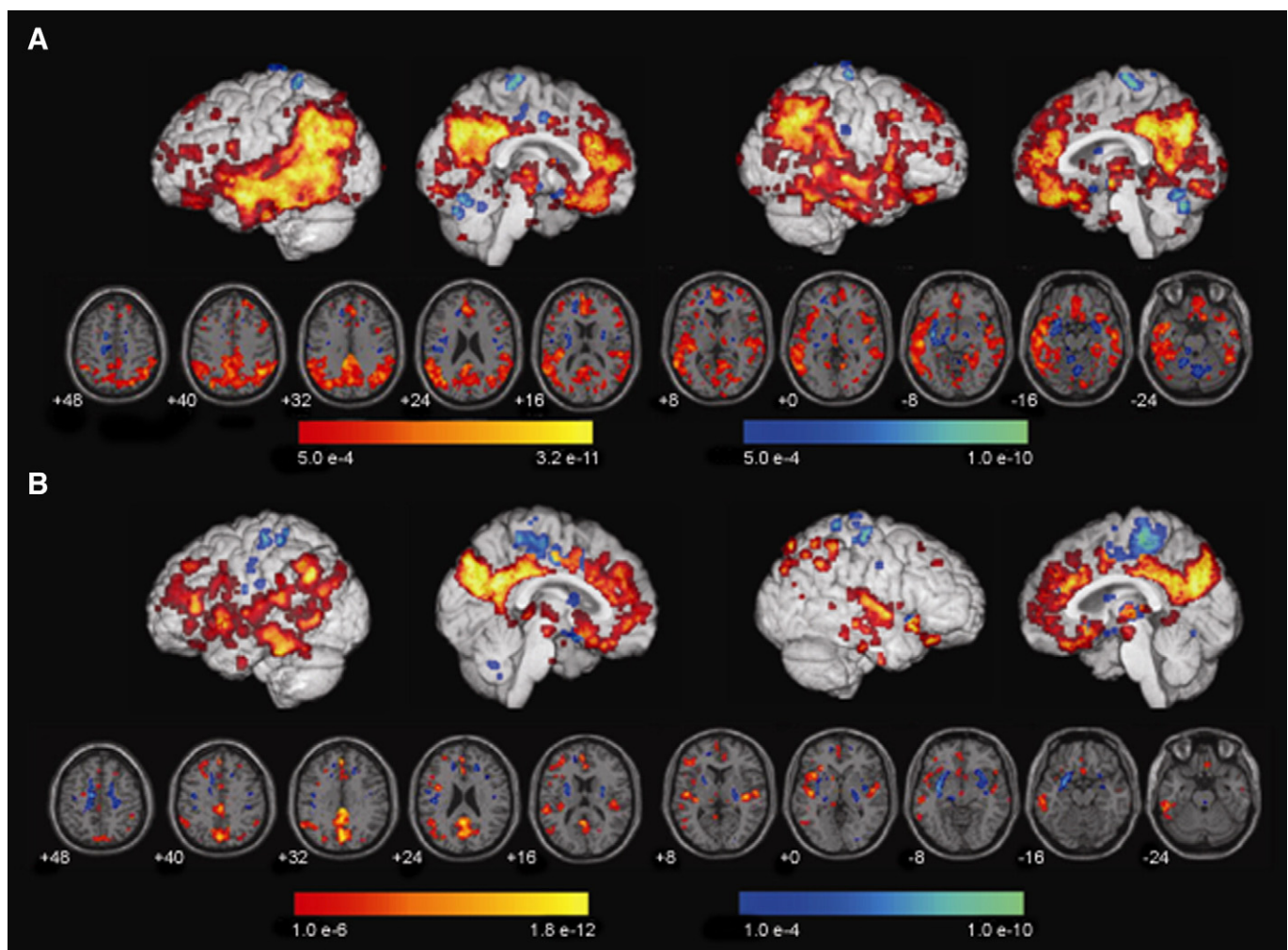

Fig. 2. Statistical brain maps of the spared sROI (in the blue color scale) and twelve-month CMRgl decline sROI (in the red-to-yellow color scale) in A) probable AD patients and B) MCI patient groups. The brain maps were generated using baseline and follow-up images acquired in the each patient group's training set (except for those studied on the HRRT and HiRez scanners) using images with a spatial resolution of 8 mm FWHM. The spared and decline sROIs in each group were subsequently used to estimate the number of patients needed to test AD-slowing treatments in the independent test set, shown in Tables 3 and 4.

Table 2
Location and magnitude of the most significant twelve-month regional-to-whole-brain CMRgl decline.

Brain region		Atlas coordinates[a]			T value	Voxel p[b] uncorrected	Extent (voxels)	Cluster p uncorrected
		X	Y	Z				
AD (n = 69)								
Posterior cingulate		−2	−37	28	7	5.9×10^{-10}	310	0.004
Precuneus	Left	0	−56	40	6.4	9.9×10^{-9}	1763	2.2×10^{-8}
	Right	8	−41	39	7.5	9.4×10^{-11}	1916	8.1×10^{-9}
Parietal	Left	−46	−46	47	6.1	3.1×10^{-8}	1683	3.8×10^{-8}
	Right	57	−50	39	7.6	6.2×10^{-11}	1671	4.1×10^{-8}
Temporal	Left	−59	−53	−8	6.7	2.5×10^{-9}	6044	4.0×10^{-18}
	Right	61	−53	−4	6.5	6.2×10^{-9}	8103	7.0×10^{-22}
Frontal	Left	−32	55	5	6.4	9.6×10^{-9}	5146	2.3×10^{-16}
	Right	30	18	51	6.6	3.9×10^{-9}	7179	3.0×10^{-20}
Medial temporal	Left	−44	−55	−11	4.1	5.6×10^{-5}	518	4.1×10^{-4}
	Right	32	−41	−8	5.8	9.5×10^{-8}	1191	1.3×10^{-6}
Occipital	Left	−28	−70	37	5.3	7.4×10^{-7}	2636	9.8×10^{-11}
	Right	32	−74	33	6.2	2.3×10^{-8}	3437	1.1×10^{-12}
MCI (n = 154)								
Posterior cingulate		−2	−37	28	6.7	1.3×10^{-10}	308	0.003
Precuneus	Left	−2	−60	36	7.4	5.4×10^{-12}	1566	2.3×10^{-8}
	Right	12	−65	29	6.1	3.7×10^{-9}	1438	6.0×10^{-8}
Parietal	Left	−59	−39	39	6.9	7.1×10^{-11}	1561	2.3×10^{-8}
	Right	48	−54	54	6.3	1.5×10^{-9}	1293	1.9×10^{-7}
Temporal	Left	−50	−50	−21	6.7	2.2×10^{-10}	5966	2.2×10^{-19}
	Right	61	−32	−22	5.4	1.5×10^{-7}	4347	7.8×10^{-16}
Frontal	Left	−28	50	23	5.5	9.6×10^{-8}	5449	2.7×10^{-18}
	Right	30	40	31	5.8	2.3×10^{-8}	3159	6.2×10^{-13}
Medial temporal	Left	−22	−34	12	3.3	6.2×10^{-4}	308	0.003
	Right	36	−43	−5	3.95	5.9×10^{-5}	75	0.098
Occipital	Left	−28	−74	30	5.2	3.8×10^{-7}	1320	1.5×10^{-7}
	Right	32	−70	33	4.6	4.4×10^{-6}	859	7.5×10^{-6}

[a] Atlas coordinates for the maximal CMRgl reductions were obtained from Talairach and Tournoux (1988). X is the distance to the right (+) or left (−) of the midline, Y is the distance anterior (+) or posterior (−) to the anterior commissure, and Z is the distance superior (+) or inferior (−) to a horizontal plane through the anterior and posterior commissures.

[b] The reported significance levels in this column are uncorrected for multiple comparisons.

the probable AD and MCI groups were similar but not identical in spatial distribution, and the magnitude of respective sROI-to-spared CMRgl reductions was greater in the probable AD group than in the MCI group.

As shown in Fig. 3, the mean twelve-month sROI-to-spared CMRgl declines were significantly different in the three subject groups (probable AD>MCI>NC) using the decline and spared sROIs defined in either the probable AD group (Fig. 3A) or the MCI group (Fig. 3B) (ANOVA $p < 0.001$).

In the probable AD group, individual twelve-month sROI-to-spared declines were correlated with clinical declines on the CDR-SB ($r = −0.25$, $p = 0.04$), but not on the MMSE ($r = 0.23$, $p = 0.06$) or

ADAS-Cog ($r = −0.14$, $p = 0.26$). In the MCI group, individual twelve-month sROI-to-spared sROI declines were correlated with clinical declines on the CDR-SB ($r = −0.19$, $p = 0.02$) and the MMSE ($r = 0.22$, $p = 0.0059$), but not on the ADAS-Cog ($r = −0.03$ $p = 0.69$).

The estimated number of patients per group needed to detect 25% AD-slowing treatment effects in twelve-month placebo-controlled RCTs with 80% power and two-tailed $\alpha = 0.05$ using FDG PET and our sROI strategy to characterize regional CMRgl declines and CDR-SB, ADAS-Cog and MMSE scores to characterize clinical decline are shown for the probable AD group in Table 3 and for the MCI group in Table 4. Estimates are based on the independent test dataset, with and without inclusion of the HRRT and HiRez scanners, and without any
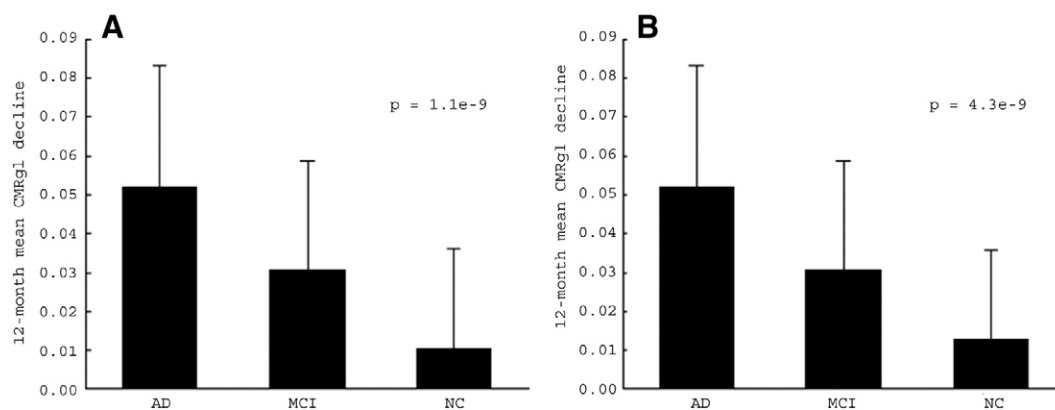

Fig. 3. Magnitude and standard deviation of twelve-month CMRgl-to-spared sROI decline in the probable AD, MCI, and NC groups using A) the probable AD sROI and B) the MCI sROI. The ROIs were empirically pre-defined using the relevant patient group's data set and the magnitude of twelve-month decline-to-spared CMRgl decline in these sROIs were characterized in the independent test dataset from the same patient group. In each comparison, the twelve-month decline-to-spared CMRgl declines were significantly different in the three subject groups (AD>MCI>NC).

Table 3
Estimated number of probable AD patients per treatment group needed to detect a 25% treatment effect with a two-sided test, $\alpha = 0.05$ and 80% power in a twelve-month placebo-controlled RCT.

	All scanners	Excluding HRRT and HiRez
sRsROI	92	66
CDR-SB	918	861
ADAS-Cog	505	353
MMSE	563	629

correction for the mean changes associated with normal aging. Using FDG PET, SPM5 and our empirically pre-defined spared and decline sROI's and no correction for normal aging effects, we estimate the need for about 92 probable AD patients (estimated CMRgl decline in AD 0.0514 ± 0.0309) or about 226 MCI patients (estimated CMRgl decline in MCI 0.0293 ± 0.0279) per group in a twelve-month RCT, a fraction of those needed using the clinical endpoints included in our analysis. By comparison, we estimate the need for about 7–10 times as many probable AD patients and about 6–26 times as many MCI patients in the same RCTs using those clinical endpoints. When the normal aging effects, estimated from the NC group (CMRgl mean decline 0.009), were figured into the power estimates, we estimate the need for about 138 probable AD patients or about 475 MCI patients per group in a twelve-month RCT. While the estimated number of needed patients was lower when the HRRT and HiRez scanners were excluded from the analysis (Tables 3 and 4), the sample size estimates were similar.

Discussion

In this study, we used FDG PET images from ADNI to demonstrate twelve-month CMRgl declines in probable AD and MCI. Using SPM5 to analyze FDG PET data from this unusually large multi-center study, mild probable AD patients and amnestic MCI patients each had twelve-month regional-to-whole-brain CMRgl declines bilaterally in posterior cingulate, medial and lateral parietal, medial and lateral temporal, and occipital regions. We also used these data to introduce the use of an empirically pre-defined sROI, in this case composed of the set of voxels consistently associated with CMRgl decline in an independent training data set, to evaluate AD-slowing treatment effects with improved statistical power and freed from the multiple regional comparison concerns. Using SPM5 and the sROI empirically pre-defined in each patient group's training data set to compute the statistical power of FDG PET to detect AD-slowing treatment effects in the respective patient group's test data set, we estimated the need for about 92 probable AD patients or about 226 MCI patients per group, or 66/217 probable AD/MCI patients when excluding data acquired from HRRT and Biograph HiRez scanners, to detect a 25% AD-slowing treatment effect with 80% power, and two-tailed $\alpha = 0.05$ in a twelve-month parallel group, placebo-controlled multi-center RCT, a fraction of the estimated numbers needed using different clinical endpoints. While this report describes the use of different significance thresholds to empirically pre-define the sROI in the training set, we have also used bootstrap resampling, along with the percentage of bootstrap

Table 4
Estimated number of MCI patients per treatment group needed to detect a 25% treatment effect with a two-sided test, $\alpha = 0.05$ and 80% power in a twelve-month placebo-controlled RCT.

	All scanners	Excluding HRRT and HiRez
sROI	227	218
CDR-SB	1448	1110
ADAS-Cog	4026	4219
MMSE	7245	5267

analyses with decline at different significance thresholds to define the sROI with quite similar results (Reiman et al., 2008). The comparability between the computationally labor intensive bootstrap with resampling method and the significance threshold method will be described in a separate methodological report.

Based on the publicly available analysis of test set data using different imaging modalities and data-analysis techniques on a subset of the individuals included in this analysis (http://www.adni-info.org/index.php?option=com_content&task=view&id=89&Itemid=44), the estimated number of probable AD or MCI patients needed in twelve-month RCTs was smallest in magnitude using FDG PET and our sROI method across the candidate FDG PET markers. As documented on the ADNI website (http://www.adni-info.org/index.php?option=com_content&task=view&id=89&Itemid==44), the estimated statistical power to detect an AD-slowing treatment effect using FDG PET was significantly better using measurements in our sROI than the estimates using other methods, including ROIs defined using a meta-analysis of the maximal CMRgl reductions in previous studies of AD. In comparisons with MRI measures, the estimates using FDG PET and our sROI method were comparable to those obtained from some FreeSurfer volumetric ROIs, boundary shift integral techniques, and tensor-based morphometry measures (Ho et al., 2009; Hua et al., 2009).

After defining sROIs to characterize the set of voxels associated with the twelve-month CMRgl decline and CMRgl sparing using the training set data from each patient group, twelve-month decline-to-spared CMRgl ratios were found to be significantly associated with categorical measurements of clinical disease severity (i.e., AD>MCI>NC), correlated with twelve-month declines in some but not all of their clinical ratings (e.g., CDR-SB but not ADAS-Cog), and associated with about one-tenth the number of patients needed to detect AD-slowing twelve-month treatment effects in twelve-month multi-center RCTs in comparison to these clinical endpoints. Each of these findings supports the value of FDG PET and this image-analysis technique in AD and MCI RCTs, and we have reason to believe that they may have even greater value in evaluating pre-symptomatic AD-slowing treatments in cognitively normal subjects at increased risk of AD (Reiman and colleagues, unpublished data) due in part to the unusually large samples and treatment durations needed to do so using clinical endpoints.

Our proposed sROI strategy has several advantages for assessing AD-slowing treatment effects in RCTs. 1) *Statistical power*: since the sROI consists of the voxels most consistently associated with CMRgl decline in the relevant patient group during the relevant between-scan interval it is likely to be associated with greater statistical power than ROIs defined using anatomical or other landmarks, which may not capture the brain regions associated with this decline due to their size or location. 2) *Freedom from the problem of multiple regional comparisons*: since the sROI provides a single endpoint for the imaging modality of interest, and since it is empirically pre-defined using an independent data set prior to the performance of an RCT, this strategy would not require any statistical correction for multiple regional comparisons and, we would argue, is likely to be accepted by regulatory agencies in future pivotal trials. By comparison, the use of multiple ROIs would require statistical correction for multiple comparisons, and the use of a voxel/brain atlas coordinate associated with maximal CMRgl declines in an independent data set would require a trade-off between the size of the search region (which would be needed to be sufficiently large to capture a treatment effect) and the number of regional comparisons in that search region. 3) *Face validity*: as our study shows—and just as one would predict—the sROI associated with twelve-month CMRgl declines in both the AD and MCI patient groups corresponds well to the brain regions implicated in both cross-sectional and longitudinal FDG PET studies of AD (Alexander et al., 2002; de Leon et al., 1983; Foster et al., 1983; Haxby et al., 1990; Jagust et al., 1988; Langbaum et al., 2009; Minoshima et al., 1994; Mosconi et al., 2005; Mosconi et al., 2009), as well as the pattern of synaptic loss (the best predictor of clinical

decline) in neuropathological studies (Selkoe, 2002). Thus, sROI CMRgl decline meets one of the regulatory agency requirements for a surrogate endpoint (i.e., the measurement should reflect a process in the disease pathway and is likely to be relevant to a patient's clinical course). 4) *Customizability*: since the trajectory of longitudinal changes in regional brain-imaging measurements may be non-linear (e.g., relatively early CMRgl declines in posterior cingulate and precuneus, which may level off in the later clinical stages of AD and relatively late CMRgl declines in frontal cortex, which may be most strongly correlated with clinical disease severity in the more severe stages of dementia), the sROI can be customized to the patient group and treatment duration proposed in the RCT, using data from a comparable subject group (e.g., in the study of more severely affected AD patients) and treatment duration as the training set needed to empirically pre-define the sROI. ADNI provides the opportunity to customize the sROI and power estimates to different patient groups (e.g., more severely affected AD patients or MCI patients with significant fibrillar Aβ burden) and treatment durations. While our preliminary findings suggest that roughly the same number of mild probable AD or MCI patients may be needed for 18- and 24-month clinical trials whether the sROI is defined using the baseline and 12-, 18-, or 24-month follow-up scans, respectively, some of the patients in this study have not yet completed their 24-month follow-up scans. To be clear, we would recommend empirically pre-defining the sROI from the most appropriate longitudinal data set (e.g., from ADNI, another longitudinal study, or placebo group data from a prior clinical trial) prior to the design or analysis of data from the clinical trial of interest. We do *not* recommend using any of the data from the clinical trial of interest to define the sROI itself. 5) *Reproducibility*: the estimated number of patients needed to evaluate AD-slowing treatment effects using the test data set was only modestly (and not significantly) greater than the estimate using the training data set. It will be important to extend our findings to other patient groups, including the placebo groups assessed using FDG PET in ongoing RCTs. 6) *Generalizability to other imaging modalities and voxel-based data-analysis techniques*: as noted in the next paragraph, the sROI strategy described here has the potential to be used using other imaging modalities and voxel-based data-analysis techniques.

Researchers from Dr. Paul Thompson's laboratory used the strategy reported here to characterize twelve-month brain shrinkage in an empirically pre-defined sROI using Tensor Based Morphometry (TBM) (Hua et al., 2009). The sROI was empirically pre-defined in ADNI's training data set to include a set of voxels in temporal cortex associated with the most significant brain shrinkage in the training data set from ADNI. When the sROI was applied to the independent test set to estimate, they estimated that as few as 48 probable AD patients or 88 MCI patients would be needed to detect a 25% AD-slowing effect (with 80% power, a two-sided test, and $\alpha = 0.05$), a fraction of the number needed to using clinical endpoints. In a follow-up analysis also using the sROI strategy (Ho et al., 2009), sample size estimates were found to be comparable for 1.5 T and 3 T MRI, in a study of 110 patients scanned longitudinally at both field strengths. Our proposed sROI strategy may be applicable to a wide range of imaging modalities, voxel-based image-analysis techniques, and clinical studies, offering a single imaging endpoint and better statistical power than other regional measurements.

While we found associations between twelve-month declines and sROI/spared ROI CMRgl and certain clinical ratings, the correlations were relatively modest. One possible explanation for the relatively modest correlations may be the relatively small magnitude and large variance in the assessed measurements of clinical decline during this time frame. Indeed, we have previously shown 24-month CMRgl declines in cognitively normal late-middle-aged APOE ε4 carriers in the absence of any significant clinical or neuropsychological declines. One might ask what implications this dissociation may have for the future use of these measurements as "reasonably likely surrogate

endpoints" for the accelerated regulatory agency approval of treatments based solely on biomarkers alone (Reiman and Langbaum, 2009)? In order for biomarkers endpoints to serve as reasonably likely surrogate endpoints in clinical trials, regulatory agencies may ultimately require evidence from other clinical trials that an AD-slowing treatment's effects on one or more biomarkers predict a clinical benefit. We have proposed strategies to demonstrate the relationship between a treatment's short-term biomarker effects and its longer-term cognitive and clinical effects to help provide this kind of evidence, even in assessment of pre-symptomatic AD treatments (Reiman and Langbaum, 2009); (Reiman et al., 2010).

Still, the present study has several limitations, which need to be addressed in future studies. First, our findings were derived from relatively small training and testing data sets, so the sROIs associated with CMRgl decline and sparing in each subject group should be confirmed in independent data sets, including the placebo data sets acquired in ongoing RCTs, extended to other pre-symptomatic and clinical stages of AD, customized to the selection criteria and treatment duration being considered in a RCT, and using clinical or population-based samples that optimally address the scientific question at hand. Second, since our training data set analyses were confined to SPM5, selected smoothing settings, and selected sROI significance thresholds, findings from this study could be extended to other voxel-based image-analysis techniques (perhaps using different registration methods) for a more exhaustive comparison of image-analysis settings. However, as we will report in a separate article, the sROI can be pre-defined using either significance thresholds or a more time-consuming bootstrap with replacement strategy, with roughly comparable results. Since the significance thresholding strategy was comparable and easier for other groups to replicate, only findings using this approach are reported here. Third, it remains to be seen whether pharmaceutical companies and regulatory agencies would accept the idea of customizing the sROI to the selection criteria and treatment duration to be used in an RCT, even though the sROI would again be empirically pre-defined prior to the trial. Finally, power estimates remain to be computed in the sub-group of probable AD or MCI patients defined using ε4 carrier status, evidence of amyloid-β pathology, or other biomarker measurements that might be used to enrich a clinical trial for those individuals most likely to demonstrate AD-related decline.

This study also has limitations that apply to the suitability of any imaging modality and any image-analysis technique as a surrogate endpoint in future RCTs. We do not yet know the extent to which different AD-slowing treatments budge different brain-imaging or other biomarker measurements, the extent to which an AD-slowing treatment's effects on one or more of these biomarkers predicts a clinical benefit at different clinical and pre-symptomatic stages of AD, the extent to which the treatment might have a confounding biomarker measurement (e.g., an effect on brain activity, synaptic activity, or brain swelling) independent of an AD-slowing effect. For these reasons, it will be important to embed the range of imaging and non-imaging biomarkers in RCTs of putative AD-slowing treatments, not just to help evaluate the treatment at hand, but to provide the data needed to find one or more biomarkers that meet regulatory agency criteria as surrogate endpoints for the accelerated approval of treatments in the earliest symptomatic and pre-symptomatic stages of the disease.

References

Alexander, G.E., Chen, K., Pietrini, P., Rapoport, S.I., Reiman, E.M., 2002. Longitudinal PET evaluation of cerebral metabolic decline in dementia: a potential outcome measure in Alzheimer's disease treatment studies. Am. J. Psychiatry 159, 738–745.

Bookheimer, S.Y., Strojwas, M.H., Cohen, M.S., Saunders, A.M., Pericak-Vance, M.A., Mazziotta, J.C., Small, G.W., 2000. Patterns of brain activation in people at risk for Alzheimer's disease. N. Engl. J. Med. 343, 450–456.

Brookmeyer, R., Gray, S., Kawas, C., 1998. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. Am. J. Public Health 88, 1337–1342.

de Leon, M.J., Ferris, S.H., George, A.E., Reisberg, B., Christman, D.R., Kricheff, I.I., Wolf, A.P., 1983. Computed tomography and positron emission transaxial tomography evaluations of normal aging and Alzheimer's disease. J. Cereb. Blood Flow Metab. 3, 391–394.

Doraiswamy, P.M., Farmer, M., Holub, R., Johnson, K., Reiman, E.M., Sadowsky, C., Safirstein, B., Spicer, K., Carpenter, A., Clark, C.M., Joshi, A., Veeraraj, C., Pontecorvo, M.J., Skovronsky, D., 2009. Relationship between regional amyloid levels and cognitive performance in healthy controls, MCI subjects, and patients with AD: phase II results from a florpiramine F18 PET imaging study. Alzheimer's Dement. 5, P77.

Drzezga, A., Lautenschlager, N., Siebner, H., Riemenschneider, M., Willoch, F., Minoshima, S., Schwaiger, M., Kurz, A., 2003. Cerebral metabolic changes accompanying conversion of mild cognitive impairment into Alzheimer's disease: a PET follow-up study. Eur. J. Nucl. Med. Mol. Imaging 30, 1104–1113.

Evans, D.A., Funkenstein, H.H., Albert, M.S., Scherr, P.A., Cook, N.R., Chown, M.J., Hebert, L.E., Hennekens, C.H., Taylor, J.O., 1989. Prevalence of Alzheimer's disease in a community population of older persons. Higher than previously reported. JAMA 262, 2551–2556.

Fagan, A.M., Mintun, M.A., Mach, R.H., Lee, S.Y., Dence, C.S., Shah, A.R., LaRossa, G.N., Spinner, M.L., Klunk, W.E., Mathis, C.A., DeKosky, S.T., Morris, J.C., Holtzman, D.M., 2006. Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid Ab42 in humans. Ann. Neurol. 59, 512–519.

Fagan, A.M., Roe, C.M., Xiong, C., Mintun, M.A., Morris, J.C., Holtzman, D.M., 2007. Cerebrospinal fluid tau/b-amyloid42 ratio as a prediction of cognitive decline in nondemented older adults. Arch. Neurol. 64, 343–349.

Fagan, A.M., Head, D., Shah, A.R., Marcus, D., Mintun, M., Morris, J.C., Holtzman, D.M., 2009. Decreased cerebrospinal fluid Abeta(42) correlates with brain atrophy in cognitively normal elderly. Ann. Neurol. 65, 176–183.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198.

Foster, N.L., Chase, T.N., Fedio, P., Patronas, N.J., Brooks, R.A., Di Chiro, G., 1983. Alzheimer's disease: focal cortical changes shown by positron emission tomography. Neurology 33, 961–965.

Fox, N.C., Warrington, E.K., Rossor, M.N., 1999. Serial magnetic resonance imaging of cerebral atrophy in preclinical Alzheimer's disease. Lancet 353, 2125.

Fox, N.C., Cousens, S., Scahill, R., Harvey, R.J., Rossor, M.N., 2000. Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease: power calculations and estimates of sample size to detect treatment effects. Arch. Neurol. 57, 339–344.

Hansson, O., Zetterberg, H., Buchhave, P., Londos, E., Blennow, K., Minthon, L., 2006. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. Lancet Neurol. 5, 228–234.

Hansson, O., Buchhave, P., Zetterberg, E., Blennow, K., Minthon, L., Warkentin, S., 2009. Combined rCBF and CSF biomarkers predict progression from mild cognitive impairment to Alzheimer's disease. Neurobiol. Aging 30, 165–173.

Haxby, J.V., Grady, C.L., Koss, E., Horwitz, B., Heston, L., Schapiro, M., Friedland, R.P., Rapoport, S.I., 1990. Longitudinal study of cerebral metabolic asymmetries and associated neuropsychological patterns in early dementia of the Alzheimer type. Arch. Neurol. 47, 753–760.

Hebert, L.E., Beckett, L.A., Scherr, P.A., Evans, D.A., 2001. Annual incidence of Alzheimer disease in the United States projected to the years 2000 through 2050. Alzheimer Dis. Assoc. Disord. 15, 169–173.

Herholz, K., Salmon, E., Perani, D., Baron, J.C., 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., W.D., 2002. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. Neuroimage 17, 302–316.

Ho, A.J., Hua, X., Lee, S., Leow, A.D., Yanovsky, I., Gutman, B., Dinov, I.D., Lepore, N., Stein, J.L., Toga, A.W., Jack, C.R., Jr., Bernstein, M.A., Reiman, E.M., Harvey, D.J., Kornak, J., Schuff, N., Alexander, G.E., Weiner, M.W., Thompson, P.M., 2009. Comparing 3 T and 1.5 T MRI for tracking Alzheimer's disease progression with tensor-based morphometry. Human Brain Mapping. doi:10.1002/hbm.20882.

Hua, X., Lee, S., Yanovsky, I., Leow, A.D., Chou, Y.Y., Ho, A.J., Gutman, B., Toga, A.W., Jack Jr., C.R., Bernstein, M.A., Reiman, E.M., Harvey, D., Kornak, J., Schuff, N., Alexander, G.E., Weiner, M.W., Thompson, P.M., 2009. Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: an ADNI study of 515 subjects. Neuroimage 48, 668–681.

Jack Jr., C.R., Lowe, V.J., Senjem, M.L., Weigand, S.D., Kemp, B.J., Shiung, M.M., Knopman, D.S., Boeve, B.F., Klunk, W.E., Mathis, C.A., Petersen, R.C., 2008. 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. Brain 131, 665–680.

Jack Jr., C.R., Lowe, V.J., Weigand, S.D., Wiste, H.J., Senjem, M.L., Knopman, D.S., Shiung, M.M., Gunter, J.L., Boeve, B.F., Kemp, B.J., Weiner, M., Petersen, R.C., the Alzheimer's Disease Neuroimaging Initiative, 2009. Serial PIB and MRI in normal, mild cognitive impairment and Alzheimer's disease: implications for sequence of pathological events in Alzheimer's disease. Brain 132, 1355–1365.

Jagust, W.J., Friedland, R.P., Budinger, T.F., Koss, E., Ober, B., 1988. Longitudinal studies of regional cerebral metabolism in Alzheimer's disease. Neurology 38, 909–912.

Johnson, A.E., Jeppsson, F., Sandell, J., Wensbo, D., Neelissen, J.A., Jureus, A., Strom, P., Norman, H., Farde, L., Svensson, S.P., 2009. AZD2184: a radioligand for sensitive detection of beta-amyloid deposits. J. Neurochem. 108, 1177–1186.

Jones, R.W., Schwam, E., Wilkinson, D., Waldemar, G., Feldman, H.H., Zhang, R., Albert, K., Schindler, R., 2009. Rates of cognitive change in Alzheimer disease: observations across a decade of placebo-controlled clinical trials with donepezil. Alzheimer Dis. Assoc. Disord. 23, 357–364.

Klunk, W.E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D.P., Bergstrom, M., Savitcheva, I., Huang, G.F., Estrada, S., Ausen, B., Debnath, M.L., Barletta, J., Price, J.C., Sandell, J., Lopresti, B.J., Wall, A., Koivisto, P., Antoni, G., Mathis, C.A., Langstrom, B., 2004. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Ann. Neurol. 55, 306–319.

Landau, S.M., Harvey, D., Madison, C.M., Koeppe, R.A., Reiman, E.M., Foster, N.L., Weiner, M.W., Jagust, W.J., 2009. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. Neurobiol. Aging. doi:10.1016/j.neurobiolaging.2009.07.002.

Langbaum, J.B.S., Chen, K., Lee, W., Reschke, C., Bandy, D., Fleisher, A.S., Alexander, G.E., Foster, N.L., Weiner, M.W., Koeppe, R.A., Jagust, W.J., Reiman, E.M., 2009. Categorical and correlational analyses of baseline fluorodeoxyglucose positron emission tomography images from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Neuroimage 45, 1107–1116.

Li, G., Sokal, I., Quinn, J.F., Leverenz, J.B., Brodey, M., Schellenberg, G.D., Kaye, J.A., Raskind, M.A., Zhang, J., Peskind, E.R., Montine, T.J., 2007. CSF tau/Abeta42 ratio for increased risk of mild cognitive impairment: a follow-up study. Neurology 69, 631–639.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E.M., 1984. 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 34, 939–944.

Mielke, R., Herholz, K., Grond, M., Kessler, J., Heiss, W.D., 1994. Clinical deterioration in probable Alzheimer's disease correlates with progressive metabolic impairment of association areas. Dementia 5, 36–41.

Minoshima, S., Foster, N.L., Kuhl, D.E., 1994. Posterior cingulate cortex in Alzheimer's disease. Lancet 344, 895.

Minoshima, S., Frey, K.A., Foster, N.L., Kuhl, D.E., 1995. Preserved pontine glucose metabolism in Alzheimer's disease: a reference region for functional brain image (PET) analysis. J. Comput. Assist. Tomogr. 19, 541–547.

Morris, J.C., 1993. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 43, 2412–2414.

Mosconi, L., Tsui, W.H., De, S.S., Li, J., Rusinek, H., Convit, A., Li, Y., Boppana, M., de Leon, M.J., 2005. Reduced hippocampal metabolism in MCI and AD: automated FDG-PET image analysis. Neurology 64, 1860–1867.

Mosconi, L., Mistur, R., Switalski, R., Tsui, W.H., Glodzik, L., Li, Y., Pirraglia, E., De, S.S., Reisberg, B., Wisniewski, T., de Leon, M.J., 2009. FDG-PET changes in brain glucose metabolism from normal cognition to pathologically verified Alzheimer's disease. Eur. J. Nucl. Med. Mol. Imaging 36, 811–822.

Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C.R., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L., 2005a. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). Alzheimers Dement. 1, 55–66.

Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L., 2005b. The Alzheimer's disease neuroimaging initiative. Neuroimaging Clin. N. Am. 15, 869–877 xi.

Nyberg, S., Jonhagen, M.E., Cselenyi, Z., Halldin, C., Julin, P., Olsson, H., Freund-Levi, Y., Andersson, J., Varnas, K., Svensson, S., Farde, L., 2009. Detection of amyloid in Alzheimer's disease with positron emission tomography using [(11)C]AZD2184. Eur. J. Nucl. Med. Mol. Img. 36, 1859–1863.

Petersen, R.C., Smith, G.E., Waring, S.C., Ivnik, R.J., Tangalos, E.G., Kokmen, E., 1999. Mild cognitive impairment: clinical characterization and outcome. Arch. Neurol. 56, 303–308.

Reiman, E.M., Langbaum, J.B.S., 2009. Brain imaging in the evaluation of putative Alzheimer's disease slowing, risk-reducing and prevention therapies. In: Jagust, W.J., D'Esposito, M. (Eds.), Imaging the Aging Brain. Oxford University Press, New York, pp. 319–350.

Reiman, E.M., Caselli, R.J., Yun, L.S., Chen, K., Bandy, D., Minoshima, S., Thibodeau, S.N., Osborne, D., 1996. Preclinical evidence of Alzheimer's disease in persons homozygous for the e4 allele for apolipoprotein E. N. Engl. J. Med. 334, 752–758.

Reiman, E.M., Caselli, R.J., Chen, K., Alexander, G.E., Bandy, D., Frost, J., 2001. Declining brain activity in cognitively normal apolipoprotein E e4 heterozygotes: a foundation for using positron emission tomography to efficiently test treatments to prevent Alzheimer's disease. Proc. Natl. Acad. Sci. U. S. A. 98, 3334–3339.

Reiman, E.M., Chen, K., Alexander, G.E., Caselli, R.J., Bandy, D., Osborne, D., Saunders, A.M., Hardy, J., 2004. Functional brain abnormalities in young adults at genetic risk for late-onset Alzheimer's dementia. Proc. Natl. Acad. Sci. U. S. A. 101, 284–289.

Reiman, E.M., Chen, K., Alexander, G.E., Caselli, R.J., Bandy, D., Osborne, D., Saunders, A.M., Hardy, J., 2005. Correlations between apolipoprotein E e4 gene dose and brain-imaging measurements of regional hypometabolism. Proc. Natl. Acad. Sci. U. S. A. 102, 8299–8302.

Reiman, E.M., Chen, K., Ayutyanont, N., Lee, W., Bandy, D., Reschke, C., Alexander, G.E., Weiner, M.W., Koeppe, R.A., Foster, N.L., Jagust, W.J., 2008. Twelve-month cerebral metabolic declines in probable Alzheimer's disease and amnestic mild cognitive impairment: preliminary findings from the Alzheimer's disease neuroimaging initiative (ADNI). Alzheimers Dement. 4, T110–T111.

Reiman, E.M., Langbaum, J.B., Tariot, P.N., 2010. Alzheimer's Prevention Initiative: a proposal to evaluate presymptomatic treatments as quickly as possible. Biomark. Med. 4, 3–14.

Rosen, W.G., Mohs, R.C., Davis, K.L., 1984. A new rating scale for Alzheimer's disease. Am. J. Psychiatry 141, 1356–1364.

Rosner, B., 1990. Fundamentals of Biostatistics, Third ed. PWS-Kent Publishing Company, Boston, MA.

Schuff, N., Woerner, N., Boreta, L., Kornfield, T., Shaw, L.M., Trojanowski, J.Q., Thompson, P.M., Jack Jr., C.R., Weiner, M.W., the Alzheimer's, Disease Neuroimaging Initiative, 2009. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. Brain 132, 1067–1077.

Selkoe, D.J., 2002. Alzheimer's disease is a synaptic failure. Science 298, 789–791.

Shaw, L.M., Korecka, M., Clark, C.M., Lee, V.M., Trojanowski, J.Q., 2007. Biomarkers of neurodegeneration for diagnosis and monitoring therapeutics. Nat. Rev. Drug Discov. 6, 295–303.

Shoghi-Jadid, K., Small, G.W., Agdeppa, E.D., Kepe, V., Ercoli, L.M., Siddarth, P., Read, S., Satyamurthy, N., Petric, A., Huang, S.C., Barrio, J.R., 2002. Localization of neurofibrillary tangles and beta-amyloid plaques in the brains of living patients with Alzheimer disease. Am. J. Geriatr. Psychiatry 10, 24–35.

Small, G.W., Mazziotta, J.C., Collins, M.T., Baxter, L.R., Phelps, M.E., Mandelkern, M.A., Kaplan, A., La Rue, A., Adamson, C.F., Chang, L., 1995. Apolipoprotein E type 4 allele and cerebral glucose metabolism in relatives at risk for familial Alzheimer disease. JAMA 273, 942–947.

Small, G.W., Kepe, V., Ercoli, L.M., Siddarth, P., Bookheimer, S.Y., Miller, K.J., Lavretsky, H., Burggren, A.C., Cole, G.M., Vinters, H.V., Thompson, P.M., Huang, S.C., Satyamurthy, N., Phelps, M.E., Barrio, J.R., 2006. PET of brain amyloid and tau in mild cognitive impairment. N. Engl. J. Med. 355, 2652–2663.

Sunderland, T., Mirza, N., Putnam, K.T., Linker, G., Bhupali, D., Durham, R., Soares, H., Kimmel, L., Friedman, D., Bergeson, J., Csako, G., Levy, J.A., Bartko, J.J., Cohen, R.M., 2004. Cerebrospinal fluid beta-amyloid1-42 and tau in control subjects at risk for Alzheimer's disease: the effect of APOE epsilon4 allele. Biol. Psychiatry 56, 670–676.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme Medical Publishers, New York.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.

Wechsler, D., 1981. Manual for the Wechsler Adult Intelligence Scale — Revised. The Psychological Corporation, New York.

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.O., Nordberg, A., Backman, L., Albert, M., Almkvist, O., Arai, H., Basun, H., Blennow, K., de,, L.M., DeCarli, C., Erkinjuntti, T., Giacobini, E., Graff, C., Hardy, J., Jack, C., Jorm, A., Ritchie, K., van, D.C., Visser, P., Petersen, R.C., 2004. Mild cognitive impairment—beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J. Intern. Med. 256, 240–246.