NeuroImage 113 (2015) 61–69

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Accelerated vs. unaccelerated serial MRI based TBM-SyN measurements for clinical trials in Alzheimer's disease

Prashanthi Vemuri [a,*], Matthew L. Senjem [a,b], Jeffrey L. Gunter [a,b], Emily S. Lundt [c], Nirubol Tosakulwong [c], Stephen D. Weigand [c], Bret J. Borowski [a], Matt A. Bernstein [a], Samantha M. Zuk [a], Val J. Lowe [a], David S. Knopman [d], Ronald C. Petersen [d], Nick C. Fox [e], Paul M. Thompson [f,h,i,j,k,l], Michael W. Weiner [g,m], Clifford R. Jack Jr. [a], For the Alzheimer's Disease Neuroimaging Initiative [1]

[a] Departments of Radiology, MN, USA
[b] Information Technology, MN, USA
[c] Health Sciences Research, MN, USA
[d] Neurology Mayo Clinic Rochester, MN, USA
[e] Dementia Research Center, UCL Institute of Neurology, London, UK
[f] Imaging genetics Center, Institute for Neuroimaging and Informatics, Department of Neurology, University of Southern California, Los Angeles, CA, USA
[g] University of California at San Francisco, Department of Veterans Affairs Medical Center, San Francisco, CA, USA
[h] Department of Psychiatry, University of Southern California, Los Angeles, CA, USA
[i] Department of Radiology, University of Southern California, Los Angeles, CA, USA
[j] Department of Pediatrics, University of Southern California, Los Angeles, CA, USA
[k] Department of Engineering, University of Southern California, Los Angeles, CA, USA
[l] Department of Opthalmology , University of Southern California, Los Angeles, CA, USA
[m] Center for Imaging of Neurodegenerative Diseases, Department of Veterans Affairs Medical Center, San Francisco, CA, USA

ARTICLE INFO

Article history:
Received 3 July 2014
Accepted 12 March 2015
Available online 20 March 2015

ABSTRACT

Objective: Our primary objective was to compare the performance of unaccelerated vs. accelerated structural MRI for measuring disease progression using serial scans in Alzheimer's disease (AD).
Methods: We identified cognitively normal (CN), early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI) and AD subjects from all available Alzheimer's Disease Neuroimaging Initiative (ADNI) subjects with usable pairs of accelerated and unaccelerated scans. There were a total of 696 subjects with baseline and 3 month scans, 628 subjects with baseline and 6 month scans and 464 subjects with baseline and 12 month scans available. We employed the Symmetric Diffeomorphic Image Normalization method (SyN) for normalization of the serial scans to obtain tensor based morphometry (TBM) maps which indicate the structural changes between pairs of scans. We computed a TBM-SyN summary score of annualized structural changes over 31 regions of interest (ROIs) that are characteristically affected in AD. TBM-SyN scores were computed using accelerated and unaccelerated scan pairs and compared in terms of agreement, group-wise discrimination, and sample size estimates for a hypothetical therapeutic trial.
Results: We observed a number of systematic differences between TBM-SyN scores computed from accelerated and unaccelerated pairs of scans. TBM-SyN scores computed from accelerated scans tended to have overall higher estimated values than those from unaccelerated scans. However, the performance of accelerated scans was comparable to unaccelerated scans in terms of discrimination between clinical groups and sample sizes required in each clinical group for a therapeutic trial. We also found that the quality of both accelerated vs. unaccelerated scans were similar.
Conclusions: Accelerated scanning protocols reduce scan time considerably. Their group-wise discrimination and sample size estimates were comparable to those obtained with unaccelerated scans. The two protocols did not produce interchangeable TBM-SyN estimates, so it is arguably important to use either accelerated pairs of scans or unaccelerated pairs of scans throughout the study duration.

* Corresponding author at: Mayo Clinic and Foundation, 200 First Street SW, Rochester, MN 55905, USA. Fax: +1 507 284 9778.
E-mail address: vemuri.prashanthi@mayo.edu (P. Vemuri).

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

http://dx.doi.org/10.1016/j.neuroimage.2015.03.026
1053-8119/

Introduction

Of all currently available disease biomarkers, atrophy of brain structures has been found to track best with change in cognitive impairment in Alzheimer's disease (AD) (Fox et al., 1999; Frisoni et al., 2010, 2013; Jack et al., 2013; Mormino et al., 2009; Savva et al., 2009). Structural changes measured on MRI therefore may be a useful outcome measures in therapeutic clinical trials (Leow et al., 2009; Leung et al., 2010a; Risacher et al., 2010; Schott et al., 2010; Vemuri et al., 2010). MRI image acquisition and data processing are the two main components when considering MRI as an outcome measure in clinical trials.

Structural MRI (sMRI) scanning has greatly advanced in the last couple of decades. Current sMRI protocols yield high-resolution brain images with approximately 1 mm^3 resolution with excellent gray–white matter contrast. The stability and reproducibility of sMRI acquisitions in large multi-center trials such as Alzheimer's Disease Neuroimaging Initiative (ADNI) (Jack et al., 2008), implemented across multiple different vendors and scanners, provide evidence that large scale therapeutic trials are feasible using sMRI as outcome measures (Cover et al., 2011; Fleisher et al., 2009; Kruggel et al., 2010). A recent improvement to MRI acquisition protocols has been the implementation of parallel imaging typically reducing MRI acquisition times by half or more (Blaimer et al., 2004; Griswold et al., 2000, 2002; Pruessmann et al., 1999; Sodickson and Manning, 1997). The acceleration of scans due to parallel imaging leads to more time- and cost-efficient acquisitions and therefore greater patient acceptance and fewer motion artifacts which is of particular importance with more cognitively impaired patients. However, these benefits come at the cost of decreased signal to noise ratio, and potentially increased image artifacts due to the reconstruction of images using data acquired in less time. The primary objective of this paper is to compare the performance of unaccelerated versus accelerated sMRI for measuring disease progression using serial scans. The images used for the comparison were the pairs of accelerated and unaccelerated 3 Tesla sMRI images acquired in ADNI-GO and ADNI-2 which included cognitively normal (CN), early mild cognitive impairment (EMCI), late MCI (LMCI) and AD subjects.

AD-related MRI processing of images has traditionally focused on regions of interest (ROIs) that are preferentially affected by the disease process, e.g. hippocampus. Cross-sectional methods can generate a summary measure or region of interest measure from sMRI at every time point. These measures have unnecessary variability due to differences in the ROI definition on each image. Specific longitudinal techniques can extract tissue loss information from serial sMRI scans (e.g. Boundary shift integral (BSI) (Freeborough and Fox, 1997) and tensor based morphometry (TBM) (Hua et al., 2008, 2013)).

In these techniques all pairs of sMRI scans are registered or warped to each other and brain loss between scans is quantified which reduces measurement variability. High accuracy is the metric that is often considered in the selection of the warping algorithms between serial scans; however recently it has been shown that symmetric registration between serial scans is crucial for obtaining bias-free longitudinal measurements (Fox et al., 2011; Holland et al., 2012). Specifically, it was shown that asymmetric warping can cause biologically implausible deceleration of atrophy and introduces bias into longitudinal measurements (Thompson and Holland, 2011). Symmetric Normalization algorithm (SyN) developed by Avants et al. (Avants et al., 2008) provides symmetric diffeomorphic normalization between serial scans and also has a high degree of accuracy when compared to manual measurements, and in comparison to other nonlinear deformation algorithms (Klein et al., 2009). This warping methodology is an ideal solution for bias-free warping and tracking disease progression in AD and other neurodegenerative diseases, so we used a TBM-SyN based methodology to evaluate the primary objective of this study.

Methods

Selection of participants and image acquisition

Two groups of subjects were analyzed in this study. The first group of subjects, selected from the Mayo Clinic Study of Aging (MCSA) and Mayo Alzheimer's Disease Research Center (ADRC), were used as a training dataset to determine which ROIs to include in the TBM-SyN summary score. The second group of subjects was the ADNI data set, identified from ADNI-GO and ADNI-2 (details below), to compare accelerated and unaccelerated sMRI scans in terms of TBM-SyN. All ADNI images were acquired on 3 T scanners using both accelerated (2× acceleration resulting in roughly half the scan time) and unaccelerated sMRI T1-weighted images as described elsewhere (Jack et al., 2010); detailed protocol parameters are available at http://adni.loni.usc.edu/methods/documents/mri-protocols/. The protocols on GE scanners were acquired using IR-FSPGR and other vendors using MPRAGE. In particular, the accelerated images used a slightly larger field of view (270 mm vs. 260 mm) to compensate for the SNR loss due to parallel imaging. Acquisition time varied slightly among vendors and across software releases, but typical values were 9:14 min for the unaccelerated acquisition, and 5:12 min with acceleration. Mayo images were acquired on 3 T scanners using an unaccelerated MPRAGE protocol similar to that used in ADNI.

The ADNI data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The 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 has been 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 mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop 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, MD, 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. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. To date these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow-up duration of each group is specified in the protocols for ADNI-1, ADNI-2 and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see www.adni-info.org.

Mayo training dataset

In the development of longitudinal measurements, statistically significant ROIs are often determined by analyzing a training set consisting of both patients and matched controls. For detecting AD specific changes, we identified an independent training set of AD and CN subjects with longitudinal MRI scans, drawn from the Mayo ADRC and MCSA. In total, there were 51 AD subjects, and 51 CN subjects (50 who were amyloid negative as measured by Pittsburgh compound B, PiB-PET imaging and one matched CN was APOE4 negative because PIB scan was not available). The PiB-negative status of the CN subjects was defined as global PiB SUVR < 1.4. The AD and CN subjects were matched on age, sex and education. The criterion for selection for AD dementia was based on the 1984 clinical criteria for probable AD which is virtually identical to the new criteria (McKhann et al., 1984, 2011). Each subject

had two serial usable (passed quality control) unaccelerated sMRI scans that were used to develop the TBM-SyN summary score. To maintain a clean training dataset, we took the following additional steps: all subjects were required to maintain the same clinical primary diagnosis at both the serial scans; and the baseline age of all subjects was restricted to ≥63 years. The subject characteristics are described in Table 1.

ADNI dataset

We identified all CN, EMCI, LMCI and AD subjects from ADNI-GO/2 with usable pairs of accelerated and unaccelerated scans. We identified all subjects with preprocessed accelerated and unaccelerated MRI data available on LONI that had passed quality control. This included 703 subjects with baseline and 3 month scans, 643 subjects with baseline and 6 month scans, and 478 subjects with baseline and 12 month scans. We excluded 36 pairs of scans because of within-subject differences in model and/or manufacturer of the MRI scanner used for the serial acquisitions. The final ADNI dataset used for this paper therefore included 696 subjects with baseline and 3 month scans, 628 subjects with baseline and 6 month scans, and 464 subjects with baseline and 12 month scans. Subject demographics are described in Table 2. We have provided the entire list of patients that were used in this study in the Supplemental material.

Image preprocessing for each individual image

The software packages used to develop the TBM-SyN scores were MATLAB R2013a (Mathworks, Natwick, MA), ANTs 1.9.x (Penn Image Computing and Science Lab, University of Pennsylvania, PA), and SPM5 (Wellcome Trust Center for Neuroimaging, UCL, UK). For each subject in this study, we began with the "N3m" preprocessed datasets that are "N3" intensity homogeneity corrected (Boyes et al., 2008) and corrected for gradient field non-linearity images (Gunter et al., 2009) that were additionally run through SPM5 bias correction (indicated by the suffix "m").

For each original T1 image both in the Mayo and ADNI dataset, we ran it through our standard pre-processing pipeline. Briefly, this pipeline first applies gradient unwarping to correct for gradient distortions, followed by spm5 unified segmentation using custom tissue priors in a custom template space (as defined in STAND400 (Vemuri et al., 2008)). The initial gray matter (GM) and white matter (WM) masks from the spm5 unified segmentation are then combined and binarized to form an initial brain mask. This brain mask is then dilated and hole filled, and used for the next step, N3 correction. The N3 correction is computed over those voxels in the dilated mask. Next, the N3 corrected image is segmented once again using spm5 unified segmentation. The native-space segmentations from this step are combined and binarized, to form a more accurate brain mask. Additionally, a mask of the third and lateral ventricles are propagated from custom template space to subject native space, using the spatial normalization parameters from SPM5 unified segmentation, followed by a sequence of morphological operations to clean up the masks. The final pass of spm5 unified segmentation also produces a bias corrected version of the input image, which we call the "N3m" image. This sequence of pre-processing is applied to all ADNI and Mayo scans in our lab.

Table 1
Mayo training dataset demographics.

	CN (n = 51)	AD (n = 51)	P-value*
Men, n (%)	23 (45.1)	23 (45.1)	0.84
Age at baseline, years	81 (66, 88)	81 (64, 92)	1.00
Education, years	15 (8, 20)	14 (7, 20)	1.00
CDR sum of boxes	0.0 (0.0, 0.5)	4.5 (0.5, 9.0)	0.001
MMSE	28 (24, 30)	21 (13, 28)	0.001
APOE ε4 carrier, n (%)	6 (11.8)	35 (68.6)	<0.001
Scan interval, years	1.3 (1.1, 2.0)	1.1 (0.9, 2.0)	1.00

Note. Unless otherwise indicated, values shown are median (minimum, maximum).
Abbreviations: CDR, Clinical Dementia Rating; MMSE, mini-mental state exam.
* From two-sided Wilcoxon rank-sum tests or chi-squared tests.

Table 2
ADNI dataset subject demographics. All subjects had a pair of accelerated and unaccelerated scans.

	CN	EMCI	LMCI	AD
Baseline to 3 months				
Subjects, n	173	278	147	98
Men, n (%)	89 (51%)	157 (56%)	79 (54%)	60 (61%)
Age at scan, years	72 (56, 88)	70 (55, 88)	72 (55, 91)	75 (55, 90)
Education, years	16 (12, 20)	16 (11, 20)	16 (11, 20)	16 (9, 20)
CDR sum of boxes	0 (0, 1)	1.0 (0.5, 4.0)	1.5 (0.5, 5.5)	4.5 (1, 10)
MMSE	29 (24, 30)	29 (23, 30)	28 (24, 30)	23 (19, 26)
APOE ε4 carrier, n (%)	48 (29%)	119 (43%)	80 (55%)	64 (74%)
Baseline to 6 months				
Subjects, n	164	250	138	76
Men, n (%)	80 (49%)	139 (56%)	74 (54%)	45 (59%)
Age at scan, years	72 (56, 88)	70 (55, 88)	72 (55, 91)	75 (55, 90)
Education, years	16 (12, 20)	16 (10, 20)	16 (11, 20)	16 (9, 20)
CDR sum of boxes	0 (0, 1)	1.0 (0.5, 4.0)	1.5 (0.5, 5.5)	4.5 (2, 10)
MMSE	29 (25, 30)	29 (23, 30)	28 (24, 30)	23 (19, 26)
APOE ε4 carrier, n (%)	50 (31%)	109 (44%)	76 (56%)	53 (74%)
Baseline to 12 months				
Subjects, n	132	211	89	32
Men, n (%)	67 (51%)	113 (54%)	49 (55%)	22 (69%)
Age at scan, years	72 (62, 88)	70 (55, 88)	72 (55, 91)	78 (55, 90)
Education, years	16 (12, 20)	16 (10, 20)	17 (12, 20)	16 (12, 20)
CDR sum of boxes	0 (0, 1)	1.0 (0.5, 4.0)	1.5 (0.5, 5.5)	4.5 (2, 10)
MMSE	29 (24, 30)	29 (23, 30)	28 (24, 30)	22 (19, 26)
APOE ε4 carrier, n (%)	36 (28%)	86 (41%)	50 (57%)	22 (71%)

Note. Unless otherwise indicated, values shown are median (minimum, maximum).
Abbreviations: CDR, Clinical Dementia Rating; MMSE, mini-mental state exam.

Within subject initial co-registration

After the initial pre-processing was completed, we formed an initial mean image of each subject's pre-processed "N3m" images. Using SPM5-based mutual information co-registration without reslicing, we iteratively registered each individual time-point N3m image to the mean, applying the transformations to the corresponding brain and ventricle masks, forming a new mean after each iteration, allowing it to continue until the mean image did not change from one iteration to the next, or until a preset maximum of 10 iterations was reached. After the final iteration, the set of all images were co-registered one last time to the first time-point image, to ensure no registration failures.

Within subject intensity balancing

After the initial registration completed, we dilated and hole-filled the brain mask from the first time point adding dominantly CSF voxels to the mask, and fit a Gaussian distribution to the histogram of these voxels. In a separate procedure we eroded the brain mask to obtain a collection of voxels dominated by white matter, and fit a separate Gaussian distribution to their intensity. In both cases the fitting was done iteratively, excluding intensities more than two sigma from the mean. We then scaled image intensities, mapping the WM and CSF spectrum peaks (that were estimated as described in (Gunter et al., 2003)) to constant arbitrary values of 20,000 and 5000, respectively.

Within subject secondary co-registration

Next, although the images had been previously co-registered, in order to ensure robust performance under fully automated conditions, we implemented another rigid registration routine into the algorithm. Using Aladin (http://sourceforge.net/projects/niftyreg/), we rigid-body (6DOF) co-registered each image to the subject's baseline image, restricting the cost function with an intracranial mask to eliminate variability in neck positioning. We computed the mean within-subject transformation by averaging the quaternion representations of the transformations and resampled the gray scale image and masks into

this mean space with 1 mm isotropic resolution using cubic spline, and linear interpolation, respectively. We next formed a new registration target by re-computing the mean gray scale image from the resampled images, and forming a new intracranial mask by applying dilation and hole filling to the union of the resampled brain and ventricle masks. We then performed affine (9DOF) registration of each time point image to the mean image, and finally resampled all images and masks into the target space at 1 mm isotropic resolution.

Within subject intensity re-balancing and differential bias correction (DBC)

We next balanced intensities and performed differential bias correction (DBC). As before we determined the WM and CSF spectrum peak intensities. Next we formed the collection of voxels that were spatially located between those selected for the CSF and WM samples, forming a collection of mostly GM with a "contamination" of WM voxels. We then fit the histogram of these voxels with the sum of two Gaussians; one Gaussian having center and width fixed at the values determined by the WM fit with arbitrary amplitude, and the parameters for the other Gaussian distribution being fitted. In contrast to the linear intensity remapping previously employed, we used a spline-based intensity re-mapping to bring each image's GM, WM and CSF peak intensities into agreement with those of the mean image.

The DBC was carried out using the collection of voxels that was consistently near CSF peak intensity or consistently near the WM peak intensity, and inside a hole-filled brain mask. Using only points inside the collection, we created a log transformed ratio image of each time point image to the mean image. Since the point collection is sparse in space, we used a tri-linear 3D interpolation to create a dense field, requiring it go to zero at the edges of the image. We then smoothed the dense field with a 20 mm isotropic Gaussian kernel, exponentiated the resulting field and finally applied the result to arrive at the final preprocessed image series for each subject.

TBM-SyN image processing

Starting with the preprocessed scans for each subject as described above, we computed the SyN deformations between each pair of images (Avants et al., 2008), in both directions explicitly, saving an image of the log transformed Jacobian determinants for each. We formed an "annualized" log Jacobian map by dividing each log Jacobian voxel by the intrascan time interval, measured in years. We then applied each deformation to the corresponding moving image, and create a "soft-mean" of the "fixed" and the "moved" image. We then applied SPM5 unified segmentation to each soft-mean image, and propagated ROI masks from the template space to the soft-mean space, to obtain mean annualized log Jacobian measurements in the various ROIs in the in-house modified AAL custom template.

TBM-SyN summary score

For selection of ROIs to be included in the TBM-SyN summary score, we used a two-sample t-test to select the top ROIs (with right and left ROIs combined) that were significantly different between the CN and AD from the Mayo Dataset as described later in this paragraph. We used an in-house modified atlas of 119 GM regions and one ventricular region. Cross-sectional GM volumes of each GM ROI in the atlas computed from the baseline image, as well as the longitudinal mean annualized log Jacobian from each ROI in the atlas between the two serial scans, were compared between the AD and CN subjects in the Mayo Training set. We used both cross-sectional GM volumes as well as longitudinal log Jacobians for selection of the ROIs because some ROIs such as the parahippocampal gyrus atrophy early in the disease process and do not show highly significant annual change between AD and CN but are important to retain because MCI and preclinical CN subjects will show early serial changes in those regions. For selection of the ROIs — we first ranked all the ROIs by the strength of the t-tests of all cross-sectional GM density and longitudinal mean annualized Jacobians. Next, we determined the number of ROIs that, when averaged together, would give the highest degree of AD-CN separation accuracy in the training data set. The ROIs that were selected based on this criterion are shown in Fig. 1. The union of ROIs selected from both the longitudinal and cross-sectional data included 15 bilateral GM ROIs [or 30 left and right ROIs] and one ventricular ROI. Since 15 of the ROIs were GM ROIs from the atlas, which show volume shrinkage, and one of the ROIs is the ventricle, which shows expansion, we inverted the sign of the ventricle log Jacobian determinant before combining it with the values from the cortical GM ROIs. The 15 bilateral GM ROIs included were: medial temporal lobes (consisting of amygdala, hippocampal, parahippocampal and entorhinal cortices), angular, precuneus, temporal lobes (fusiform, superior, mid and inferior temporal gyri and superior and mid temporal poles), and occipital lobes (superior, mid and inferior ROIs). For each image, a TBM-SyN score was created as a sum of the median annualized log Jacobian determinant in these 15 bilateral GM ROIs, and the negative median annualized log Jacobian determinant of the ventricles.

Comparison of accelerated and unaccelerated scans

We performed two main analyses to compare the accelerated and unaccelerated scans. The purpose of the first set of analyses was to test the interchangeability of TBM-SyN scores from accelerated and unaccelerated scans. We tested for differences between accelerated and unaccelerated TBM-SyN values within each diagnostic group using paired t-tests. Linear regression was used to evaluate the relationship between accelerated (y) versus unaccelerated (x) TBM-SyN values separately for each scan interval. To evaluate agreement between the two techniques we included Bland–Altman plots. The purpose of the second set of analyses was to compare the performance of accelerated and unaccelerated scans in terms of group-wise discrimination. We calculated the area under the receiver operating characteristic curve (AUROC) as a nonparametric measure of effect size (Acion et al., 2006) and calculated 95% confidence intervals for each AUROC estimate (Newcombe, 2006). We tested group-wise differences in TBM-SyN using the Wilcoxon rank sum test, equivalent to evaluating whether the AUROC was different from 0.5. We also directly tested group-wise discrimination as summarized by the AUROC for accelerated versus unaccelerated scans (DeLong et al., 1988). Finally, we estimated sample sizes for a hypothetical trial in which the intervention resulted in a 25% improvement over the placebo TBM-SyN rate via the standard t-test

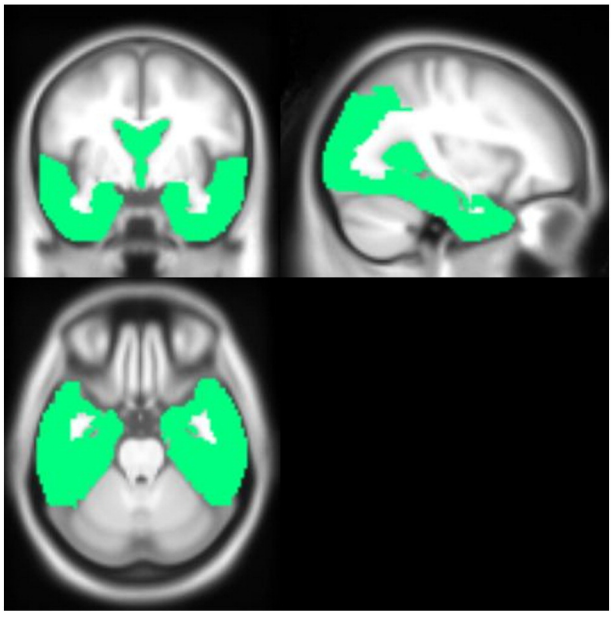

Fig. 1. The region of interest selected based on t-test differences between the 51 AD and 51 PiB negative CN subjects to compute the TBM-SyN score.

based sample size formula: $n = 2\sigma 2(z1\text{-}\alpha/2 + z\beta)2/\Delta 2$. These sample sizes refer to a 25% reduction in absolute atrophy rate and assume as is normally the case that the hypothetical treatment does not alter the variance in the rate. Here σ^2 is estimated by the sample variance, $z_{1-\alpha/2} = z_{1-0.05/2} \approx 1.960$, $z\beta = z_{0.80} \approx 0.842$, and Δ represents the minimum detectible difference in means (Rosner, 2011). We used bootstrap resampling with 3000 replicates to get 95% CIs for the sample size estimates. We also used the bootstrap with 3000 replicates to obtain interval estimates of the difference in sample sizes required for the two methods.

We performed additional analyses 1) to compare the quality control of all the subjects scanned in ADNI2 and ADNI GO (before the selection of usable scans for the comparison analyses); and 2) to compare the overall tissue segmentations behavior differences using SPM based paired t-test between baseline accelerated and baseline unaccelerated scans of cognitively normal individuals.

Results

Fig. 2 shows box plots comparing the distribution of TBM-SyN summary scores, computed from accelerated versus unaccelerated scan pairs within each clinical group. In the baseline to 3-month comparisons, none of the clinical groups had significantly different TBM-SyN scores generated from accelerated versus unaccelerated scan pairs ($p > 0.05$), possibly because there are minimal sMRI based differences detectable in a 3 month interval. In the baseline to 6-month comparisons, the TBM-SyN scores from accelerated scans were significantly greater than unaccelerated scans in both the CN and AD groups ($p = 0.03$), with a trend in the LMCI and EMCI groups ($p < 0.1$). In the baseline to 12-month comparisons, the TBM-SyN scores from accelerated scans were significantly greater from unaccelerated scans in both the EMCI and AD groups ($p < 0.01$).

For further clarity on the subject of equivalency of the two scan types, we show scatter plots by scan interval in Fig. 3 and Bland–Altman plots in Fig. 4. Our first observations from the scatter plots were that for each interval the regression line does not pass through the origin ($p < 0.01$) and its slope is less than unity ($p < 0.01$) indicating that the two protocols do not produce equivalent results. However, the coefficient of determination, R^2, increased with the interscan duration indicating that the scans with larger intervals (possibly 2 years) are likely to be equivalent because the sMRI changes over longer periods of time are much greater compared to the variability due to the MRI acquisition differences. This decreasing variability and decreasing difference between the accelerated versus unaccelerated TBM-SyN scores can also

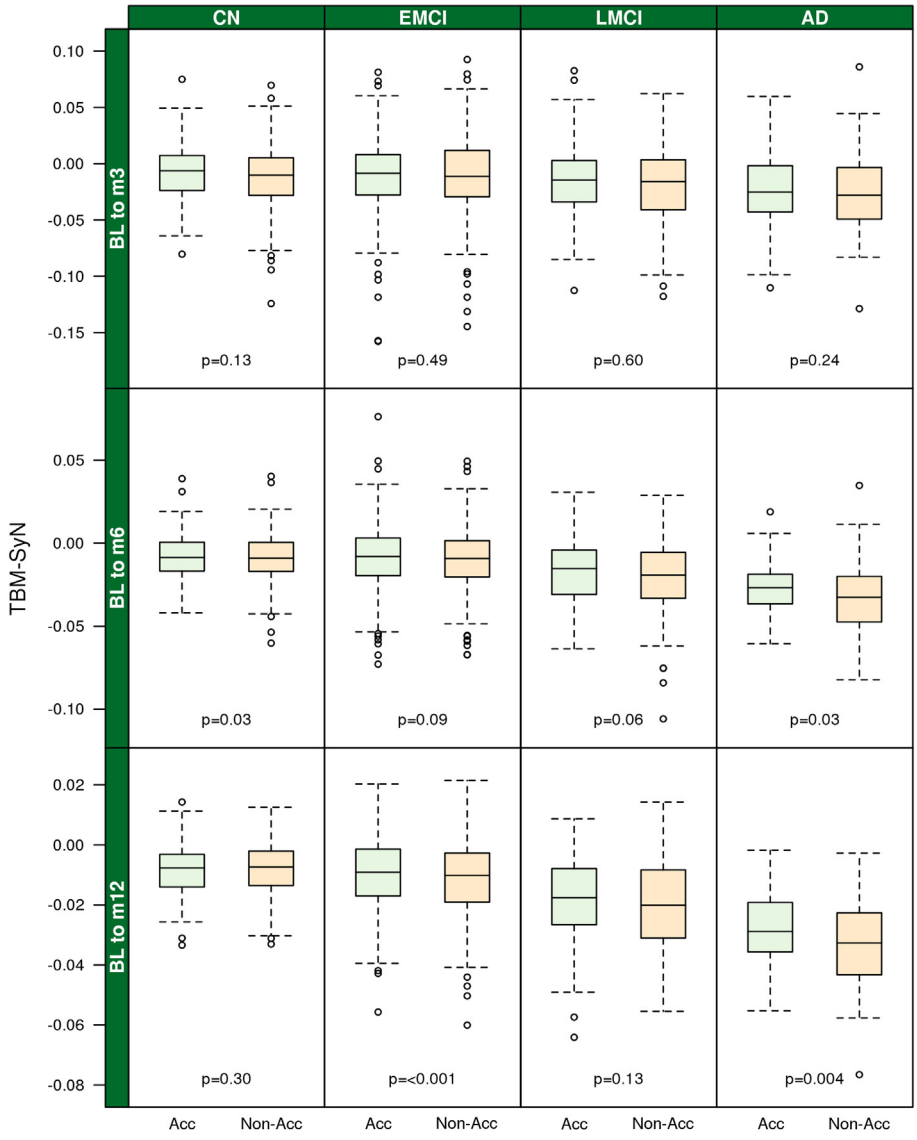

Fig. 2. Box plots of accelerated and unaccelerated TBM-SyN values by disease group and scan interval. P-values are from a paired t-test comparing accelerated to unaccelerated values.

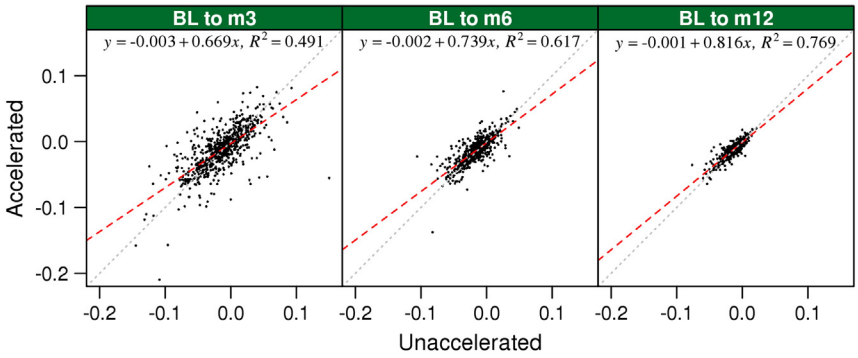

Fig. 3. Scatter plots of accelerated vs unaccelerated TBM-SyN with regression lines in red. The equations and the coefficient of determination, R^2, are at the top of each panel. The dotted line represents the identity line and the dashed line represents the fit.

be confirmed from the Bland–Altman plots. Again, with increasing scan interval the TBM-SyN scores from accelerated and unaccelerated scans show increasing agreement. However, with the intervals used in this study the values are not interchangeable.

The estimated AUROC and confidence intervals for distinguishing different clinical groups for each scan type are shown in Fig. 5. The p-values indicate if the performance of accelerated or unaccelerated scans was significantly different in terms of ability to distinguish the two clinical groups. The AUROC point estimates and corresponding p-values comparing accelerated versus unaccelerated are shown in Table 3. Though the AUROC point estimates for the unaccelerated scans appeared to be a little better overall than those for accelerated scans, the performance was not statistically different, except for the CN vs. EMCI discrimination for the change in 0–12 months.

Sample size estimates with bootstrap 95% CIs to detect a 25% reduction in atrophy with 80% power and two-sided $\alpha = 0.05$ are shown in Table 4. Sample sizes vary in the expected way with smaller samples needed for longer interscan durations and with more impaired clinical groups. The columns labeled "Accelerated" and "Unaccelerated" indicate the sample size and 95% CI for each of the scan types, and the column labeled "Difference" indicates the estimated sample size difference and 95% CI. Positive numbers in this column indicate that the sample sizes with unaccelerated scans are lower. In order for the differences to be statistically significant at $p < 0.05$ level, the confidence intervals cannot include zero. Using this criterion, the differences were only significant in the case of EMCI measurements between baseline and 12-month intervals, where the sample sizes necessary to detect 25% reduction in the TBM-SyN using accelerated scans were 361 compared to a much lower 272 required using unaccelerated scans. There were similar indications favoring unaccelerated scans in this group with six-month interval, although sample sizes were very large for both methods.

Here we present the quality control comparison results. For both accelerated and unaccelerated scans, there were only a very small proportion of scans that were considered unusable (i.e., failed quality control). Comparing accelerated to unaccelerated scans, the failure rates were 0.6% ($n = 5$) versus 0.7% ($n = 6$) at baseline, 0% ($n = 0$) versus 0% ($n = 0$) at three months, 0.5 ($n = 3$) vs. 0% ($n = 0$) at six months, and 0.2% ($n = 1$) vs. 0.4% ($n = 2$) at 12 months. Based on McNemar's test of correlated proportions, we found no significant differences in failure rates ($p > 0.99$ for baseline; $p = 0.24$ for 6 months; $p > 0.99$ for 12 months). Since the failure rates were zero for both accelerated and unaccelerated scans at 3 months, McNemar's test could not be performed. Scans were also graded on a numeric scale and based on paired signed rank tests, we found no evidence that accelerated scans received lower quality scores than unaccelerated ($p = 0.5$ for baseline; $p = 0.4$ for 3 months; $p = 0.7$ for 6 months; $p = 0.7$ for 12 months). Additionally, we compared the failure rate for accelerated and unaccelerated scans by diagnosis and found that both AD and non-AD cases had similar failure rates between unaccelerated and accelerated scans for all the scans ($p > 0.1$ for Wilcoxon signed rank test for baseline, 3 month, 6 month and 12 month scans).

When we compared the overall tissue segmentations behavior by computing the paired t-test between baseline accelerated and baseline unaccelerated scans of cognitively normal individuals (Supplemental Fig. 1). We found that there were no fundamental differences between tissue segmentation properties of both types of scans in the areas of the regions that were used for computing the TBM-SyN scores.

Discussion

In this study we applied a SyN based TBM methodology to measure structural changes between pairs of MRI scans for the comparison of accelerated scans with unaccelerated scans. TBM-SyN scores from both sets of scans are not equivalent but the overall performance of both types of scans for clinical discrimination and clinical trials was similar. Additionally, one scan type does not significantly outperform another scan type. We therefore conclude that, for longitudinal studies involving sMRI as an outcome measure, it is important to maintain

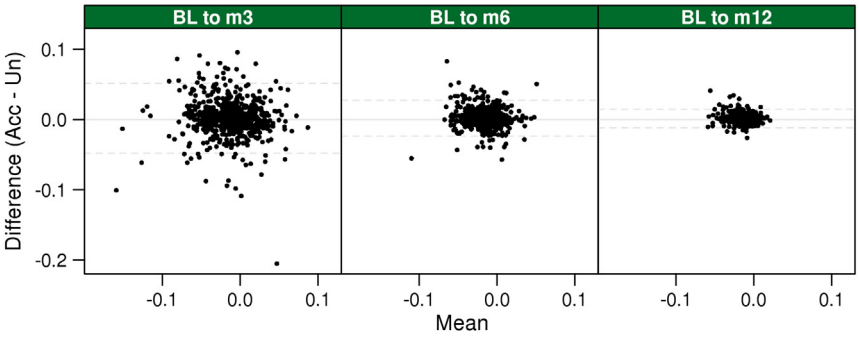

Fig. 4. Bland–Altman agreement plots to assess interchangeability. The x-axis indicates the mean of the accelerated and unaccelerated TBM-SyN values while the y-axis indicates their difference.

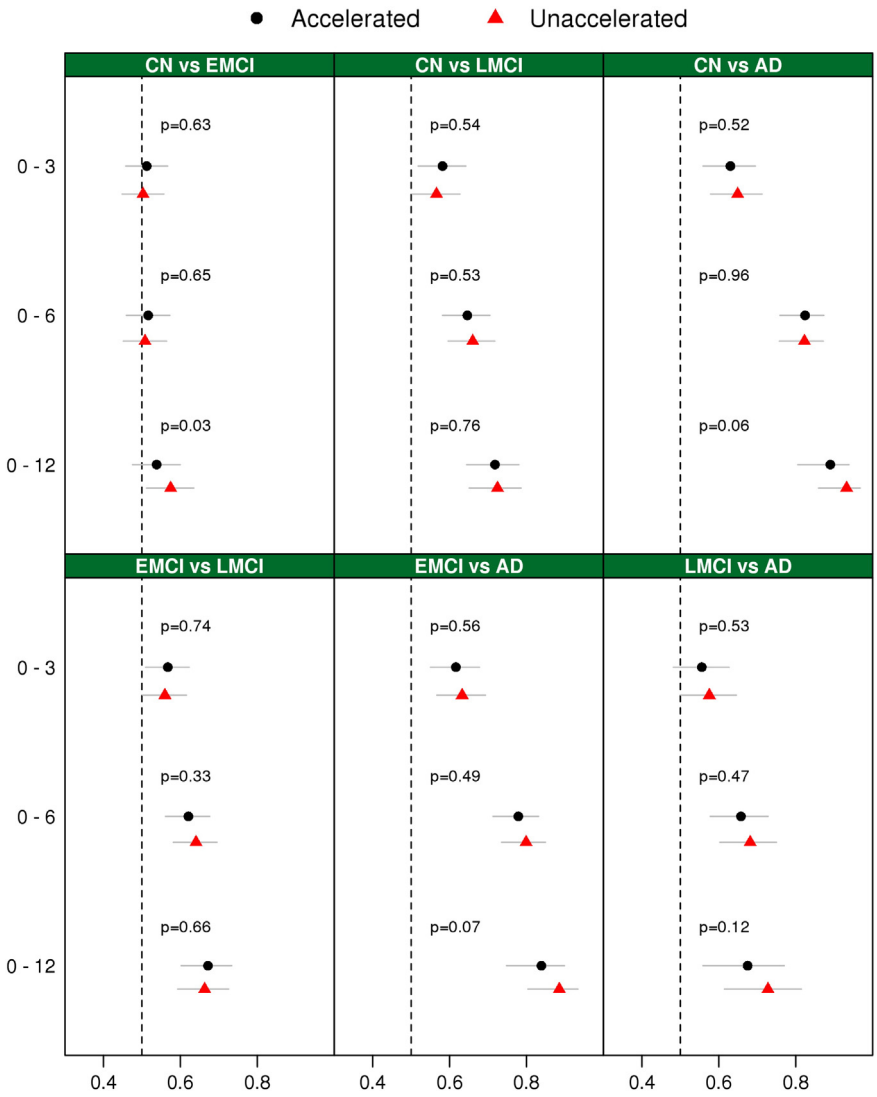

Fig. 5. Estimated AUROC and confidence intervals summarizing group-wise discrimination. P-values are from a Chi-squared test of whether accelerated or unaccelerated TBM-SyN estimates provide better group-wise discrimination.

Table 3
AUROC point estimates and p-value for a test of whether accelerated or unaccelerated TBM-SyN estimates provide better group-wise discrimination.

Groups compared	AUROC			
	Scan interval, months	Accelerated	Unaccelerated	P-value
CN vs EMCI	0–3	0.51	0.50	0.63
	0–6	0.52	0.51	0.65
	0–12	0.54	0.57	0.03
CN vs LMCI	0–3	0.58	0.57	0.54
	0–6	0.65	0.66	0.53
	0–12	0.72	0.72	0.76
CN vs AD	0–3	0.63	0.65	0.52
	0–6	0.82	0.82	0.96
	0–12	0.89	0.93	0.06
EMCI vs LMCI	0–3	0.57	0.56	0.74
	0–6	0.62	0.64	0.33
	0–12	0.67	0.66	0.66
EMCI vs AD	0–3	0.62	0.63	0.56
	0–6	0.78	0.8	0.49
	0–12	0.84	0.89	0.07
LMCI vs AD	0–3	0.56	0.58	0.53
	0–6	0.66	0.68	0.47
	0–12	0.67	0.73	0.12

Table 4
Estimated sample sizes with bootstrap 95% CIs to detect a 25% reduction in TBM-SyN with 80% power and two-sided $\alpha = 0.05$.

	Accelerated n (95% CI)	Unaccelerated n (95% CI)	Difference* n (95% CI)
Baseline to 3 months			
CN	2238 (1089, 8984)	1729 (897, 4596)	509 (−751, 5349)
EMCI	2850 (1507, 7799)	2673 (1409, 7865)	177 (−3641, 2687)
LMCI	1015 (584, 2303)	841 (499, 1754)	174 (−373, 1068)
AD	593 (328, 1332)	438 (254, 1009)	155 (−103, 761)
Baseline to 6 months			
CN	749 (446, 1524)	667 (421, 1332)	82 (−152, 446)
EMCI	1213 (728, 2532)	898 (580, 1605)	315 (−7, 1297)
LMCI	297 (213, 451)	286 (202, 428)	11 (−95, 132)
AD	133 (74, 271)	107 (70, 192)	26 (−50, 138)
Baseline to 12 months			
CN	259 (181, 412)	276 (200, 404)	−17 (−84, 46)
EMCI	361 (265, 523)	272 (205, 375)	89 (35, 182)
LMCI	157 (113, 219)	154 (108, 230)	3 (−39, 36)
AD	56 (33, 97)	51 (30, 93)	5 (−25, 32)

* Difference defined as the accelerated sample size minus the unaccelerated sample size. Confidence intervals that exclude zero may be considered significant at $\alpha = 0.05$.

either accelerated or unaccelerated MRI scanning protocols for all subjects throughout the study. Potentially important scan-time reductions may be obtained with accelerated protocols without unduly sacrificing group-wise discrimination or clinical trial efficiency.

The results from this paper are encouraging because the performance from accelerated scans which were acquired in roughly half the scan time was comparable to the performance from unaccelerated scans for both clinical discrimination and sample size estimates. The slight reduction in the signal to noise ratio and contrast to noise did not significantly reduce the performance of the accelerated scans. We also presented the data on scan quality ratings on visual inspection given by trained analysts, which were not significantly different between the two scan types supporting the use of accelerated scans for more cost and time efficient AD therapeutic trials. It is important to note that the comparison of accelerated to unaccelerated scans based on performance is specific to the results from TBM-SyN and other studies have investigated the differences between the scan types using different methodologies for computing change between serial scans and have found similar results (Ching et al., in press; Krueger et al., 2012).

The most surprising result from this study was the non-equivalent nature (i.e. significant differences) of the TBM-SyN scores computed from accelerated and unaccelerated scans. We hypothesize that two noise components may significantly contribute to the non-equivalence of the TBM-SyN scores from the two scan types: 1) the noise of the measurement itself, i.e. higher noise in measurements from shorter inter-scan intervals versus longer intervals; and 2) differences due to systematically different acquisition and reconstruction schemes used in accelerated scans compared to sMRI unaccelerated scans. The differences due to the second component are attributable to the subtle differences in the contrast between gray and WM (though not visible to the naked eye) as well as reconstruction approximation assumptions made in parallel imaging acquisitions where reconstruction of the entire image is done with half the amount of data needed for reconstruction of an unaccelerated image. Though these two noise components may not be completely independent, Figs. 3 and 4 illustrate that the effect of the first component gets comparably lower with increasing time and the regression line approaches the identity line. Additionally, we can ascertain that the noise from the second component is probably smaller than the first component because if it were not, then the regression would not approach the identity line with increasing inter-scan time intervals. These results suggest that with larger time intervals between scans (perhaps 2 years or more), it may be possible to mix data from accelerated and unaccelerated scans; however with the typical scan durations of one year or less, as used in this study, our data suggests that any given study should be performed either with all scans accelerated or all scans unaccelerated.

In the methods, we noted that the accelerated protocols used a slightly increased field of view (270 mm vs. 260 mm) to compensate for the SNR loss due to parallel imaging. That increase was needed because many of the sites did not have access to a 32-channel head coil, but instead were using 8- or 12-channel models. As 32-channel head coils become more prevalent, we expect the slight increase in FoV will become unnecessary for protocols accelerated by a factor of two (Krueger et al., 2012).

A primary motivation behind the development of TBM-SyN based methodology was to avoid bias seen in longitudinal measurements due to asymmetric warping (Thompson and Holland, 2011). Additionally we used a strategy to select ROIs that are specific to AD pathology by comparing AD subjects to a group of CN who were amyloid negative or at least not APOE4 carriers. The ROIs selected correspond well with the well-known Braak staging of neurofibrillary tangles in AD (Braak and Braak, 1997). We believe that employing both these strategies enabled us to establish a method that performs very well in comparison to the existing methodologies for longitudinal measurements using sMRI (Holland et al., 2012; Hua et al., 2013; Leung et al., 2010b; Vemuri et al., 2010), and the sample size estimates show that TBM-SyN scores would be a reasonable metric to use in therapeutic trials.

There are two limitations of this study. First, we did not conduct the experiment where two different scan types were used in the pair of serial scans while computing the TBM-SyN scores (i.e. accelerated first scan and unaccelerated second scan and vice versa). Second, we did not take into account the type of parallel imaging acquisition and reconstruction strategy used by the different MRI vendors. These analyses will be the subject of future work.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.03.026.

Acknowledgments

This work was supported by NIH grants RC2 AG036535, R00 AG37573, R01 AG11378, P50 AG16574, P50 AG16574/P1, U01 AG06786; the Alexander Family Alzheimer's Disease Research Professorship of the Mayo Foundation, the Elsie and Marvin Dekelboum Family Foundation, U.S.A. and Opus building NIH grant C06 RR018898. The funding sources were not involved in the manuscript review or approval.

Data collection and sharing (ADNI data) for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Disclosure

The authors report no conflicts of interest.

References

Acion, L., Peterson, J.J., Temple, S., Arndt, S., 2006. Probabilistic index: an intuitive non-parametric approach to measuring the size of treatment effects. Stat. Med. 25, 591–602.

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. 12, 26–41.

Blaimer, M., Breuer, F., Mueller, M., Heidemann, R.M., Griswold, M.A., Jakob, P.M., 2004. SMASH, SENSE, PILS, GRAPPA: how to choose the optimal method. Top. Magn. Reson. Imaging 15, 223–236.

Boyes, R.G., Gunter, J.L., Frost, C., Janke, A.L., Yeatman, T., Hill, D.L., Bernstein, M.A., Thompson, P.M., Weiner, M.W., Schuff, N., Alexander, G.E., Killiany, R.J., DeCarli, C., Jack, C.R., Fox, N.C., 2008. Intensity non-uniformity correction using N3 on 3-T scanners with multi-channel phased array coils. NeuroImage 39, 1752–1762.

Braak, H., Braak, E., 1997. Frequency of stages of Alzheimer-related lesions in different age categories. Neurobiol. Aging 18, 351–357.

Ching, C.R.K., Hua, X., Hibar, D.P., Ward, C.P., Gunter, J.L., Bernstein, M.A., Jack, C.R., Weiner, M., Thompson, P.M., 2015 Jan. Does MRI scan acceleration affect power to track brain change? Neurobiol. Aging 36 (Suppl 1), S167–S177.

Cover, K.S., van Schijndel, R.A., van Dijk, B.W., Redolfi, A., Knol, D.L., Frisoni, G.B., Barkhof, F., Vrenken, H., 2011. Assessing the reproducibility of the SienaX and Siena brain

atrophy measures using the ADNI back-to-back MP-RAGE MRI scans. Psychiatry Res. 193, 182–190.
DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44, 837–845.
Fleisher, A.S., Donohue, M., Chen, K., Brewer, J.B., Aisen, P.S., 2009. Applications of neuroimaging to disease-modification trials in Alzheimer's disease. Behav. Neurol. 21, 129–136.
Fox, N.C., Scahill, R.I., Crum, W.R., Rossor, M.N., 1999. Correlation between rates of brain atrophy and cognitive decline in AD. Neurology 52, 1687–1689.
Fox, N.C., Ridgway, G.R., Schott, J.M., 2011. Algorithms, atrophy and Alzheimer's disease: cautionary tales for clinical trials. NeuroImage 57, 15–18.
Freeborough, P.A., Fox, N.C., 1997. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. IEEE Trans. Med. Imaging 16, 623–629.
Frisoni, G.B., Fox, N.C., Jack Jr., C.R., Scheltens, P., Thompson, P.M., 2010. The clinical use of structural MRI in Alzheimer disease. Nat. Rev. Neurol. 6, 67–77.
Frisoni, G.B., Bocchetta, M., Chetelat, G., Rabinovici, G.D., de Leon, M.J., Kaye, J., Reiman, E.M., Scheltens, P., Barkhof, F., Black, S.E., Brooks, D.J., Carrillo, M.C., Fox, N.C., Herholz, K., Nordberg, A., Jack Jr., C.R., Jagust, W.J., Johnson, K.A., Rowe, C.C., Sperling, R.A., Thies, W., Wahlund, L.O., Weiner, M.W., Pasqualetti, P., Decarli, C., 2013. Imaging markers for Alzheimer disease: which vs how. Neurology 81, 487–500.
Griswold, M.A., Jakob, P.M., Nittka, M., Goldfarb, J.W., Haase, A., 2000. Partially parallel imaging with localized sensitivities (PILS). Magn. Reson. Med. 44, 602–609.
Griswold, M.A., Jakob, P.M., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A., 2002. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn. Reson. Med. 47, 1202–1210.
Gunter, J.L., Shiung, M.M., Manduca, A., Jack Jr., C.R., 2003. Methodological considerations for measuring rates of brain atrophy. J. Magn. Reson. Imaging 18, 16–24.
Gunter, J.L., Bernstein, M.A., Borowski, B.J., Ward, C.P., Britson, P.J., Felmlee, J.P., Schuff, N., Weiner, M., Jack, C.R., 2009. Measurement of MRI scanner performance with the ADNI phantom. Med. Phys. 36, 2193–2205.
Holland, D., McEvoy, L.K., Dale, A.M., 2012. Unbiased comparison of sample size estimates from longitudinal structural measures in ADNI. Hum. Brain Mapp. 33, 2586–2602.
Hua, X., Leow, A.D., Parikshak, N., Lee, S., Chiang, M.C., Toga, A.W., Jack Jr., C.R., Weiner, M.W., Thompson, P.M., 2008. Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: an MRI study of 676 AD, MCI, and normal subjects. NeuroImage 43, 458–469.
Hua, X., Hibar, D.P., Ching, C.R., Boyle, C.P., Rajagopalan, P., Gutman, B.A., Leow, A.D., Toga, A.W., Jack Jr., C.R., Harvey, D., Weiner, M.W., Thompson, P.M., 2013. Unbiased tensor-based morphometry: improved robustness and sample size estimates for Alzheimer's disease clinical trials. NeuroImage 66, 648–661.
Jack Jr., C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., J, L.W., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27, 685–691.
Jack Jr., C.R., Bernstein, M.A., Borowski, B.J., Gunter, J.L., Fox, N.C., Thompson, P.M., Schuff, N., Krueger, G., Killiany, R.J., Decarli, C.S., Dale, A.M., Carmichael, O.W., Tosun, D., Weiner, M.W., 2010. Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. Alzheimers Dement. 6, 212–220.
Jack Jr., C.R., Knopman, D.S., Jagust, W.J., Petersen, R.C., Weiner, M.W., Aisen, P.S., Shaw, L.M., Vemuri, P., Wiste, H.J., Weigand, S.D., Lesnick, T., Pankratz, V.S., Donohue, M., Trojanowski, J.Q., 2013. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol. 12, 207–216.
Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R.V., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage 46, 786–802.
Krueger, G., Granziera, C., Jack Jr., C.R., Gunter, J.L., Littmann, A., Mortamet, B., Kannengiesser, S., Sorensen, A.G., Ward, C.P., Reyes, D.A., Britson, P.J., Fischer, H., Bernstein, M.A., 2012. Effects of MRI scan acceleration on brain volume measurement consistency. J. Magn. Reson. Imaging 36, 1234–1240.
Kruggel, F., Turner, J., Muftuler, L.T., 2010. Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort. NeuroImage 49, 2123–2133.
Leow, A.D., Yanovsky, I., Parikshak, N., Hua, X., Lee, S., Toga, A.W., Jack Jr., C.R., Bernstein, M.A., Britson, P.J., Gunter, J.L., Ward, C.P., Borowski, B., Shaw, L.M., Trojanowski, J.Q., Fleisher, A.S., Harvey, D., Kornak, J., Schuff, N., Alexander, G.E., Weiner, M.W., Thompson, P.M., 2009. Alzheimer's disease neuroimaging initiative: a one-year follow up study using tensor-based morphometry correlating degenerative rates, biomarkers and cognition. NeuroImage 45, 645–655.
Leung, K.K., Barnes, J., Ridgway, G.R., Bartlett, J.W., Clarkson, M.J., Macdonald, K., Schuff, N., Fox, N.C., Ourselin, S., 2010a. Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. NeuroImage 51, 1345–1359.
Leung, K.K., Clarkson, M.J., Bartlett, J.W., Clegg, S., Jack Jr., C.R., Weiner, M.W., Fox, N.C., Ourselin, S., 2010b. Robust atrophy rate measurement in Alzheimer's disease using multi-site serial MRI: tissue-specific intensity normalization and parameter selection. NeuroImage 50, 516–523.
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.
McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack Jr., C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., Morric, J.C., Rosser, M.N., Scheltens, P., Thies, W., Weintraub, S., Phelps, C.H., 2011. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging and the Alzheimer's Assocation Workgroup. Alzheimers Dement. 7, 263–269.
Mormino, E.C., Kluth, J.T., Madison, C.M., Rabinovici, G.D., Baker, S.L., Miller, B.L., Koeppe, R.A., Mathis, C.A., Weiner, M.W., Jagust, W.J., 2009. Episodic memory loss is related to hippocampal-mediated beta-amyloid deposition in elderly subjects. Brain 132, 1310–1323.
Newcombe, R.G., 2006. Confidence intervals for an effect size measure based on the Mann–Whitney statistic. Part 2: asymptotic methods and evaluation. Stat. Med. 25, 559–573.
Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P., 1999. SENSE: sensitivity encoding for fast MRI. Magn. Reson. Med. 42, 952–962.
Risacher, S.L., Shen, L., West, J.D., Kim, S., McDonald, B.C., Beckett, L.A., Harvey, D.J., Jack Jr., C.R., Weiner, M.W., Saykin, A.J., 2010. Longitudinal MRI atrophy biomarkers: relationship to conversion in the ADNI cohort. Neurobiol. Aging 31, 1401–1418.
Rosner, B., 2011. Fundamentals of Biostatistics. Brooks/Cole, Cengage Learning, Boston.
Savva, G.M., Wharton, S.B., Ince, P.G., Forster, G., Matthews, F.E., Brayne, C., 2009. Age, neuropathology, and dementia. N. Engl. J. Med. 360, 2302–2309.
Schott, J.M., Bartlett, J.W., Barnes, J., Leung, K.K., Ourselin, S., Fox, N.C., 2010. Reduced sample sizes for atrophy outcomes in Alzheimer's disease trials: baseline adjustment. Neurobiol. Aging 31, 1452–1462 (1462 e1451-1452).
Sodickson, D.K., Manning, W.J., 1997. Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays. Magn. Reson. Med. 38, 591–603.
Thompson, W.K., Holland, D., 2011. Bias in tensor based morphometry Stat-ROI measures may result in unrealistic power estimates. NeuroImage 57, 1–4.
Vemuri, P., Gunter, J.L., Senjem, M.L., Whitwell, J.L., Kantarci, K., Knopman, D.S., Boeve, B.F., Petersen, R.C., Jack Jr., C.R., 2008. Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. NeuroImage 39, 1186–1197.
Vemuri, P., Wiste, H.J., Weigand, S.D., Knopman, D.S., Trojanowski, J.Q., Shaw, L.M., Bernstein, M.A., Aisen, P.S., Weiner, M., Petersen, R.C., Jack Jr., C.R., Initiative, A.S.D.N., 2010. Serial MRI and CSF biomarkers in normal aging, MCI and AD. Neurology 75, 143–151.