Mapping Genetic Influences on Brain Shape using Multi-Atlas Fluid Image Alignment

Meena Mani[1,2]
meena.mani@ucla.edu

Yi-yu Chou[1]
yiyu.chou@loni.ucla.edu

Natasha Leporé[1]
natasha.lepore@loni.ucla.edu

Agatha Lee[1]
alee@loni.ucla.edu

Jan de Leeuw[2]
deleeuw@stat.ucla.edu

Katie McMahon[3]
katie.mcmahon@qimr.edu.au

Margie Wright[3]
margie.wright@qimr.edu.au

Arthur Toga[1]
toga@loni.ucla.edu

Paul M. Thompson[1]
thompson@loni.ucla.edu

[1]*Laboratory of NeuroImaging*
University of California–Los Angeles
Los Angeles, USA.

[2]*Department of Statistics*
University of California–Los Angeles
Los Angeles, USA.

[3]*Centre for Magnetic Resonance*
University of Queensland
Brisbane, Australia.

Abstract

In this pilot study, we developed a set of computer vision based surface segmentation and statistical shape analysis algorithms to study genetic influences on brain structure in a database of brain MRI scans of normal twins. A set of manually delineated 3D parametric surfaces, representing the lateral ventricles, was deformed, using a Navier-Stokes fluid image registration algorithm, onto all the images in the database. The geometric transformations thus obtained were used to propagate the segmentation labels to all the other images. 3D radial distance maps were derived to encode anatomical shape differences. The proportion of shape variance attributable to genetic factors, known as the heritability, was estimated from the shape models using a restricted maximum likelihood method to increase statistical power. Segmentation errors associated with projecting labels onto new images were greatly reduced through multi-atlas averaging. The resulting algorithms provide a convenient and sensitive tool to recover and analyze small intra-pair image differences, and will make it easier to detect genetic influences on brain structure.

1 Introduction

Accurate automated segmentation and labeling of 3D brain images is an extremely challenging problem in computer vision, but is urgently needed for a vast range of medical applications, including clinical trials and population studies of disease. Many population-based studies of Alzheimer's disease, multiple sclerosis, and schizophrenia [27] now use sequential imaging to examine brain changes over time in hundreds or even thousands of subjects, making automated image analysis essential.

There is also great scientific and commercial interest in identifying treatments, genes, environmental or demographic factors that influence brain integrity, as quantified by computational measures of the shape and volume of anatomical substructures. Even so, the discovery of factors that influence brain structure has been held back by the lack of algorithms to automatically identify and compare models of brain structures on a large scale. Automated brain segmentation systems have been proposed, but they are not yet in wide use for large-scale population studies, and many structures are so complex and variable in shape that they are still most commonly traced manually by human experts.

Although many algorithms have been proposed to partially automate brain structure extractions, they still require some user input and therefore some amount of expert knowledge. This can include selecting several points

by hand on the structure boundary to initialize a deformable template close to the structure, prior to high-dimensional fluid registration. Hogan et al. [20] used a fluid registration model to deform a template hippocampal surface model into new subjects' scans, yielding a set of models that were analyzed for group differences in shape in Alzheimer's disease (AD) [1] and shape asymmetries [31]. Level-set or active surface methods, which use partial differential equations to evolve a deforming template under image-derived forces, often require some interactive initialization for accurate label propagation [32].

Some groups have attempted fully automated subcortical segmentation, such as [10] and [15], who used a Markov Random Field model to encode statistical prior information on the expected intensities and adjacencies of structures. Zhou et al. [33] used fuzzy templates to automatically segment different brain structures based on information extracted from a set of training images. Artificial neural networks were used by Ferrarini et al. [14] to study ventricular shape variations in healthy elderly and AD subjects, generating a control average surface and a cloud of corresponding nodes across a data set. Heckemann et al. [19] performed label propagation using free-form deformations and decision fusion to provide automated anatomical segmentations of the brain.

Computational anatomic studies of large brain MRI databases have led to significant neuroscientific discoveries regarding brain changes throughout life and in disease. Even so, the effects of genetics and environment on brain structure are still largely unknown as their detection requires large image databases, automated shape or volumetric analysis, and computational methods to detect gene effects on structures extracted from populations of 3D images.

Here we combine several powerful computer vision approaches with genetic modeling techniques to detect subtle and localized effects of genetic factors on brain structure. First, we combine a Navier-Stokes 3D fluid image registration approach with a parametric surface modeling method to create robust and accurate shape models of a complex brain structure, the lateral ventricles. This structure is a sensitive index of brain change and disease progression in Alzheimer's disease [29], HIV/AIDS [28], and schizophrenia [22].

We extracted parametric surface models from a database of brain MRIs from identical and fraternal twins. Using this set of shape models, we fitted a restricted maximum likelihood model to isolate effects of genetic versus non-genetic factors on brain structure. We were able to isolate genetic sources of variance from computationally derived shape and geometric brain measures, because random genetic polymorphisms (variations across subjects) influence brain shape and are inherited identically in identical twins, while fraternal twins share on average only half of these polymorphisms. To our knowledge, this is the first study to apply genetic variance modeling to anatomical shapes extracted from an MRI brain database with a robust computer vision system. The resulting system aims to identify brain structural features that are most strongly influenced by genes, and ultimately will be used to screen and test for effects of specific candidate genes (polymorphic alleles) that are thought to influence brain morphology (such as apolipoprotein E4, and brain-derived neurotrophic factor, BDNF, among others).

The lateral ventricles, Figure 1, are essentially fluid-filled cavities in the brain, and are enlarged in brain disorders such as Alzheimer's, schizophrenia and bipolar illness. Ventricular shape changes indirectly reflect atrophy in surrounding structures, but surface-based maps of the ventricles provide sensitive assessments of brain integrity, as well as volumetric measures that correlate with cognitive deterioration in Alzheimer's disease [6, 29], HIV/AIDS [28], schizophrenia [22, 23, 25]. Ventricular size and shape are altered as each of these diseases progresses [21], and offer a potential approach to evaluate disease progression in large-scale drug trials. Those at heightened genetic risk for schizophrenia (e.g., a patient.s unaffected twin or sibling) have enlarged ventricles compared to healthy controls [2, 25], as do those at genetic risk for Alzheimer's disease [7] compared to elderly controls. There is great interest in identifying the anatomical selectivity of these gene effects and, once identified, the specific genes that cause them [4, 5].

To better understand how genes influence brain morphology, it would be useful to identify specific structures under genetic influence and pinpoint locations within them that contain genetically-influenced signals. These signals can then serve as a proxy to assess the influence of specific genes. Shape variation in a population can be encoded by using 3D shape representations, for instance Procrustes analysis of set of 3D anatomical landmarks [17] or by using medial axis representations, which describe the boundary of a closed shape in terms of the distance and angle of boundary points relative to a medial curve or surface threading through the center of the structure [24]. A recent twin study [16, 25] examined the genetic effects of neurodegenerative disease on the lateral ventricles using two surface-based shape measures, one derived from a medial representation and the other involving spherical harmonics. Spherical harmonics provide a complete orthonormal series (i.e., basis function) expansion for closed shapes that are homotopic to a sphere, and, if used with statistical dimension reduction techniques such as principal component analysis, their coefficients can be used to distil principal modes of variation or perform pattern classification. The study's findings suggest that shape statistics can uncover group differences not detectable using simple volume measurements.

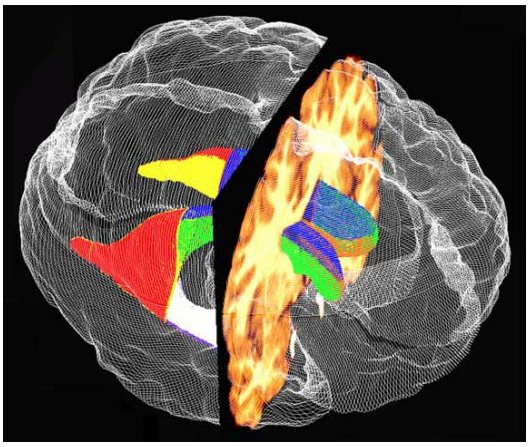

Figure 1. Shape models of the lateral ventricles in a brain MRI scan. Here a set of 3D parametric mesh surfaces is used to represent the lateral ventricles, a fluid-filled structure in the brain. Their complex surface geometry is affected not just in individuals with degenerative disease but also in those at genetic risk for disease. A coronal section from the corresponding 3D MRI scan, and a model of the brain surface (white mesh) are also shown. The anterior horns of the ventricles are shown in blue (top surface) and green (bottom surface), and the posterior horns are shown in red (top surface) and yellow (bottom surface).

In this paper, we extract a surface model of the ventricles based on the medial axis representation, and use it to assess shape differences in a population of subjects. The brain images are first rigidly and fluidly registered [9] to an individual subject's brain on which the ventricles have been manually segmented. This manual label is propagated to each of the brain images using the displacement fields from the registration step. The procedure is performed on a set of 3 reference brains, whose labellings are then combined by averaging. Finally, the radial distance to the medial axis is computed. The details of this algorithm can be found in [7, 8]. The multi-atlas averaging technique dramatically reduces random error introduced through automatic label propagation from a single reference image. Furthermore, this approach eliminates the problem of disconnected voxels that are commonly found in other bottom-up segmentation methods that independently classify each individual voxel as belonging to the structure or not, based on a feature set. Other multi-atlas approaches are described in [12, 15, 19, 26].

Using the radial distance as a basis for comparison, we seek to establish a genetic continuum by examining degrees of relatedness in various pairings as measured by the intraclass correlation (ICC). To assess heritability, we pool both MZ and DZ twin data and solve a restricted maximum likelihood (REML) expression using Fisher's scoring routine. This restricts the number of parameters being estimated, thereby increasing statistical power and the stability of the model.

2 Methods

2.1 Subjects

3D anatomical brain imaging data was acquired from 32 subjects, 10 MZ (6 males, 4 females) and 6 same-sex DZ (4 males, 2 female) twin pairs as part of a 5-year research study of 700 pairs of twins. Informed consent was obtained from all participants and the study was approved by the institutional review board. The subjects ranged in age from 22-25 years and all were healthy. Zygosity was established objectively by typing nine independent DNA microsatellite polymorphisms (PIC > 0.7) by using standard polymerase chain reaction (PCR) methods and genotyping. These results were crossed checked with blood group (ABO, MNS, and Rh), and phenotypic data (hair, skin and eye color), giving an overall probability of correct zygosity assignment of greater than 99.99%.

2.2 MRI image acquisition and preprocessing

3D T1-weighted images were acquired from all subjects on a 4 Tesla Bruker Medspec whole body scanner, located at the CMR and Wesley Hospital MRI Research Facility in Australia. Subjects were scanned using a customized MP-RAGE 3D T1-weighted sequence to resolve anatomy at high resolution (0.9 mm isotropic resolution); TR=2500 ms; TE=3.83 ms; T1=1500 ms; pulse angle=15; coronal orientation; FOV $= 230 \times 230 \times 230$ mm^3; acquisition matrix $= 256 \times 256 \times 256$. All the images were spatially normalized to the International Consortium for Brain Mapping (ICBM-53) average brain imaging template [11] using 9 parameter registration (i.e., 3 rotations, 3 translations, 3 scalings).

2.3 Automated segmentation using multi-atlas fluid image alignment

For this study, we evaluated the left lateral ventricles by tracing consecutive coronal slices of the anterior and posterior horns. Data from the inferior horns were not used because it is difficult to identify consistently in young healthy subjects. For further anatomical specificity, we divided

these two horns into 2 sections, top and bottom. The fluid registration approach we follow, which was first described in [9], treats the deforming image as a compressible viscous fluid whose motion is governed by a linearized version of the Navier-Stokes partial differential equation (PDE). A 3D convolution filter, implementing the Green's function of the differential operator governing the flow [3, 18], was used to speed up the algorithm. In this case, a matching function based on least squares difference between the image intensities was sufficient for good registration, rather than a mutual information or entropy-based similarity measure, due to the high contrast between ventricular CSF and brain tissues. The 3D deformation fields from the registration were applied to the manually segmented image, and the contours were applied to the unlabeled images and tri-linearly interpolated. We randomly picked 3 initial templates (i.e., expertly hand-labeled images) to perform the averaging procedure.

2.4 Shape parameter extraction

The contours of the lateral ventricles, segmented as described above, are resampled in order to obtain a regular parametric 3D grid of $150 \times 100 = 15,000$ surface points [29, 30]. The medial axis, which curves through the cross-sectional centroids of the 3D shape, is the reference from which the perpendicular distance to the grid points at the surface is measured. These radial distances allow statistical comparisons of local surface contractions or expansions across subjects to be made.

2.5 Genetic analysis

A phenotype is a specific biological trait such as eye color, that can be measured in a population; such a trait is typically influenced by genes and the environment. Heritability is the proportion of phenotypic variance attributable to genetic variance. It ranges in value from 0.0 (where genes do not contribute at all to individual phenotypic differences) to 1.0 (individual differences are entirely attributable to genetic differences). Heritability also depends on the range of typical environments in the population studied. If the environment of the population is fairly uniform, as in the case of twins reared together, heritability can be estimated from phenotypic measures in identical and fraternal twins, as each type of twin differs in genetic similarity. The twin design provides a mechanism to study the relative contribution of genes to phenotypic variability. MZ twins are genetically identical, while DZ twins share, on average, 50% of their genes. By extending comparisons to pairs with varying degrees of kinship, we design a genetic continuum from which to assess heritability of brain substructure shape.

2.5.1 Intraclass correlation calculation

The standard approach to measure the degree of relationship for twin or other unordered data pairs is to use the intraclass correlation (ICC). The ANOVA (analysis of variance) formulation

$$r = \frac{MS_{between} - MS_{within}}{MS_{between} + MS_{within}} \tag{1}$$

treats each pair as a random effect and the data from the members of each pair are viewed as measurement errors; this model is widely used in twin study calculations. Due to the variability inherent in estimates derived from a small sample, the computed ICC values may be negative. Increasing the sample size typically gives positive ICC values. This can be explained by the fact that the $MS_{between}$ and the MS_{within} are estimates of the population variance σ^2, but the $MS_{between}$ is calculated from sample means and the MS_{within} is calculated from sample variances. Adding more twin pairs to the study will not affect the within-pair variance, but it will affect the distribution of the means if there are differences between twin pairs. The net affect is for the estimated ICC to become positive as additional twin pairs are included.

For this study, we use the restricted maximum likelihood (REML) method, which gives an unbiased ICC estimate. The non-negative REML formula is given by

$$r = \max\left(0,\ \frac{\frac{n}{n-1}MS_{between} - MS_{within}}{\frac{n}{n-1}MS_{between} + MS_{within}}\right) \tag{2}$$

where n is the number of twin pairs.

2.5.2 Heritability calculation

The simplest expression for heritability is

$$h^2 = \frac{r}{G}, \tag{3}$$

where G is the degree of genetic similarity (assumed to be 0.5 for DZ and 1.0 for MZ twins) and r is the ICC. From this, we can obtain independent estimates of heritability from either group of twins. Heritability may also be estimated by combining information from both sets of twins using Falconer's formulation [13]

$$h^2 = 2(r_{MZ} - r_{DZ}). \tag{4}$$

Since the number of pairs is small, we specify a more restrictive likelihood model with a smaller number of parameters. We do this by pooling the MZ and DZ data. We assume that both groups are distributed with the same mean, μ, and variance, σ^2, and that they differ only

by their covariances, ω^2_{MZ} and ω^2_{DZ}. The negative log-likelihood $L(\mu, \sigma^2, \omega^2_{MZ}, \omega^2_{DZ})$ must be minimized over $\sigma^2 > \omega^2_{MZ} > \omega^2_{DZ}$ using an iterative algorithm. This ensures that the heritability estimates are non-negative. We compute initial values from method of moments estimates and use Fisher's scoring algorithm given by:

$$\mathbf{\Theta}_{i+1} = \mathbf{\Theta}_i + I(\Theta)^{-1} V(\Theta) \tag{5}$$

where i is the iteration step size, I is the Fisher information and $V(\Theta)$ is the score function. This is an alternative to the Newton-Raphson method where the Hessian is replaced by the expected value of the Hessian.

3 Results

For this study, we evaluated the left lateral ventricle tracing consecutive coronal slices of the anterior and posterior horn. To further increase resolution, we divided each of these two horns into 2 sections, top and bottom. The number of left ventricular locations sampled ($15,000 \times 4 = 60,000$ points) provides an intraclass correlation (ICC) statistic and heritability coefficient for each surface grid point; these have been displayed in the form of a maps (Figures 2 and 3). For summary purposes, these were tabulated as a mean and standard deviation for all the sampled grid points (Tables 1 and 2).

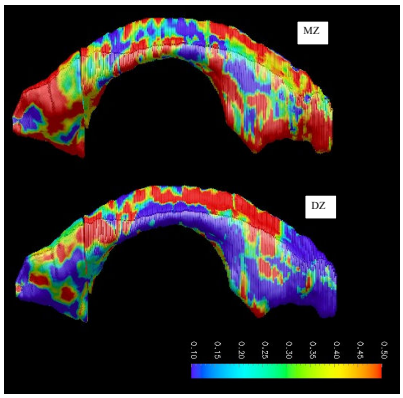

Figure 2. ICC maps show that genetically similar MZ twins(top) have greater intra-pair similarity than do DZ twins (bottom). The increase in ICC is indicated by the increased presence of red in the top diagram ($r > 0.5$.)

As shown in Table 1, the intraclass correlation coefficients were numerically greater in identical twins than in randomly selected non-related pairs of subjects, for all

Table 1. ICC, 1 atlas

	MZ(n=10)		DZ(n=6)		NR(n=10)	
	Mean	SD	Mean	SD	Mean	SD
Top Anter.	0.35	0.20	0.23	0.26	0.11	0.15
Bot Anter.	0.38	0.21	0.20	0.22	0.07	0.11
Top Post.	0.42	0.22	0.25	0.22	0.03	0.06
Bot Post.	0.46	0.23	0.21	0.21	0.05	0.10

of the measures chosen. Although quantitative testing of the differences in correlations for identical versus fraternal twins would require a large sample to confirm, this pilot sample shows that ICCs are numerically greater for twins with greater genetic affinity. As in prior studies of heritable traits, there is evidence for a genetic continuum in which similarity is greatest for MZ twins, somewhat less for DZ twins, and zero for randomly selected unrelated subjects.

The effects of multi-atlas segmentation and labeling of the images is also evident from Table 2. In Chou et al. [7, 8], the Hausdorff error of ventricular labelings, estimated by comparing binary maps of manual versus automated segmentations, decreased as the number of propagated templates increased. When 3 ventricular templates are propagated into each brain scan and averaged, the net effect, relative to using a single template, is to reduce a source of methodological error, namely the error associated with labeling of the ventricles. Because this source of labeling error is diminished, Table 2 shows that in general all the ICC

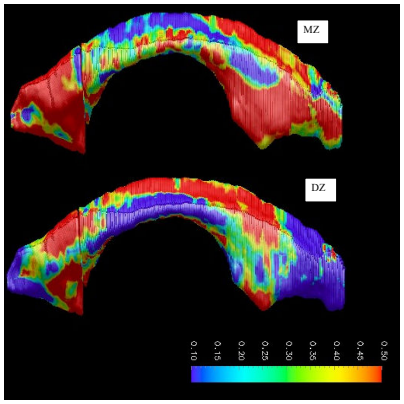

Figure 3. ICC maps after 3 atlas averaging. The ICC values (MZ-top, DZ-bottom) increase with the use of additional reference atlases (compare with Figure 2) suggesting that a source of error has been reduced.

Table 2. ICC, 3 atlas registration: $ICC_{MZ} > ICC_{DZ} \gg ICC_{NR}$.

	MZ(n=10)		DZ(n=6)		NR(n=10)	
	Mean	SD	Mean	SD	Mean	SD
Top Anter.	0.37	0.22	0.27	0.26	0.15	0.18
Bot Anter.	0.40	0.24	0.26	0.23	0.10	0.13
Top Post.	0.51	0.19	0.32	0.28	0.02	0.04
Bot Post.	0.58	0.20	0.30	0.26	0.05	0.10

coefficients increase for every measure chosen, and for each type of twin. This is reasonable, as the sources of labeling error include hand digitization errors in the templates, as well as minor errors in boundary correspondence due to imperfect fluid image registration. Because the magnitude of these errors is not likely to be correlated between members of a twin pair (or an unrelated pair), their removal results in all the ICCs increasing. In some respects, the averaging of parametric surface models from multiple high-dimensional registrations could be generalized to a random effects model of shape variance in a population, where the errors of measurement in a single subject are reduced by repeated measures and hierarchical modeling rather than direct averaging.

4 Discussion

In this study, we combined several novel computer vision algorithms for automated brain image segmentation, fluid image registration, surface parameterization, shape statistics and quantitative genetics to automatically assess how genes influence brain structure in a medical image database. We performed fluid segmentation of the lateral ventricles using a 3D Navier-Stokes registration model in an MRI database of twins, modeling surface shape variation using surface-based statistics derived from a medial axis transform. The segmentation approach is novel as it uses multi-template averaging and a hybrid surface- and volume-based high-dimensional image registration to improve accuracy. In initial studies, some other groups [16, 25] have proposed approaches for modeling genetic influences on brain shape and volume in twins. Here we use a stable and robust restricted maximum likelihood method to compute genetic effects on brain structure, as well as proportions of variance attributable to genes, including tests of reliability (via the use of different anatomic templates for labeling) and estimates of heritability computed in an expanding MRI database. Even in this pilot study, which is intended to provide a proof of concept for the approach, this benefit of multi-atlas versus single-atlas image segmentation is substantial, with ICCs for identical twins rising from 0.4 to 0.5 in some cases with the addition of multiple segmentation templates (or atlases; see Tables 1 and 2). This increased genetic signal-to-noise ratio may be of substantial value if these image-derived measures are to serve as a quantitative endophenotype to search for the effects of individual candidate genes on brain structure, as the database is greatly expanded.

References

[1] J. Ashburner, J. Csernansky, C. Davatzikos, N. Fox, G. Frisoni, and P. Thompson. Computer-Assisted Imaging to Assess Brain Structure in Healthy and Diseased Brains. *Lancet Neurology*, 2(2), February 2003. 2

[2] W. F. C. Baaré, C. J. van Oel, H. E. Hulshoff Pol, H. G. Schnack, S. Durston, M. M. Sitskoorn, and R. S. Kahn. Volumes of Brain Structures in Twins Discordant for Schizophrenia. *Arch Gen Psychiatry*, 58(1):33–40, 2001. 2

[3] M. Bro-Nielsen and C. Gramkow. Fast Fluid Registration of Medical Images. In *VBC*, pages 267–276, 1996. 4

[4] T. Cannon, W. Hennah, T. van Erp, P. Thompson, J. Lonnqvist, M. Huttunen, T. Gasperoni, Tuulio-Henriksson, T. Pirkola, A. Toga, J. Kaprio, J. Mazziotta, and L. Peltonen. DISC1/TRAX Haplotypes Associate with Schizophrenia, Reduced Prefrontal Gray Matter, and Impaired Short- and Long-Term Memory. *Archives of General Psychiatry*, 62(11):1205–13, November 2005. 2

[5] T. Cannon, P. Thompson, T. van Erp, M. Huttunen, J. Lonnqvist, J. Kaprio, and A. Toga. Mapping Heritability and Molecular Genetic Associations with Cortical Features Using Probabilistic Brain Atlases: Methods and Initial Applications to Schizophrenia. *Neuroinformatics*, 4(1):5–19, 2006. 2

[6] O. Carmichael, H. Aizenstein, S. Davis, J. Becker, P. Thompson, C. Meltzer, and Y. Liu. Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. *NeuroImage*, 27(4):979–990, 2005. 2

[7] Y. Chou, N. Leporé, G. de Zubicaray, S. Rose, O. Carmichael, J. Becker, A. Toga, and P. Thompson. Automated 3D Mapping and Shape Analysis of the Lateral Ventricles via Fluid Registration of Multiple Surface Based Atlases. In *ISBI*, 2007. 2, 3, 5

[8] Y. Chou, N. Leporé, G. de Zubicaray, S. Rose, O. Carmichael, J. Becker, A. Toga, and P. Thompson. Automated 3D Mapping and Shape Analysis of the Lateral Ventricles via Fluid Registration of Multiple Surface Based Atlases. *NeuroImage, submitted*, 2007. 3, 5

[9] G. E. Christensen, R. D. Rabbitt, and M. I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, 1996. 3, 4

[10] M. Chupin, R. Mukuna-Bantumbakulud, D. Hasboun, E. Bardinet, S. Baillet, S. Kinkingnehun, L. Lemieux, B. Dubois, and L. Garner. Anatomically constrained region deformation for the automated segmentation of the hippocampus and the amygdala: Method and validation on con-

trols and patients with Alzheimer's disease. *NeuroImage*, 34:996–1019, 2007. 2

[11] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr*, 18(2):192–205, 1994. 3

[12] B. Dawant, S. Hartmann, J. Thirion, D. Vandermeulen, and P. Demaerel. Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations: Part I, methodology and validation on normal subjects. *IEEE Transactions on Medical Imaging*, 18:909–916, 1999. 3

[13] D. S. Falconer. *Introduction to Quantitative Genetics*. Longman, second edition, 1981. 4

[14] L. Ferrarini, W. Palm, H. Olofsen, M. Buchem, J. H. Reiber, and F. Admiraal-Behloul. Shape differences of the brain ventricles in Alzheimer's disease. *NeuroImage*, 32(3):1060–1069, 2006. 2

[15] B. Fischl, D. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, N. Rosen, and A. Dale. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002. 2, 3

[16] G. Gerig, M. Styner, M. E. Shenton, and J. A. Lieberman. Shape versus Size: Improved Understanding of the Morphology of Brain Structures. In *MICCAI*, pages 24–32, 2001. 2, 6

[17] C. Goodall. Procrustes Methods in the Statistical Analysis of Shape. *Journal of the Royal Statistical Society, Series B*, 53(2):285–339, 1991. 2

[18] C. Gramkow and M. Bro-Nielsen. Comparison of three filters in the solution of the navier-stokes equation in registration. In *Proceedings of the Scandinavian Conference on Image Analysis-SCIA'97*, pages 795–802, 1997. 4

[19] R. Heckemann, J. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage*, pages 115–126, July 2006. 2, 3

[20] R. Hogan, K. D. Mark, L. Wang, S. Joshi, M. Miller, and R. Bucholz. Mesial Temporal Sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients. *Radiology*, 216:291–297, 2000. 2

[21] D. H. Mathalon, E. V. Sullivan, K. O. Lim, and A. Pfefferbaum. Progressive Brain Volume Changes and the Clinical Course of Schizophrenia in Men: A Longitudinal Magnetic Resonance Imaging Study. *Arch Gen Psychiatry*, 58(2):148–157, 2001. 2

[22] K. Narr, P. Thompson, T. Sharma, J. Moussai, R. Blanton, B. Anvar, A. Edris, R. Krupp, J. Rayman, M. Khaledy, and A. Toga. 3D Shape Characterization and Mapping of Temporo-Limbic Regions and the Lateral Ventricles in Schizophrenia. *Biological Psychiatry*, 50(2):84–97, 2001. 2

[23] K. Narr, P. Thompson, P. Szeszko, D. Robinson, S. Jang, R. Woods, S. Kim, K. Hayashi, D. Asunction, A. Toga, and B. R.M. Regional specificity of hippocampal volume reductions in first-episode schizophrenia. *NeuroImage*, 21:1563–1575, 2004. 2

[24] M. Styner and G. Gerig. Medial models incorporating shape variability, 2001. 2

[25] M. Styner, J. A. Lieberman, R. K. McClure, D. R. Weinberger, D. W. Jones, and G. Gerig. Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors. *PNAS*, 102(13):4872–4877, 2005. 2, 6

[26] C. Svarer, K. Madsen, S. G. Hasselbalch, L. H. Pinborg, S. Haugbøl, V. G. Frøkjær, S. Holm, O. B. Paulson, and G. M. Knudsen. MR-based automatic delineation of volume of interest in human brain PET imaging using probability maps. *NeuroImage*, 24(4):969–979, February 2005. 3

[27] P. Thompson, G. Bartzokis, K. Hayashi, A. Klunder, P. Lu, N. Edwards, M. Hong, M. Yu, J. Geaga, A. Toga, C. Charles, D. Perkins, J. McEvoy, R. Hamer, M. Tohen, G. Tollefson, and J. Lieberman. Time-Lapse Mapping Reveals Different Disease Trajectories in Schizophrenia depending on Antipsychotic Treatment. *Journal of Neuroscience, submitted*, 2007. 1

[28] P. Thompson, R. Dutton, K. Hayashi, A. Lu, S. Lee, J. Lee, O. Lopez, H. Aizenstein, A. Toga, and J. Becker. 3D mapping of ventricular and corpus callosum abnormalities in HIV/AIDS. *NeuroImage*, 31(1):12–23, 2006. 2

[29] P. Thompson, K. Hayashi, G. de Zubicaray, A. Janke, S. Rose, J. Semple, M. Hong, D. Herman, D. Gravano, D. Doddrell, and A. Toga. Mapping hippocampal and ventricular change in Alzheimer's disease. *NeuroImage*, 22(4):1754–1756, August 2004. 2, 4

[30] P. Thompson, C. Schwartz, and A. Toga. High-resolution random mesh algorithms for creating a probabilistic 3D surface atlas of the human brain. *NeuroImage*, 3(1):19–34, February 1996. 4

[31] Y. Wang, X. Gu, K. Hayashi, T. Chan, P. Thompson, and S. Yau. Surface Parameterization using Riemann Surface Structure. In *ICCV2005*, pages 1061–1066, 2005. 2

[32] P. Yushkevich, J. Piven, C. Hazlett, H. Smith, G. Smith, R. Ho, S. Ho, J. Gee, and G. Gerig. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage*, 31:1116–1128, 2006. 2

[33] J. Zhou and J. Rajapakse. Segmentation of subcortical brain structures using fuzzy templates. *NeuroImage*, 28(4):915–924, 2005. 2

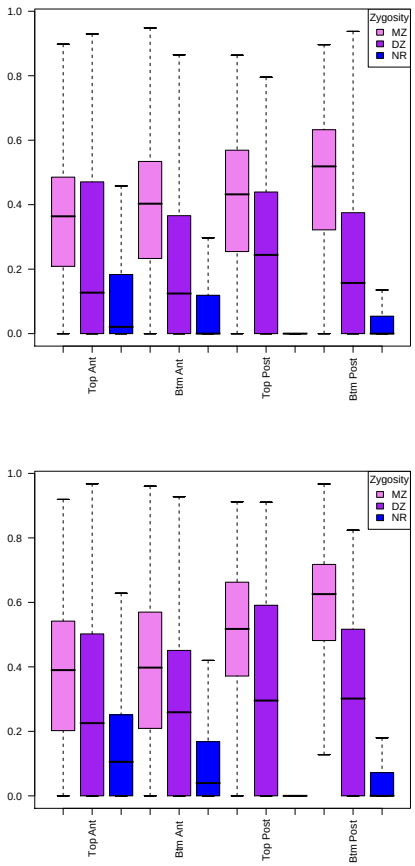

Figure 4. Top: 1 atlas; Bottom: 3 atlas registration. Boxplots comparing ICC values for MZ, DZ and NR pairs for the Top Anterior, Bottom Anterior, Top Posterior and Bottom Posterior locations of the lateral ventricle. $ICC_{MZ} > ICC_{DZ} \gg ICC_{NR}$. **The null hypothesis that there is no significant difference between the means for any given pairing (MZ vs NR; DZ vs NR; MZ vs DZ) was rejected.**

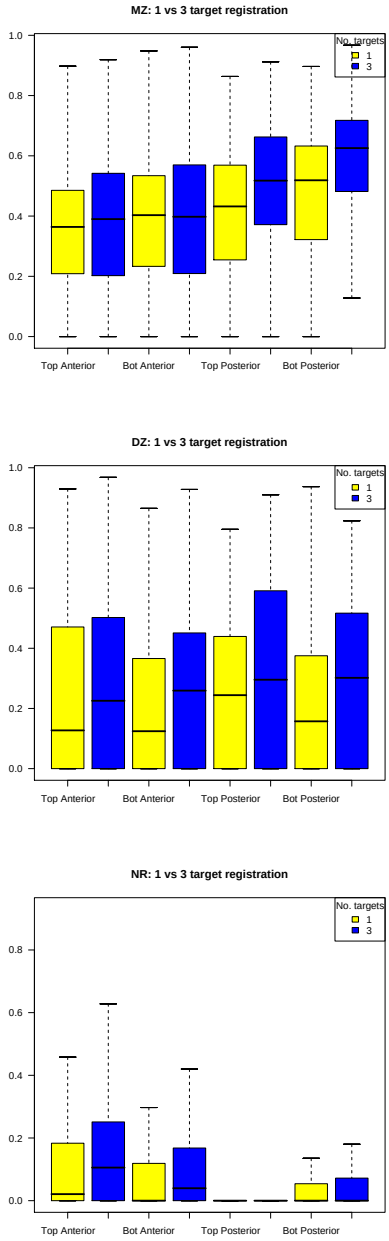

Figure 5. Top: MZ; Center: DZ; Bottom: NR; Here we compare the ICCs for measures derived from anatomical segmentations that use a single deformable surface template, versus the generally more accurate segmentations derived from averaging the results of three deformable surface segmentations. The higher ICCs obtainable with more templates suggest that the anatomical labeling error has been reduced and would otherwise be a major source of methodological error, depleting power to assess genetically influenced shape differences.