*Biostatistics* (2010), **11**, 4, *pp.* 609–630
doi:10.1093/biostatistics/kxq016
Advance Access publication on March 29, 2010



# Surface shape analysis with an application to brain surface asymmetry in schizophrenia

CHRISTOPHER J. BRIGNELL

*School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK*

IAN L. DRYDEN*

*School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK and Department of Statistics, University of South Carolina, Columbia, SC 29208, USA*
ian.dryden@nottingham.ac.uk

S. ANTONIO GATTONE

*Department SEFeMeQ, University of Rome "Tor Vergata", 00133 Rome, Italy*

BERT PARK, STUART LEASK

*School of Community Health Sciences, University of Nottingham, Nottingham, NG3 6AA, UK*

WILLIAM J. BROWNE

*Department of Clinical Veterinary Science, University of Bristol, Bristol, BS40 5DU, UK*

SEAN FLYNN

*Department of Psychiatry, University of British Columbia, Vancouver, BC V3C 4J2, Canada*

## SUMMARY

Some methods for the statistical analysis of surface shapes and asymmetry are introduced. We focus on a case study where magnetic resonance images of the brain are available from groups of 30 schizophrenia patients and 38 controls, and we investigate large-scale brain surface shape differences. Key aspects of shape analysis are to remove nuisance transformations by registration and to identify which parts of one object correspond with the parts of another object. We introduce maximum likelihood and Bayesian methods for registering brain images and providing large-scale correspondences of the brain surfaces. Brain surface size-and-shape analysis is considered using random field theory, and also dimension reduction is carried out using principal and independent components analysis. Some small but significant differences are observed between the the patient and control groups. We then investigate a particular type of asymmetry called torque. Differences in asymmetry are observed between the control and patient groups, which add strength to other observations in the literature. Further investigations of the midline plane location in the 2 groups and the fitting of nonplanar curved midlines are also considered.

*Keywords*: Bayesian; Independent components analysis; Laplace distribution; Likelihood; Principal components analysis; Size and shape; Star-shaped; Symmetry; Torque.

*To whom correspondence should be addressed.

## 1. INTRODUCTION

Statistical shape analysis has developed rapidly since the pioneering work of Kendall (1984) and Bookstein (1986). The primary focus has been the analysis of landmark data where points of correspondence (landmarks) are located on each object, and shape analysis of the points is invariant under certain geometrical transformations of the object. The most commonly considered geometrical invariances are translation, rotation, and scaling.

There is growing interest in developing methodology for analysis of the shapes of curves, surfaces and volumes, as well as landmark data. Applications include brain shape in medical image analysis, face shape, cell shape in biomedical science, molecule shape in bioinformatics, and many other examples (e.g. Grenander and Miller, 1994; Kent *and others*, 1994, 2000; Stoyan and Molchanov, 1997; Joshi *and others*, 1997; Bookstein, 1997; Van Essen *and others*, 1998; Morris *and others*, 1999; Thompson *and others*, 2000; Fischl *and others*, 2001; Hobolth *and others*, 2002; Hobolth, 2003; Klassen *and others*, 2003; Chung *and others*, 2003; Dryden, 2005; Bock and Bowman, 2006; Qiu *and others*, 2006; Barry and Bowman, 2008).

There are several general approaches for carrying out shape analysis of objects. One method is to register the objects using simple global transformations which are treated as nuisance parameters (such as translation and rotation), and statistical analysis is then carried out on the "residuals" after matching. This approach will be our main focus. An alternative is to match the objects closely using sophisticated high-dimensional nonlinear deformations and then carry out statistical analysis using the deformation parameters and/or residuals. For this type of approach, we shall also consider nonlinear deformations with the popular technique of voxel-based morphometry (VBM) (Ashburner and Friston, 2000). Further techniques include orthogonal basis function modeling and then carrying out statistical analysis on the fitted parameters.

Our motivating application is the analysis of brain surface shape in neuroscience. The study involves the investigation of whether there are large-scale shape differences between schizophrenia patients and controls. The cortical surface of the brain tends to exhibit asymmetry, and in particular the right frontal region is larger than the left and the left occipital region is larger than the right on average. This particular asymmetry is often called "brain torque." We aim to investigate the asymmetry in the brain surfaces from magnetic resonance (MR) images, as well as other aspects of shape difference between schizophrenia patients and controls.

Of fundamental importance in shape analysis is the choice of registration used to remove the unwanted transformations and the labeling used to ensure correspondence of parts between objects. For example, patients in an MR scanner are not located in exactly the same position or orientation, and so there is a need to carry out some form of registration to remove these rigid body transformations as part of the analysis. In our data set, if we were to ignore the registrations, we would end up making meaningless comparisons of different parts of brains.

The labeling of each brain specifies which parts of one brain correspond to which parts of another. Correspondence can be either in terms of biological homology or in a geometrical sense. The geometrical properties of a labeled object that are invariant under a registration group of transformations are called the "shape" of the object. Specific examples include size and shape where the registration group is the rigid body transformation group of translation and rotation (as in our application) and similarity shape where the registration group includes translation, rotation, and scaling (e.g. for object recognition from photographs at different scales). There are a variety of ways to deal with registration invariance and label correspondence, including estimating optimal registration and labeling parameters using a statistical model. In our case study, there are 3 main stages to the analysis:

1. Segmentation of the brain surface from each image.

2. Estimation of the midline (midsagittal cross-section) and internal landmarks so that each brain image can be translated and rotated into a standardized frame of reference.
3. Inference about the shape and asymmetry of the brain surface.

At the first stage, the brain surface itself is identified in each image, and then the second stage involves registration of each brain image by translation and rotation. Once registration has been carried out, the labeling of the surface can be defined, and we take a large number of points located on the surface using polar coordinates. Finally, the shape and asymmetry analysis of the high-dimensional point sets can be carried out.

In Section 2, we provide more detail about the application that forms the main part of the work. We describe the data set, the preprocessing used to extract the brain surface boundary, and introduce some notation and the parameters of interest. In Section 3, we describe our mathematical representations of a continuous surface size and shape. In Section 4, we investigate a statistical model for the MR images, specifying the likelihood in independent regions of the image. We consider maximum likelihood estimation (MLE) for registering the images, and we also consider a Markov chain Monte Carlo (MCMC) method for simulating from the posterior distribution in a Bayesian approach. In Section 5, we investigate different aspects of shape analysis—brain surface shape analysis, asymmetry and torque, and curved midline analysis. We conclude with a brief discussion.

## 2. Brain surface data

### 2.1 *The data*

Our main motivation for this work is a study in neuroscience investigating shape and symmetry in the surface of the brain in schizophrenia patients and controls. The data set consists of $n = 68$ 3D MR images of the brain from 29 male healthy controls, 25 male schizophrenia patients, 9 female healthy controls, and 5 female schizophrenia patients. The mean ages are male controls (36.6), male patients (33.2), female controls (33.9), and female patients (33.4). The MR images are proton density–weighted images and were collected by Sean Flynn at the University of British Columbia, Canada. All scans in our case study were from individuals under 50 and nearly all the subjects were right-handed (writing hand) except 1 male patient and 1 male control. Each volunteer's image consists of $256 \times 256 \times 256$ voxels (3D pixels of size $1\text{mm}^3$).

### 2.2 *Preprocessing: brain surface segmentation*

There are many sophisticated image analysis programs available to assist with common tasks in medical image analysis. We use a medical image analysis tool to extract the brain surface boundary, which is the boundary between the gray matter and the cerebral spinal fluid (CSF). Since we are interested in large-scale shape differences, we wish to extract the bounding region of the cortical surface, which is an envelope over the cortical surface itself. The actual cortical surface is very complicated geometrically, containing folds, sulci and gyri, and methods using high-dimensional spherical basis functions such as spherical harmonics or spherical wavelets (Shen *and others*, 2004; Chung *and others*, 2007; Yu *and others*, 2007) aim to model such structure. However, in our case, we are interested in large-scale shape differences in a bounding region of the cortical surface, and so we use the brain extraction tool (BET) of Smith (2002) which is available in the FSL library (Smith *and others*, 2004). The tool requires a tuning parameter to be set (between 0 and 1—the default is 0.5) and fits a balloon-like template through an energy minimization scheme. The preprocessing parameter is chosen manually by comparing the BET image and the magnetic resonance imaging scan to see if any CSF has been included or if not enough gray matter has been

included. For our images, it was often necessary to raise the tuning parameter to values in the range 0.5–0.7, as too much of the CSF was being included. We call the output a "BET image," which is an image of the same size as the original but with a mask of zero value on voxels placed outside the brain, with the brain surface at the boundary. Note that any error in estimating the brain surface is considerably less than the variability between different brains, and so it is a reasonable practical approach to treat the brain surface boundary as part of the data (i.e. as known).

### 2.3 *Parameterization*

Each brain has been approximately oriented in the scanner but there are differences in translation and rotation that we wish to remove so that we can compare the brain surfaces of different individuals. We use the following Euclidean coordinate system throughout: $x$-axis: posterior $\rightarrow$ anterior (back to front), $y$-axis: inferior $\rightarrow$ superior (bottom to top), $z$-axis: right $\rightarrow$ left (n.b. "left" = patient's left). The sagittal plane is the $x$-$y$ plane, the coronal plane is the $y$-$z$ plane, and the axial plane is the $x$-$z$ plane.

The registration of each brain is obtained by estimating rigid body transformations, with translation represented by $\xi = (\xi_x, \xi_y, \xi_z)^T \in \mathbb{R}^3$ and a rotation using 3 Eulerian angles: $\theta_{\mathrm{p}}$ (pitch angle about $x$ axis), $\theta_{\mathrm{r}}$ (roll angle about $y$ axis), $\theta_{\mathrm{y}}$ (yaw angle about $z$ axis). We write $\phi = (\xi_x, \xi_y, \xi_z, \theta_{\mathrm{p}}, \theta_{\mathrm{r}}, \theta_{\mathrm{y}})^T$ for the registration parameters.

We consider registration of each brain into Talairach space (Talairach and Tournoux, 1988) and this procedure involves locating a midline plane (which is the midsagittal cross-section joining the 2 hemispheres), locating 2 landmarks called the anterior commissure (AC) and posterior commissure (PC) in this plane, and then finally rotating and translating the brain such that the line joining the commissures (AC–PC line) in the midline plane is horizontal with the midpoint of the AC–PC line at the origin. The locations of the AC and PC landmarks in the original image are determined by the registration parameters $\phi$ and an additional parameter $\xi_c > 0$ specifying the distance between them.

## 3. SURFACE SIZE AND SHAPE

We can regard a surface in $\mathbb{R}^d$ as a function $\{X(t) \in \mathbb{R}^d : t \in \mathcal{D}\}$ on some domain $\mathcal{D}$. For our application $d = 3$. In general, the function $X(t)$ contains the shape and registration information of the surface. We can derive a shape function by placing the surface in a standardized registration, for example, by matching to a template, or registering on particular landmarks, or standardizing to a particular plane. In our application, we will estimate the midline of the brain, the AC and PC, and then translate and rotate the brain into a standard frame of reference. The registration method is described in detail in Section 4.

The labeling of an object is a one-to-one and onto function (i.e. a diffeomorphism) $h(u) : \mathcal{D} \rightarrow \mathcal{D}$ which assigns the correspondence between parts of surfaces, so a relabeled surface is $X(h(u))$. In our application, we will have $h(u) = u$, after registering each brain and identifying a large number of points on the brain surface.

We focus on the particular case of star-shaped objects in $\mathbb{R}^d$ which have a surface represented by a radial function $R(t)$ from an internal point, where $R(t) > 0$ and $t \in \mathcal{D} \subseteq S^{d-1}$ are suitable spherical coordinates. Here $S^{d-1}$ is the unit radius sphere in $d$-dimensions. Star-shaped objects are very common in many applications, and the simple univariate polar representation lends itself to convenient mathematical and statistical analysis. In our application, we consider radial vectors from a central point of the brain to the surface, after it has been registered to a standardized position using translation and rotation. The radial functions $R(t)$ then represent the size and shape of the continuous surface. In order to compare different star-shaped objects, we will first need to obtain a standardized registration and then compare the radial functions. Possible models for the size-and-shape functions include Gaussian processes, and an alternative

is to work with $\log R(t)$, since $R(t) > 0$. If shape (with scale invariance) is required, we could work with $U(t) = R(t)/S$ for a suitable size measure $S$ (Dryden, 2005), but in our application, we wish to retain size.

In practice, a discrete set of points will be available, and we assume that the points are identified on a surface for $k$ values of $t_i \in \mathcal{D}, i = 1, \ldots, k$ (e.g. on regularly spaced rays from the origin). Assuming the object is in a standardized registration, the collection of radii $R = (R_1, \ldots, R_k)^T$ measures the size and shape of the surface. Suitable models in this case include multivariate normal distributions, either for $R$ or $(\log R_1, \ldots, \log R_k)^T$. Given the high-dimensional nature of the data, some form of dimension reduction such as principal components analysis (PCA) or independent components analysis (ICA) may be useful. We take $t_i$ to be regularly spaced in angles on the sphere and the radii $R(t_i)$ are the distances from the origin to the brain surface when a ray fired from outside the head at direction $t_i$ first hits the brain surface boundary ($i = 1, \ldots, k$). In our case, $k = 62\,501$. Note that the $R(t_i)$ are deterministic functions of the registration parameters conditional on the brain surface boundary being known (having been estimated by BET in our case).

## 4. Brain registration

### 4.1 *Maximum likelihood registration*

In order to carry out statistical inference for the parameters $(\phi, \xi_c)$, we first need to specify the likelihood function for the voxels in the preprocessed BET images. Consider a single BET image $Y$, with $Y_t$ denoting the voxel value at location $t = (t_x, t_y, t_z) \subseteq \mathcal{S} = \{1, \ldots, 256\}^3$. The center of the image is denoted $(O_x, O_y, O_z)$ and in our case this is equal to $(128.5, 128.5, 128.5)$. After preprocessing, the voxels in the BET image outside the brain surface are identically zero. We construct the likelihood for the nonzero voxels only. We shall partition the voxel grid into distinct regions which depend on the parameters of interest and then specify independent distributions for each region.

For our analysis, the brain surface labeling is determined by the registration of the images. We shall consider registration based on the approximate symmetrical structure in the brain in the close vicinity of the midline, and so a very important part of the model is the distribution of the voxel values in the midline region.

We shall divide each BET image into 5 parts:

1. The midline region $\mathcal{M}$ (defined as the nonzero voxels within distance $\epsilon_M$ of the midline plane, whose mirror image about the midline is also nonzero).
2. The AC region $\mathcal{A}$ (defined as the voxels within $\epsilon_A \leqslant \epsilon_M$ in the $x, y$, or $z$ direction of the AC landmark).
3. The PC region $\mathcal{P}$ (defined as the voxels within $\epsilon_P \leqslant \epsilon_M$ in the $x, y$, or $z$ direction of the PC landmark).
4. The remaining nonzero pixels $\mathcal{O}$ of the BET image.
5. The region outside the brain surface containing zero-value pixels (which do not contribute to the likelihood).

Note that $\mathcal{A} \subseteq \mathcal{M}$ and $\mathcal{P} \subseteq \mathcal{M}$. In our application, we take $\epsilon_{\mathrm{M}} = 15, \epsilon_{\mathrm{A}} = 10 = \epsilon_{\mathrm{P}}$ (in millimeters/voxels). When registering into Talairach space, the procedure consists of first finding a midline and then finding the AC and PC landmarks. We shall specify our likelihood in 2 stages as well, first of all specifying the likelihood for the midline region and then for the AC/PC regions.

The first stage of registration involves a rigid body transformation with rotation by $\theta_{\mathrm{p}}, \theta_{\mathrm{r}}$ and translation in $z$ to $\xi_z + O_z$. Each voxel $t$ is sent to a new location $s = s(t, \theta) = (s_x, s_y, s_z)$ (rounded to the nearest integer). The midline plane is given by the new coordinates $(s_x, s_y, \xi_z + O_z)$, and we take the

midline as lying at a half-integer position for $\xi_z + O_z$ (lying halfway between 2 planes of voxels). Let $s$ indicate voxels which are left of the midline ($s_z > \xi_z + O_z$). For any registration, the new registered image is defined on the same voxel grid as the old image using a suitable wrap around the edges of the image.

Let $s'$ be the reflection of $s = (s_x, s_y, s_z)$ about the midline, that is, $s' = (s_x, s_y, \xi_z + O_z - |s_z - \xi_z - O_z|)$. We shall regard the voxels at sites $s$ and $s'$ as paired voxels. The midline region is $\mathcal{M} = \{s: |s_z - \xi_z - O_z| \leqslant \epsilon_M\}$, where both pairs of voxels are nonzero. Each pair of voxel values $(Y_s, Y_{s'})$ in $\mathcal{M}$ is transformed to $U_s = (Y_s - Y_{s'})/2$, $V_s = (Y_s + Y_{s'})/2$ (with Jacobian 1/2). We assume that all $V_s$ are independent of all $U_s$, and the likelihood of $V_s$ is assumed constant with respect to $(\xi_z, \theta_p, \theta_r)$.

For voxels in $\mathcal{M}$, we consider a model where the expected values of the gray levels $Y_s$ are symmetrical about the midline, that is, $E[Y_s] = \mu_s = E[Y_{s'}]$, and so $E[U_s] = 0$. For simplicity, we assume that the $U_s$ are independent. Note that $V_s$ in $\mathcal{M} \setminus \{\mathcal{A} \cup \mathcal{P}\}$ are not necessarily independent, but below we do require $V_s$ in $\mathcal{A} \cup \mathcal{P}$ are independent. If the dependencies in $V_i$ are strong in the regions $\mathcal{A}$ or $\mathcal{P}$ then we are using a pseudo-likelihood approximation in that small part (Besag, 1986).

Note that we are not assuming independence of most of the voxel intensities themselves, but rather independence of the $U_s$ is assumed, which is a much less restrictive assumption. If a model with dependent $U_s$ really was required, then we could still use our method, but it would again be a pseudo-likelihood approach. We have investigated the autocorrelation of $U_i$ with neighboring values, and it soon decreases to zero after a very short distance. Thus, uncorrelated $U_i$ seems a reasonable assumption. Independence is of course stronger, but we think it is not unreasonable.

Exploratory data analysis was carried out by examining the histogram of $U_s$ for a good choice of midline for some example MR scans. We observed that a Laplace (double exponential) distribution fits well to the data in $\mathcal{M}$ for the images, as seen in the example in the middle plot of Figure 1. The Laplace model seems appropriate in many image problems perhaps due to the presence of occasional large errors (e.g. segmentation differences at boundaries of structures). Although by far the majority of voxels match up well for a good choice of midline plane, there will be a few large discrepancies (due to tissue boundaries or other nonmatching parts). A model that can accommodate some large positive or negative values is more appropriate here than a short-tailed distribution, such as the Gaussian.

Therefore, a suitable model for $U_s$ with $s(t, \theta) \in \mathcal{M}$ has density

$$f(u_s) = \frac{\psi w_s}{2} \exp(-\psi w_s |u_s|),$$

where $w_s = w_{s'}$ are predetermined weights, that is, a Laplace distribution with scale parameter $\psi w_s$ for $s \in \mathcal{M}$. The weight of the $s$th voxel at a perpendicular distance $s_z$ from the midline is taken as $w_s = \max\{(10.5 - |s_z - \xi_z - O_z|)/10, 0.5\}$.

In order to specify a model in the commissure regions $\mathcal{A}$ and $\mathcal{P}$, we use some training data obtained by manually locating the AC and PC on the midplane of $n_t = 7$ scans. After translating and rotating the images into their final registration the data, the $V_s$, in the region $\mathcal{A} \cup \mathcal{P}$, are standardized for each scan to a common mean and variance by means of a transformation of the form $V_s^* = \beta_j V_s + \gamma_j$, $j = 1, \ldots, n_t$. For each voxel, $s$, in $\mathcal{A}$ and $\mathcal{P}$, we can calculate an estimate for the mean and variance of $V_s^*$ from the training data. The training data here define 2 templates, which represent typical distributions of voxels in the AC and PC regions.

To locate the AC for each scan, we use the model $\beta V_s + \gamma \sim N(\mu_s^A, (\sigma_s^A)^2/w_s^A)$, $s \in \mathcal{A}$, independently and to locate the PC, we model $\beta V_s + \gamma \sim N(\mu_s^P, (\sigma_s^P)^2/w_s^P)$, $s \in \mathcal{P}$, independently, where $\beta$ and $\gamma$ specify a linear transformation of the voxel values to match the mean and variance of the templates in $\mathcal{A} \cup \mathcal{P}$. Note that the mean and variance parameters do depend on voxel location $s$. The $w_s^A$ and $w_s^P$ are weights based on the distance of voxel $s$ from each commissure, and we take the weights as $1/\{1 + (c_x - s_x)^2 + (c_y - s_y)^2 + (c_z - s_z)^2\}$, where $(c_x, c_y, c_z)$ is the location of a commissure. The additional parameters $\beta$ and $\gamma$ are to be estimated based on the pixel values in $\mathcal{A} \cup \mathcal{P}$ at each possible

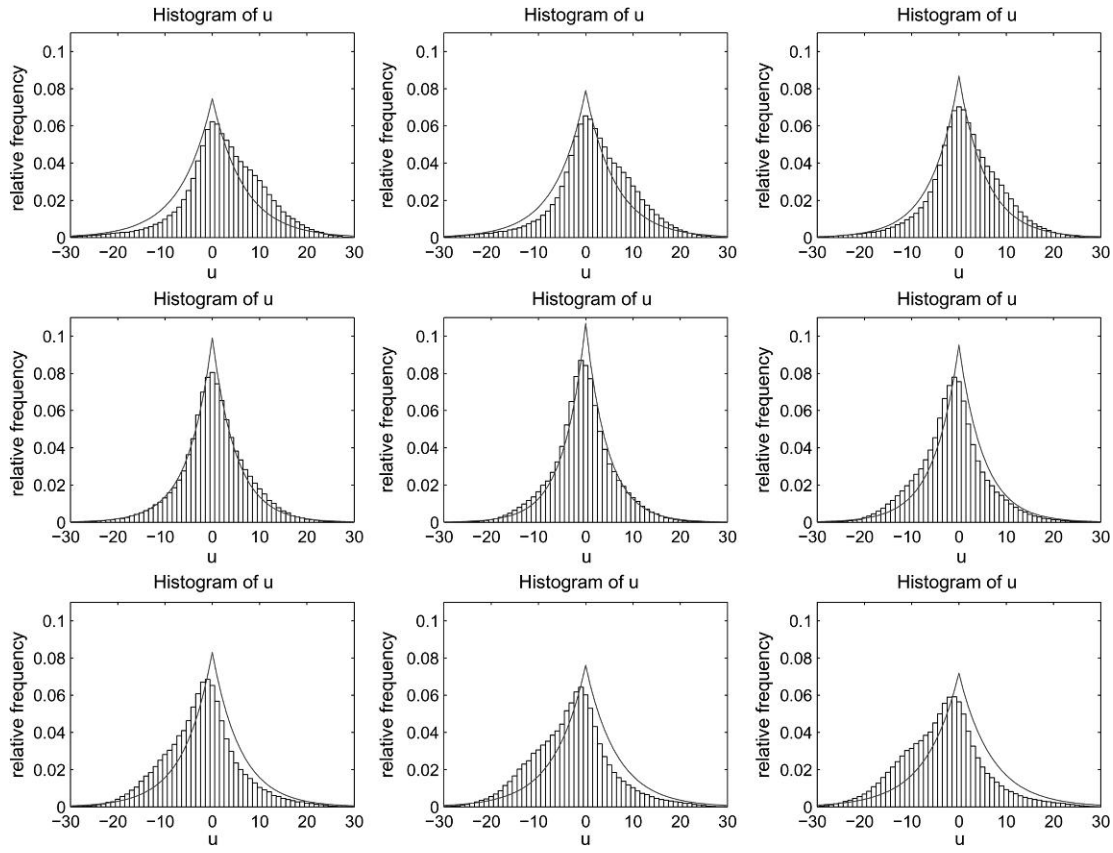


Fig. 1. Histograms of $u_s$ for several choices of midline and fitted Laplace distributions. From left to right, the top row shows the histogram corresponding to $\xi_z$ equal to $-2, -1, 0$; the middle row shows histograms corresponding to $\xi_z$ equal to $1, 2, 3$; and the bottom shows the histograms for $\xi_z$ equal to $4, 5, 6$. The middle plot with $\xi_z = 2$ appears to be the best choice here, which is confirmed by MLE.

registration. Recall that the location of the AC and PC landmarks also requires an additional parameter $\xi_c$ which is the length between the landmarks. Finally, we simply take the likelihood contribution from the remaining voxels in $\mathcal{O}$ to be constant with respect to the model parameters.

Hence, the log-likelihood for the nonzero voxels in an individual scan is taken to be

$$\begin{aligned}\log L(U, V|\phi, \psi, \xi_c) = &\sum_{s\in\mathcal{M}} \left\{ \log\left(\frac{\psi w_s}{2}\right) - \psi w_s |u_s| \right\} \\ &+ \sum_{s\in\mathcal{A}} \left\{ \log\beta - \frac{w_s^{\mathrm{A}}}{2(\sigma_s^{\mathrm{A}})^2}(\beta v_s + \gamma - \mu_s^{\mathrm{A}})^2 \right\} \\ &+ \sum_{s\in\mathcal{P}} \left\{ \log\beta - \frac{w_s^{\mathrm{P}}}{2(\sigma_s^{\mathrm{P}})^2}(\beta v_s + \gamma - \mu_s^{\mathrm{P}})^2 \right\} + \text{constant}.\end{aligned}$$

Note that the mean parameters $\mu_s^{\mathrm{A}}$, $\mu_s^{\mathrm{P}}$ and variance parameters $(\sigma_s^{\mathrm{A}})^2$, $(\sigma_s^{\mathrm{P}})^2$ in the AC and PC templates are regarded as known (having been estimated from training data).

The profile MLE of $\psi^{-1}$ given the other parameters is obtained by solving

$$\frac{\partial \log L}{\partial \psi} = 0 = \frac{n_{\mathrm{M}}}{\psi} - \sum_{s \in \mathcal{M}: s_z > 0} w_s |u_s|, \Rightarrow \hat{\psi}^{-1} = \frac{1}{n_{\mathrm{M}}} \sum_{s \in \mathcal{M}: s_z > 0} w_s |u_s|$$

and $n_{\mathrm{M}}$ is the number of voxels in the midline. In order to maximize over the other parameters we consider 4 stages.

MLE Algorithm

1. The midline plane is estimated with a simple grid search at steps of 0.01 radians for $\theta_{\mathrm{p}}, \theta_{\mathrm{r}}$ and unit steps for $\xi_z$.
2. Approximate estimates of $\xi_x$, $\xi_y$, $\theta_y$, $\xi_c$, $\beta$ and $\gamma$ are found by maximizing the likelihood of the region $\mathcal{A} \cup \mathcal{P}$ given the midline plane estimates in step 1.
3. The estimates of $\xi_x$ and $\xi_y$ are refined by maximizing the likelihood of the region $\mathcal{A}$.
4. The estimates of $\theta_y$, a rotation about the AC obtained in step 3, and $\xi_c$, the intercommissure distance, are finalized by maximizing the likelihood of the region $\mathcal{P}$.

In each case, a simple grid search over the parameters is performed, with angular steps of 0.01 radians and unit translations. Although we could consider iterating the steps in the algorithm, it made no difference when we tried it in some example images.

EXAMPLE We evaluate the log-likelihood for an example image using a grid of values with rotations 0.01 radians apart and translations at 1 mm intervals. We find that the approximate maximum likelihood estimators at step 1 of the algorithm are $\hat{\theta}_{\mathrm{p}} = -0.07$, $\hat{\theta}_{\mathrm{r}} = -0.06$, and $\hat{\xi}_{\mathrm{z}} = 2.0$, with $\hat{\psi}^{-1} = 3.3613$. In the upper row of Figure 2, we see the image transformed from its original orientation to the MLE registration of the midline. Note that after the transformation, the crosshairs bisect the brain's 2 hemispheres. In Figure 1, we see histograms of the voxels in $\mathcal{M}$ and the fitted Laplacian density for different choices of $\xi_z$ with $\xi_z = 2.0$ in the middle plot. Proceeding to register the image on the AC and PC, we find the approximate discretized maximum likelihood estimates are $\hat{\xi}_x = -4.0$, $\hat{\xi}_y = 41.0$, $\hat{\theta}_y = 0.29$, and $\hat{\xi}_c = 26$. In the lower row of Figure 2, we see the image translated and rotated from its midline registration such that the origin coincides with the AC and the AC–PC line is horizontal.

### 4.2 *Bayesian registration*

An alternative approach involves Bayesian inference for the registration parameters. The priors for the parameters $\phi$, $\xi_c$, $\beta$ and $\gamma$ are taken as independent and uniform as we have no prior knowledge. We also take $\psi \sim \Gamma(\alpha_0, \beta_0)$ independently and choose $\alpha_0$ and $\beta_0$ to give a fairly vague prior. The log-posterior density is given by

$$\log \pi(\phi, \psi, \xi_c, \beta, \gamma | U, V) = \log L(U, V | \phi, \psi, \xi_c, \beta, \gamma) + (\alpha_0 - 1) \log \psi - \psi \beta_0 + \text{constant}.$$

One can simulate from the posterior using an MCMC algorithm. In particular, we use Metropolis–Hastings updates for the registration parameters, and a Gibbs step for $\psi$ since

$$(\psi | \phi, \xi_c, \beta, \gamma, U, V) \sim \Gamma\left(n_{\mathrm{M}} + \alpha_0, \sum_{s \in \mathcal{M}: s_z > 0} w_s |u_s| + \beta_0\right).$$

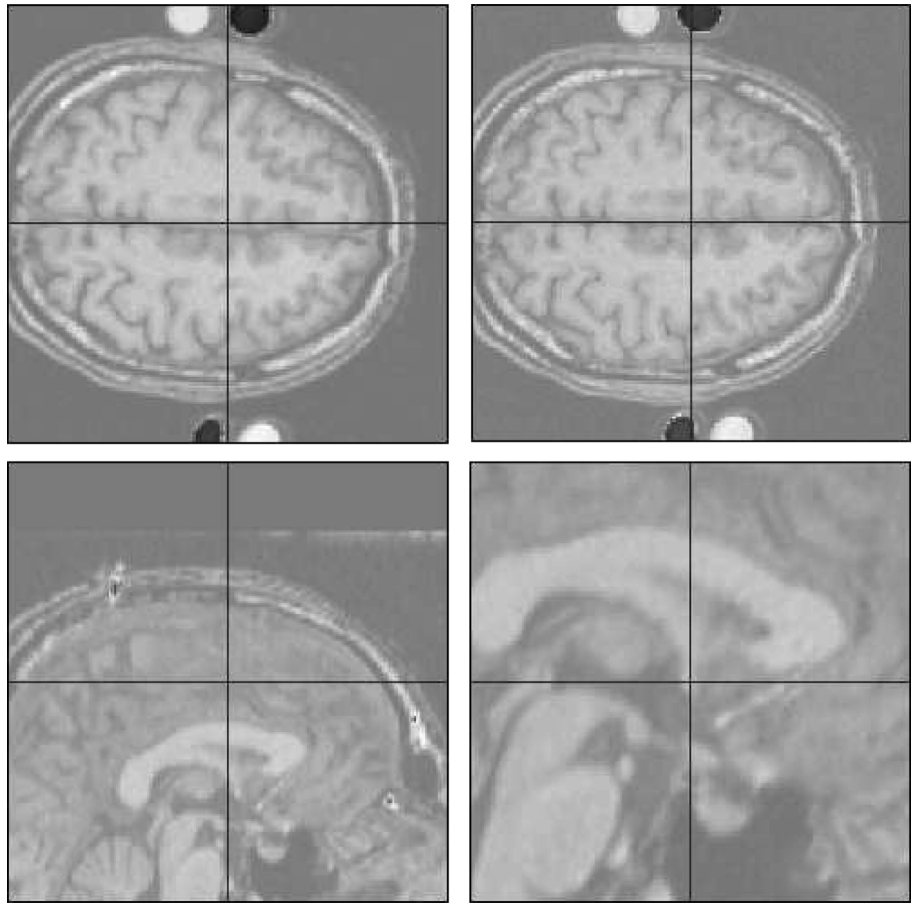

Fig. 2. An image transformed from its original orientation (top left) to the MLE registration of the midline (top right), using axial views. The image in the coronal view at the midline (bottom left) is then translated and rotated to have the crosshair on the AC, and the AC–PC line horizontal in the final registration (zoomed in—bottom right).

We took the conditional modes of $\beta$ and $\gamma$ at each iteration, and our prior hyperparameters were $\alpha_0 = 0.1$ and $\beta_0 = 1/3$, which results in a fairly vague prior.

EXAMPLE We implement the MCMC algorithm for the same image as for the MLE example. Figure 3 shows the parameters after 15 000 iterations. The starting value was taken as the approximate maximum likelihood estimate, as obtained in the previous example. The "maximum *a posteriori*" (MAP) estimate is $\hat{\xi}_x = -3.9, \hat{\xi}_y = 40.8, \hat{\xi}_z = 2.0, \hat{\theta}_{\rm p} = -0.070, \hat{\theta}_{\rm r} = -0.057, \hat{\theta}_y = 0.289, \hat{\xi}_c = 25.79$ and $\hat{\psi}^{-1} = 2.8868$. Over the first 2000 iterations, we use an adapting stage (see Browne and Draper, 2000) to choose the variances for the sampling distributions, and we take the next 2000 iterations as the burn-in period. The adapting stage involves changes in the proposal variances with the aim of achieving 50% acceptance for each parameter. The proposal variances are fixed once the acceptance rate remains in the range 40–60%. Similar results after burn-in were obtained using MCMC simulations with starting values away from the maximum likelihood estimate.

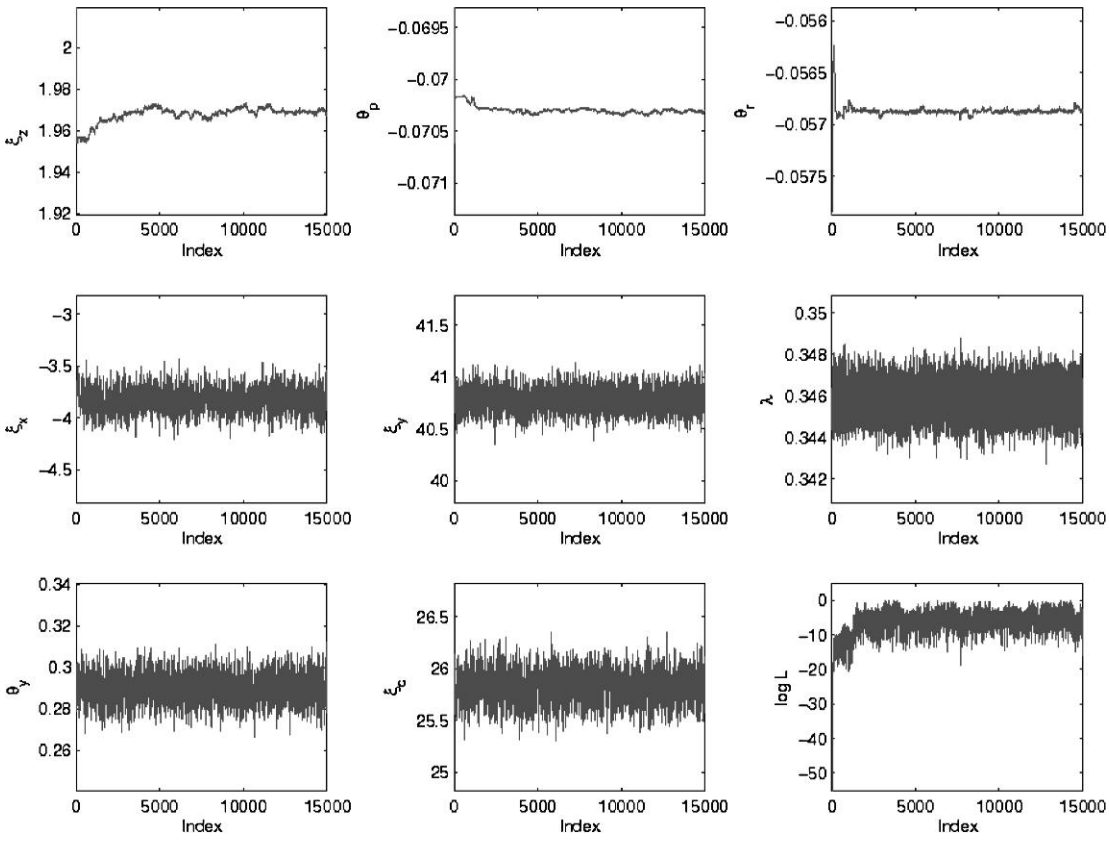


Fig. 3. Plots of parameter values and the log-likelihood from the MCMC algorithm over the first 15 000 iterations. There are 2000 iterations for the adapting stage, 2000 iterations for burn-in, and 11 000 iterations post burn-in.

As expected, the approximate MLE and MCMC-based estimates are similar, and the posterior variability is small. We observed similar close agreements between the MLE and MCMC-based estimates in many further images, with small posterior variability. For example, in Table 1, we give the MAP and approximate MLEs for a set of 7 further example images and there is close agreement. Hence, we shall proceed with shape analysis of the brain surfaces by fixing the registration at the estimated registration given by the approximate MLE, as the MLE is much faster to compute.

We compared the automatically located AC and PC from the MLE procedure with manually located landmarks in all the images. The landmarks using both methods were in similar positions (exactly the same position in 30 scans and very small departures in the rest).

## 5. SHAPE ANALYSIS

### 5.1 *Labeling*

Each brain is registered to Talairach space using the maximum likelihood method given in Section 4.1, and we now consider the labeling of each brain. The actual cortical surface of the brain is complicated and highly variable between individuals, and so estimating a natural labeling of the actual cortical surface is difficult. Our approach using correspondence of the BET image surface would not allow detailed shape

Table 1. *The MAP estimates from the MCMC algorithm and the approximate MLEs for an example set of 7 brain images. In the ML case, the $\xi$ measurements are obtained to the nearest millimeter except for $\xi_c$ which is to the nearest* 0.5 *mm. There is close agreement between the estimates*

| Image | Method | $\xi_x$ | $\xi_y$ | $\xi_z$ | $\theta_p$ | $\theta_r$ | $\theta_y$ | $\xi_c$ |
|---|---|---|---|---|---|---|---|---|
| 1 | MAP | −11.72 | 34.26 | 7.58 | −0.05 | −0.08 | 0.25 | 28.82 |
| 1 | MLE | −12 | 34 | 8 | −0.06 | −0.08 | 0.25 | 29 |
| 2 | MAP | −15.06 | 41.38 | 3.2 | −0.01 | −0.01 | 0.30 | 26.95 |
| 2 | MLE | −15 | 41 | 3 | −0.01 | −0.01 | 0.32 | 27 |
| 3 | MAP | −12.07 | 39.05 | −1.97 | 0.01 | 0.01 | 0.23 | 27.59 |
| 3 | MLE | −12 | 39 | −2 | 0.01 | 0.02 | 0.24 | 27.5 |
| 4 | MAP | −12.42 | 35.4 | −4.26 | 0.02 | −0.07 | 0.26 | 25.1 |
| 4 | MLE | −12 | 35 | −4 | 0.02 | −0.07 | 0.25 | 24.5 |
| 5 | MAP | −13.75 | 34.95 | −0.68 | 0.01 | 0.01 | 0.16 | 27 |
| 5 | MLE | −14 | 35 | −1 | 0.02 | 0.00 | 0.16 | 27 |
| 6 | MAP | −7.78 | 30.15 | 3.74 | −0.07 | 0.00 | 0.07 | 25.51 |
| 6 | MLE | −7 | 29 | 4 | −0.07 | 0.00 | 0.11 | 25 |
| 7 | MAP | −11.72 | 40.77 | 2.28 | −0.03 | −0.03 | 0.22 | 27.69 |
| 7 | MLE | −12 | 41 | 2 | −0.03 | −0.03 | 0.23 | 28.5 |

analysis of small regions of the surface (since corresponding parts would not match precisely) but it does allow the comparison of global large-scale shapes.

We have located $k$ radii lengths in a group of $n$ people, where the lengths are denoted by $R_{ij}$, $j = 1, \ldots, n$, for person $j$ at equally spaced angles $t_i, i = 1, \ldots, k$, emanating from a central point (AC–PC midpoint). We decided to restrict our analysis to the part of the brain surface lying above the axial plane in which the AC–PC lies, as the labeling in this part of the surface gives a reasonable meaningful correspondence between parts of the brain surface. The domain $\mathcal{D}$ in our application is therefore the upper hemisphere $S^2_+$. The number of radii taken on the upper hemisphere is $k = 62\,501$. Note that this labeling gives a sensible approximate correspondence.

Alternatively more complicated transformations could be used to match up parts of a surface if desired, for example, Fischl *and others* (2001) consider an ontogenetically informed deformation method involving matching homologous structures on a flattened gray–white matter interface. For the broad large-scale shape analysis that we consider our approach seems reasonable.

### 5.2 *Brain surface analysis*

Since $R_{ij}$ are positive lengths, it makes sense to consider the geometric mean of the quantities, calculated by averaging the logarithms of the radii and then taking the exponential. Hence, we carry out statistical analysis on the log-radii $x_{ij} = \log R_{ij}$, where $i = 1, \ldots, k$, $j = 1, \ldots, n$ which are radii observed at fixed directions on the unit upper hemisphere in 3 dimensions $t_i \in S^2_+, i = 1, \ldots, k$. An estimate of the mean form is the geometric mean of each of the radii

$$\{\exp(\bar{x}_i)t_i : i = 1, \ldots, k\},$$

where $\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} \log R_{ij}$. We could alternatively have used the actual radii rather than taking logarithms (and in fact it makes little practical difference in this particular application as the mean is large relative to the standard deviation). Both the radii and the logarithms look reasonably normal at each point.

To investigate group differences first of all we provide an explanatory plot in Figure 4(a) of the mean control brain and highlight the few areas where the mean patient brain radii are more than 2.5% larger

Fig. 4. The mean brain surface for the control group (a) (left) and patient group (b) (right) are shown in light gray. In (a) we indicate in dark gray which radii on the surface are at least 2.5% larger for the mean patient and in (b) we show in gray which radii are at least 2.5% larger for the mean control brain surfaces. The dark gray radii have *t*-statistic at least 3, and this area is significantly smaller in the patient group.

than the mean control. These areas are primarily on the top of the brain around the midline. We also provide a plot in Figure 4(b) of the patient mean brain surface and plot the areas where the control mean is more than 2.5% larger than the mean patient. These areas are on the right temporal lobes. We now wish to examine whether these features are statistically significant.

We carry out an analysis of covariance for each log-radius, with 2 factors (patient/control and sex) and 1 covariate (age). For each log-radius, we fit a linear model with 4 parameters, and we are primarily interested in the patient/control effect. Under the null hypothesis ($H_0$) that there is no difference in mean between the patients and controls, we consider the usual *t*-test on the patient/control parameter after fitting the linear model for each log-radius. A positive *t*-statistic is obtained when the mean control is larger at a location. We have $n_p = 30$ and $n_c = 38$ for the schizophrenia and control sample sizes.

Since we have a very large number of tests ($k = 62\,501$), we have to take into account the multiple testing problem. Under $H_0$, we assume that we have a smooth stationary, isotropic Student's *t* distributed random field on a hemisphere, and so we can use the results of Worsley (1994) in order to provide a corrected *p*-value for the test. There are $n_p + n_c - 4 = 64$ degrees of freedom. We need to estimate the roughness $\lambda$, which is the variance of the derivative of the random field in any direction. We fit a local polynomial spline (Wand and Jones, 1995) to great circles of neighboring *t*-statistics around the hemisphere at a particular distance apart $\mathrm{d}\phi$. By choosing the bandwidth parameter by visual inspection, we then compute the sample variance of the estimated first derivatives obtained from the KernSmooth library in R (Wand and Ripley, 2009; R Development Core Team, 2008). We arrive at a final estimate $\hat{\lambda} \approx 3$ using local polynomials of degree 1 and bandwidth $10\,\mathrm{d}\phi$.

We consider the *p*-value based on the maximum $T_{\max}$ of the random field and a threshold $z$. The corrected *p*-value for a hemisphere is given by

$$P(T_{\max} > z) \approx \rho_0(z) + \pi\rho_1(z) + 2\pi\rho_2(z),$$

where

$$\rho_0(z) = P(t_\nu > z), \tag{1}$$

$$\rho_1(z) = \frac{\lambda^{1/2}}{2\pi}\left(1+\frac{z^2}{\nu}\right)^{(1-\nu)/2}, \tag{2}$$

$$\rho_2(z) = \frac{\lambda}{(2\pi)^{3/2}}\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\left(\frac{\nu}{2}\right)^{1/2}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{z^2}{\nu}\right)^{(1-\nu)/2} z \tag{3}$$

(see Cao and Worsley, 2001). For a threshold of $z = 3$, we have a corrected $p$-value of 0.065 for a one-sided test with $\lambda = 3$. The most extreme $t$-statistic in the data is 3.28 and there is a cluster of $t$-statistics more than 3 in the right temporal lobe (in the dark gray cluster in Figure 4(b)), where the schizophrenia patients' brains are smaller. It is rather difficult to estimate the roughness parameter, and so if $\lambda$ was larger the evidence would be weaker (although the opposite is true if $\lambda$ is smaller). In conclusion, we may have some weak evidence for a mean reduction in the right temporal lobe for the schizophrenia group.

### 5.3 *Dimension reduction*

Due to the very high-dimensional nature of the data, it is also worthwhile considering dimension reduction techniques. We consider PCA in order to examine important lower-dimensional projections of the data. The procedure involves obtaining the eigenvalues and eigenvectors of the sample covariance matrix of the $x_{ij}$'s. Let $\hat{\gamma}_1, \ldots, \hat{\gamma}_M$ be the eigenvectors corresponding to eigenvalues $\hat{\lambda}_1, \ldots, \hat{\lambda}_M$, where $M = \min(k, n-1)$ and write $\hat{\gamma}_{il}$ for the $i$th element of $\hat{\gamma}_l$, $l = 1, \ldots, M$. If $n << k$ these high-dimensional eigenvectors can be computed in $O(n^3)$ steps (see e.g. Dryden, 2005).

The effect of the the $l$th principal component (PC$l$) can be examined by viewing

$$\{\exp(\bar{x}_i \pm c\hat{\lambda}_l^{1/2}\hat{\gamma}_{il})t_i : i = 1, \ldots, k\}$$

for various values of $c$. Here PCA has been carried out on the log-radii $x_{ij}$ and then the exponential function is used to bring the vector back to the scale of the original surface. From now on we use the abbreviation PC to mean 'principal component'.

Investigating the variability in size and shape, we see that PC1 is highly significantly different between the sexes, and given that this measures overall size this is expected (see Figure 5). PC2 is also displayed in Figure 5, and we see that it shows the effect of larger frontal and temporal lobes versus smaller rear brain surface. PC1 and PC2 account for 53.6% and 10.1% of the variability in size and shape in the data.

The first 20 PCs account for 88.8% of the size-and-shape variability, and we investigate group differences in this linear subspace. We carry out a two sample $t$-test on each of the first 20 PC scores, and then address the fact that we have multiple tests using the false discovery rate (FDR) of Benjamini and Hochberg (1995). The smallest $p$-values are 0.015 for PC11 and PC19 but these are not significant at FDR = 0.1. We also carry out symmetrical PC analysis, where each residual vector from the overall mean has its reflection also included in an augmented data set (see Theobald *and others*, 2004 ). This procedure forces the PCs to either be symmetrical or asymmetrical. The first 20 PCs explain 85.9% of the size-and-shape variability, and again we restrict our analysis to this linear subspace. We see that PC10 is highly significantly different between patients and controls, with $p$-value 0.001, which is also significant at FDR = 0.03, taking account the multiple tests. This PC is a symmetrical PC that picks up the size of the temporal lobes as well as other effects including the height of the brain surface. The effect of this PC is shown in Figure 5. PCs 1–9 and 11–20 do not show significant differences in the groups, taking into account multiple comparisons (with smallest unadjusted $p$-value 0.046 for PC11).

Independent components analysis (ICA) (Hyvärinen *and others*, 2001) is an alternative dimension reduction technique that can sometimes reveal differences between groups by projecting data into directions of maximum non-Gaussianity. We also carry out ICA of the surfaces after first reducing to 10 PCs, using

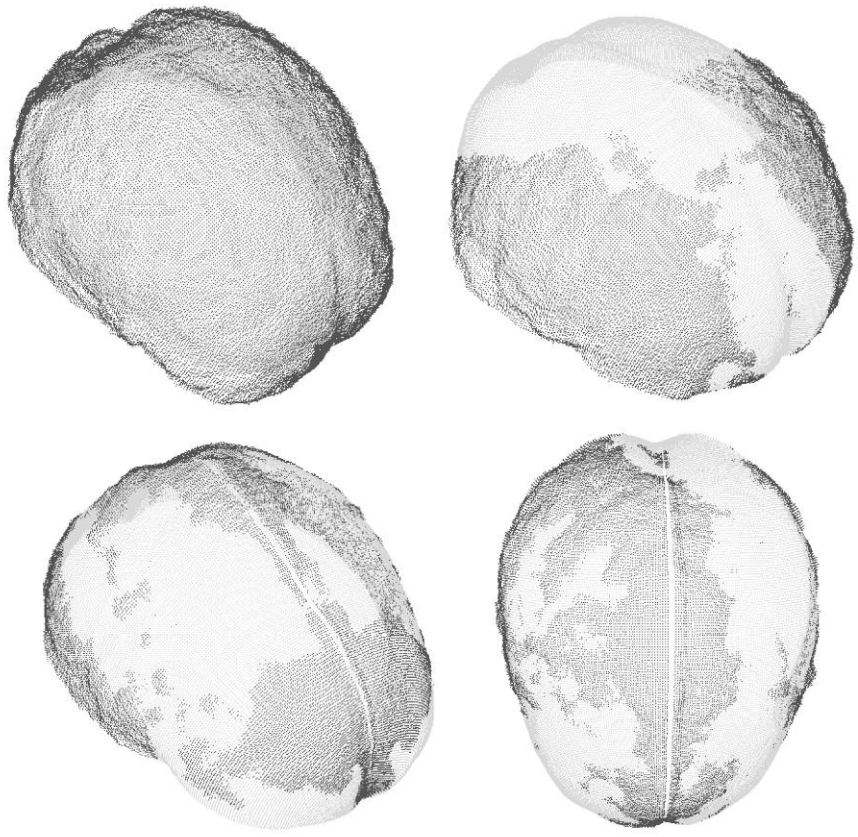

Fig. 5. The effects of the PCs. We display the overall mean brain surface in light gray, and in dark gray we display $\exp(\bar{x} + 6\hat{\lambda}_j^{1/2}\hat{\gamma}_j)$ for all values with radius greater than the mean. The choice of 6 is made to exaggerate the effect of each component. The plots are (top left) PC1, (top right) PC2, (bottom left) symmetrical PC10, (bottom right) an IC.

symmetrical PCA again. We use the fast ICA implementation in the package R (Marchini *and others*, 2003; R Development Core Team, 2008). We find that one of the independent component (IC) scores is significantly different in the 2 groups. The effect of this IC is shown in Figure 5 and it includes asymmetry in the temporal lobe, as well as other effects such as the height of the brain surface. In particular, the schizophrenia patients have slightly higher brain surfaces around the midline on average.

A major disadvantage of the PC and ICA is that each PC contains a number of effects, and so it is difficult to disentangle which effects are significant and which effects are noise. Given the small differences, one most be cautious about the practical importance of the differences despite the statistical significance. There is considerable overlap between the schizophrenia and control distributions, so these measures alone would be poor discriminators but may be helpful as part of a classification method using data from different modalities.

### 5.4 *Asymmetry and torque*

The final part of the investigation into shape is a study of asymmetry in the brain surface. In each registered scan $n_S = 100$ equally spaced axial slices are taken, and we estimate the volumes contained within brain

surface boundaries in each slice and above the horizontal plane containing the commissures for the the left-hand side ($V_{rj}^{\mathrm{L}}$) and right-hand side ($V_{rj}^{\mathrm{R}}$), $r = 1, \ldots, n_S$ of the $j$th scan ($j = 1, \ldots, n$). The asymmetry function has components

$$\eta_{rj} = (V_{rj}^{\mathrm{R}} - V_{rj}^{\mathrm{L}})/V_j, \quad r = 1, \ldots, n_S,$$

where $V_j$ is the maximum slice volume in the $j$th scan, $-1 < \eta_{rj} < 1$. Write $\eta_j = \left(\eta_{1j}, \ldots, \eta_{n_S j}\right)^T$ for the asymmetry function for the $j$th person. We perform a small amount of smoothing with a Loess smoother (with fraction $f = 0.05$). Smoothing is commonly carried out at a preliminary stage in functional data analysis (see Ramsay and Silverman, 2005). Note that similar measures of asymmetry were described in outline by Chance *and others* (1999) and in more detail by Barrick *and others* (2005) based on the difference in volume in the 2 hemispheres in a series of slices but without normalizing by the maximum slice volume in the $j$th scan.

In Figure 6, we see a plot of the mean smoothed asymmetry function and the loadings of the first 3 PCs from the pooled sample of $n = 68$ smoothed asymmetry functions. It is clear in the mean asymmetry function that there is rightward asymmetry toward the frontal region of the brain. From the plots of the PC loadings, it seems clear from PC1 that the main source in variability is in the occipital region. PC2 shows a gradual increase in variability nearer the front. PC3, however, highlights a more general twisting in the brain and will best detect regions where the control group is more asymmetric. We shall focus further on this PC below.

In Figure 7, we see the results of conducting a $t$-test of $H_0$: $\mu_{\mathrm{c}} - \mu_{\mathrm{p}} = 0$ at each slice, where $\mu_{\mathrm{c}}$ and $\mu_{\mathrm{p}}$ are the means of the control and patient group. Taking into account multiple comparisons, these results are not statistically significant, but it is worth noting that the controls have greater rightward asymmetry between slices 83 and 87 in the sample.

We consider fitting a linear regression model with response PC score 3. We see that there are statistically significant differences in PC score 3 between patients and controls ($p$-value $= 0.049$) but there is no significant association with age or sex. The result provides some weak evidence that controls are more likely to display torque, with greater rightward asymmetry than patients in the frontal lobe. This effect reverses in the rear half of the brain. The number of females in this study is particularly small, so particular care should be taken with interpreting that twisting is not associated with sex. We also consider ICA of the asymmetry functions, after first reducing to 10 PCs. None of these ICs was significantly different between the 2 groups.

### 5.5 *Midline plane and curve analysis*

In all our analysis so far we have used a flat midline plane to register the brain images and to calculate the symmetry functions. However, it is somewhat of a simplification to assume that the join of the left and right brain hemispheres is the flat plane $\xi_z = 0$. In reality, inspection of the scans shows a tendency for the midline to curve, especially at anterior and posterior extremities and, to a lesser extent, in superior regions. We therefore also wish to investigate what difference a curved midline would make to our interpretations about asymmetry. In particular, we estimate a curved midline plane separating to the 2 hemispheres and adjust the left and right slice volumes for this correction in the midline location.

To estimate the location of the curved midline, we first of all register each brain with a flat midline using MLE, as in Section 4.1. We fix all parameters at the MLE except $\xi_z$ which is allowed to vary throughout the the midline region $\mathcal{M}$. We consider localized regions in $\mathcal{M}$ and estimate a possible different $\xi_z$ in each region. The size of the localized region, $\mathcal{L}$, was chosen large enough to avoid detecting local symmetries not centered on the join and reflect general movement in the join, and we fix $\epsilon_{\mathrm{M}} = 15$ mm. Each $\mathcal{L}$ was centered on locations at 5 mm intervals in the $x$-$y$ plane and $\hat{\xi}_z$ recorded, after a discrete grid search in unit steps, as the displacement from $\xi_z = 0$ at that location.

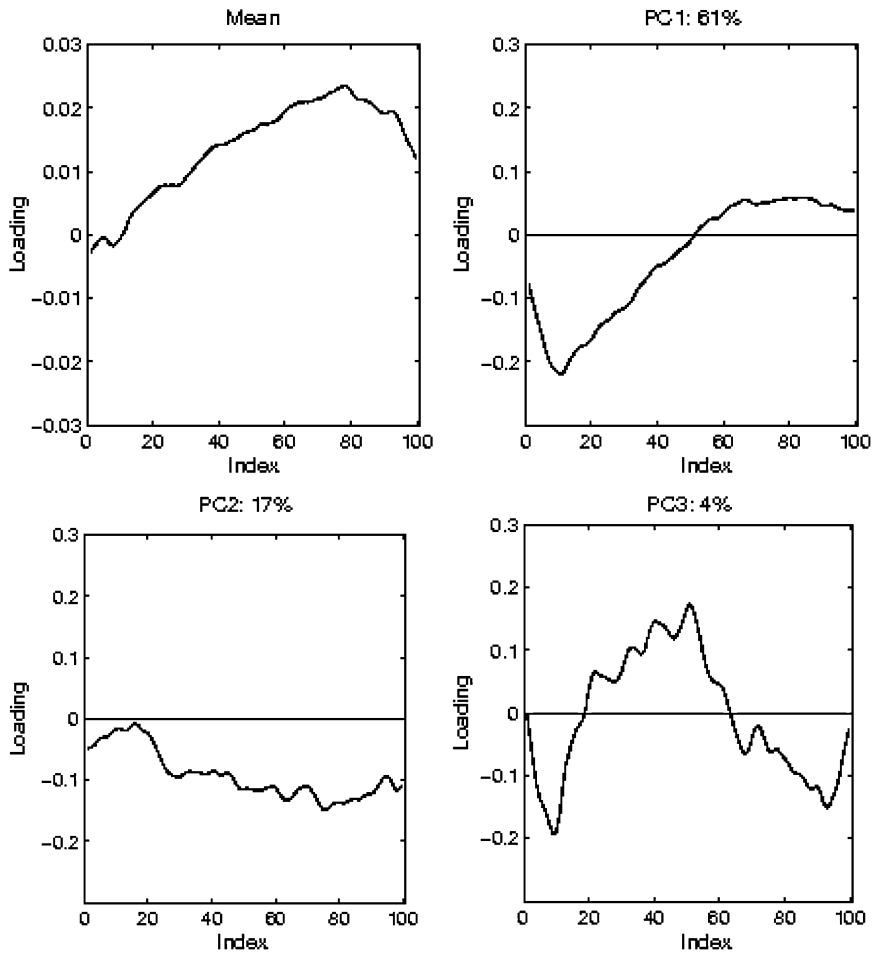


Fig. 6. The mean (top left) and loadings for PCs 1–3 for the smoothed asymmetry functions. A low index (left end of each picture) corresponds to slices in the occipital region of the brain and a high index (right end of each picture) corresponds to the frontal region of the brain.

The maximum likelihood estimates of $\xi_z$ at each location were analyzed using $t$-tests of $H_0$: $\mu_c - \mu_p = 0$, where $\mu_c$ and $\mu_p$ are the mean leftward displacements of the control and patient groups, respectively. In the region between 25 and 50 mm posterior of the AC and extending up to 60 mm above the axis containing the commissures the midline is further to the right in the patient group, as seen in Figure 8. Note that the patient curved midline is more to the right in the patients compared to the controls, between the temporal lobes. This feature might explain the reduced rightward asymmetry seen in controls, compared to patients, seen between slices 34 to 37 (approximately) in Figure 7. A further difference between the 2 groups is observed in the occipital region.

To incorporate a curved midline in the symmetry analysis, a curved midplane is fitted to the midplane coordinates by means of a thin-plate smoothing spline for each scan (see Dryden and Mardia, 1998, Chapter 10). The thin-plate smoothing spline is fitted to the estimated midline points $\xi_z$ in each of the local regions. The fitted thin-plate spline for one of the scans is shown in Figure 9. It clearly demonstrates that a curved midline fits the interhemispherical join better than a straight midline.

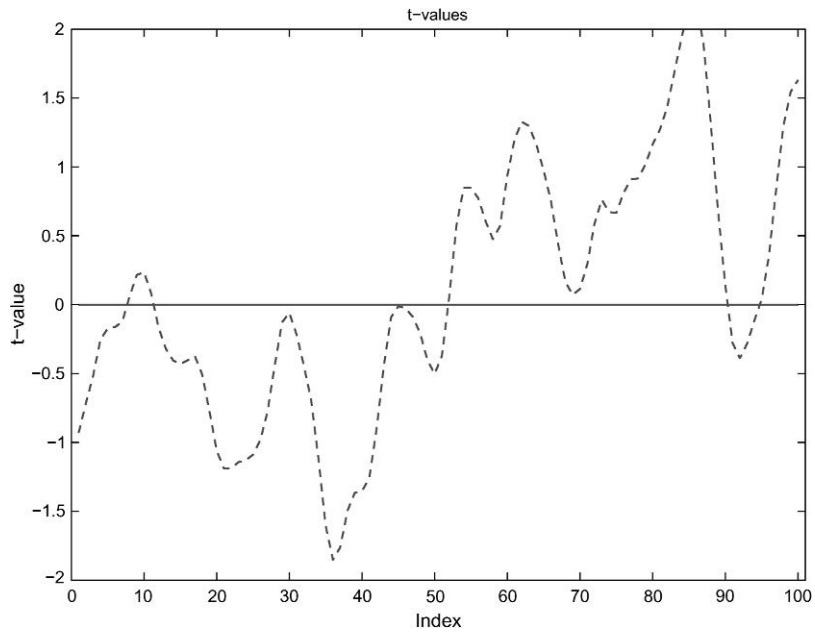


Fig. 7. The *t*-statistics from *t*-tests between control and patient groups at each slice. High *t*-values indicates greater rightward asymmetry in the control group.

Repeating the symmetry analysis with a curved midline showed no significant differences between the 4 subgroups (male/female and control/patient combinations). This suggests that there is possibly another explanation for differences between the control and patients. Either the differences in asymmetry are due to a difference about a flat midline (as found in Section 5.4) or the differences could be due to differences in the curvature of the midline plane. As is always the case in shape analysis, the interpretation of a difference depends on the choice of registration, and this is demonstrated well in this example. We can choose a simple registration method and more variability is retained in the shapes or we can choose a more sophisticated registration method which reduced the variability in the shapes. The choice to be made is very much up to the user.

The simpler planar midline registration is the more conventional method and with its use we have conferred with findings in the literature. The curved midline does provide a more highly parameterized and unusual method of registration, and it is worth further exploration in future studies.

## 6. Discussion

We have explored a number of aspects of the shape analysis of brain surfaces. Our main findings have been that there is weak evidence for a reduction in size in the right temporal lobe on average in the patients. A reduction in the temporal lobes has been noted in the literature previously (see e.g. Lawrie *and others*, 2002), and so our study provides further evidence for this effect. Dimension reduction analysis did produce some significant differences including a slightly higher brain surface in the patient group. However, it is rather difficult to disentangle multiple effects in the components, and so we are cautious about interpreting this difference. This difference was not deemed significant in the random field method.

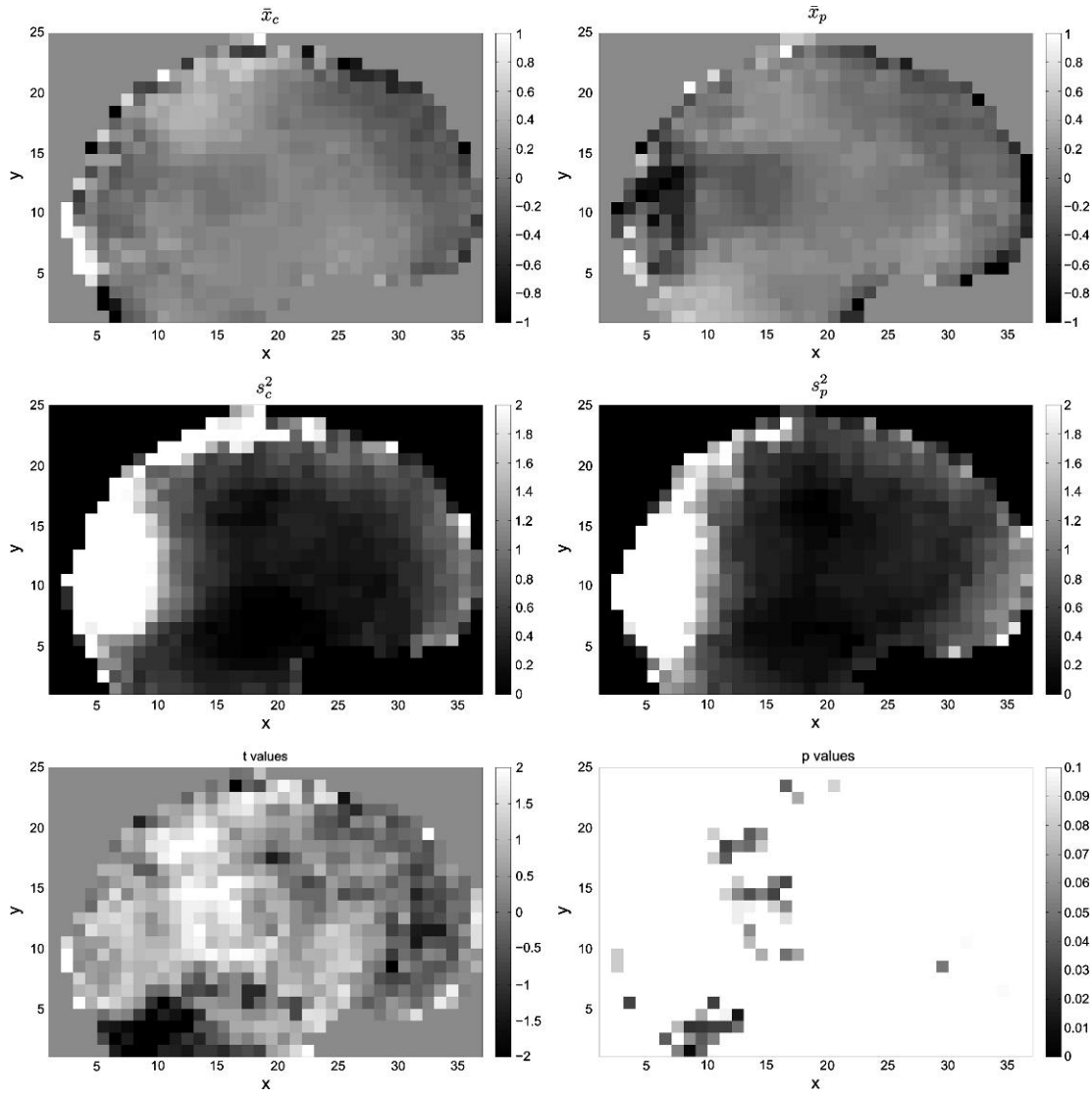


Fig. 8. Top row: The mean of the interhemispherical join's displacement from the plane $\xi_z = 0$ at 5 mm intervals in the x–y plane for the control group (left) and the patient group (right), with darker areas indicating a displacement to the "right." Middle row: The variance of displacement at each location for the control group (left) and patient group (right), with darker areas indicating low variance. Bottom left: *t*-values for the difference in the 2 groups, with darker areas indicating the control group displaced further to the right than the patient group. Bottom right: *p*-values for each *t*-test thresholded at $p = 0.1$. Darker areas indicate higher significance.

We have found significant differences in average brain torque: the feature where the right frontal region is larger than the left and the left occipital region is larger than the right. It has been observed in the literature that brain torque differs with handedness and gender (Kertesz *and others*, 1990), with females tending to exhibit less torque on average. It has been suggested that schizophrenia patients also tend to

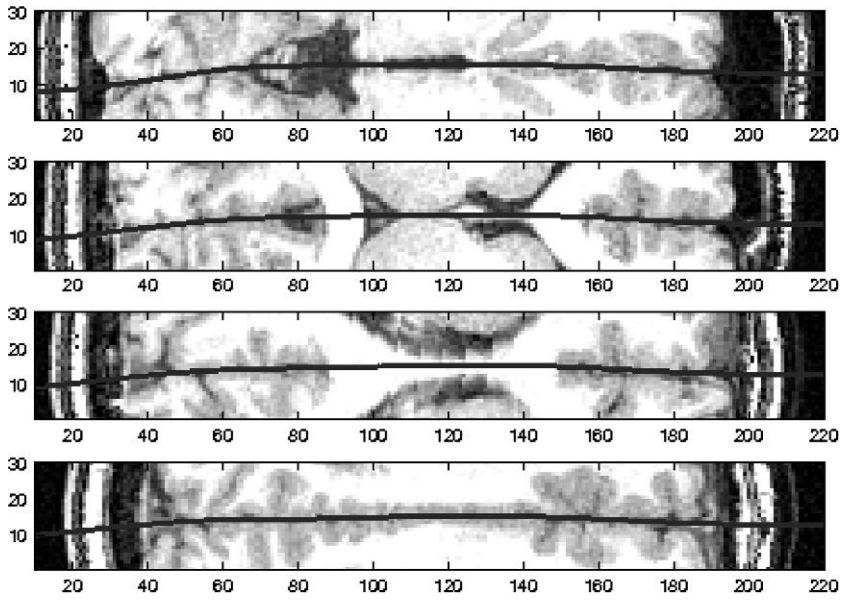


Fig. 9. The interhemispherical join for an example brain image and the fitted curved midline (dark gray line) on axial slices at $y = 0, 10, 20, 30$ mm above the AC–PC line. Clear curvature of the fitted midline can be seen.

have less torque (Bilder *and others*, 1994; Barrick *and others*, 2005). Our own findings have confirmed small differences between the schizophrenia patients and the controls, with the schizophrenia patients having less torque. Our samples were dominated by male right-handers, and so we could not investigate handedness. We did not pick up any significant asymmetry differences with gender, although care should be taken with this interpretation as the number of females in our study was small.

An alternative and very popular technique for comparing brain images is VBM in SPM2 (Ashburner and Friston, 2000). The optimized VBM protocol (Good *and others*, 2001) has been carried out on our data set and details of the analysis are given in Brignell (2007). The VBM procedure is very different from our own in that it works on the gray matter intensities after nonlinear registration to an average template brain. A significant difference was observed in a small part of the left superior temporal lobe, with the patients having less concentrated gray matter (or less volume) than the controls.

It is noteworthy that small significant differences have been obtained with several different methods in our study: the random field method, the surface shape PCA and ICA methods, the curved midline analysis and VBM. All analyses are complementary, being consistent with small differences in the temporal lobes, but the interpretation of the differences is dependent on the method of registration. The application illustrates that the choice of registration is fundamental in interpreting size and shape differences. This feature is an unavoidable but sometimes overlooked aspect of shape analysis.

## Acknowledgments

We are grateful to 2 anonymous referees for their comments which were very helpful in improving the paper. *Conflict of Interest:* None declared.

## Funding

Engineering and Physical Sciences Research Council; Leverhulme Trust.

## REFERENCES

ASHBURNER, J. AND FRISTON, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage* **11**, 805–821.

BARRICK, T. R., MACKAY, C. E., PRIMA, S., MAES, F., VANDERMEULEN, D., CROW, T. J. AND ROBERTS, N. (2005). Automatic analysis of cerebral asymmetry: an exploratory study of the relationship between brain torque and planum temporale asymmetry. *Neuroimage* **24**, 678–691.

BARRY, S. J. E. AND BOWMAN, A. W. (2008). Linear mixed models for longitudinal shape data with applications to facial modeling. *Biostatistics* **9**, 555–565.

BENJAMINI, Y. AND HOCHBERG, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B* **57**, 289–300.

BESAG, J. E. (1986). On the statistical analysis of dirty pictures (with discussion). *Journal of the Royal Statistical Society, Series B* **48**, 259–302.

BILDER, R. M., WU, H., BOGERTS, B., DEGREEF, G., ASHTARI, M., ALVIR, J. M., SNYDER, P. J. AND LIEBERMAN, J. A. (1994). Absence of regional hemispheric volume asymmetries in first-episode schizophrenia. *The American Journal of Psychiatry* **151**, 1437–1447.

BOCK, M. T. AND BOWMAN, A. W. (2006). On the measurement and analysis of asymmetry with applications to facial modelling. *Journal of the Royal Statistical Society, Series C* **55**, 77–91.

BOOKSTEIN, F. L. (1986). Size and shape spaces for landmark data in two dimensions (with discussion). *Statistical Science* **1**, 181–242.

BOOKSTEIN, F. L. (1997). Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis* **1**, 225–243.

BRIGNELL, C. J. (2007). Shape analysis and statistical modelling in brain imaging, [PhD Thesis]. University of Nottingham, Nottingham.

BROWNE, W. J. AND DRAPER, D. (2000). Implementation and performance issues in the Bayesian and likelihood fitting of multilevel models. *Computational Statistics* **15**, 391–420.

CAO, J. AND WORSLEY, K. J. (2001). Applications of random fields in human brain mapping. In: Moore, M. (editor), *Spatial Statistics: Methodological Aspects and Applications*. Lecture Notes in Statistics, Volume 159. New York: Springer, pp. 169–182.

CHANCE, S. A., CRAVEN, R. M., GONZALEZ BALLESTER, M. A. AND CROW, T. J. (1999). A new method for MRI assessment of torque in schizophrenic brains. *Schizophrenia Research [abstracts]* **6**, 193.

CHUNG, M. K., DALTON, K. M., SHEN, L., EVANS, A. C. AND DAVIDSON, R. J. (2007). Weighted Fourier series representation and its application to quantifying the amount of gray matter. *IEEE Transactions on Medical Imaging* **26**, 566–581.

CHUNG, M. K., WORSLEY, K. J., ROBBINS, S. AND EVANS, A. C. (2003). Tensor-based brain surface modeling and analysis. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Volume 1. Washington, DC: IEEE Computer Society, pp. 467–476.

DRYDEN, I. L. (2005). Statistical analysis on high-dimensional spheres and shape spaces. *Annals of Statistics* **33**, 1643–1665.

DRYDEN, I. L. AND MARDIA, K. V. (1998). *Statistical Shape Analysis*. Chichester, UK: Wiley.

FISCHL, B., LIU, A. AND DALE, A. (2001). Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Transactions on Medical Imaging* **20**, 70–80.

GOOD, C., JOHNSRUDE, I. S., ASHBURNER, J., HENSON, R. N. A., FRISTON, K. J. AND FRACKOWIAK, R. S. J. (2001). A voxel-based morphometric study of ageing in 465 normal adult brains. *Neuroimage* **14**, 21–36.

GRENANDER, U. AND MILLER, M. I. (1994). Representations of knowledge in complex systems (with discussion). *Journal of the Royal Statistical Society, Series B* **56**, 549–603.

HOBOLTH, A. (2003). The spherical deformation model. *Biostatistics* **4**, 583–595.

HOBOLTH, A., KENT, J. T. AND DRYDEN, I. L. (2002). On the relation between edge and vertex modelling in shape analysis. *Scandinavian Journal of Statistics* **29**, 355–374.

HYVÄRINEN, A., KARHUNEN, J. AND OJA, E. (2001). *Independent Components Analysis*. New York: Wiley.

JOSHI, S. C., MILLER, M. I. AND GRENANDER, U. (1997). On the geometry and shape of brain sub-manifolds. *International Journal of Pattern Recognition and Artificial Intelligence* **11**, 1317–1343.

KENDALL, D. G. (1984). Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society* **16**, 81–121.

KENT, J. T., DRYDEN, I. L. AND ANDERSON, C. R. (2000). Using circulant symmetry to model featureless objects. *Biometrika* **87**, 527–544.

KENT, J. T., MARDIA, K. V. AND RABE, S. (1994). Face description from laser range data. In: Bookstein, F. L., Duncan, J. S., Lange, N. and Wilson, D. C. (editors), *Mathematical Methods in Medical Imaging III*. SPIE Proceedings, Volume 2299. Bellingham, WA, pp. 32–45.

KERTESZ, A., POLK, M., BLACK, S. E. AND HOWELL, J. (1990). Sex, handedness, and the morphometry of cerebral asymmetries on magnetic resonance imaging. *Brain Research* **530**, 40–48.

KLASSEN, E., SRIVASTAVA, A., MIO, W. AND JOSHI, S. H. (2003). Analysis of planar shapes using geodesic paths on shape spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**, 372–383.

LAWRIE, S. M., WHALLEY, H. C., ABUKMEIL, S. S., KESTELMAN, J. N., MILLER, P., BEST, J. J., OWENS, D. G. AND JOHNSTONE, E. C. (2002). Temporal lobe volume changes in people at high risk of schizophrenia with psychotic symptoms. *The British Journal of Psychiatry* **181**, 138–143.

MARCHINI, J., HEATON, C. AND RIPLEY, B. D. (2003). *fastICA package in R. R package*. http://cran.r-project.org.

MORRIS, R. J., KENT, J. T., MARDIA, K. V., FIDRICH, M., AYKROYD, R. G. AND LINNEY, A. (1999). Analysing growth in faces. In: Arabnia, H. R. (editor), *Proceedings of Conference on Imaging Science, Systems and Technology, Las Vegas, 1999*. Bogart, GA: Computer Science Research, Education and Applications (CSREA) Press, pp. 404–409.

QIU, A., BITOUK, D. AND MILLER, M. I. (2006). Smooth functional and structural maps on the neocortex via orthonormal bases of the Laplace-Beltrami operator. *IEEE Transactions on Medical Imaging* **25**, 1296–1306.

R DEVELOPMENT CORE TEAM (2008). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

RAMSAY, J. O. AND SILVERMAN, B. W. (2005). *Functional Data Analysis*, 2nd edition. New York: Springer.

SHEN, L., FORD, J., MAKEDON, F. AND SAYKIN, A. J. (2004). A surface-based approach for classification of 3D neuroanatomic structures. *Intelligent Data Analysis* **8**, 519–542.

SMITH, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping* **17**, 143–155.

SMITH, S. M., JENKINSON, M., WOOLRICH, M. W., BECKMANN, C. F., BEHRENS, T. E. J., JOHANSEN-BERG, H., BANNISTER, P. R., LUCA, M. D., DROBNJAK, I., FLITNEY, D. E. *and others* (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23**, S208–S219.

STOYAN, D. AND MOLCHANOV, I. S. (1997). Set-valued means of random particles. *Journal of Mathematical Imaging and Vision* **7**, 111–121.

TALAIRACH, P. AND TOURNOUX, J. (1988). *A Stereotactic Coplanar Atlas of the Human Brain*. New York: Thieme.

THEOBALD, C. M., GLASBEY, C. A., HORGAN, G. W. AND ROBINSON, C. D. (2004). Principal component analysis of landmarks from reversible images. *Journal of the Royal Statistical Society, Series A* **53**, 163–175.

THOMPSON, P. M., GIEDD, J. N., WOODS, R. P., MACDONALD, D., EVANS, A. C. AND TOGA, A. W. (2000). Growth patterns in the developing human brain detected using continuum-mechanical tensor mapping. *Nature* **404**, 190–193.

VAN ESSEN, D. C., DRURY, H. A., JOSHI, S. AND MILLER, M. I. (1998). Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 788–795.

WAND, M. P. AND JONES, M. C. (1995). *Kernel Smoothing*. London: Chapman and Hall.

WAND, M. P. AND RIPLEY, B. D. (2009). *KernSmooth 2.23. Functions for Kernel Smoothing Corresponding to the Book: Wand, M.P. and Jones, M.C. (1995) "Kernel Smoothing". R package*. http://cran.r-project.org.

WORSLEY, K. J. (1994). Local maxima and the expected Euler characteristic of excursion sets of $\chi^2$, $F$ and $t$ fields. *Advances in Applied Probability* **26**, 13–42.

YU, P., GRANT, P. E., QI, Y., HAN, X., SÉGONNE, F., PIENAAR, R., BUSA, E., PACHECO, J., MAKRIS, N., BUCKNER, R. L. *and others* (2007). Cortical surface shape analysis based on spherical wavelets. *IEEE Transactions on Medical Imaging* **26**, 582–597.

[*Received May 9, 2008; revised May 7, 2009; accepted for publication February 23, 2010*]