

VALIDATING UNBIASED REGISTRATION ON LONGITUDINAL MRI SCANS FROM THE ALZHEIMER'S DISEASE NEUROIMAGING INITIATIVE (ADNI)

Igor Yanovsky¹, Paul M. Thompson², Stanley J. Osher¹, Xue Hua²,
David W. Shattuck², Arthur W. Toga², Alex D. Leow²
and the Alzheimer's Disease Neuroimaging Initiative

¹Department of Mathematics, University of California, Los Angeles, CA, USA
²Laboratory of Neuro Imaging, UCLA School of Medicine, Los Angeles, CA, USA

ABSTRACT

This paper examines the power of different nonrigid registration models to detect changes in tensor based morphometry (TBM), and their stability when no real changes are present. Specifically, we investigate an asymmetric version of a recently proposed unbiased registration method, using mutual information as the matching criterion. We compare matching functionals (sum of squared differences and mutual information), as well as large deformation registration schemes (viscous fluid registration versus symmetric and asymmetric unbiased registration) for detecting changes in serial MRI scans of 10 elderly normal subjects and 10 patients with Alzheimer's Disease scanned at 2-week and 1-year intervals. We demonstrated that the unbiased methods, both symmetric and asymmetric, have higher reproducibility. The unbiased methods were also less likely to detect changes in the absence of any real physiological change. Moreover, they measured biological deformations more accurately by penalizing bias in the corresponding statistical maps.

Index Terms— Mutual information, image registration.

1. INTRODUCTION

In recent years, computational neuroimaging has become an exciting interdisciplinary field with many applications in functional and anatomic brain mapping, image-guided surgery, and multimodality image fusion. In this paper, we introduce a novel Asymmetric Unbiased model (by contrast with the Symmetric Unbiased model) and, for the first time, we analyze unbiased models with mutual information based matching functionals (prior work has focused on the case where the summed squared intensity difference is used as the criterion for registration). Most importantly, we aim to provide quality calibrations for different non-rigid registration techniques in TBM. In particular, we compare two common matching functionals: L^2 , or the sum of squared intensity differences, versus mutual information, and three regularization techniques (fluid registration versus the Asymmetric Unbiased and Symmetric Unbiased techniques). Our experiments are designed to decide which registration method is more reproducible, more reliable, and offers less artifactual variability in regions of homogeneous image intensity. The foundation of our calibrations is based on the assumption that, by scanning healthy normal human subjects twice over a 2-week period using the same protocol, serial MRI scan pairs should not show any systematic biological change. Therefore, any regional structural differences detected using TBM over such a short interval may be assumed to be errors. We apply statistical analysis to the profile of these errors,

providing information on the reliability, reproducibility and variability of different registration techniques. Moreover, serial images of 10 subjects from the ADNI follow-up phase (images acquired one year apart) were analyzed in a similar fashion and compared to the ADNI baseline data. In images collected one year apart, real anatomical changes are present; neurobiological changes due to aging and dementia include widespread cell shrinkage, regional gray and white matter atrophy and expansion of fluid-filled spaces in the brain. Thus, a good computational technique should be able to differentiate between longitudinal image pairs collected for the ADNI baseline (2-week) and follow-up (1-year) phases.

2. UNBIASED IMAGE REGISTRATION

Let Ω be an open and bounded domain in \mathbb{R}^n , for arbitrary n . Without loss of generality, assume that the volume of Ω is 1, i.e. $|\Omega| = 1$. Let $I_1, I_2 : \Omega \rightarrow \mathbb{R}$ be the two images to be registered. We seek to find the transformation $\mathbf{g} : \Omega \rightarrow \Omega$ that maps the source image I_2 into correspondence with the target image I_1 . In this paper, we will restrict this mapping to be differentiable, one-to-one, and onto. We denote the Jacobian matrix of a deformation \mathbf{g} to be $D\mathbf{g}$, with Jacobian denoted by $|D\mathbf{g}(\mathbf{x})|$. The displacement field $\mathbf{u}(\mathbf{x})$ from the position \mathbf{x} in the deformed image $I_2 \circ \mathbf{g}(\mathbf{x})$ back to $I_2(\mathbf{x})$ is defined in terms of the deformation $\mathbf{g}(\mathbf{x})$ by the expression $\mathbf{g}(\mathbf{x}) = \mathbf{x} - \mathbf{u}(\mathbf{x})$.

We now describe the construction of the Unbiased Large-Deformation Image Registration. We associate three probability density functions (PDFs) to \mathbf{g} , \mathbf{g}^{-1} , and the identity mapping id : $P_{\mathbf{g}}(\mathbf{x}) = |D\mathbf{g}(\mathbf{x})|$, $P_{\mathbf{g}^{-1}}(\mathbf{x}) = |D\mathbf{g}^{-1}(\mathbf{x})|$, $P_{\text{id}}(\mathbf{x}) = 1$. By associating deformations with their corresponding global density maps, we can now apply information theory to quantify the magnitude of deformations. In our approach, we choose the Kullback-Leibler (KL) divergence and symmetric Kullback-Leibler (SKL) distance. The KL divergence between two probability density functions, X and Y , is defined as $KL(X, Y) = \int_{\Omega} X \log X/Y dx \geq 0$. We define the SKL distance as $SKL(X, Y) = KL(X, Y) + KL(Y, X)$. The Unbiased method solves for the deformation \mathbf{g} (or, equivalently, for the displacement \mathbf{u}) minimizing the energy functional E , consisting of the image matching term F and the regularizing term R which is based on KL divergence or SKL distance. The general minimization problem can be written as: $\inf_{\mathbf{u}} \{E(\mathbf{u}) = F(\mathbf{u}) + \lambda R(\mathbf{u})\}$. Here, $\lambda > 0$ is a weighting parameter.

To quantify the magnitude of deformation \mathbf{g} , in this paper we introduce a new regularization term R_{KL} , which is an asymmetric measure between P_{id} and $P_{\mathbf{g}}$: $R_{KL}(\mathbf{g}) = KL(P_{\text{id}}, P_{\mathbf{g}})$. This regularization term can be shown to be

$$R_{KL}(\mathbf{g}) = \int_{\Omega} P_{\text{id}} \log \frac{P_{\text{id}}}{P_{\mathbf{g}}} dx = \int_{\Omega} -\log |D\mathbf{g}(\mathbf{x})| dx. \quad (1)$$

Funded by P41 RR13642, R21 EB001561, U54 RR021813

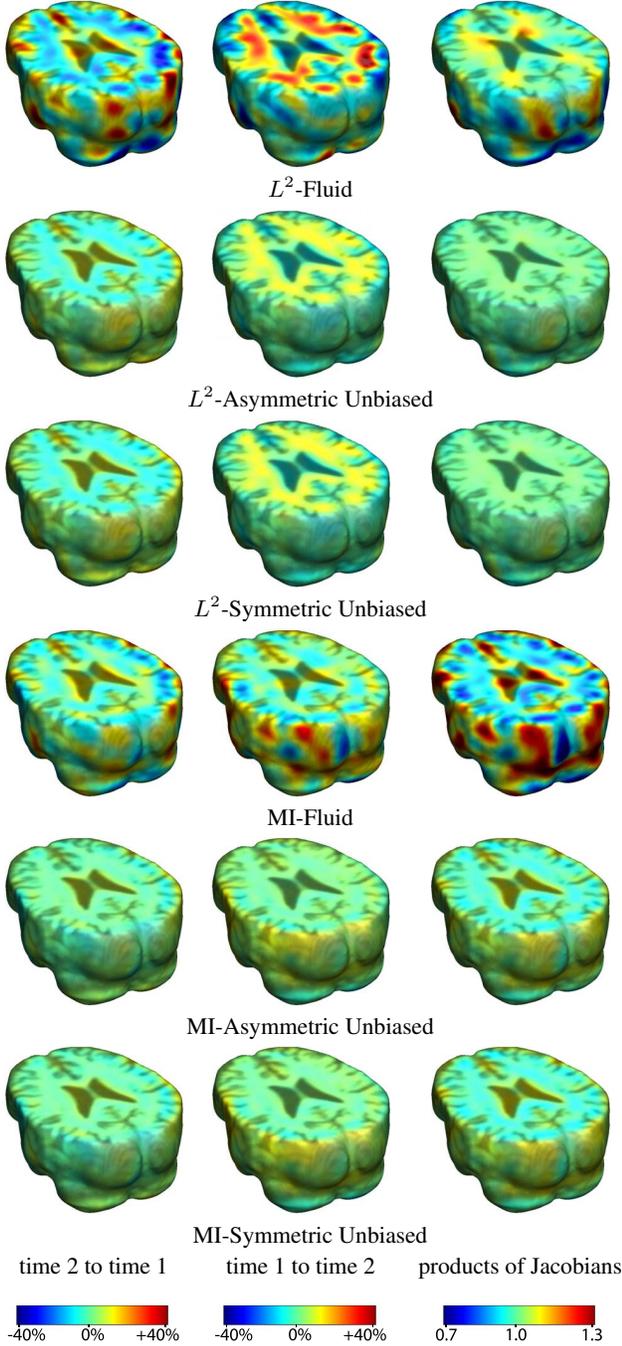


Fig. 1. Nonrigid registration was performed on an image pair of one of the subjects from the ADNI baseline study (serial MRI images acquired two weeks apart). Jacobian maps of deformations from time 2 to time 1 (column 1) and time 1 to time 2 (column 2) are superimposed on the target volumes. The unbiased methods generate less noisy Jacobian maps with values closer to 1; this shows the greater stability of the approach when no volumetric change is present. Column 3 examines the inverse consistency of deformation models. Products of Jacobian maps are shown for the forward and backward directions. For the unbiased methods, the products of the Jacobian maps are less noisy, with values closer to 1, showing better inverse consistency.

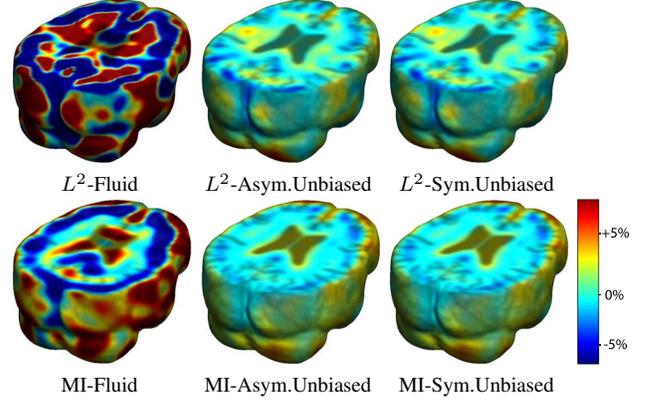


Fig. 2. Nonrigid registration was performed on the ADNI baseline study (serial MRI images acquired two weeks apart) of ten normal elderly subjects. For each method, the mean of the resulting 10 Jacobian maps is superimposed on one of the brain volumes. Visually, Fluid registration generates noisy mean maps, while maps generated using the Asymmetric Unbiased and Symmetric Unbiased methods are less noisy with values closer to 1.

The regularization functional based on the symmetric KL distance between $P_{\mathbf{id}}$ and $P_{\mathbf{g}}$ is $R_{SKL}(\mathbf{g}) = SKL(P_{\mathbf{id}}, P_{\mathbf{g}})$. As shown in [1, 2], the regularization term is linked to statistics on Jacobian maps as follows

$$R_{SKL}(\mathbf{g}) = \int_{\Omega} (|D\mathbf{g}(\mathbf{x})| - 1) \log |D\mathbf{g}(\mathbf{x})| d\mathbf{x}. \quad (2)$$

Notice that the symmetric unbiased regularizing functional is point-wise nonnegative, while the asymmetric unbiased regularizer in (1) can take either positive or negative values locally.

3. FIDELITY METRICS

In this paper, the matching functional F takes two forms: the L^2 -norm (the sum of squared differences) and MI (mutual information). The L^2 distance between the deformed image $I_2 \circ \mathbf{g}(\mathbf{x}) = I_2(\mathbf{x} - \mathbf{u})$ and image $I_1(\mathbf{x})$ is $F_{L^2}(\mathbf{u}) = \frac{1}{2} \int_{\Omega} (I_2(\mathbf{x} - \mathbf{u}) - I_1(\mathbf{x}))^2 d\mathbf{x}$.

To define the mutual information between the deformed image $I_2(\mathbf{x} - \mathbf{u})$ and the target image $I_1(\mathbf{x})$, we denote by $p_{\mathbf{u}}^{I_1}$ and $p_{\mathbf{u}}^{I_2}$ the intensity distributions estimated from $I_1(\mathbf{x})$ and $I_2(\mathbf{x} - \mathbf{u})$, respectively. An estimate of their joint intensity distribution is denoted as $p_{\mathbf{u}}^{I_1, I_2}$. We also let $i_1 = I_1(\mathbf{x})$, $i_2 = I_2(\mathbf{x} - \mathbf{u}(\mathbf{x}))$ denote intensity values at point $\mathbf{x} \in \Omega$. Given the displacement field \mathbf{u} , the mutual information computed from I_1 and I_2 is provided by

$$MI_{\mathbf{u}}^{I_1, I_2} = \int_{\mathbb{R}^2} p_{\mathbf{u}}^{I_1, I_2}(i_1, i_2) \log \frac{p_{\mathbf{u}}^{I_1, I_2}(i_1, i_2)}{p^{I_1}(i_1)p^{I_2}(i_2)} di_1 di_2. \quad (3)$$

We seek to maximize $MI_{\mathbf{u}}^{I_1, I_2}$, or minimize $F_{MI}(\mathbf{u}) = -MI_{\mathbf{u}}^{I_1, I_2}$.

4. MINIMIZATION OF ENERGY FUNCTIONALS

In general, we expect minimizers of the energy functional $E(\mathbf{u})$ to exist. Computing the first variation of this functional, we obtain the gradient of $E(\mathbf{u})$, namely $\partial_{\mathbf{u}}E(\mathbf{u})$. We define the force field \mathbf{f} , which drives I_2 into registration with I_1 , as

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \partial_{\mathbf{u}}E(\mathbf{u}) = \partial_{\mathbf{u}}F(\mathbf{u}) + \lambda \partial_{\mathbf{u}}R(\mathbf{u}). \quad (4)$$

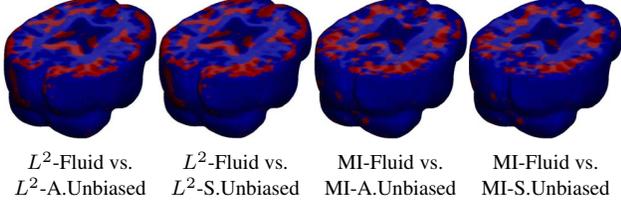


Fig. 3. Voxel-wise paired t test for the deviation gain S empirically thresholded at 2.82 ($p = 0.005$ on the voxel level with 9 degrees of freedom), showing where Asymmetric Unbiased and Symmetric Unbiased registration outperform Fluid registration (regions in red) with statistical significance on a voxel level.

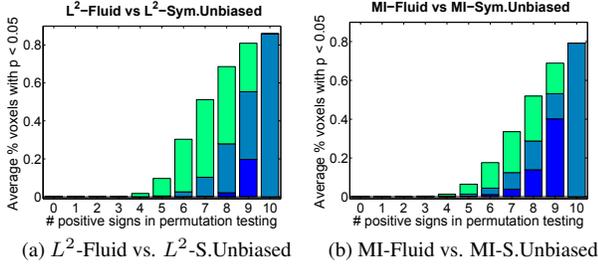


Fig. 4. Multiple Comparison Analysis using permutation testing on the deviation gain S of (a) L^2 -Fluid over L^2 -Sym.Unbiased and (b) MI-Fluid over MI-Sym.Unbiased, both for ADNI baseline dataset. Different colors signify minimum, average, and maximum percentage of voxels with $p < 0.05$. The result indicates that unbiased regularization technique outperforms Fluid methods with $p < 0.001$.

Here, $R(\mathbf{u})$ is either $R_{KL}(\mathbf{u})$ or $R_{SKL}(\mathbf{u})$. We minimize $E(\mathbf{u})$ using the fluid flow proposed in [3]. Given the velocity field \mathbf{v} , the following partial differential equation can be solved for \mathbf{u} :

$$\partial \mathbf{u} / \partial \tau = \mathbf{v} - \mathbf{v} \cdot \nabla \mathbf{u}. \quad (5)$$

The instantaneous velocity as in [4] is obtained by convolving \mathbf{f} with Gaussian kernel G_σ of variance σ^2 , $\mathbf{v} = G_\sigma * (-\mathbf{f}(\mathbf{x}, \mathbf{u}))$.

To avoid possible confusion, we summarize the methods we will be referring to in our subsequent analyses. In later discussions, minimization of the following energies

$$E(\mathbf{u}) = F_{L^2}(\mathbf{u}) + \lambda R_{KL}(\mathbf{u}), \quad (6)$$

$$E(\mathbf{u}) = F_{L^2}(\mathbf{u}) + \lambda R_{SKL}(\mathbf{u}) \quad (7)$$

using equations (4), (5) will be referred to as L^2 -Asymmetric Unbiased and L^2 -Symmetric Unbiased models, respectively. The model above, provided $\lambda = 0$, will be referred to as the L^2 -Fluid model.

Similarly, minimization of

$$E(\mathbf{u}) = F_{MI}(\mathbf{u}) + \lambda R_{KL}(\mathbf{u}), \quad (8)$$

$$E(\mathbf{u}) = F_{MI}(\mathbf{u}) + \lambda R_{SKL}(\mathbf{u}) \quad (9)$$

will be referred to as the MI-Asymmetric Unbiased and MI-Symmetric Unbiased models, respectively. Such models, with $\lambda = 0$, define the MI-Fluid model.

5. STATISTICAL ANALYSIS

5.1. Voxel-wise t test

Based on the authors' approach in [2], we observe that, given that there is no systematic structural change within two weeks, any deviation of the Jacobian map from one should be considered error. Thus,

we expect that a better registration technique would yield $\log |Dg|$ values closer to 0 (i.e., smaller log Jacobian deviation translates into better methodology). Mathematically speaking, one way to test the performance is to consider the deviation map dev of the logarithmically transformed Jacobian away from zero, defined at each voxel as $dev(\mathbf{x}) = |\log |Dg(\mathbf{x})||$. For two different registration methods A and B , we define the voxel-wise deviation gain of A over B (denoted by $S^{A,B}$) as $S^{A,B}(\mathbf{x}) = dev^A(\mathbf{x}) - dev^B(\mathbf{x})$.

For the ADNI baseline dataset, we perform a group paired t test across subjects, by computing a voxel-wise t -map of deviation gains. In order to statistically compare the performance of two registration methods, we rely on the standard t test on the voxel mean of S . To construct a suitable null hypothesis, we notice that the following relation would hold, assuming B outperforms A : $S^{A,B} > 0$. Thus, the null hypothesis in this case would be testing if the mean deviation gain is zero: $H_0 : \mu_{S^{A,B}} = 0$. To determine the ranking of A and B , we have to consider one-sided alternative hypotheses. For example, when testing if B outperforms A , we use the following alternative hypothesis: $H_1 : \mu_{S^{A,B}} > 0$.

For both the ADNI follow-up and baseline datasets, we create a voxel-wise t map using the local log Jacobian values of the ten subjects, allowing us to test the validity of the zero mean assumption.

5.2. Permutation Testing to Correct Multiple Comparisons

To determine the overall global effects of different registration methods on the deviation of log Jacobian maps throughout the brain, we performed permutation tests to adjust for multiple comparisons. Following the analyses in [5], we resampled the observations by randomly flipping the sign of $S_i^{A,B}$ ($i = 1, 2, \dots, n$) under the null hypothesis. For each permutation, voxelwise t tests are computed. We then compute the percentage of voxels inside the chosen ROI (in this case the ICBM mask) with T statistics exceeding a certain threshold. The multiple comparisons- corrected p value may be determined by counting the number of permutations whose above-defined percentage exceeds that of the un-permuted observed data. For example, we say that sequence B outperforms A on the whole brain if this corrected p value is smaller than 0.05 (that is, less than 5% of all permutations have the above-defined percentage greater than that of the original data). All possible ($2^{10} = 1024$) permutations were considered in determining the final corrected p value.

To visually assess the global significance level of the voxel-wise t tests on deviation gains and log-Jacobian values, we also employed the cumulative distribution function (CDF) plot. In brief, we plot observed cumulative probabilities against the theoretical distribution under the null hypothesis. In the case of deviation gains S of a worse technique A over a better technique B in the ADNI baseline data, we expect a CDF curve to lie above the Null line, in the sense that a better technique exhibits less systematic changes. In the case of log-Jacobian values, a better registration technique, on the other hand, should be able to separate the CDF curves between ADNI baseline and follow-up phases.

6. RESULTS

In this section, we tested the Asymmetric Unbiased and Symmetric Unbiased models and compared the results to those obtained using the Fluid registration model [3, 4]. Of note, both Asymmetric Unbiased and Symmetric Unbiased methods performed equally well. For each regularization technique, we employed both L^2 and mutual information matching functionals (see (6)-(9)). In order to obtain a fair comparison, re-gridding was not employed.

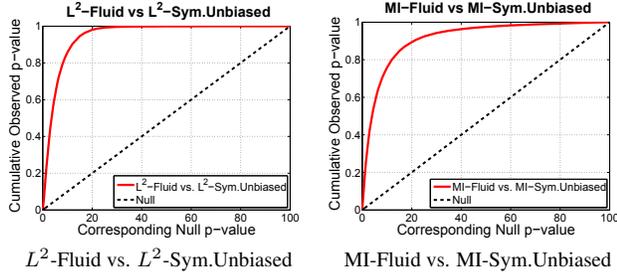


Fig. 5. Cumulative distribution of p -values for the deviation gain S of L^2 -Fluid over L^2 -Symmetric Unbiased and MI-Fluid over MI-Symmetric Unbiased. Here, the ADNI baseline dataset is used. In both figures, the CDF line is well above the Null line ($y = x$), indicating that symmetric unbiased method outperforms Fluid method.

First, nonlinear registration was performed on the ADNI baseline dataset (each scan is 128 by 160 by 128). Here, we compared Fluid, Asymmetric Unbiased, and Symmetric Unbiased methods coupled with both L^2 and mutual information matching ($\lambda = 500$ and $\lambda = 1000$ for L^2 -Symmetric Unbiased and L^2 -Asymmetric Unbiased algorithms; $\lambda = 5$ and $\lambda = 10$ for MI-Symmetric Unbiased and MI-Asymmetric Unbiased methods, respectively). Since the Asymmetric Unbiased model quantifies only the forward deformation, the weight of the corresponding regularization functional is half the magnitude of that of the Symmetric Unbiased model, and hence, a weighting parameter twice as large should be used.

Figure 1 shows the results of registering a pair of serial MRI images for one of the subjects. The deformation was computed in both directions (time 2 to time 1, and time 1 to time 2) using all three regularization methods based on L^2 and mutual information matching. Results indicate the Asymmetric Unbiased and Symmetric Unbiased methods outperform Fluid method, generating more stable inverse consistent maps [6] with less variability.

Figure 2 shows the mean Jacobian maps of ten subjects obtained using Fluid, Asymmetric Unbiased, and Symmetric Unbiased registration algorithms coupled with both L^2 and mutual information matching. Jacobian maps generated using unbiased models have values closer to 1, whereas Fluid model generated noisy mean maps. Figures 3 and 4 demonstrate the Unbiased regularization technique outperforming Fluid registration with statistical significance.

To emphasize the differences between the distributions of log Jacobian values for Fluid and unbiased methods, in Figure 5, we plotted the cumulative distribution function of the p -values in deviation gains ($S^{A,B}$). For a null distribution, this cumulative plot falls along the line $y = x$. Larger upward inflections of the CDF curve near the origin are associated with significant deviation gains, indicating that the unbiased technique outperforms Fluid method in being less likely to exhibit structural changes in the absence of systematic biological changes.

Second, we analyzed the ADNI follow-up phase dataset (each scan is 220 by 220 by 220). As the images are now one year apart, real anatomical changes are present, which enables the comparison of methods in the presence of biological changes. In Figure 6 we observe the Fluid method generating noisy mean Jacobian maps, while maps generated using unbiased methods suggest a volume reduction in gray matter as well as ventricular enlargement. Figure 7 displays the cumulative distribution of p -values for the voxel-wise log Jacobian t -maps for both ADNI baseline and follow-up datasets. We expect a better method to separate these two CDF curves, indicating that a real biological change has occurred between the two time

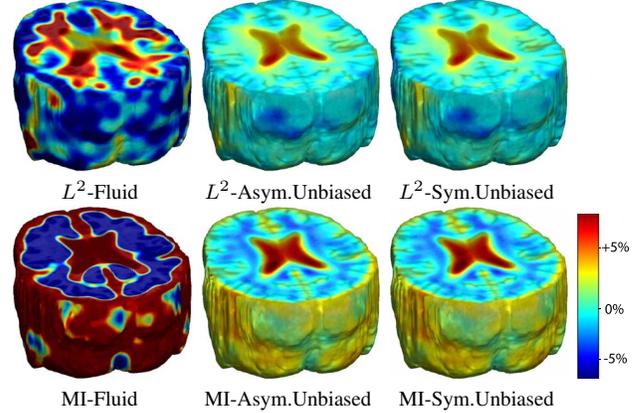


Fig. 6. Nonrigid registration was performed on the ADNI follow-up study (serial MRI images acquired 12 months apart). For each method, the mean of the resulting 10 Jacobian maps is superimposed on one of the brain volumes. Fluid registration generates noisy mean maps, while maps generated using unbiased methods suggest a volume reduction in gray matter as well as ventricular enlargement.

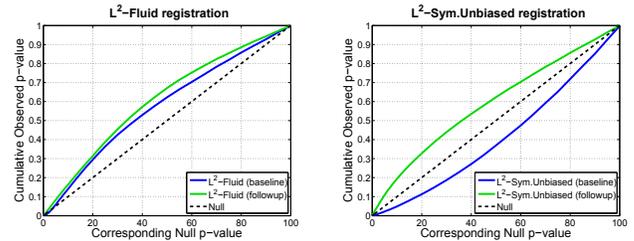


Fig. 7. Cumulative distribution of p -values for the voxelwise log Jacobian t -maps for both ADNI baseline (in blue) and follow-up (in green) using L^2 -Fluid and L^2 -Symmetric Unbiased methods. Here, a better method should separate these two CDF plots (see Section 5.2) with the Null line in between, indicating a real biological change has occurred between these two time points.

points. A greater separation is accomplished when Symmetric Unbiased method is used, while the Fluid method does not differentiate between the two datasets.

7. REFERENCES

- [1] I. Yanovsky et al., “Topology preserving log-unbiased nonlinear image registration: Theory and implementation,” *CVPR*, 2007.
- [2] A. Leow et al., “Statistical properties of Jacobian maps and the realization of unbiased large-deformation nonlinear image registration,” *IEEE Trans. Med. Im.*, vol. 26, 2007.
- [3] G. Christensen et al., “Deformable templates using large deformation kinematics,” *IEEE Trans. Im. Proc.*, vol. 5, 1996.
- [4] E. D’Agostino et al., “A viscous fluid model for multimodal non-rigid image registration using mutual information,” *Med. Im. Anal.*, vol. 7, pp. 565–575, 2003.
- [5] A. Leow et al., “Longitudinal stability of MRI for mapping brain change using tensor-based morphometry,” *NeuroImage*, 2006.
- [6] G. Christensen and H. Johnson, “Consistent image registration,” *IEEE Trans. Med. Im.*, vol. 20, 2001.