



Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: [www.elsevier.com/locate/ynimg](http://www.elsevier.com/locate/ynimg)

## Consistent multi-time-point brain atrophy estimation from the boundary shift integral

Kelvin K. Leung<sup>a,b,\*</sup>, Gerard R. Ridgway<sup>a,c</sup>, Sébastien Ourselin<sup>a,b,1</sup>, Nick C. Fox<sup>a,1</sup>  
and The Alzheimer's Disease Neuroimaging Initiative<sup>2</sup>

<sup>a</sup> Dementia Research Centre (DRC), UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK

<sup>b</sup> Centre for Medical Image Computing (CMIC), Department of Medical Physics and Bioengineering, University College London, WC1E 6BT, UK

<sup>c</sup> Wellcome Trust Centre for Neuroimaging, 12 Queen Square, London, WC1N 3BG, UK

### ARTICLE INFO

#### Article history:

Received 28 June 2011

Revised 12 October 2011

Accepted 17 October 2011

Available online xxxx

#### Keywords:

Boundary shift integral

Symmetric registration

Differential bias correction

Brain atrophy

Alzheimer's disease

### ABSTRACT

Brain atrophy measurement is increasingly important in studies of neurodegenerative diseases such as Alzheimer's disease (AD), with particular relevance to trials of potential disease-modifying drugs. Automated registration-based methods such as the boundary shift integral (BSI) have been developed to provide more precise measures of change from a pair of serial MR scans. However, when a method treats one image of the pair (typically the baseline) as the reference to which the other is compared, this systematic asymmetry risks introducing bias into the measurement. Recent concern about potential biases in longitudinal studies has led to several suggestions to use symmetric image registration, though some of these methods are limited to two time-points per subject. Therapeutic trials and natural history studies increasingly involve several serial scans, it would therefore be useful to have a method that can consistently estimate brain atrophy over multiple time-points. Here, we use the log-Euclidean concept of a within-subject average to develop affine registration and differential bias correction methods suitable for any number of time-points, yielding a longitudinally consistent multi-time-point BSI technique. Baseline, 12-month and 24-month MR scans of healthy controls, subjects with mild cognitive impairment and AD patients from the Alzheimer's Disease Neuroimaging Initiative are used for testing the bias in processing scans with different amounts of atrophy. Four tests are used to assess bias in brain volume loss from BSI: (a) inverse consistency with respect to ordering of pairs of scans 12 months apart; (b) transitivity consistency over three time-points; (c) randomly ordered back-to-back scans, expected to show no consistent change over subjects; and (d) linear regression of the atrophy rates calculated from the baseline and 12-month scans and the baseline and 24-month scans, where any additive bias should be indicated by a non-zero intercept. Results indicate that the traditional BSI processing pipeline does not exhibit significant bias due to its use of windowed sinc interpolation, but with linear interpolation and asymmetric registration, bias can be pronounced. Either improved interpolation or symmetric registration alone can greatly reduce this bias, and our proposed method combining both aspects shows no significant bias in any of the four experiments.

© 2011 Elsevier Inc. All rights reserved.

### Introduction

Cerebral atrophy is a characteristic and relentlessly progressive feature of Alzheimer's disease (AD) (Fox and Schott, 2004). Atrophy correlates with neuronal loss at autopsy (Brun and Glund, 2002)

and with cognitive decline in life (Fox et al., 1999; Jack et al., 2004). For these reasons, there is great interest in using rate of cerebral atrophy, measured from serial MRI, as an outcome measure in trials of potential disease-modifying therapies (Jack et al., 2008b; Hampel et al., 2010). Automated image analysis techniques are increasingly used in the measurement of brain atrophy from MRI (Freeborough et al., 1997; Smith et al., 2001; Boyes et al., 2006; Desikan et al., 2008; Hua et al., 2009; Avants et al., 2010; Leung et al., 2010). Automated measures have obvious attractions in terms of avoiding time-consuming manual measurements. There is also the potential for improved precision and reliability, however it is essential that automated measures do not themselves introduce a bias (Fox et al., 2011; Reuter and Fischl, 2011) (to distinguish between consistent errors in measurements and RF-inhomogeneity-induced intensity changes in MR scans, we refer to the latter as 'bias field'). Methodological biases

\* Corresponding author at: Dementia Research Centre (DRC), UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK.

E-mail address: [kk.leung@ucl.ac.uk](mailto:kk.leung@ucl.ac.uk) (K.K. Leung).

<sup>1</sup> Denotes equal senior author.

<sup>2</sup> Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([www.loni.ucla.edu/ADNI](http://www.loni.ucla.edu/ADNI)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. ADNI investigators included (complete listing available at [http://adni.loni.ucla.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Autorship\\_List.pdf](http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Autorship_List.pdf)).

could confound placebo-treatment comparisons in trials, and biased measurements from a pilot study could lead to under-powering or over-powering of subsequent trials, with ethical as well as scientific and financial implications (Thompson et al., 2011). One potential source of bias arises if the measurement process is not independent of the order of different time points. For example, Yushkevich et al. (2010) showed that deformation-based morphometry (DBM) introduces bias in the measurement of hippocampal atrophy if there is asymmetry in the application of global transformations between serial MRI scans. Asymmetry here refers to the baseline and repeat scans being treated differently in some way, e.g. when only the repeat scans are transformed and interpolated. Furthermore, Yushkevich et al. (2010) suggested that the 'asymmetry in the application of the global transformation between serial MRI images was the leading contributor to bias, whereas the asymmetry in the high-dimensional deformable transformation is less implicated in the bias'.

Registration-based methods of calculating atrophy (e.g. structural image evaluation, using normalization, of atrophy (SIENA) (Smith et al., 2001) and the boundary shift integral (BSI) (Freeborough et al., 1997; Leung et al., 2010)) offer more precise measurements of brain atrophy from serial MRI scans than manual measurements. The BSI involves global registration of repeat scans to the baseline scans, followed by differential bias correction (DBC) (Lewis and Fox, 2004) and the calculation of the BSI itself. However, the global registrations used in both Freeborough et al. (1997) and Leung et al. (2010) were asymmetric, with the global transformations being applied to the repeat scans, with consequent risk of bias (Yushkevich et al., 2010). In addition, although DBC applies the differential bias field symmetrically to two time points, the current formulation does not extend to more than two time points. For example, given a baseline scan  $I_1$  (acquired at time  $t_1$ ) and two repeat scans ( $I_2$  and  $I_3$  acquired at time  $t_2$  and  $t_3$ ), DBC is often applied separately to the registered scan-pair  $I_1$  and  $I_2$  for calculating atrophy between  $t_1$  and  $t_2$ , and to the registered scan-pair  $I_1$  and  $I_3$  for calculating atrophy between  $t_1$  and  $t_3$ . The baseline scan  $I_1$  is thus processed differently depending on which pair of scans are being analysed, which may introduce inconsistencies in the measurement of brain atrophy in studies with more than two time points. Here, we aim to make these steps as unbiased as possible by ensuring that the results in each step are inverse- and transitive-consistent over multiple time points.

Inverse inconsistency has long been recognised as a problem in image registration (Christensen and Johnson, 2001): the transformation that maps a scan  $I_1$  to another scan  $I_2$  does not generally equal the inverse of the transformation that maps  $I_2$  to  $I_1$ . Many non-linear registration algorithms exist that estimate consistent forward and backward transformations (Christensen and Johnson, 2001; Leow et al., 2005; Ye and Chen, 2009), or estimate midway transformations that warp  $I_1$  and  $I_2$  to a middle space (Rogelj and Kovacic, 2006; Beg and Khan, 2007; Avants et al., 2008; Yang et al., 2008). For affine registration, Smith et al. (2001) applied the 'square root' of affine matrix to transform  $I_1$  and  $I_2$  to a halfway position, so that 'both images are subjected to a similar degree of interpolation-related blurring'. In addition, when registering multiple scans ( $I_1$ ,  $I_2$  and  $I_3$ ) such as those acquired in a longitudinal study, transitive inconsistency refers to the problem that the transformation that maps  $I_1$  to  $I_3$  does not equal the composition of the transformations that map  $I_1$  to  $I_2$  and  $I_2$  to  $I_3$ . Geng et al. (2005) extended the work of Christensen and Johnson (2001) to include a penalty term for minimising transitive errors, whereas Skrinjar et al. (2008) constructed inverse- and transitive-consistent transformations from the pairwise transformations to a reference scan. Other related works include groupwise registration, which aims to register multiple scans into an unbiased middle space (for example Woods et al., 1998; Aljabar et al., 2006). Recently, Reuter et al. (2010) proposed a robust and highly accurate inverse-consistent linear registration algorithm that maps two scan  $I_1$  and  $I_2$  to a midway space.

In this paper, we set out a multi-time-point and symmetric method for calculating atrophy, and assess the technique for bias. We used MR scans of healthy controls, subjects with mild cognitive impairment (MCI) and patients with Alzheimer's disease (AD), downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) website ([www.loni.ucla.edu/ADNI](http://www.loni.ucla.edu/ADNI)) as the test data.

## Methods

### Overview

We first describe the processing pipeline of the boundary shift integral, and then the differential bias correction and symmetric global linear registration for multiple time points. The key mathematical basis of this work is the link between scalar or matrix multiplication and the geometric mean. Because the bias field in an MR image is typically modelled as a multiplicative scaling of the image intensities, Lewis and Fox (2004) proposed to apply half of the differential bias symmetrically to each image by multiplying one image by the square root of the relative multiplicative bias field and the other by the inverse of this. This idea extends naturally to the geometric mean, which can be expressed as an arithmetic mean in log-space, minimising the sum of squared 'distances' (in terms of log ratios) from the mean to the set of images. Spatial affine transformations can be expressed in terms of matrix multiplications in homogeneous coordinates, and one can use the simple and efficient log-Euclidean distance metric and associated mean (Arsigny et al., 2006). This directly generalises the log-space mean used in DBC to provide a geometric mean of affine transformations using the matrix logarithm (Alexa, 2002; Aljabar et al., 2006). In the following sections, we will illustrate our methods using scans from three time points ( $I_1$ ,  $I_2$  and  $I_3$  acquired at time  $t_1$ ,  $t_2$  and  $t_3$ ). In addition, key results will be stated for the general case of  $n$  images  $\{I_i\}_{i=1}^n$ .

We evaluate our novel multi-time-point symmetric BSI pipeline using MR scans of healthy controls, subjects with MCI and patients with AD, downloaded from the ADNI website. The use of different disease groups allows assessment of potential interaction between bias and amount of atrophy. We assess the bias in the results arising from asymmetric and symmetric DBC and registration schemes with different choices of interpolation using four tests. The first test examines the inverse consistency of the atrophy calculated from the baseline and 24-month scans (Christensen and Johnson, 2001; Fox et al., 2011). The calculated atrophy should be independent of the order of the time points specified in the image processing pipeline. The second test examines the transitivity of the atrophy calculated from the baseline, 12-month and 24-month scans (Fox et al., 2011). The sum of atrophy calculated from the baseline and 12-month scans, and the 12-month and 24-month scans should equal the atrophy calculated from the baseline and 24-month scan. The third test uses the availability of two back-to-back scans at each time point from ADNI to perform a direct test of bias (Yushkevich et al., 2010), as the mean atrophy between the two scans is expected to be zero. The fourth test uses the intercept of a regression line fitted to the atrophy rates calculated from the baseline and 12-month scans and the baseline and 24-month scans (Yushkevich et al., 2010; Thompson et al., 2011; Hua et al., 2011). Under the assumption of exactly linear progression of atrophy, additive bias would be indicated by a non-zero intercept.

### Boundary shift integral (BSI)

The processing pipeline of BSI involves the following steps (Freeborough et al., 1997; Leung et al., 2010):

1. *Brain segmentation.* The whole brains in the baseline and repeat scans are automatically and independently delineated from dura, scalp and other non-brain tissue using a multi-atlas propagation

and segmentation (MAPS) technique (Leung et al., 2011). MAPS uses non-linear registration of the top 19 best-matched templates from a manually-segmented library of 682 brain scans to generate multiple segmentations, and combines them using shape based averaging (Rohlfing and Maurer, 2007) to create a reliable and accurate brain segmentation.

2. *Global linear registration.* The brains on the scans are registered using a global linear transformation.
3. *Differential bias correction (DBC).* DBC is applied to correct for the differences in intensity inhomogeneity between the registered scans.
4. *Calculation of brain volume difference.* KN-BSI (Leung et al., 2010) is used to calculate the brain volume difference from the DBC-corrected scans.

We describe the DBC and global linear registration in more details in the following sections.

#### Differential bias correction (DBC)

This section describes the simple extension of the formulation of DBC (Lewis and Fox, 2004) from two time points into multiple time points. The formation of an MR image is assumed to be

$$v(x) = u(x)b(x) + n(x), \quad (1)$$

where  $v$  is the measured intensity,  $b$  is the bias field,  $u$  is the true signal intensity and  $n$  is the noise. Given two images  $v_1$  and  $v_2$ , the ratio of the bias fields, called the differential bias field, is then estimated via log-space median-filtering of the ratio of the observed images, under the assumption that the median-filter will remove the higher-spatial frequency differences in anatomical structure (e.g. atrophy and noise) in the ratio, leaving the low spatial frequency intensity inhomogeneity:

$$r_{12} = b_1/b_2 \approx \exp(\text{median-filter}(\log(v_1) - \log(v_2))). \quad (2)$$

Using the square root of the differential bias field to create the corrected images,  $v'_1$  and  $v'_2$ , results in both images having the same midway bias field ( $\sqrt{b_1 b_2}$ )

$$\begin{aligned} v'_1 &= v_1/\sqrt{r_{12}} = u_1\sqrt{b_1 b_2} + n_1/\sqrt{r_{12}}, \\ v'_2 &= v_2/\sqrt{r_{21}} = u_2\sqrt{b_1 b_2} + n_2/\sqrt{r_{21}}. \end{aligned}$$

To extend to the case of multiple time points, with  $v_i = u_i b_i + n_i$ , the pairwise differential bias fields are

$$r_{ij} = b_i/b_j = 1/r_{ji}, \quad (3)$$

and we create the  $i$ -th corrected image by dividing it by the geometric mean of its differential bias fields with respect to the set of  $n$  images:

$$v'_i = v_i/r_i^{\text{average}}, \quad (4)$$

where

$$r_i^{\text{average}} = \sqrt[n]{\prod_{j=1}^n r_{ij}} = \exp\left(\frac{1}{n} \sum_{j=1}^n \log r_{ij}\right). \quad (5)$$

To illustrate this with three images, the geometric means of the pairwise differential bias fields are

$$\begin{aligned} r_1^{\text{average}} &= \sqrt[3]{r_{11}r_{12}r_{13}} = \sqrt[3]{b_1^2/(b_2 b_3)}, \\ r_2^{\text{average}} &= \sqrt[3]{r_{21}r_{22}r_{23}} = \sqrt[3]{b_2^2/(b_1 b_3)}, \\ r_3^{\text{average}} &= \sqrt[3]{r_{31}r_{32}r_{33}} = \sqrt[3]{b_3^2/(b_1 b_2)}. \end{aligned}$$

and the corrected images are

$$\begin{aligned} v'_1 &= v_1/r_1^{\text{average}} = u_1\sqrt[3]{b_1 b_2 b_3} + n_1/r_1^{\text{average}}, \\ v'_2 &= v_2/r_2^{\text{average}} = u_2\sqrt[3]{b_1 b_2 b_3} + n_2/r_2^{\text{average}}, \\ v'_3 &= v_3/r_3^{\text{average}} = u_3\sqrt[3]{b_1 b_2 b_3} + n_3/r_3^{\text{average}}. \end{aligned}$$

All of them having a common bias field  $\sqrt[3]{b_1 b_2 b_3}$ .

#### Global linear registration

##### Asymmetric registration

Let  $I_1(T)$  be the resampled image of a scan  $I_1$  under a transformation  $T$ , and  $\text{similarity}(I_1, I_2)$  be a similarity measure between two scans  $I_1$  and  $I_2$ . The transformation  $T_{\text{asym}}(I_1, I_2)$  that aligned  $I_1$  and  $I_2$  using an asymmetric registration scheme (see Fig. 1(a)) is given by

$$T_{\text{asym}}(I_1, I_2) = \arg \max_T \text{similarity}(I_1, I_2(T)). \quad (6)$$

i.e.  $T_{\text{asym}}(I_1, I_2)$  is applied to resample scan  $I_2$  to the space of scan  $I_1$ .

##### Symmetric registration

In order to remove the dependency on the order of different time points in the measurement process, we registered the baseline

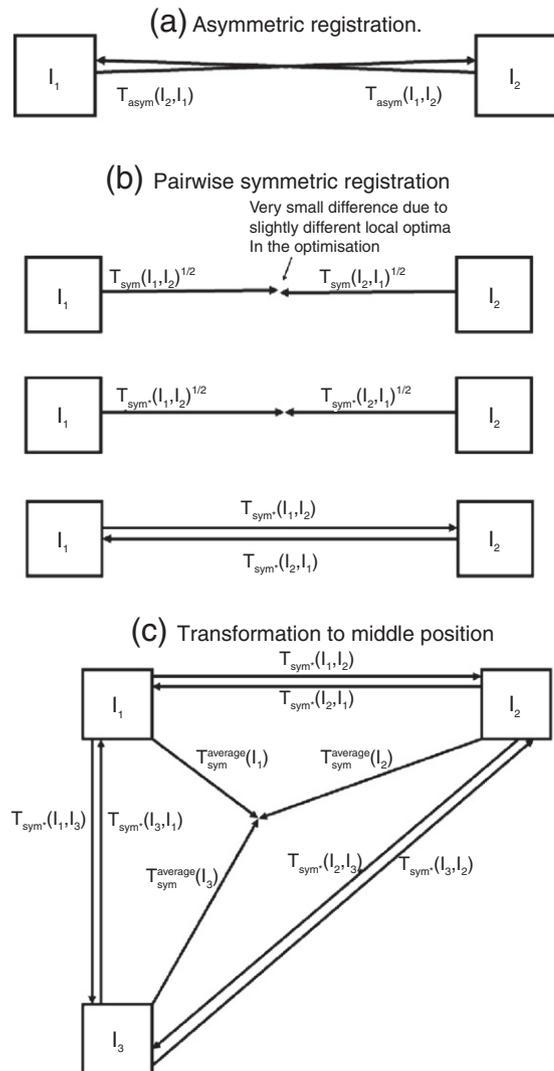


Fig. 1. Asymmetric and symmetric registration schemes.

and repeat scans using symmetric registration (Reuter et al., 2010) and groupwise registration (Aljabar et al., 2006) schemes. As shown in Fig. 1, after the pairwise symmetric registrations were performed to find the transformations between the scans, the positions of the registered scans were calculated by averaging the results of the pairwise registrations. The details of each step are described below:

1. *Pairwise symmetric registration.* Our aim was to estimate a global linear transformation  $T_{sym}(I_1, I_2)^{1/2}$  which aligned  $I_1$  and  $I_2$  to a middle space. Therefore,  $T_{sym}(I_1, I_2)^{1/2}$  is given by

$$T_{sym}(I_1, I_2)^{1/2} = \arg \max_T \text{similarity}(I_1(T), I_2(T^{-1})). \quad (7)$$

Ideally,  $T_{sym}(I_1, I_2)^{1/2} = (T_{sym}(I_2, I_1)^{1/2})^{-1}$ , meaning the registration is symmetric and inverse-consistent. In practice, however, they might not be exactly the same, because the optimisation process might reach slightly different local optima of the objective function for  $T$  and  $T^{-1}$ . Therefore, we calculate the average of the two transformations (denoted as  $T_{sym}(I_1, I_2)$ ) (Alexa, 2002) (see Fig. 1(b)):

$$T_{sym^*}(I_1, I_2)^{1/2} = \text{expm} \frac{\logm T_{sym}(I_1, I_2)^{1/2} + \logm (T_{sym}(I_2, I_1)^{1/2})^{-1}}{2}, \quad (8)$$

where  $\text{expm}$  and  $\logm$  are matrix exponential and logarithm. This averaged transformation is now symmetric, as can be shown by noting that the familiar property  $\log(1/x) = -\log(x)$  generalises to the matrix case,  $\logm(X^{-1}) = -\logm X$ ,

$$\begin{aligned} T_{sym^*}(I_2, I_1)^{1/2} &= \text{expm} \frac{\logm T_{sym}(I_2, I_1)^{1/2} + \logm (T_{sym}(I_1, I_2)^{1/2})^{-1}}{2} \\ &= \text{expm} \frac{-\logm (T_{sym}(I_2, I_1)^{1/2})^{-1} - \logm T_{sym}(I_1, I_2)^{1/2}}{2} \\ &= \text{expm} \frac{\logm T_{sym}(I_1, I_2)^{1/2} + \logm (T_{sym}(I_2, I_1)^{1/2})^{-1}}{2} \\ &= (T_{sym^*}(I_1, I_2)^{1/2})^{-1}. \end{aligned}$$

2. *Transformation to the average position.* We perform all the pairwise symmetric registrations to obtain  $T_{sym}(I_1, I_2)$ ,  $T_{sym}(I_1, I_3)$  and  $T_{sym}(I_2, I_3)$ . The scans are then transformed to the average position (Alexa, 2002; Aljabar et al., 2006) by the geometric means of the pairwise transformations (see Fig. 1(c)):

$$\begin{aligned} T_{sym}^{average}(I_1) &= \text{expm} \frac{\logm T_{sym^*}(I_1, I_1) + \logm T_{sym^*}(I_1, I_2) + \logm T_{sym^*}(I_1, I_3)}{3} \\ &= \text{expm} \frac{\logm T_{sym^*}(I_1, I_2) + \logm T_{sym^*}(I_1, I_3)}{3}, \\ T_{sym}^{average}(I_2) &= \text{expm} \frac{\logm T_{sym^*}(I_2, I_1) + \logm T_{sym^*}(I_2, I_2) + \logm T_{sym^*}(I_2, I_3)}{3} \\ &= \text{expm} \frac{\logm T_{sym^*}(I_2, I_1) + \logm T_{sym^*}(I_2, I_3)}{3}, \\ T_{sym}^{average}(I_3) &= \text{expm} \frac{\logm T_{sym^*}(I_3, I_1) + \logm T_{sym^*}(I_3, I_2) + \logm T_{sym^*}(I_3, I_3)}{3} \\ &= \text{expm} \frac{\logm T_{sym^*}(I_3, I_1) + \logm T_{sym^*}(I_3, I_2)}{3}. \end{aligned}$$

In general, for  $n$  images, the average position of the  $i$ -th image is given by the following transformation:

$$T_{sym}^{average}(I_i) = \text{expm} \frac{1}{n} \sum_{j=1}^n \logm T_{sym^*}(I_i, I_j), \quad (9)$$

which has the same form as Eq. (5).

The pairwise transformations between  $I_1$ ,  $I_2$  and  $I_3$  under this averaging scheme are then given by composing transformations to and from the average:

$$T(I_1, I_2) = T_{sym}^{average}(I_1) \circ T_{sym}^{average}(I_2)^{-1}, \quad (10)$$

$$T(I_2, I_3) = T_{sym}^{average}(I_2) \circ T_{sym}^{average}(I_3)^{-1}, \quad (11)$$

$$T(I_1, I_3) = T_{sym}^{average}(I_1) \circ T_{sym}^{average}(I_3)^{-1}. \quad (12)$$

Importantly, these can be shown to satisfy transitivity exactly. For example:

$$T(I_1, I_2) \circ T(I_2, I_3) = (T_{sym}^{average}(I_1) \circ T_{sym}^{average}(I_2)^{-1}) \circ (T_{sym}^{average}(I_2) \circ T_{sym}^{average}(I_3)^{-1}) \quad (13)$$

$$= T_{sym}^{average}(I_1) \circ T_{sym}^{average}(I_3)^{-1} \quad (14)$$

$$= T(I_1, I_3). \quad (15)$$

#### Implementation details

We use a two-stage global linear registration in the BSI processing pipeline. The first stage is used to recover scanner scaling artefacts by performing a 12-degree of freedom (DOF) registration over 8-voxel dilated brain regions (obtained by performing binary morphological dilation on the brain regions 8 times using a cross structuring element of size  $3 \times 3 \times 3$ ) (Clarkson et al., 2009). We use 12-DOF instead of the 9-DOF used by Clarkson et al. (2009) because 9-DOF registration is inherently asymmetric – allowing only scalings that are aligned with the reference axes, not the source. If either image can be scaled anisotropically along its own axes, and the images are acquired such that these axes need to be rotated to align anatomy, then the separate scalings together with the rotation between the pairs of axes effectively allow skews. More formally, 12-DOF transformations form a matrix Lie group with an associated semi-Riemannian manifold (Woods, 2003) so their inverses and compositions are also 12-DOF; this is not generally true of 9-DOF transformations, whose inverses or compositions are only guaranteed to be within the broader 12-DOF group.

We parameterise the 12-DOF transformation as 3 translations, 3 rotations in Euler angles, 3 scaling factors and 3 skew factors, applied in the reverse of that order (i.e. skews first, translations last). The second stage is then to accurately register the brains, without further changing the scalings or skews – to avoid the registration masking any global component of atrophy. This is performed by re-optimising the 3 translations and 3 rotations of the transformation from the first stage over 2-voxel dilated brain regions. The dilation is intended to restore potentially missing brain tissue and to capture edge information for the registration.

During image registration, the scans  $I_1$  and  $I_2$  in Eq. (7) are transformed into a middle space by the transformations  $T$  and  $T^{-1}$  of the current iteration using tri-linear interpolation. Apodized normalised cross correlation (Jenkinson et al., 2002) is used as the similarity measure between the resampled scans. The apodization removes small local minima in the normalised cross correlation ‘formed by the changes of the number of voxels in the overlapping field of view with changing transformation parameters’ (Jenkinson et al., 2002). In order to minimise computational cost, we use a distance transform to obtain the distances of each voxel from the edges of the dilated brain regions in  $I_1$  and  $I_2$ , and use tri-linear interpolation to calculate the distance to the edge at any location during image registration. The distance of each voxel within the overlapping dilated brain region is then given by the minimum of the distances from the edges of dilated

brain regions. The distance threshold was empirically chosen here to be 2 mm.

The optimisation is performed using Powell's method (Press et al., 1992), a derivative-free direction-set procedure, which avoids the non-trivial analytical computation of the derivatives of the apodized normalised cross correlation and the complication that the derivatives with respect to the parameters of the forward transformation could be numerically inconsistent with the gradient of the inverse transform. The optimisation stopped when the RMS deviation (Jenkinson, 1999) within a spherical volume of radius 100 mm between two successive transformations was less than 0.001 mm.

The final resampling of the transformed scan was performed either by (a) tri-linear interpolation or (b) windowed sinc interpolation with a Welch window,  $1 - x^2/m^2$ , of radius  $m=5$  (Meijering et al., 1999). Although a high-order interpolation is preferred in the resampling of images, it is not used in all situations (such as during the parameter-estimation stage of image registration) due to its high computational cost. We therefore compared the results generated from the linear and the windowed sinc interpolation.

For DBC, we followed the suggestion in the original paper (Lewis and Fox, 2004) to use a median filter of radius 5 (i.e. a  $11 \times 11 \times 11$  box).

All the algorithms in this manuscript were implemented using the Insight Segmentation and Registration Toolkit ([www.itk.org](http://www.itk.org)).

## Experiments

### Image data

Our test data from ADNI were chosen on the basis that subjects had baseline, 12-month and 24-month 1.5T  $T_1$ -weighted volumetric MRI scans. In total, MRI scans of 439 subjects (138 controls, 211 MCI and 90 AD) were downloaded from the ADNI database, and a list of the ADNI subject IDs is provided in the supplementary material to facilitate direct comparison or replication of results. The mean (SD) age of the groups were: 76.1 (4.6) years for the controls, 75.1 (6.9) years for the MCI subjects, and 75.8 (7.1) years for the AD subjects. The mean (SD) of the MMSE (a cognitive test with a maximum score of 30) of the controls was 29.2 (0.9), which was higher than the MCI subjects (27.1 (1.7)) or the AD subjects (23.3 (1.9)). Representative imaging parameters were TR=2300 ms, TI=1000 ms, TE=3.5 ms, flip angle=8°, field of view=240×240 mm and 160 sagittal 1.2 mm-thick-slices and a 192×192 matrix yielding a voxel resolution of 1.25×1.25×1.2 mm, or 180 sagittal 1.2 mm-thick-slices with a 256×256 matrix yielding a voxel resolution of 0.94×0.94×1.2 mm. The full details of the ADNI MR imaging protocol are described by Jack et al. (2008a), and are listed on the ADNI website (<http://www.loni.ucla.edu/ADNI/Research/Cores/>). Two back-to-back  $T_1$ -weighted scans were acquired during each scanning session. Each scan underwent a quality control evaluation at the Mayo Clinic (Rochester, MN, USA). Quality control included inspection of each incoming image file for protocol compliance, clinically significant medical abnormalities, and image quality. Out of the two back-to-back  $T_1$ -weighted scans, the scan with better quality was processed using the standard ADNI image processing pipeline, which included post-acquisition correction of gradient warping (Jovicich et al., 2006), B1 non-uniformity correction (Narayana et al., 1988) depending on the scanner and coil type, intensity non-uniformity correction (Sled et al., 1998) and phantom based scaling correction (Gunter et al., 2006) with the geometric phantom scan having been acquired with each patient scan. In addition to the processed scans, we also downloaded the unprocessed back-to-back  $T_1$ -weighted scans of the baseline time point. We then applied intensity non-uniformity correction (Sled et al., 1998) to the unprocessed scans inside brain masks as described by Boyes et al. (2008).

### Assessment of bias

For ease of understanding and interpretation, we decided to report absolute brain volume loss in millilitres (ml), instead of relative brain volume loss as a percentage of (a chosen time-point's) whole brain volume, although the latter (using the baseline) is more commonly used in the literature to correct for the differences in brain volumes between different subjects. Furthermore, we note that the adjustment of nuisance variables including whole brain volume and age may be better performed in a single statistical model (Schott et al., 2010). We do not expect the use of percentage of whole brain volume to have a significant effect on the results, given that the calculation of the whole brain volume itself is not biased.

### Asymmetric and symmetric registration schemes

We now briefly summarise each registration scheme:

- *Brain volume loss using asymmetric registration scheme ( $L^{asym}$ )*. The scans were registered and transformed to a single time point (Section "Asymmetric registration"). DBC was applied to each pair of images as in Lewis and Fox (2004) before calculation of the BSI.
- *Brain volume loss using symmetric registration scheme ( $L^{sym}$ )*. All the scans were registered and transformed to a middle position (Section "Symmetric registration"). Symmetric DBC was applied (Section "Differential bias correction (DBC)") before calculation of the BSI.

### Experimental details

We used the following experiments to assess bias in brain volume loss introduced by the different BSI processing pipelines with symmetric and asymmetric registration and DBC schemes:

1. *Test of inverse consistency using two time points*. Volume difference should be independent of the order of the scans given to the BSI processing pipelines. We compared brain volume loss between the baseline and 24-month scans calculated from the asymmetric registration scheme:

(a)  $L^{asym}$ .

- i.  $L_{0-24}^{asym}$ . The 24-month scan was transformed to the baseline scan.
- ii.  $L_{24-0}^{asym}$ . The baseline scan was transformed to the 24-month scan.

We model the measured (biased) brain volume loss as the sum of an additive bias and the brain volume loss without it. If  $\varepsilon$  be the additive bias and  $\hat{L}$  the true brain volume loss, we have

$$L_{0-24}^{asym} = \hat{L}_{0-24} + \varepsilon_{0-24}, \quad (16)$$

$$L_{24-0}^{asym} = \hat{L}_{0-24} - \varepsilon_{0-24}. \quad (17)$$

$\varepsilon_{0-24}$  is given by half of the difference between  $L_{0-24}^{asym}$  and  $L_{24-0}^{asym}$ . We used a paired  $t$ -test to assess the differences between  $L_{0-24}^{asym}$  and  $L_{24-0}^{asym}$ .

2. *Test of transitivity using three time points*. We calculated brain volume loss using the asymmetric and symmetric registration scheme between the baseline and 12-month scans ( $L_{0-12}$ ), between the 12-month and 24-month scans ( $L_{12-24}$ ), and between the baseline and 24-month scans ( $L_{0-24}$ ). If there were no bias in brain volume loss, the sum of  $L_{0-12}$  and  $L_{12-24}$  should equal  $L_{0-24}$ .

(a)  $L^{asym}$ .

- i.  $L_{0-12}^{asym}$ . The 12-month scan was transformed to the baseline scan.
- ii.  $L_{12-24}^{asym}$ . The 24-month scan was transformed to the 12-month scan.
- iii.  $L_{0-24}^{asym}$ . The 24-month scan was transformed to the baseline scan.

(b)  $L^{sym}$ .

From our model of additive bias in brain volume loss, we have

$$L_{0-12} = \hat{L}_{0-12} + \varepsilon_{0-12}, \tag{18}$$

$$L_{12-24} = \hat{L}_{12-24} + \varepsilon_{12-24}, \tag{19}$$

$$L_{0-24} = \hat{L}_{0-24} + \varepsilon_{0-24}. \tag{20}$$

Assuming that  $\hat{L}_{0-12} + \hat{L}_{12-24} = \hat{L}_{0-24}$  and the registration scheme introduces a similar amount of bias in each case (i.e.  $\varepsilon_{0-12} \approx \varepsilon_{12-24} \approx \varepsilon_{0-24}$ ), we estimated  $\varepsilon_{0-24}$  by calculating the difference between  $L_{0-12} + L_{12-24}$  and  $L_{0-24}$ . A paired *t*-test was used to test for any differences between  $L_{0-12} + L_{12-24}$  and  $L_{0-24}$ .

3. *Direct test of bias using the back-to-back scans of the baseline time point.* We used the two back-to-back  $T_1$ -weighted scans to test bias introduced by the symmetric and asymmetric registration schemes as in (Yushkevich et al., 2010). The back-to-back scans were acquired immediately after one another and so any differences are unlikely to be due to atrophy, but may represent noise, artefacts, movement, or potential physiological fluctuations. Any non-zero mean value indicated that there was an additive bias in the processing pipeline. In order to avoid any systematic differences between the first and the second of the two back-to-back scans, we randomly assigned the two scans as 'baseline' and 'repeat'. We computed brain volume loss from BSI using the following schemes:

(a)  $L^{asym}$ .

(a)  $L^{asym-forward}$ . The 'repeat' scan was transformed to the 'baseline' scan.

(b)  $L^{asym-backward}$ . The 'baseline' scan was transformed to the 'repeat' scan.

(b)  $L^{sym}$ .

A *t*-test was used to assess the non-zero mean value of brain volume loss.

4. *Test of bias using the intercept of the regression line from three time points.* We calculated brain volume loss  $L_{0-12}$  from the baseline and 12-month scans and  $L_{0-24}$  from the baseline and 24-month scans using the symmetric and asymmetric schemes listed below. A regression line was fitted to the brain volume loss. If the rate of brain volume loss was constant over the 24 months, the intercept of the regression line would be expected to be zero. Gradual acceleration would give a negative intercept, deceleration a positive one.

(a)  $L^{asym}$ .

i.  $L^{asym-forward}$ . The 12-month and 24-month scans were both transformed to the baseline scan.

ii.  $L^{asym-backward}$ . The baseline and 12-month scans were both transformed to the 24-month scan.

(b)  $L^{sym}$ .

Results and discussion

Test of inverse consistency using two time points

Table 1 shows that brain volume loss from BSI using the asymmetric registration scheme is not inverse-consistent ( $p < 0.001$ , all tests). Up to 4.3 ml difference in brain volume loss ( $L_{0-24}^{asym} > L_{24-0}^{asym}$ ) was detected between the forward and backward registration in controls when using tri-linear interpolation with the asymmetric registration scheme. The bias in brain volume loss was greatly reduced to less than 1 ml ( $L_{0-24}^{asym} < L_{24-0}^{asym}$ ) in all subject groups when windowed sinc interpolation was used. However, when using a symmetric registration scheme, brain volume loss calculated using tri-linear and

Table 1

Test of inverse consistency using two time points. The table shows the mean (SD) of brain volume loss (ml) between the baseline and 24-month scans calculated using asymmetric and symmetric registration scheme. It also shows the 95% CI (in square brackets) and *p*-value in the comparison between  $L_{0-24}^{asym}$  and  $L_{24-0}^{asym}$ .

	Control (n = 138)	MCI (n = 211)	AD (n = 90)
<i>(a) Linear interpolation</i>			
$L_{0-24}^{asym}$	18.14 (8.33)	24.21 (12.79)	32.78 (14.22)
$L_{24-0}^{asym}$	9.62 (8.62)	17.67 (13.14)	28.24 (13.35)
$(L_{0-24}^{asym} - L_{24-0}^{asym})/2$	4.26 (2.71)	3.27 (2.80)	2.27 (2.85)
	[3.80, 4.71],	[2.89, 3.65],	[1.67, 2.87],
	$p < 0.001$	$p < 0.001$	$p < 0.001$
$L_{0-24}^{sym}$ and $L_{24-0}^{sym}$	13.61 (8.10)	20.66 (12.57)	30.09 (13.65)
<i>(b) Sinc interpolation</i>			
$L_{0-24}^{asym}$	13.63 (8.01)	20.67 (12.75)	30.34 (13.77)
$L_{24-0}^{asym}$	14.25 (8.15)	21.48 (12.77)	31.13 (13.64)
$(L_{0-24}^{asym} - L_{24-0}^{asym})/2$	-0.31 (0.58)	-0.40 (0.59)	-0.39 (0.66)
	[-0.41, -0.21],	[-0.48, -0.32],	[-0.53, -0.25],
	$p < 0.001$	$p < 0.001$	$p < 0.001$
$L_{0-24}^{sym}$ and $L_{24-0}^{sym}$	13.93 (8.06)	21.01 (12.74)	30.64 (13.84)

windowed sinc interpolation was similar. This suggested that brain volume loss using the symmetric registration scheme was not severely affected by the choice of interpolation.

There is always a possible tradeoff between bias and variance; for example, independently performed measurements of brain volume from separate time-points (e.g. using MAPS; Leung et al., 2011) would have low or zero bias, but would have substantially higher variance than the BSI. Results in Table 1 show that the SD of BSI in each group is almost unchanged, while the bias is reduced when using the symmetric registration scheme.

Table 1 also shows that bias in  $L^{asym}$  with tri-linear interpolation (difference in brain volume loss in  $L_{0-24}^{asym}$  and  $L_{24-0}^{asym}$ ) is larger in controls and smaller in AD subjects. This is a surprising result, since the presence of multiplicative bias (something that is difficult to evaluate accurately without ground truth) would be expected to result in greater biases for the more affected subjects. Instead, in a post-hoc analysis using linear regression, there seems to be a slight negative correlation (regression slope [95% CI] = -0.026 [-0.046, -0.006],  $p = 0.01$ ) between the bias and brain volume loss from  $L_{0-24}^{asym}$  with windowed sinc interpolation (see Fig. 2). The intercept [95% CI] of

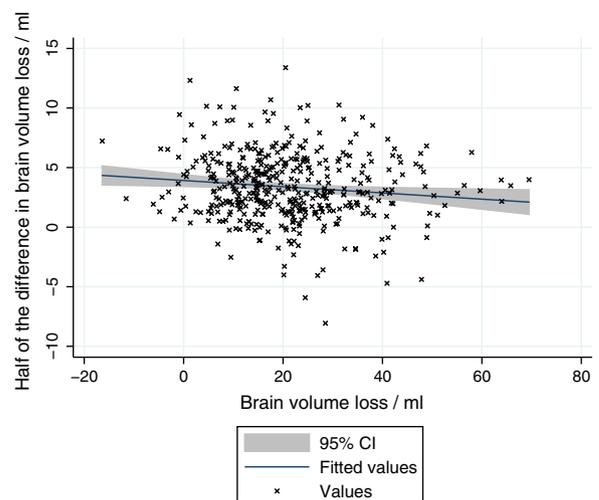


Fig. 2. Regression of bias in brain volume loss ( $(L_{0-24}^{asym} - L_{24-0}^{asym})/2$ ) against brain volume loss in the inverse consistency test using two time points. The figure shows the regression of bias in brain volume loss ( $(L_{0-24}^{asym} - L_{24-0}^{asym})/2$ ) from the asymmetric registration scheme with tri-linear interpolation in the inverse consistency test against brain volume loss from the symmetric registration scheme with windowed sinc interpolation. We found a negative correlation (-0.026,  $p = 0.001$ ) between the bias in brain volume loss and the brain volume loss.

the regression line was 3.92 [3.42, 4.42] ( $p < 0.001$ ), meaning that the difference between forward and backward  $L^{asym}$  was estimated to be 3.92 ml if there is no atrophy. No significant correlation was found between the bias and baseline brain volume (regression slope = 0.001 [-0.001, 0.003],  $p = 0.4$ ). BSI is calculated over the exclusive OR region of the 1-voxel dilated union and 1-voxel eroded intersection regions from the baseline and repeat brain regions as described in the following steps:

1. Perform binary morphological dilation on the baseline and repeat brain regions once by 1 voxel;
2. Calculate the union of the dilated regions;
3. Perform binary morphological erosion on the baseline and repeat brain regions by 1 voxel;
4. Calculate the intersection of the eroded regions;
5. Calculate the exclusive OR region from the union and intersection regions.

Therefore, the greater change in brain regions between baseline and repeat scans in AD subjects means that there is a larger exclusive OR region for more of the bias to get averaged out (see Fig. 3), though this would need to be explored more thoroughly in future work.

Test of transitivity using three time points

As shown in Table 2, brain volume loss from BSI using tri-linear interpolation with the asymmetric registration scheme was not transitive-consistent, with the sum  $L_{0-12} + L_{12-24}$  being 2-4 ml greater than  $L_{0-24}$ . When using tri-linear interpolation with the asymmetric registration scheme, the amount of transitive bias was similar to the amount of inverse bias from the corresponding entries in Table 1. Marginal statistical significance ( $p = 0.04$ ) in the transitive bias was found in controls when using tri-linear interpolation with the symmetric registration scheme, although the amount of bias was very small (0.018 ml,  $L_{0-12} + L_{12-24} < L_{0-24}$ ). Nonetheless, the symmetric registration scheme was able to reduce the bias when compared to the asymmetric registration scheme.

Small, but statistically significant ( $p < 0.001$ ), transitive bias of 0.31 ml ( $L_{0-12} + L_{12-24} < L_{0-24}$ ) was found in controls when using

Table 2

Test of transitivity using three time points. The table shows the mean (SD) of brain volume loss (ml) between the baseline and 12-month scans ( $L_{0-12}$ ), the 12-month and 24-month scans ( $L_{12-24}$ ) and the baseline and 24-month scans ( $L_{0-24}$ ). It also shows the 95% CI (in square brackets) and  $p$ -value in the comparison between  $L_{0-12} + L_{12-24}$  and  $L_{0-24}$ .

	Control (n = 138)	MCI (n = 211)	AD (n = 90)
<i>(a) Linear interpolation</i>			
$L_{0-12}^{asym}$	10.75 (7.54)	14.17 (8.68)	17.94 (8.25)
$L_{12-24}^{asym}$	11.69 (6.74)	13.27 (9.36)	17.41 (11.07)
$L_{0-24}^{asym}$	18.14 (8.33)	24.21 (12.79)	32.78 (14.22)
$L_{0-12}^{asym} + L_{12-24}^{asym} - L_{0-24}^{asym}$	4.30 (2.85)	3.23 (2.98)	2.57 (3.14)
	[3.82, 4.78], $p < 0.001$	[2.82, 3.63], $p < 0.001$	[1.91, 3.23], $p < 0.001$
$L_{0-12}^{sym}$	6.22 (6.73)	10.52 (8.45)	14.85 (7.92)
$L_{12-24}^{sym}$	7.37 (6.50)	10.14 (8.79)	15.23 (9.79)
$L_{0-24}^{sym}$	13.61 (8.10)	20.66 (12.57)	30.09 (13.65)
$L_{0-12}^{sym} + L_{12-24}^{sym} - L_{0-24}^{sym}$	-0.02 (0.10)	-0.00 (0.12)	-0.01 (0.14)
	[-0.04, -0.00], $p = 0.04$	[-0.02, 0.01], $p = 0.6$	[-0.04, 0.02], $p = 0.6$
<i>(b) Sinc interpolation</i>			
$L_{0-12}^{sym}$	5.98 (6.81)	10.56 (8.62)	15.15 (7.97)
$L_{12-24}^{sym}$	7.34 (6.49)	9.98 (8.97)	15.19 (9.86)
$L_{0-24}^{sym}$	13.63 (8.01)	20.67 (12.75)	30.34 (13.77)
$L_{0-12}^{sym} + L_{12-24}^{sym} - L_{0-24}^{sym}$	-0.31 (0.86)	-0.13 (1.05)	-0.00 (1.03)
	[-0.45, -0.16], $p < 0.001$	[-0.27, 0.01], $p = 0.07$	[-0.22, 0.21], $p = 1.0$
$L_{0-12}^{asym}$	6.24 (6.67)	10.75 (8.54)	15.29 (8.01)
$L_{12-24}^{asym}$	7.68 (6.38)	10.27 (8.88)	15.37 (9.69)
$L_{0-24}^{asym}$	13.93 (8.06)	21.01 (12.74)	30.64 (13.84)
$L_{0-12}^{asym} + L_{12-24}^{asym} - L_{0-24}^{asym}$	-0.02 (0.13)	0.00 (0.14)	0.02 (0.18)
	[-0.04, 0.01], $p = 0.1$	[-0.02, 0.02], $p = 0.7$	[-0.02, 0.06], $p = 0.3$

windowed sinc interpolation with the asymmetric registration scheme. When using windowed sinc interpolation with the symmetric registration scheme, there was no transitive bias ( $p \geq 0.1$ , all tests).

The standard deviation (SD) of the difference between  $L_{0-12} + L_{12-24}$  and  $L_{0-24}$  indicated the absolute agreement between them at a pairwise level. The SD of the difference for the symmetric

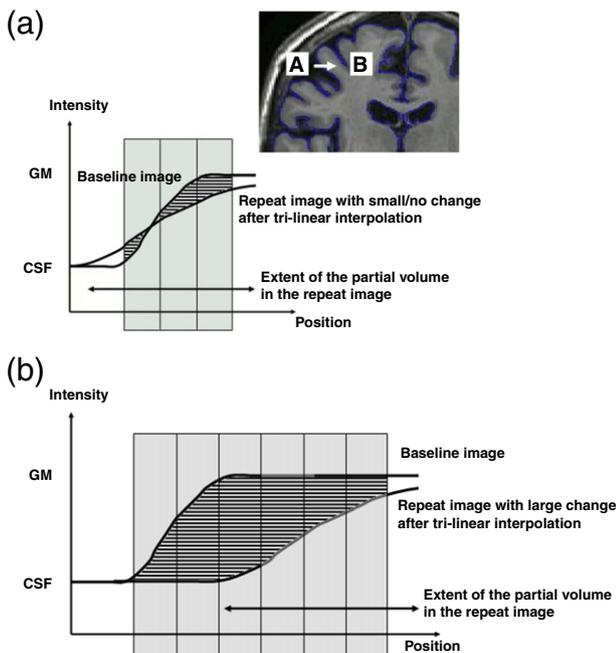


Fig. 3. Illustration of the averaging of partial volume effect in subjects with small and large brain volume changes.

Table 3

Post-hoc comparison between  $L^{sym*}$  (brain volume loss calculated from the transformation of the 0-, 12- and 24-month scans to the average position of each pair of the scans using  $T_{sym}(I_1, I_2)^{1/2}$  from Eq. (8)) and  $L^{sym}$  (brain volume loss calculated from the transformation to the average position of all scans using  $T_{sym}^{average}(I_i)$  from Eq. (9)). The table shows the mean (SD) of brain volume loss (ml) between the baseline and 12-month scans ( $L_{0-12}$ ), the 12-month and 24-month scans ( $L_{12-24}$ ) and the baseline and 24-month scans ( $L_{0-24}$ ). It also shows SD ratio between brain volume loss from the two methods. Numbers in square brackets denote 95% CI.

	Control (n = 138)	MCI (n = 211)	AD (n = 90)
<i>(a) Sinc interpolation</i>			
$L_{0-12}^{sym*}$	6.27 (6.72)	10.83 (8.62)	15.34 (8.03)
$L_{12-24}^{sym*}$	7.76 (6.33)	10.27 (8.90)	15.50 (9.84)
$L_{0-24}^{sym*}$	13.85 (8.08)	20.96 (12.76)	30.65 (13.78)
$L_{0-12}^{sym*} + L_{12-24}^{sym*} - L_{0-24}^{sym*}$	-0.18 (0.97)	0.14 (0.94)	0.19 (0.90)
	[0.01, 0.34], $p = 0.03$	[0.01, 0.26], $p = 0.04$	[0.00, 0.38], $p = 0.05$
$L_{0-12}^{sym}$	6.24 (6.67)	10.75 (8.54)	15.29 (8.01)
$L_{12-24}^{sym}$	7.68 (6.38)	10.27 (8.88)	15.37 (9.69)
$L_{0-24}^{sym}$	13.93 (8.06)	21.01 (12.74)	30.64 (13.84)
$L_{0-12}^{sym} + L_{12-24}^{sym} - L_{0-24}^{sym}$	-0.02 (0.13)	0.00 (0.14)	0.02 (0.18)
	[-0.04, 0.01], $p = 0.1$	[-0.02, 0.02], $p = 0.7$	[-0.02, 0.06], $p = 0.3$
<i>SD ratio</i>			
$L_{0-12}^{sym*}$ and $L_{0-12}^{sym}$	1.01 [0.99, 1.03], $p = 0.4$	1.01 [1.00, 1.02], $p = 0.05$	1.00 [0.99, 1.02], $p = 0.7$
$L_{12-24}^{sym*}$ and $L_{12-24}^{sym}$	1.00 [0.98, 1.01], $p = 0.4$	1.00 [0.99, 1.01], $p = 0.6$	1.02 [1.00, 1.03], $p = 0.02$
$L_{0-24}^{sym*}$ and $L_{0-24}^{sym}$	1.00 [0.99, 1.01], $p = 0.7$	1.00 [0.99, 1.01], $p = 0.7$	1.00 [0.99, 1.01], $p = 0.4$

**Table 4**

Direct test of bias using the back-to-back  $T_1$ -weighted scans in symmetric and asymmetric registration schemes. The table shows the mean (SD) of brain volume loss (ml), and the 95% CI (in square brackets) and  $p$ -value when compared to the mean value of zero.

	Control (n = 138)	MCI (n = 211)	AD (n = 90)
<i>(a) Linear interpolation</i>			
$L^{asym-forward}$	5.30 (4.01) [4.62, 5.97], $p < 0.001$	3.95 (3.51) [3.48, 4.43], $p < 0.001$	3.99 (3.50) [3.25, 4.72], $p < 0.001$
$L^{asym-backward}$	-4.89 (3.49) [-5.48, -4.30], $p < 0.001$	-4.16 (3.69) [-4.66, -3.66], $p < 0.001$	-3.63 (3.73) [-4.41, -2.85], $p < 0.001$
$L^{sym}$	0.02 (2.37) [-0.38, 0.42], $p = 0.9$	-0.12 (2.10) [-0.41, 0.16], $p = 0.4$	0.18 (2.09) [-0.26, 0.61], $p = 0.4$
<i>(b) Sinc interpolation</i>			
$L^{asym-forward}$	0.21 (2.45) [-0.20, 0.63], $p = 0.3$	-0.23 (2.06) [-0.51, 0.05], $p = 0.1$	0.09 (2.13) [-0.36, 0.53], $p = 0.7$
$L^{asym-backward}$	0.23 (2.45) [-0.18, 0.64], $p = 0.3$	-0.05 (2.12) [-0.34, 0.24], $p = 0.7$	0.19 (2.04) [-0.24, 0.61], $p = 0.4$
$L^{sym}$	0.11 (2.53) [-0.31, 0.54], $p = 0.6$	-0.15 (2.09) [-0.44, 0.13], $p = 0.3$	0.03 (2.09) [-0.41, 0.46], $p = 0.9$

registration scheme was much lower than the asymmetric registration scheme using either tri-linear or windowed sinc interpolation. This implied that brain volume loss using the symmetric registration scheme was more consistent than the asymmetric registration scheme. We feel that it is important to use the symmetric registration scheme if BSI is used to quantify the brain volume change of a single individual subject across multiple time-points.

In the particular situation that the subject has nearly the same position in the first and second scans, and a 'vastly' different position in the third scan, the transformation of all the scans into the average position may lead to unnecessarily strong interpolation in the first and second scans. We think that this would not produce suboptimal BSI result. This is because similar interpolation errors will be introduced in the resampled images of the first and second scans, and will mostly be cancelled out as BSI directly compares the two resampled images. For example, consider an extreme case that the first and second scans are two identical scans: BSI of the two scans will be zero, no matter how strong the interpolation is. Furthermore, compared with the asymmetric transformation of the third scan to the first/second scan, the symmetric transformation of MR scans into an average position may reduce the interpolation errors in the first/second scan and the third scan. To provide experimental evidence, we performed an additional post-hoc experiment that separately considered each pair of scans from the 0-, 12- and 24-month time-points, transforming both scans to the average position of each pair using  $T_{sym}(I_1, I_2)^{1/2}$  from Eq. (8) and calculating the BSI. We used Pitman's method to compare the SD of the BSI calculated from this pair-wise method

**Table 5**

Test of bias using the intercept of the regression line from three time points (baseline, 12-month and 24-month). The table shows the intercept [95% CI] of the regression line fitted to the brain volume loss (ml) over the scan interval.

	Control (n = 138)	MCI (n = 211)	AD (n = 90)
<i>(a) Linear interpolation</i>			
$L^{asym-forward}$	3.41 [1.40, 5.43], $p = 0.001$	3.49 [1.73, 5.26], $p < 0.001$	1.73 [-0.98, 4.43], $p = 0.2$
$L^{asym-backward}$	-5.73 [-7.63, -3.84], $p < 0.001$	-3.23 [-5.08, -1.37], $p = 0.001$	-4.17 [-7.24, -1.10], $p = 0.008$
$L^{sym}$	-1.13 [-2.98, 0.73], $p = 0.2$	-0.13 [-1.85, 1.60], $p = 0.9$	-1.69 [-4.53, 1.15], $p = 0.2$
<i>(b) Sinc interpolation</i>			
$L^{asym-forward}$	-1.61 [-3.47, 0.24], $p = 0.09$	-0.04 [-1.80, 1.72], $p = 1.0$	-1.41 [-4.23, 1.42], $p = 0.3$
$L^{asym-backward}$	-0.79 [-2.68, 1.10], $p = 0.4$	0.46 [-1.31, 2.23], $p = 0.6$	-0.91 [-3.77, 1.94], $p = 0.5$
$L^{sym}$	-1.42 [-3.25, 0.40], $p = 0.1$	-0.01 [-1.74, 1.73], $p = 1.0$	-1.45 [-4.29, 1.39], $p = 0.3$

and our multi-time-point method. The SD ratio shown in Table 3 suggests that the variability of BSI calculated using our method is very similar to the pair-wise method, showing no evidence that our method is suboptimal. On the other hand, a small but statistically significant transitivity error is found with the pair-wise method. The higher SD in  $L_{0-12}^{sym} + L_{12-24}^{sym} - L_{0-24}^{sym}$  in Table 3 also indicates the higher inconsistency in transitivity when compared to our method.

*Direct test of bias using back-to-back scans from the same baseline time point*

The direct test indicated that a bias of between 3.6 and 5.3 ml was introduced when using tri-linear interpolation with the asymmetric registration scheme ( $p < 0.001$ , all tests) (see Table 4). Results also show that the bias depends on the direction of the transformation when using tri-linear interpolation with the asymmetric registration scheme: given two back-to-back scans, a positive bias (apparent brain volume loss) is detected when transforming the repeat scans, and a negative bias (apparent brain volume gain) is detected when transforming the baseline scans. Recall that either the first or the second of the back-to-back scans was randomly assigned to be the 'baseline', with the other assigned to be the 'repeat'. Furthermore, the difference (between  $L^{asym-forward}$  and  $L^{asym-backward}$  was around the value (4.3 ml) estimated in the post-hoc analysis in the Section "Test of inverse consistency using two time points" for scans with no atrophy. No bias was found in other cases. Both the symmetric registration scheme and windowed sinc interpolation were able to drastically reduce the bias.

A bias of 4 ml is typically less than 0.4% of whole brain volume. However, it is important when trying to determine difference in rate of brain volume loss between treatment and placebo groups in a clinical trial, as the total change may only be around 1–2% over the study.

The SD of brain volume loss indicated the scan and re-scan reproducibility of BSI. When using windowed sinc interpolation with the symmetric registration scheme, we obtained a SD of 2.2 ml or 0.22% of whole brain volume for all the subjects ( $n = 439$ ). A previous study using a different cohort of subjects showed that the SD of percentage brain volume loss from classic-BSI using back-to-back scans was 0.47% ( $n = 60$ , 38 AD and 22 controls) (Boyes et al., 2006). The improvement in reproducibility may be due to the improvement in KN-BSI (Leung et al., 2010) over classic-BSI or greater consistency of acquisition of the MR scans from ADNI.

*Test of bias using the intercept of the regression line from three time points*

Tri-linear interpolation with the asymmetric registration scheme introduced a bias of 3.41 – 5.73 ml in the intercept of the regression line in controls (forward and backward registration,  $p \leq 0.001$ , both tests), MCI subjects (forward and backward registration,  $p \leq 0.001$ , both tests) and AD subjects (only backward registration,  $p = 0.008$ )

as shown in Table 5 and Fig. 4. No bias was found in other cases. Again, either the symmetric registration scheme or windowed sinc interpolation was able to drastically reduce the bias.

#### Related work

There has recently been substantial interest and concern regarding potential biases in longitudinal image analysis (Thomas et al., 2009; Thompson et al., 2011; Yushkevich et al., 2010; Reuter and Fischl, 2011; Fox et al., 2011; Holland et al., 2011). As mentioned in the introduction, we followed Yushkevich et al. (2010) in using the back-to-back scans and the intercepts of the regression line fitted to

the brain volume loss estimated from two time points, to assess bias in the BSI processing pipeline. In these two tests of bias, we reached the same conclusion that a symmetric global registration scheme was able to remove bias present in an asymmetric scheme.

Yushkevich et al. (2010) suggested that asymmetric interpolation/smoothing from an asymmetric global/affine transformation may cause bias in non-linear registration based atrophy quantification, because the result of a non-linear registration is strongly influenced by the initial gradient of the image similarity metric. For all nonlinear registration algorithms based on voxel-similarity (the majority, though HAMMER is a notable exception; Shen and Davatzikos, 2002), the computation of the metric and its gradient requires paired voxels and hence interpolated images, calculated from the starting point of the initial affine transformation. Algorithms that do not require a separate a priori transformation and reslicing of images have the advantage that this internal voxel-similarity optimisation requires only one interpolation step, rather than accumulating errors from both the prior and the internal interpolation. However, as shown by Yushkevich et al. (2010), who used ANTS (Avants et al., 2008) and IRTK (Rueckert et al., 1999) without prior transformation/interpolation, even the single interpolation step is able to introduce bias via the nonlinear optimisation process. Our results showed that windowed sinc interpolation dramatically reduced the bias in the BSI when using asymmetric registration. This suggests that for asymmetric non-linear registration based measurement, it might be preferable to resample the images after the affine registration with a high-order interpolation method and use them in the non-linear registration, rather than to avoid the double interpolation at the cost of being more detrimentally affected by lower-order interpolation in the non-linear optimisation process. Note that such interpolation prior to estimating the non-linear registration does not mean that images warped using the non-linear transformation have to suffer two interpolation steps, since the transformations can still easily be composed and applied to the original image, even if this image was not used to estimate the non-linear transformation.

Yushkevich et al. (2010) found that the effect of longitudinal bias was small as long as sample sizes were calculated relative to the atrophy rate in controls. However, a control group may not be available in some studies. It is important to note that this breaks down if the amount of longitudinal bias differs between controls and patients, as it did for the asymmetric registration with linear interpolation (Fig. 2).

Our formulation of symmetric registration and DBC of multiple scans is based on existing DBC (Lewis and Fox, 2004), symmetric registration (Reuter et al., 2010) and groupwise schemes (Alexa, 2002; Aljabar et al., 2006). In this paper, we suggest a consistent formulation for more than two scans, and highlight the important connection between the ideas of the (scalar and matrix) square root and the geometric (scalar and matrix log-Euclidean) average.

#### Strength of the study

In addition to using back-to-back scans and the intercept-based experiments, we assessed inverse- and transitivity-consistency of the BSI processing pipeline as suggested by Reuter and Fischl (2011). As mentioned in the last section, we also investigated the effect of tri-linear and windowed sinc interpolation on bias in the pipeline. Furthermore, the study utilised serial MR scans of 439 subjects acquired from multiple sites. The large number of subjects provided enough statistical power to detect subtle bias caused by an asymmetric registration scheme.

#### Limitations

Our study shares many of the limitations described by Yushkevich et al. (2010). Since the true atrophy is zero in the back-to-back scans,

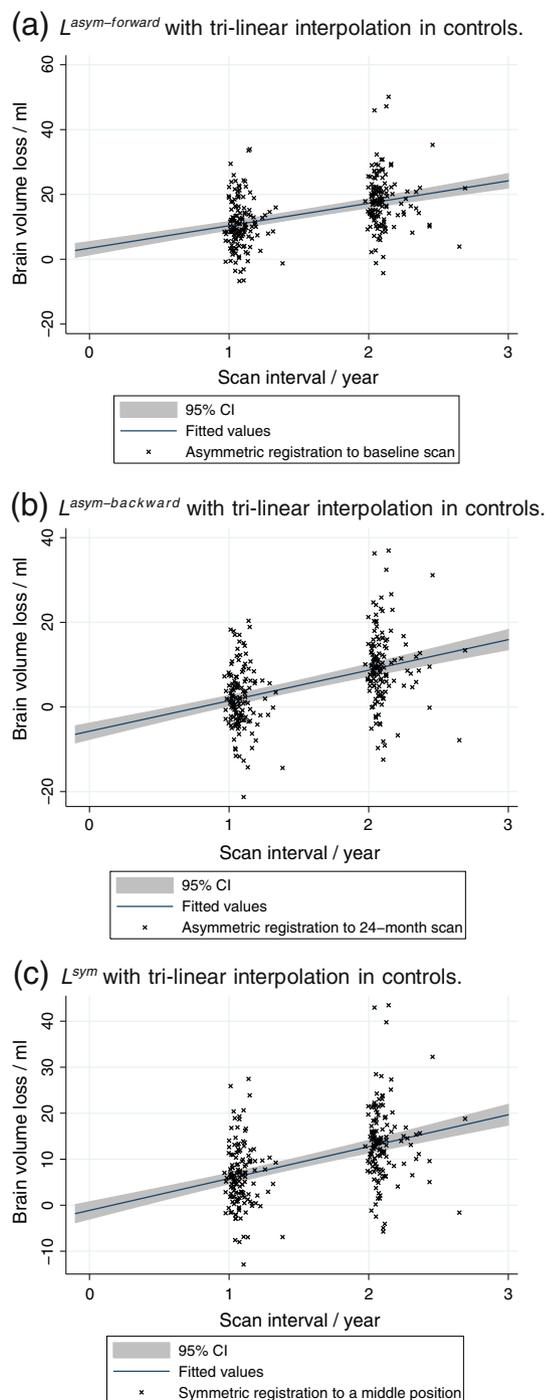


Fig. 4. Test of bias using the intercept of linear regression fitted to the brain volume loss against scan interval in controls.

multiplicative bias (which refers to bias that is proportional to the amount of true atrophy, in contrast to simple 'additive bias', which is a constant addition independent of the true value) is not present in the back-to-back scans and cannot be detected. In the intercept-based experiment, multiplicative bias has the same effect on the brain volume loss between the baseline and 12-month scans, and between the baseline and 24-month scans, so that it does not affect the intercept. More importantly, the intercept-based experiment assumes that the brain volume loss is linear over a two-year period. However, some studies have shown the acceleration of brain atrophy in normal ageing, mild cognitive impairment and Alzheimer's disease (Scahill et al., 2003; Chan et al., 2003; Jack et al., 2008c). When acceleration in rate of atrophy is occurring, one would expect the mean atrophy rate to be higher over the second time period and therefore the intercept would be slightly negative. Indeed, we found this numerically to be the case (Table 5), although the intercept was not significantly different from zero. It is also far from clear whether hypothetically uniform (non-accelerating) atrophy would be defined as linear in either absolute or percentage terms – since it cannot be linear in both.

One criticism of constructing a middle transformation from the pairwise transformations is that a poorly registered pair of images can adversely affect the constructed middle transformation (Christensen and Johnson, 2001). Visual checks should be performed to make sure that all the image pairs are well aligned.

## Conclusions

We described an unbiased brain atrophy estimation method, that can consistently model changes over more than two time-points, by combining symmetric registration and differential bias correction techniques through the concept of a geometric mean. Using MR scans of 439 subjects downloaded from the ADNI database, bias in the BSI processing pipeline was assessed using four tests: (a) inverse consistency, (b) transitive consistency, (c) back-to-back scans, and (d) intercept of the regression line fitted to the brain volume loss over time. Results suggested that the use of windowed sinc interpolation or a symmetric registration scheme consistently reduced the bias in the BSI processing pipeline. No bias was detected in any of the four tests when using windowed sinc interpolation with the proposed symmetric procedure.

## Acknowledgments

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 5-year public-private partnership. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians in developing new treatments and monitoring their effectiveness, as well as lessening the time and cost of clinical trials. The Principal Investigator is Michael W. Weiner, M.D., VA Medical Center and University of California – San Francisco. ADNI is the result of efforts of many co-investigators and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research – approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years. For up-to-date information, see [www.adni-info.org](http://www.adni-info.org).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the

following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://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 California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

This work was undertaken at UCLH/UCL with funding from the Department of Health's National Institute of Health Research Centres funding scheme, the Medical Research Council and Alzheimer's Research UK. The Dementia Research Centre is an Alzheimer's Research UK Co-ordinating Centre and has also received equipment funded by the Alzheimer's Research UK.

The authors would like to thank all the image analysts and the research associates in the Dementia Research Centre for their help in the study. In particular, the authors would like to thank Dr. Jonathan Bartlett for statistical analysis. The implementations of registration algorithms used the Insight Segmentation and Registration Toolkit (ITK), an open source software developed as an initiative of the U.S. National Library of Medicine and available at [www.itk.org](http://www.itk.org). The authors would particularly like to thank the ADNI study subjects and investigators for their participation.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at [doi:10.1016/j.neuroimage.2011.10.068](https://doi.org/10.1016/j.neuroimage.2011.10.068).

## References

- Alexa, M., 2002. Linear combination of transformations. *ACM Transactions on Graphics* 21, 380–387 (July).
- Aljabar, P., Bhatia, K., Hajnal, J., Boardman, J., Srinivasan, L., Rutherford, M., Dyet, L., Edwards, A., Rueckert, D., 2006. Analysis of growth in the developing brain using non-rigid registration. *Biomedical Imaging: Nano to Macro*, 2006. 3rd IEEE International Symposium on, pp. 201–204.
- Arsigny, V., Commowick, O., Pennec, X., Ayache, N., 2006. A log-Euclidean polyaffine framework for locally rigid or affine registration. *Biomedical Image Registration*. In: Plum, J., Likar, B., Gerritsen, F. (Eds.), *Lecture Notes in Computer Science*, vol. 4057. Springer, Berlin/Heidelberg, pp. 120–127.
- 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 (1), 26–41 (Feb).
- Avants, B., Cook, P.A., McMillan, C., Grossman, M., Gee, J.C., 2010. Sparse unbiased analysis of anatomical variance in longitudinal imaging. *Med. Image Comput. Comput. Assist. Interv.* 13 (Pt 1), 324–331.
- Beg, M.F., Khan, A., 2007. Symmetric data attachment terms for large deformation image registration. *IEEE Trans. Med. Imaging* 26 (9), 1179–1189 (Sep).
- Boyes, R.G., Rueckert, D., Aljabar, P., Whitwell, J., Schott, J.M., Hill, D.L.G., Fox, N.C., 2006. Cerebral atrophy measurements using Jacobian integration: comparison with the boundary shift integral. *Neuroimage* 32 (1), 159–169 (Aug).
- Boyes, R.G., Gunter, J.L., Frost, C., Janke, A.L., Yeatman, T., Hill, D.L.G., 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., ADNI Study, 2008. Intensity non-uniformity correction using N3 on 3-T scanners with multichannel phased array coils. *Neuroimage* 39 (4), 1752–1762 (Feb).
- Brun, A., Englund, E., 2002. Regional pattern of degeneration in Alzheimer's disease: neuronal loss and histopathological grading. *Histopathology* 41 (3A), 40–55 (Sep).
- Chan, D., Janssen, J.C., Whitwell, J.L., Watt, H.C., Jenkins, R., Frost, C., Rossor, M.N., Fox, N.C., 2003. Change in rates of cerebral atrophy over time in early-onset Alzheimer's disease: longitudinal MRI study. *Lancet* 362 (9390), 1121–1122 (Oct).
- Christensen, G.E., Johnson, H.J., 2001. Consistent image registration. *IEEE Trans. Med. Imaging* 20 (7), 568–582 (Jul).
- Clarkson, M.J., Ourselin, S., Nielsen, C., Leung, K.K., Barnes, J., Whitwell, J.L., Gunter, J.L., Hill, D.L.G., Weiner, M.W., Jack, C.R., Fox, N.C., Alzheimer's Disease Neuroimaging

- Initiative, 2009. Comparison of phantom and registration scaling corrections using the adni cohort. *Neuroimage* 47 (4), 1506–1513 (Oct).
- Desikan, R.S., Fischl, B., Cabral, H.J., Kemper, T.L., Guttman, C.R.G., Blacker, D., Hyman, B.T., Albert, M.S., Killiany, R.J., 2008. MRI measures of temporoparietal regions show differential rates of atrophy during prodromal AD. *Neurology* 71 (11), 819–825 (Sep).
- Fox, N.C., Schott, J.M., 2004. Imaging cerebral atrophy: normal ageing to Alzheimer's disease. *Lancet* 363 (9406), 392–394 (Jan).
- Fox, N.C., Scahill, R.L., Crum, W.R., Rossor, M.N., 1999. Correlation between rates of brain atrophy and cognitive decline in AD. *Neurology* 52 (8), 1687–1689 (May).
- Fox, N.C., Ridgway, G.R., Schott, J.M., 2011. Algorithms, atrophy and Alzheimer's disease: cautionary tales for clinical trials. *Neuroimage* 57 (1), 15–18 (Jul).
- Freeborough, P.A., Fox, N.C., Kitney, R.I., 1997. Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans. *Comput. Methods Programs Biomed.* 53 (1), 15–25 (May).
- Geng, X., Kumar, D., Christensen, G.E., 2005. Transitive inverse-consistent manifold registration. *Inf. Process. Med. Imaging* 19, 468–479.
- Gunter, J.L., Bernstein, M.A., Borowski, B.J., Felmlee, J.P., Blezek, D.J., Mallozzi, R.P., Levy, J.R., Schuff, N., Jack, C.R., 2006. Validation testing of the MRI Calibration Phantom for the Alzheimer's Disease Neuroimaging Initiative Study. *ISMRM*, p. 2652.
- Hampel, H., Frank, R., Broich, K., Teipel, S.J., Katz, R.G., Hardy, J., Herholz, K., Bokde, A.L.W., Jessen, F., Hoessler, Y.C., Sanhai, W.R., Zetterberg, H., Woodcock, J., Blennow, K., 2010. Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. *Nat. Rev. Drug Discov.* 9 (7), 560–574 (Jul).
- Holland, D., McEvoy, L.K., Dale, A.M., the Alzheimer's Disease Neuroimaging Initiative, 2011. Unbiased comparison of sample size estimates from longitudinal structural measures in ADNI. *Hum. Brain Mapp.* n/a–n/a <http://dx.doi.org/10.1002/hbm.21386>.
- Hua, X., Lee, S., Yanovsky, I., Leow, A.D., Chou, Y.-Y., Ho, A.J., Gutman, B., Toga, A.W., Jack, C.R., Bernstein, M.A., Reiman, E.M., Harvey, D.J., Kornak, J., Schuff, N., Alexander, G.E., Weiner, M.W., Thompson, P.M., the Alzheimer's Disease Neuroimaging Initiative, 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 (4), 668–681 (Dec).
- Hua, X., Gutman, B., Boyle, C.P., Rajagopalan, P., Leow, A.D., Yanovsky, I., Kumar, A.R., Toga, A.W., Jack, C.R., Schuff, N., Alexander, G.E., Chen, K., Reiman, E.M., Weiner, M.W., Thompson, P.M., the Alzheimer's Disease Neuroimaging Initiative, 2011. Accurate measurement of brain changes in longitudinal MRI scans using tensor-based morphometry. *Neuroimage* 57 (1), 5–14 (Jul).
- Jack, C.R., Shiung, M.M., Gunter, J.L., O'Brien, P.C., Weigand, S.D., Knopman, D.S., Boeve, B.F., Ivnik, R.J., Smith, G.E., Cha, R.H., Tangalos, E.G., Petersen, R.C., 2004. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. *Neurology* 62 (4), 591–600 (Feb).
- Jack, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Whitwell, J.L., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L.G., 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., 2008a. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27 (4), 685–691 (Apr).
- Jack, C.R., Weigand, S.D., Shiung, M.M., Przybelski, S.A., O'Brien, P.C., Gunter, J.L., Knopman, D.S., Boeve, B.F., Smith, G.E., Petersen, R.C., 2008b. Atrophy rates accelerate in amnesic mild cognitive impairment. *Neurology* 70 (19 Pt 2), 1740–1752 (May).
- Jack, C.R., Petersen, R.C., Grundman, M., Jin, S., Gamst, A., Ward, C.P., Sencakova, D., Doody, R.S., Thal, L.J., the Alzheimer's Disease Cooperative Study (ADCS), M., 2008c. Longitudinal MRI findings from the vitamin E and donepezil treatment study for MCI. *Neurobiol. Aging* 29 (9), 1285–1295 (Sep).
- Jenkinson, M., 1999. Measuring transformation error by RMS deviation. Tech. Rep. TR99MJ1, Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB). Department of Clinical Neurology, University of Oxford. <http://www.fmriv.ox.ac.uk/analysis/techrep/>.
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17 (2), 825–841 (Oct).
- Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., Macfall, J., Fischl, B., Dale, A., 2006. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *Neuroimage* 30 (2), 436–443 (Apr).
- Leow, A., Huang, S.-C., Geng, A., Becker, J., Davis, S., Toga, A., Thompson, P., 2005. Inverse consistent mapping in 3D deformable image registration: its construction and statistical properties. *Inf. Process. Med. Imaging* 19, 493–503.
- Leung, K.K., Clarkson, M.J., Bartlett, J.W., Clegg, S., Jack, C.R., Weiner, M.W., Fox, N.C., Ourselin, S., Alzheimer's Disease Neuroimaging Initiative, 2010. Robust atrophy rate measurement in Alzheimer's disease using multi-site serial MRI: tissue-specific intensity normalization and parameter selection. *Neuroimage* 50 (2), 516–523 (Apr).
- Leung, K.K., Barnes, J., Modat, M., Ridgway, G.R., Bartlett, J.W., Fox, N.C., Ourselin, S., Alzheimer's Disease Neuroimaging Initiative, 2011. Brain MAPS: an automated, accurate and robust brain extraction technique using a template library. *Neuroimage* 55 (3), 1091–1108 (Apr).
- Lewis, E.B., Fox, N.C., 2004. Correction of differential intensity inhomogeneity in longitudinal MR images. *Neuroimage* 23 (1), 75–83 (Sep).
- Meijering, E., Niessen, W., Pluim, J., Viergever, M., 1999. Quantitative comparison of sinc-approximating kernels for medical image interpolation. *Medical Image Computing and Computer-Assisted Intervention MICCAI 1999*. In: Taylor, C., Colchester, A. (Eds.), *Lecture Notes in Computer Science*, vol. 1679. Springer, Berlin/Heidelberg, pp. 210–217.
- Narayana, P., Brey, W., Kulkarni, M., Sievenpiper, C., 1988. Compensation for surface coil sensitivity variation in magnetic resonance imaging. *Magn. Reson. Imaging* 6 (3), 271–274.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd Edition. Cambridge University Press, New York, NY, USA.
- Reuter, M., Fischl, B., 2011. Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage* 57 (1), 19–21 (Jul).
- Reuter, M., Rosas, H.D., Fischl, B., 2010. Highly accurate inverse consistent registration: a robust approach. *Neuroimage* 53 (4), 1181–1196 (Dec).
- Rogelj, P., Kovacic, S., 2006. Symmetric image registration. *Med. Image Anal.* 10 (3), 484–493 (Jun).
- Rohlfing, T., Maurer, C.R., 2007. Shape-based averaging. *IEEE Trans. Image Process.* 16 (1), 153–161 (Jan).
- Rueckert, D., Sonoda, L.L., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imaging* 18 (8), 712–721 (Aug).
- Scahill, R.L., Frost, C., Jenkins, R., Whitwell, J.L., Rossor, M.N., Fox, N.C., 2003. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Arch. Neurol.* 60 (7), 989–994 (Jul).
- Schott, J.M., Bartlett, J.W., Barnes, J., Leung, K.K., Ourselin, S., Fox, N.C., Alzheimer's Disease Neuroimaging Initiative Investigators, 2010. Reduced sample sizes for atrophy outcomes in Alzheimer's disease trials: baseline adjustment. *Neurobiol. Aging* 31 (8), 1452–1462 (Aug 1462.e1–2).
- Shen, D., Davatzikos, C., 2002. HAMMER: hierarchical attribute matching mechanism for elastic registration. *IEEE Trans. Med. Imaging* 21 (11), 1421–1439 (Nov) <http://dx.doi.org/10.1109/TMI.2002.803111>.
- Skrinjar, O., Bistoquet, A., Tagare, H., 2008. Symmetric and transitive registration of image sequences. *Int. J. Biomed. Imaging* 2008, 686875.
- Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* 17 (1), 87–97 (Feb).
- Smith, S.M., Stefano, N.D., Jenkinson, M., Matthews, P.M., 2001. Normalized accurate measurement of longitudinal brain change. *J. Comput. Assist. Tomogr.* 25 (3), 466–475.
- Thomas, A.G., Marrett, S., Saad, Z.S., Ruff, D.A., Martin, A., Bandettini, P.A., 2009. Functional but not structural changes associated with learning: an exploration of longitudinal voxel-based morphometry (VBM). *Neuroimage* 48 (1), 117–125 (Oct).
- Thompson, W.K., Holland, D., the Alzheimer's Disease Neuroimaging Initiative, 2011. Bias in tensor based morphometry Stat-ROI measures may result in unrealistic power estimates. *Neuroimage* 57 (1), 1–4 (Jul).
- Woods, R.P., 2003. Characterizing volume and surface deformations in an atlas framework: theory, applications, and implementation. *Neuroimage* 18 (3), 769–788 (Mar).
- Woods, R.P., Grafton, S.T., Watson, J.D., Sicotte, N.L., Mazziotta, J.C., 1998. Automated image registration: II. Intersubject validation of linear and nonlinear models. *J. Comput. Assist. Tomogr.* 22 (1), 153–165.
- Yang, D., Li, H., Low, D.A., Deasy, J.O., Naqa, I.E., 2008. A fast inverse consistent deformable image registration method based on symmetric optical flow computation. *Phys. Med. Biol.* 53 (21), 6143–6165 (Nov).
- Ye, X., Chen, Y., 2009. A New Algorithm for Inverse Consistent Image Registration. *Proceedings of the 5th International Symposium on Advances in Visual Computing: Part I. ISVC '09*. Springer-Verlag, Berlin, Heidelberg, pp. 855–864.
- Yushkevich, P.A., Avants, B.B., Das, S.R., Pluta, J., Altinay, M., Craige, C., Alzheimer's Disease Neuroimaging Initiative, 2010. Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: an illustration in ADNI 3 T MRI data. *Neuroimage* 50 (2), 434–445 (Apr).