Reducing Variability in Anatomical Definitions Over Time Using Longitudinal Diffeomorphic Mapping

Daniel J. Tward[1(✉)], Chelsea S. Sicat[1], Timothy Brown[1], Arnold Bakker[2], and Michael I. Miller[1]

[1] Center for Imaging Science, Johns Hopkins University, Baltimore, MD 21218, USA
{dtward,chelsea,timothy,mim}@cis.jhu.edu
[2] Department of Psychiatry and Behavioural Sciences, Johns Hopkins University, Baltimore, MD 21218, USA
abakker@jhu.edu

Abstract. We address the challenge of variability in the definition of anatomical structures over time in a single subject, using a template-based diffeomorphic mapping algorithm to filter out inconsistencies. Shape changes are parametrized through 2D surfaces, while data attachment is specified through dense 3D images. The mapping uses two geodesic trajectories through diffeomorphism space: template to baseline, and baseline through the timeseries. We apply this algorithm to a study of atrophy in the entorhinal and surrounding cortex in patients with mild cognitive impairment, characterized by rate of change of log-volume. We compare the uncertainty in atrophy rate measured from manual segmentations, to that computed with segmentations filtered using our longitudinal method, and to that computed from FreeSurfer. Our method correlates well with manual (correlation coefficient 0.9881, and results in significantly less variability than manual (p 8.86e-05) and FreeSurfer (p 1.03e-04).

1 Introduction

While post mortem analysis of plaques and tangles in the brain have long been used as the diagnostic criteria for Alzheimer's disease, structural imaging can be important in clinical studies of the disease. The earliest anatomical changes in patients with mild cognitive impairment (MCI) [12] are neuronal cell death in the entorhinal cortex (EC) [7], and have been detected through structural imaging. Atrophy biomarkers measured through neuroimaging have been shown to be predictive of disease onset [24], and are associated with reduced performance on memory related tasks [22] relevant to a patient's lifestyle.

We have been quantifying these early changes through techniques in computational anatomy known as diffeomorphometry. By identifying spatial correspondences between brain atlases and subject scans via smooth diffeomorphic mappings, we can infer information about anatomical changes through properties of theses diffeomorphisms, such as their determinant of Jacobian (see [9] for a recent review).

Estimation of these diffeomorphisms is complicated by high dimensional nuisance variables. The mappings are not uniquely specified in the homogeneous interiors of anatomical structures. Discriminating information is only present on image discontinuities. To address this problem, parsimonious representations have been developed by parametrizing shape changes through a function on the bounding surfaces of anatomical structures, an object of the natural dimension for describing these shapes. In this framework, low dimensional parametric coordinates have been developed by expanding this function in a basis determined through principal component analysis [13,21], or through eigenfunctions of the Laplace-Beltrami operator [20].

On the other hand, the majority of neurimaging data, such as T1 MR images and binary segmentations, is in the form of dense 3D volumes. We have developed a method to incorporate the advantages of the efficient surface representation such as robustness to noise [20] and reproducibility [19], with simple and realistic noise models obtained by working directly with neuroimaging data. These models include white noise (sum of square error) when working with T1 images, or multivariate Bernoulli when working with multiple segmentations [17], as opposed to less easily interpretable data attachment models that work directly with surfaces such as currents or varifolds [1].

Despite these advantages, our ability to infer properties of disease progression is limited by the inconsistency of segmentations of anatomical structures. The problem we address here is the variability in these definitions within a single subject over time in longitudinal studies. This source of variability prevents us from making inferences on an individual level, and mandates larger sample sizes in studies of populations. We address this challenge by extending our framework to map onto each segmentation in the timeseries simultaneously.

This extension has been approached in several different ways. Longitudinal Freesurfer [6,14] addresses this issue with a common initialization of optimization problems for each scan in a timeseries. It avoids modelling any growth or atrophy process with the intention of avoiding bias by privileging a given (e.g. baseline) scan, and to allow the capture of sudden changes.

Several models for growth and atrophy scenarios using flows of diffeomorphisms are discussed in [5], with a focus on modelling populations of timeseries, and describing relationships of a given growth process to a typical one. More complex statistical processes are described through higherarchical geodesic models in [15]. In [11], several parametrizations of these flows are considered, including piecewise geodesic (also used in [5]), spline based, and geodesic shooting (in order of increasing regularity in time).

For our specific problem, filtering out inconsistencies in anatomical definitions, we are less concerned with over regularization and use the shooting approach, using two geodesic trajectories through the space of diffeomorphisms: one from template to baseline, and one from baseline through the timeseries. Such an approach will be shown to significantly reduce this source of variability.

2   Method

2.1   Data

T1 brain MR images from the Alzhimer's Disease Neuroimaging Initiative (ADNI) dataset were examined.[1] Twenty patients were selected, older adults (age 72 ± 8 years), 60% male, education of 17 ± 3 years, with mild cognitive impairment, and having a continuous left collateral sulcus (the most common anatomical variant, others are described for example in [3]). Each subject was scanned up to 5 times, and at least 3 times (so that a residual can be estimated after linear fitting), over 2 years. The EC and immediately lateral cortex (the trans entorhinal cortex, TEC) were analyzed for the presence of atrophy.

Structures were delineated by manual segmentation, on T1 structural scans using the anatomical boundaries described in [8]. The emergence of EC occurs 2 mm caudal to the appearance of the limen insulae and ends 1 mm caudal to the disappearance of the uncus. To account for morphological variation in the limen insulae and uncus, the most rostral boundary of EC was defined 4 mm anterior to the hippocampal head and the most caudal boundary was defined 2 mm posterior to the disappearance of the gyrus intralimbicus, which appears at the caudal end of the uncus. With regard to the medial-lateral boundaries, segmentations were extended as far medially as discernible gray/white matter boundaries would allow and the EC/TEC boundary was delineated vertically at the midpoint of the medial bank of the collateral sulcus. For comparison with state of the art, segmentations were also performed by FreeSurfer version 5.1 that utilizes the 2010 Desikan-Killany atlas [6,14]. We did not use the longitudinal pipeline because the resulting data for the entorhinal cortex was extremely variable and did not look viable.

Imaging data for each subject was rigidly aligned to baseline by minimizing sum of square error in T1 images, and imaging data between subjects was rigidly aligned to a single subject through 4 landmarks placed automatically at the boundaries of the segmentations, minimizing sum of square distances between pairs of landmarks. For each subject i, at time t_j, we denote the rigidly aligned manual segmentation image as J^{ij}.

---

[1] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see http://www.adni-info.org.

2.2   Equations of Diffeomorphometry

The background space of an image, $\Omega \subset \mathbb{R}^3$ is deformed by a diffeomorphism $\varphi : \Omega \to \Omega$, which is generated by a flow under a smooth time varying velocity vector field $v : \Omega \to \mathbb{R}^3$

$$\dot{\varphi}_t = v_t(\varphi_t), \qquad \varphi_0 = \text{identity}, \qquad \varphi \doteq \varphi_1. \tag{1}$$

To ensure solutions are diffeomorphisms, the vector fields are modelled as belonging to a Hilbert space of smooth functions V [23], with inner product given by $\langle v, v \rangle_V \doteq \langle Lv, Lv \rangle_{L_2}$ for L a differential operator designed to give smoother functions a smaller norm. Here we implicitly choose L such that the Green's function of L^*L (for * referring to adjoint) is given by the Gaussian kernel $K(x, y) = \exp\left(-\frac{1}{2\sigma^2}|x - y|^2\right)$ for $\sigma = 6\,\text{mm}$.

This defines geodesics through the space of diffeomorphisms given by Euler's equation. Defining the momentum distribution in the dual V^* as $\mu \doteq L^*Lv$, this can be written as [10]

$$\dot{\mu} = -[Dv]^T \mu \tag{2}$$

We model the family of deformations used to study this population through initial momentum supported on the N_v vertices of a triangulated surface template, q_i for $i \in \{1, \ldots, N_v\}$, that contours our structures of interest

$$\mu_0 = \sum_{i=1}^{N_v} \delta_{q_i} p_i$$

where δ_x is the Dirac distribution centered at x, and p_i is a vector in $\mathbb{R}^3$ parametrizing our deformation. The parameters p will be estimated to model the shape of each structure in a timeseries or population. For notational convenience we write $\exp(p) = \varphi$, for p the parameters, and φ the diffeomorphism calculated by solving (2) and (1). Note that $\|v\|_V^2 \doteq \langle v, v \rangle_V = \sum_{i,j=1}^{N_v} p_i^T K(q_i, q_j) p_j$ which we write as $\|p\|_{V^*}^2$.

These diffeomorphisms act on images through their inverse $\varphi \cdot I \doteq I \circ \varphi^{-1}$, for $I : \Omega \to \mathbb{R}$ an image, which in our discrete implementation is computed through trilinear interpolation. In general we will estimate p by minimizing the sum of square error between segmentation images $\| \exp(p) \cdot I - J \|_{L_2}^2$ for I our template image and J a target image.

2.3   Algorithms

We construct a hypertemplate surface, with vertices q_h, using a restricted delaunay triangulation [2] of the isosurface of the average image of our aligned baseline segmentations. A hypertemplate image I_h is generated by filling each voxel with

its fraction inside the surface (estimated by Monte Carlo sampling). We calculate the deformation parameterized by p_h^0, minimizing the cost function

$$\frac{1}{2\sigma_{p_h^0}^2}\|p_h^0\|_{V^*}^2 + \sum_{i=1}^{N_s}\frac{1}{2\sigma_{p_h^i}^2}\|p_h^i\|_{V^*}^2 + \frac{1}{2\sigma_{I_h^i}^2}\|\exp(p_h^i)\cdot\exp(p_h^0)\cdot I_h - J^{i1}\|_{L_2}^2$$

over p_h^0 and the nuisance parameters p_h^i, for J^{i1} the target binary segmentation image of the baseline scan for the i-th out of $N_s = 20$ subjects. The $\sigma_\cdot^2$ are scalar parameters that provide the flexibility to change relative weighting between terms, but here they are each set to 1. We denote our resulting template image $I \doteq \exp(p_h^0)\cdot I_h$ and our template vertices $q = \exp(p_h^0)(q_h)$ (i.e. each vertex is transformed directly by the diffeomorphism generated by p_h^0). A diagram of this setup is shown in Fig. 1.
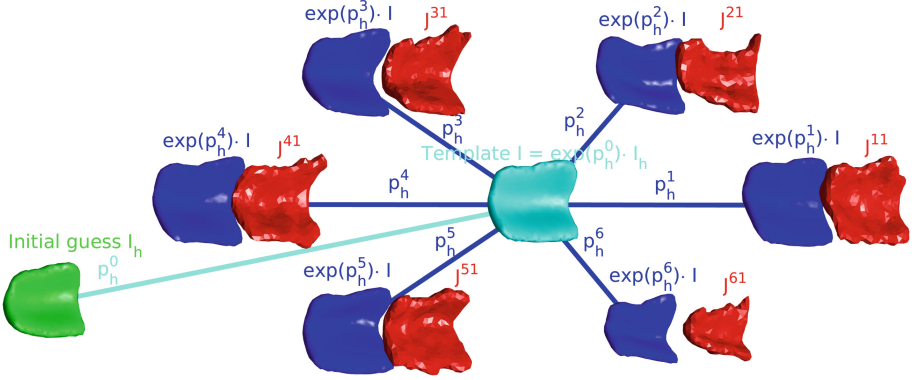
Fig. 1. Illustration of template estimation procedure for six subjects in our population. The hypertemplate is shown in green, the estimated template in cyan, the template deformed to match baseline in blue, and the target baseline scans in red (Color figure online).

Given this template (surface q and image I), we can map onto the timeseries for subject i at each time t_j for $j \in \{1, \ldots, N_{ti}\}$ by minimizing the cost function

$$\frac{1}{2\sigma_{p_0}^2}\|p_0^i\|^2 + \frac{1}{2\sigma_{p_1}^2}\|p_1^i\|_{V^*}^2(t^{iN_{ti}} - t^{i1})$$

$$+ \sum_{j=1}^{N_{ti}}\frac{1}{2\sigma_{I^j}^2}\|\exp(p_1^i(t^{ij} - t^{i1}))\cdot\exp(p_0^i)\cdot I - J^{ij}\|_{L_2}^2$$

over the parameters p_0^i and p_1^i. The $\sigma_\cdot^2$ again provide relative weighting between terms, and they are set to $\sigma_{p_0}^2 = 2$, $\sigma_{p_1}^2 = \sigma_{I^j}^2 = 1$ (chosen heuristically). For this dataset we express t^{ij} in units of 6 months. A diagram of this setup is shown in Fig. 2, where $I^{ij} \doteq \exp(p_1^i(t_j - t_1))\cdot\exp(p_0^i)\cdot I$. Essentially p_0^i represents the
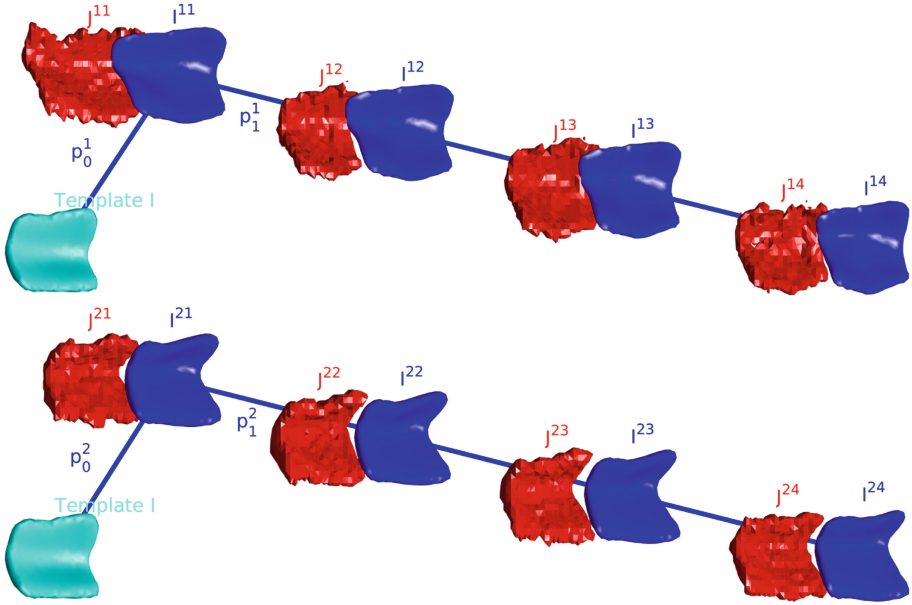
Fig. 2. Example longitudinal mapping results for two subjects. The template is shown in cyan, the deforming template in blue, and the target in red. Top: high variability example, bottom: low variability example (Color figure online).

"intercept" and p_1^i represents the "slope" of a linear regression in the space of diffeomorphisms for subject i.

Each of these minimization problems are solved by gradient descent using an adjoint algorithm. The gradient of each term in the matching cost, namely $(\varphi \cdot I - J)$, is transported backwards in time through a linearized version of the dynamics (1), (2), and $\dot{I}_t = -\nabla I_t v_t$ (optical flow), and contributes additively to the gradient of the the cost function with respect to the parameters. Details of this approach can be found in [20] (or [4] for a similar approach).

2.4   Per Subject Atrophy Rate Estimation

We use the following log-linear model to estimate volumetric atrophy rate in each subject

$$\log v^{ij} = a_0^i + a_1^i t^{ij} + \epsilon^{ij}$$

where v^{ij} is subject i's entorhinal cortex and trans entorhinal cortex volume at time t^{ij}, estimated by summing voxels in the segmentations I^{ij} or J^{ij} (for I, these take value 1 for interior voxels and values between 0 and 1 for boundary voxels) times voxel volume ($0.9375\,\mathrm{mm} \times 1.2\,\mathrm{mm} \times 0.9375\,\mathrm{mm} = 1.0574\,\mathrm{mm}^3$), a_0^i is a nuisance parameter (log volume at $t = 0$), and a_1^i is the atrophy rate

(exponential time constant). In this model, ϵ^{ij} is assumed to be independent Gaussian noise with variance σ_i^2.

Each parameter is estimated by maximum likelihood, including σ_i^2 which is the mean square error of the fit. The variance of our atrophy rate estimator is given by

$$\mathrm{Var}[\hat{a}_1^i] = \frac{\hat{\sigma}_i^2}{N_{ti}\sigma_{ti}^2} \tag{3}$$

where N_{ti} is the number of timepoints for subject i, σ_{ti}^2 is the variance in scan times, and $\hat{\sigma}_i^2$ is our estimate of the variance in ϵ^{ij}. Note that (3) agrees with the residual bootstrap variance estimator within 3.3% (root mean square percent error), but in the case that our linear model is incorrect this quantity can still be interpreted simply as a rescaling of square error after the linear fit.

2.5   Evaluation

We evaluate the accuracy of atrophy rate estimates by examining correlation with manual segmentations. We evaluate the variability by comparing the standard deviation (calculated from (3)) of this estimator measured from manual segmentations, to that measured after our filtering procedure, and to that computed from FreeSurfer.

3   Results

3.1   Mapping Results

The estimated entorhinal cortex and trans-entorhinal cortex atlas is shown in cyan in Figs. 1 and 2 (cyan). Two example longitudinal maps are shown in Fig. 2, illustrating a high variability case (top), and a low variability case (bottom). Note the difference in anterior-posterior (left-right on the figure) extent in the manual segmentations (red) for the first two timepoints for the high variability subject. This inconsitency has been filtered out by our mapping procedure (blue).

3.2   Atrophy Rate

For the two subjects shown in Fig. 2, volumetric analysis is shown in Fig. 3. Volumes of the manual segmentations are shown as red dots, while volumes of the deforming template are shown as a blue line. The volume of the deformed template corresponding to each measured timepoint is shown as a blue dot, and the volume of the template itself is shown as a cyan dot on the left. The reduction in variance due to the longitudinal mapping procedure is evident, particularly for the highly variable subject (left).

The atrophy rate estimated for each subject is shown in Fig. 4, with manual segmentations shown in red, the results of our longitudinal mapping procedure shown in blue, and results from Freesurfer shown in green for comparison with state of the art.
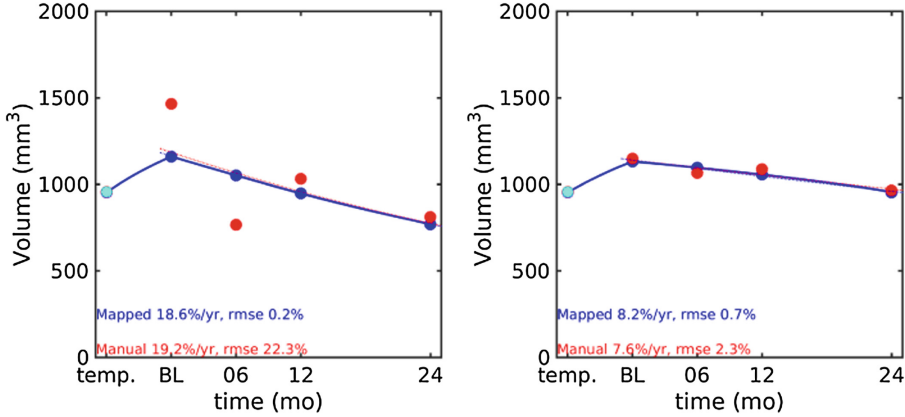
Fig. 3. The mapping procedure's stabilization of volume measurements is illustrated for left: high variability example, and right: low variability example. The horizontal axis indicates months elapsed since baseline scan.
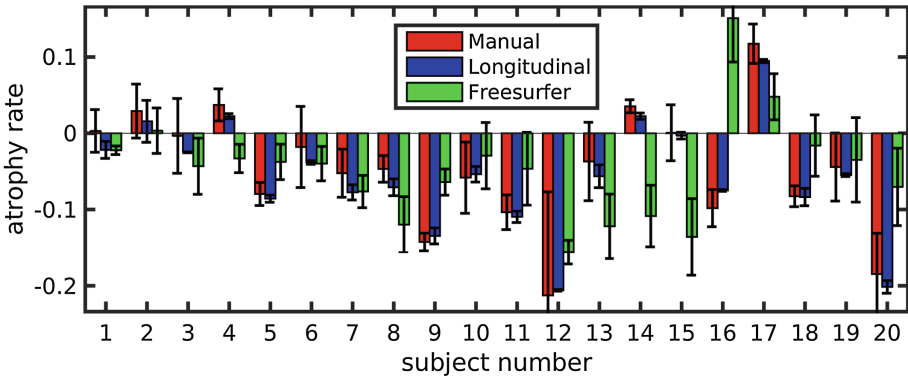

Fig. 4. Estimated atrophy rate in units of exponential time constant, for each of the 20 subjects examined (horizontal axis), is shown for each subject examined as a bar. Standard deviation of the estimator is shown as an errorbar.

3.3   Quality of Atrophy Rate Estimates

Atrophy rate estimators are quite consistent between the manual segmentations and the longitudinal maps (correlation coefficient 0.9881), and not very consistent with FreeSurfer results (correlation coefficient 0.2283), as can be seen in the scatter plot in Fig. 5.

The standard deviation of our atrophy rate estimator, computed according to the square root of (3), is shown in Fig. 6. Significant differences between the three methods are determined by pairwise signed rank tests. Variance is significantly reduced in longitudinal maps relative to manual segmentations ($p = 8.86\text{e-}05$) and relative to FreeSurfer ($p = 1.03\text{e-}04$). However, variance in FreeSurfer estimates is not significantly different from manual segmentations ($p = 6.81\text{e-}01$).
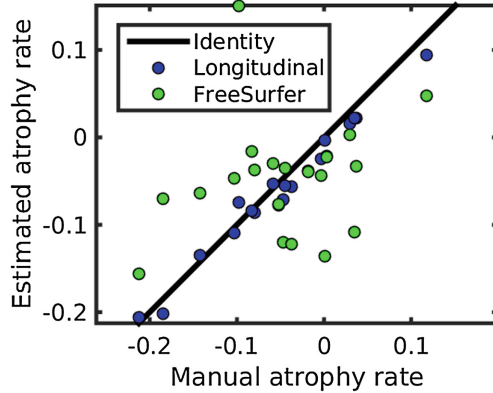
Fig. 5. Correlation between atrophy rate measured from manual segmentations (horizontal axis) and estimated with two methods (vertical axis) in units of exponential time constant is visualized with a scatter plot. Correlation coefficient for longitudinal maps: 0.9881, and for FreeSurfer: 0.2283.
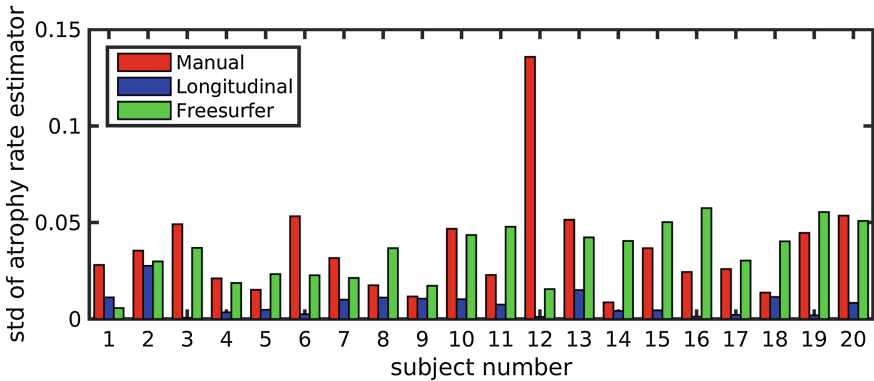

Fig. 6. Standard deviation of our atrophy rate estimator in units of exponential time constant, for each of the 20 subjects examined (horizontal axis), is shown for volume of manual segmentations (red) and volume of our template deformed by the longitudinal mapping procedure (Color figure online).

4   Discussion and Conclusion

As our population of manual segmentations expands to include healthy control subjects in addition to those with MCI, we intend to employ this procedure to identify changes that are specific to disease, as opposed to normal aging. Local modelling of tissue change based on determinant of Jacobian of our mappings will likely prove more sensitive than the volumetry presented here, and can be expanded to include volume change (determinant of 3×3 Jacobian), surface area change (determinant of the 2×2 component of the Jacobian tangent to the

template surface), and thickness change (determinant of the 1×1 component of the Jacobian normal to the template surface). An example of early work with this type of analysis can be found at [18].

Because our method treats each segmentation differently depending on its position in the timeseries, unlike the approach in longitudinal FreeSurfer as mentioned in the introduction, the potential for processing bias exists. This was estimated by reversing the order of the timeseries and repeating the experiment, showing an average overestimate in the magnitude of atrophy rate constant of 0.01. This is a small source of error relative to the inconsistencies in anatomical definitions over time we have sought to address. One simple approach for removing this source of bias is to take the average of the forwards and reversed atrophy rates. A second would be to choose the one which produces the smaller value of the cost function. These and other strategies for removing bias will be the subject of future research.

The longitudinal mapping procedure presented here is able to filter segmentation images, significantly reducing uncertainty in atrophy rate measurements, while correlating strongly with raw manual segmentation results. This procedure has important implications for clinical studies of Alzheimer's disease, where reduced variability will allow for sufficient statistical power at smaller sample sizes.

References

1. Charon, N., Trouvé, A.: The varifold representation of nonoriented shapes for diffeomorphic registration. SIAM J. Imaging Sci. **6**(4), 2547–2580 (2013)
2. Cheng, S.W., Dey, T.K., Shewchuk, J.: Delaunay Mesh Generation. CRC Press, Boca Raton (2012)
3. Ding, S.L., Van Hoesen, G.W.: Borders, extent, and topography of human perirhinal cortex as revealed using multiple modern neuroanatomical and pathological markers. Hum. Brain Mapp. **31**(9), 1359–1379 (2010)
4. Durrleman, S., Allassonnière, S., Joshi, S.: Sparse adaptive parameterization of variability in image ensembles. Int. J. Comput. Vis. **101**(1), 161–183 (2013). http://dx.doi.org/10.1007/s11263-012-0556-1
5. Durrleman, S., Pennec, X., Trouvé, A., Braga, J., Gerig, G., Ayache, N.: Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data. Int. J. Comput. Vis. **103**(1), 22–59 (2013)
6. Fischl, B.: Freesurfer. NeuroImage **62**(2), 774–781 (2012)
7. Gómez-Isla, T., Price, J.L., McKeel Jr., D.W., Morris, J.C., Growdon, J.H., Hyman, B.T.: Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. J. Neurosci. **16**(14), 4491–4500 (1996)
8. Insausti, R., Juottonen, K., Soininen, H., Insausti, A.M., Partanen, K., Vainio, P., Laakso, M.P., Pitkänen, A.: MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. Am. J. Neuroradiol. **19**(4), 659–671 (1998)
9. Miller, M.I., Younes, L., Trouve, A.: Diffeomorphometry and geodesic positioning systems for human anatomy. Technol. (Singap. World Sci.) **2**, 36 (2014). http://dx.doi.org/10.1142/S2339547814500010
10. Miller, M.I., Trouvé, A., Younes, L.: Geodesic shooting for computational anatomy. J. Math. Imaging Vis. **24**(2), 209–228 (2006)
11. Miller, M.I., Trouvé, A., Younes, L.: Hamiltonian systems and optimal control in computational anatomy: 100 years since D'arcy Thompson. Annu. Rev. Biomed. Eng. **17**, 447–509 (2015)
12. Petersen, R.C.: Mild cognitive impairment as a diagnostic entity. J. Intern. Med. **256**(3), 183–194 (2004)
13. Qiu, A., Younes, L., Miller, M.: Principal component based diffeomorphic surface mapping. IEEE Trans. Med. Imaging **31**(2), 302–311 (2012)
14. Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B.: Within-subject template estimation for unbiased longitudinal image analysis. NeuroImage **61**(4), 1402–1418 (2012)
15. Singh, N., Hinkle, J., Joshi, S., Fletcher, P.T.: Hierarchical geodesic models in diffeomorphisms. Int. J. Comput. Vis. **117**(1), 70–92 (2016)
16. Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G.D., et al.: XSEDE: accelerating scientific discovery. Comput. Sci. Eng. **16**(5), 62–74 (2014)
17. Tward, D.J., Bakker, A., Gallagher, M., Miller, M.I.: Changes in medial temporal lobe anatomy quantified using probabilistic atlas construction and surface diffeomorphometry. In: Alzheimer's Association International Conference 2015 (2015)
18. Tward, D.J., Sicat, C.C., Brown, T., Miller, E.A., Ratnanather, J.T., Younes, L., Bakker, A., Albert, M., Gallagher, M., Mori, S., Miller, M.I.: Local atrophy of entorhinal and trans-entorhinal cortex in mild cognitive impairment measured via diffeomorphometry. In: Society for Neuroscience 2016 meeting. Abstract Control Number 8556, November 2016

19. Tward, D., Jovicich, J., Soricelli, A., Frisoni, G., Trouvé, A., Younes, L., Miller, M.: Improved reproducibility of neuroanatomical definitions through diffeomorphometry and complexity reduction. In: Wu, G., Zhang, D., Zhou, L. (eds.) MLMI 2014. LNCS, vol. 8679, pp. 223–230. Springer, Heidelberg (2014). doi:10.1007/978-3-319-10581-9_28
20. Tward, D., Miller, M., Trouve, A., Younes, L.: Parametric surface diffeomorphometry for low dimensional embeddings of dense segmentations and imagery. IEEE Trans. Pattern Anal. Mach. Intell. (2016). doi:10.1109/TPAMI.2016.2578317
21. Tward, D.J., Ma, J., Miller, M.I., Younes, L.: Robust diffeomorphic mapping via geodesically controlled active shapes. Int. J. Biomed. Imaging **2013**, 19 p. (2013). Article No. 3
22. Varon, D., Loewenstein, D.A., Potter, E., Greig, M.T., Agron, J., Shen, Q., Zhao, W., Celeste Ramirez, M., Santos, I., Barker, W.: Minimal atrophy of the entorhinal cortex and hippocampus: progression of cognitive impairment. Dement. Geriatr. Cogn. Disord. **31**(4), 276–283 (2011)
23. Younes, L.: Shapes and Diffeomorphisms. Applied Mathematical Sciences, vol. 171. Springer, Heidelberg (2010)
24. Younes, L., Albert, M., Miller, M.I.: The BIOCARD research team: inferring changepoint times of medial temporal lobe morphometric change in preclinical Alzheimer's disease. NeuroImage Clin. **5**, 178–187 (2014). http://dx.doi.org/10.1016/j.nicl.2014.04.009