

Statistical Image Analysis of Longitudinal RAVENS Images: Methodology and Case Study

Seonjoo Lee¹, Vadim Zipunnikov², Brian S. Caffo², Daniel S. Reich^{2,3} and Dzung L. Pham¹

¹ *The Henry M. Jackson Foundation, Bethesda, MD 20817, USA*

² *Department of Biostatistics, Johns Hopkins University, Baltimore, MD 21205, USA*

³ *National Institutes of Health, Bethesda, MD 20892, USA*

seonjool@gmail.com

SUMMARY

Regional analysis of volumes examined in normalized space (RAVENS) are transformation images used in the study of brain morphometry. In this paper, RAVENS images are analyzed using a longitudinal variant of voxel-based morphometry (VBM) and longitudinal functional principal component analysis (LFPCA) for high-dimensional images. We demonstrate that the latter overcomes the limitations of standard longitudinal VBM analyses, which does not separate registration errors from other longitudinal changes and baseline patterns. This is especially important in contexts where longitudinal changes are only a small fraction of the overall observed variability, which is typical in normal aging and many chronic diseases. Our simulation study shows that LFPCA effectively separates registration error from baseline and longitudinal signals of interest by decomposing RAVENS images measured at multiple visits into three components: a subject-specific imaging random intercept that quantifies the cross-sectional variability, a subject-specific imaging slope that quantifies the irreversible changes over multiple visits, and a subject-visit specific imaging deviation. We describe strategies to identify baseline/longitudinal variation and registration errors combined with covariates of interest. Our analysis suggests that specific regional brain atrophy and ventricular enlargement are associated with multiple sclerosis (MS) disease progression.

Key words: Longitudinal functional principal component analysis; Regional analysis of volumes examined in normalized space; Voxel-based morphometry; Multiple sclerosis; Brain volume measurement

*To whom correspondence should be addressed

1. INTRODUCTION

Magnetic resonance imaging (MRI) is commonly used in the study of brain structure. Many studies are based on measurements of tissue volumes within a number of predefined regions of interest (ROIs); for example, see [Bartzokis *and others* \(2001\)](#) and [Bermel *and others* \(2003\)](#). Although ROI analysis can directly quantify the volume of structures and reduce the dimensionality of images, the ROIs have to be defined before the analysis is conducted. In disease studies, this can be difficult without sufficient prior knowledge about what and how various regions will be affected. Moreover, ROI based measurements can be time-consuming and labor-intensive. The results of the analysis will depend on the quality of the ROI delineation and thus depend upon the experience of the operator and accuracy of segmentation algorithms. Voxel-based morphometry (VBM) is a complementary technique that measures local brain volume changes in a normalized space and thus does not suffer from these limitations ([Ashburner and Friston, 2000, 2001](#)).

To quantify local brain volumes, we focus here on the Regional Analysis of Volumes Examined in Normalized Space (RAVENS) approach, using tissue-preserving image warping ([Goldszal *and others*, 1998](#); [Davatzikos, 1998](#); [Davatzikos *and others*, 2001](#); [Shen and Davatzikos, 2003](#)) though the method of analysis in principle extends to other morphometric images. RAVENS registers each subject to a template of anatomy so that the intensities of the RAVENS image represents regional volumes relative to those of template ([Shen and Davatzikos, 2002](#)). Figure 1 displays examples of RAVENS images in the template space. The first subject has much larger ventricles than the second subject (and template). Its RAVENS image of ventricles is displayed underneath the associated T1 image with red and blue colors representing higher and lower intensities, respectively. Subject 1 has larger ventricular volume, depicted by brighter ventricular intensities in the RAVENS image. Similarly, the second brain, having a smaller ventricle than that of the first subject, has lower intensities in its RAVENS image. Statistical VBM analysis of RAVENS images (RAVENS-VBM) is applied to the resulting spatial distributions of gray matter (GM), white matter (WM), and ventricular cerebrospinal fluid (CSF). Local atrophy or enlargement can be detected if the spatial distributions of GM, WM, and CSF significantly change across subjects.

In many disease studies, longitudinal patterns of brain structure between and within control and patient groups are of interest. Many studies have been based on ROI volume measurements followed by statistical analysis, such as a linear mixed model. Several neuroimaging software platforms, including: FSL ([Smith *and others*, 2004](#)), the SPM-VBM toolbox (available at <http://dbm.neuro.uni-jena.de/vbm>) and SurfStat ([Worsley *and others*, 2009](#)), support flexible longitudinal models. Statistical inference of the contrast between two different time points is the most commonly used approach ([Bendfeldt *and others*, 2011](#)). Numerous other approaches for longitudinal imaging data have been proposed for prediction. The methods include support vector machine classifier ([Chen and DuBois Bowman, 2011](#)) and bayesian spatial model ([Derado *and others*, 2012](#)).

Current statistical approaches have three major limitations: 1) continuous longitudinal trajectories are not fully investigated; 2) they do not distinguish among longitudinal changes, registration errors, or technological and biological scan-to-scan variability; 3) an integrated framework of cross-sectional and longitudinal analysis is not provided, i.e. cross-sectional and longitudinal analyses are applied separately.

Our main goal is to introduce a new statistical framework for longitudinal VBM analysis, particularly using RAVENS images. To overcome the limitations of the current existing statistical analysis, we consider a data-driven analysis to provide a more complete statistical framework to analyze high-dimensional longitudinal brain images. A framework that allows this conceptual partition of variability is longitudinal functional principal component analysis (LFPCA) ([Greven](#)

and others, 2011). This method was originally proposed for low to moderate dimensional functional data and was extended to high dimensional data by *Zipunnikov and others (2012)*. The main idea of high-dimensional inference is based on projecting onto the intrinsic low dimensional space spanned by high-dimensional vectors (*Di and others, 2009; Zipunnikov and others, 2011*). More precisely, we start by modeling the observed data with high-dimensional longitudinal functional principal component analysis (HD-LFPCA). Each RAVENS ventricular image is unfolded into a $p \times 1$ dimensional vector, where $p \approx 80,000$ is the number of voxels in the RAVENS ventricular image. These vectors are decomposed in their baseline, longitudinal and visit-to-visit components; each component is then estimated from the data. The method takes only a few minutes on a standard PC.

In this paper we focus on LFPCA as a useful tool for analyzing longitudinal images, particularly to quantify cross-sectional and longitudinal variability in the data. The simulation study illustrates the application of LFPCA to a simplified imaging setting. It demonstrates that LFPCA effectively separates longitudinal, cross sectional, and other variations. Notably, the simulation study shows that LFPCA can separate registration errors from baseline and longitudinal components of interest.

The rest of this paper is laid out as follows: Section 2 reviews LFPCA and estimation procedures for high-dimensional data. Section 3 illustrates how to apply LFPCA to brain imaging data and offers interpretation of results. Section 4 describes a longitudinal multiple sclerosis (MS) study with LFPCA results.

2. METHODS

In this section, we provide a description of the original LFPCA approach developed by [Greven and others \(2011\)](#) and its extension for high-dimensional data analysis ([Zipunnikov and others, 2012](#)). Throughout this section, we refer to both as LFPCA.

2.1 *Random Intercept and Random Slope Model*

Consider a longitudinal brain imaging study with subjects labeled by index i with occasion indexed by j and scan time by variable t_{ij} for $j = 1, \dots, J_i$. Each image is unfolded into a p -dimensional column vector $\mathbf{y}_{ij}(v)$; the index v of each entry corresponds to a particular location in the brain for each subject and visit in normalized space. A random slope and random intercept model is commonly used to analyze longitudinal data, which has been extended to functional ([Greven and others, 2011](#)) and imaging ([Zipunnikov and others, 2012](#)) studies as follows:

$$y_{ij}(v) = \eta(v, t_{ij}) + x_{i,0}(v) + x_{i,1}(v)t_{ij} + W_{ij}(v), \quad (2.1)$$

where $y_{ij}(v)$ denotes the image intensity at voxel v , $\eta(v, t_{ij})$ is a fixed main effect, and $x_{i,0}(v)$ and $x_{i,1}(v)$ denote the random intercept and random slope for subject i , respectively. The term $W_{ij}(v)$ is a random subject-visit specific imaging deviation, which is assumed to be a zero mean, second-order stationary random process uncorrelated with $\mathbf{X}_i(v) = (x_{i,0}(v), x_{i,1}(v))^\top$. The covariance operators of $\mathbf{X}_i(v)$ and $W_{ij}(v)$ are denoted as $\mathbf{K}^X(v_1, v_2)$ and $\mathbf{K}^W(v_1, v_2)$, respectively.

While this is a natural and relatively simple model for longitudinally observed data, the scale of the problem requires aggressive dimensionality reduction. LFPCA reduces dimensionality by projecting onto the subspaces which explain principal directions of variation in the data. In model (2.1), there are two sources of variation: subject-to-subject, captured by \mathbf{X}_i , and visit-to-visit within a subject, captured by W_{ij} and the model assumption on \mathbf{X}_i and W_{ij} in (2.1) allows us to partition the variation of the data and LFPCA models latent processes \mathbf{X}_i and W_{ij} using a Karhunen-Loeve (K-L) expansion ([Karhunen, 1947](#); [Loève, 1948](#)).

The K-L expansion decomposes the two latent processes as $\mathbf{X}_i(v) = \sum_{k=1}^{\infty} \xi_{ik} \phi_k^X(v)$ and $W_{ij}(v) = \sum_{l=1}^{\infty} \zeta_{ijl} \phi_l^W(v)$, where $\phi_k^X = (\phi_k^{X,0}, \phi_k^{X,1})$ and ϕ_l^W are the eigenfunctions of $\mathbf{K}^X(v_1, v_2)$ and $\mathbf{K}^W(v_1, v_2)$, respectively, such that

$$\mathbf{K}^X(v_1, v_2) = \begin{pmatrix} \mathbf{K}_{00}^X(v_1, v_2) & \mathbf{K}_{10}^X(v_1, v_2) \\ \mathbf{K}_{01}^X(v_1, v_2) & \mathbf{K}_{11}^X(v_1, v_2) \end{pmatrix} = \sum_{k=1}^{N_X} \lambda_k^X \phi_k^X(v_1) \left\{ \phi_k^X(v_2) \right\}^\top.$$

LFPCA truncates K-L representations and represents observed data through a linear mixed-effects model:

$$y_{ij}(v) = \eta(v, t_{ij}) + \sum_{k=1}^{N_X} \xi_{ik} \phi_k^{X,0}(v) + t_{ij} \sum_{k=1}^{N_X} \xi_{ik} \phi_k^{X,1}(v) + \sum_{l=1}^{N_W} \zeta_{ijl} \phi_l^W(v), \quad (2.2)$$

$$(\xi_{ik_1}, \xi_{ik_2}) \sim (0, 0, \lambda_{k_1}^X, \lambda_{k_2}^X, 0); (\zeta_{il_1}, \zeta_{il_2}) \sim (0, 0, \lambda_{l_1}^W, \lambda_{l_2}^W, 0),$$

where “ $\sim (\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$ ” denotes that a pair of variables has a distribution with mean (μ_1, μ_2) , variance σ_1^2, σ_2^2 , and correlation ρ . We assume that $\lambda_{k_1} \geq \lambda_{k_2}$ if $k_1 \leq k_2$. Since $\mathbf{X}_i(v)$ and $W_{ij}(v)$ are uncorrelated, the scores $\{\xi_{ik}\}_{k=1}^{\infty}$ and $\{\zeta_{ijl}\}_{l=1}^{\infty}$ are also uncorrelated. A very important characteristic of model (2.2) is that both N_X and N_W are expected to be small in most applications.

For the unfolded vector, (2.2) can be rewritten as

$$\begin{aligned} \mathbf{y}_{ij} &= \boldsymbol{\eta}(t_{ij}) + \sum_{k=1}^{N_X} \xi_{ik} \boldsymbol{\phi}_k^{X,0} + t_{ij} \sum_{k=1}^{N_X} \xi_{ik} \boldsymbol{\phi}_k^{X,1} + \sum_{l=1}^{N_W} \zeta_{ijl} \boldsymbol{\phi}_l^W \\ &= \boldsymbol{\eta}(t_{ij}) + \boldsymbol{\Phi}_X^0 \boldsymbol{\xi}_i + t_{ij} \boldsymbol{\Phi}_X^1 \boldsymbol{\xi}_i + \boldsymbol{\Phi}_W \boldsymbol{\zeta}_{ij}, \end{aligned} \quad (2.3)$$

where $\mathbf{y}_{ij} = (y_{ij}(v_1), \dots, y_{ij}(v_p))^\top$ is a $p \times 1$ dimensional vector; $\boldsymbol{\phi}_k^{X,0}, \boldsymbol{\phi}_k^{X,1}$, and $\boldsymbol{\phi}_l^W$ are $p \times 1$ eigenvectors; $\boldsymbol{\Phi}_X^s = (\boldsymbol{\phi}_1^{X,s}, \dots, \boldsymbol{\phi}_{N_X}^{X,s})$ for $s = 0, 1$; $\boldsymbol{\Phi}_W = (\boldsymbol{\phi}_1^W, \dots, \boldsymbol{\phi}_{N_W}^W)$; $\boldsymbol{\xi}_i = (\xi_{i1}, \dots, \xi_{iN_X})^\top$; $\boldsymbol{\zeta}_i = (\zeta_{i1}, \dots, \zeta_{iN_W})^\top$.

In brain imaging data analysis, LFPCA can separate biological signals from non-biological artifacts. For example, registration errors due to structural differences between subjects can be captured by baseline subject-specific components $\boldsymbol{\Phi}_X^0$ and scanner variability can be captured by visit-to-visit components $\boldsymbol{\Phi}_W$. This will be illustrated via an extensive simulation experiment in Section 3.

The fixed effect $\boldsymbol{\eta}(v, t_{ij})$ can be estimated in a number of ways (Greven *and others*, 2011). The analyses in Sections 3–4 simply use the sample mean across all the image observations. Once $\boldsymbol{\eta}(v, t_{ij})$ is estimated by the sample mean $\tilde{\boldsymbol{\eta}}(v, t_{ij})$, the longitudinal eigenanalysis is applied to the residual images $\tilde{\mathbf{y}}_{ij}(v) = y_{ij}(v) - \tilde{\boldsymbol{\eta}}(v, t_{ij})$ that are modeled as follows:

$$\tilde{\mathbf{y}}_{ij} = \boldsymbol{\Phi}_X^0 \boldsymbol{\xi}_i + t_{ij} \boldsymbol{\Phi}_X^1 \boldsymbol{\xi}_i + \boldsymbol{\Phi}_W \boldsymbol{\zeta}_{ij}. \quad (2.4)$$

2.2 Estimation

Zipunnikov *and others* (2012) modified the original approach of Greven *and others* (2011) and developed a method of moments estimator based on quadratics of $\tilde{\mathbf{y}}_{ij}$. The $p \times p$ -covariance of $\tilde{\mathbf{y}}_{ij_1}$ and $\tilde{\mathbf{y}}_{ij_2}$ is given by

$$\mathbf{E} \{ \tilde{\mathbf{y}}_{ij_1} \tilde{\mathbf{y}}_{ij_2}^\top \} = \mathbf{K}_{00}^X + T_{ij_1} \mathbf{K}_{10}^X + T_{ij_2} \mathbf{K}_{X}^{10} + T_{ij_1} T_{ij_2} \mathbf{K}_X^{11} + \delta_{j_1, j_2} \mathbf{K}^W, \quad j_1, j_2 = 1, \dots, J_i, \quad (2.5)$$

where $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ otherwise. Model (2.5) can be rewritten in terms of unfolded vectors $\mathbf{K}^v = \{ \text{vec}(\mathbf{K}_{00}), \text{vec}(\mathbf{K}_{01}), \text{vec}(\mathbf{K}_{10}), \text{vec}(\mathbf{K}_{11}), \text{vec}(\mathbf{K}^W) \}$ and $\mathbf{f}_{ij_1 j_2} = (1, T_{ij_2}, T_{ij_1}, T_{ij_1} T_{ij_2}, \delta_{j_1, j_2})^\top$ such that $\mathbf{E} \text{vec}(\tilde{\mathbf{y}}_{ij_1} \tilde{\mathbf{y}}_{ij_2}^\top) = \mathbf{K}^v \mathbf{f}_{ij_1 j_2}$. By concatenating all vectors across all subjects and visits we obtain a moment matrix identity for the $p^2 \times m$ matrix \mathbf{Y} : $\mathbf{E} \mathbf{Y} = \mathbf{K}^v \mathbf{F}$, where $m = \sum_{i=1}^N J_i^2$. Then covariance parameters \mathbf{K}^v can be unbiasedly estimated by using ordinary least squares (OLS): $\hat{\mathbf{K}}^v = \mathbf{Y} \mathbf{F}^\top (\mathbf{F} \mathbf{F}^\top)^{-1}$.

The covariance operators \mathbf{K}^X and \mathbf{K}^W are $2p \times 2p$ and $p \times p$ dimensional, respectively. For high-dimensional functional data, storing or diagonalizing these matrices is not feasible. Zipunnikov *and others* (2012) proposed HD-LFPCA, a novel estimation approach that takes advantage of an intrinsically small dimension of the space spanned by high-dimensional data vectors. First we form a $p \times J_i$ dimensional matrix $\tilde{\mathbf{y}}_i$, where column j corresponds to a demeaned-RAVENS image obtained for subject i at visit j . The $p \times J$ dimensional data matrix $\tilde{\mathbf{y}} = (\tilde{\mathbf{y}}_1; \dots; \tilde{\mathbf{y}}_n)$ is formed by column-binding the blocks of data corresponding to each subject, where $J = \sum_{i=1}^n J_i$. The data matrix can be decomposed as $\tilde{\mathbf{y}} = \mathbf{V} \mathbf{S} \mathbf{U}^\top$ using a singular value decomposition (SVD) approach. In the RAVENS application described in Section 4, $J = 148$. Equation (2.4) can be rewritten as

$$\tilde{\mathbf{y}}_{ij} = \mathbf{V} \mathbf{S} \mathbf{U}_{ij} = \boldsymbol{\Phi}_X^0 \boldsymbol{\xi}_i + t_{ij} \boldsymbol{\Phi}_X^1 \boldsymbol{\xi}_i + \boldsymbol{\Phi}_W \boldsymbol{\zeta}_{ij}. \quad (2.6)$$

By multiplying with \mathbf{V}^\top to the left, we have

$$\begin{aligned}\mathbf{S}\mathbf{U}_{ij} &= \mathbf{V}^\top \hat{\Phi}_X^0 \boldsymbol{\xi}_i + t_{ij} \mathbf{V}^\top \hat{\Phi}_X^1 \boldsymbol{\xi}_i + \mathbf{V}^\top \hat{\Phi}_W \boldsymbol{\zeta}_{ij} \\ &= \mathbf{A}_X^0 \boldsymbol{\xi}_i + \mathbf{A}_X^1 \boldsymbol{\xi}_i + \mathbf{A}_W \boldsymbol{\zeta}_{ij}.\end{aligned}\tag{2.7}$$

We estimate $\hat{\mathbf{A}}_X^0$, $\hat{\mathbf{A}}_X^1$, and $\hat{\mathbf{A}}_W$ as described earlier, and estimate $\hat{\Phi}_X^0 = \mathbf{V} \hat{\mathbf{A}}_X^0$, $\hat{\Phi}_X^1 = \mathbf{V} \hat{\mathbf{A}}_X^1$, and $\hat{\Phi}_W = \mathbf{V} \hat{\mathbf{A}}_W$. Note that multiplying by \mathbf{V}^\top in equation (2.6) reduces the model to its low-dimensional form (2.7), without losing the original correlation structure of the data. Once inference is conducted in model (2.7), then quantities of interest from model (2.6) can be estimated by pre-multiplying equation (2.7) by \mathbf{V} .

Principal scores ξ_i and ζ_{ij} are estimated via Best Linear Unbiased Predictions (BLUPs) as follows. The stacked vector of i th subject data, $\text{vec}(\tilde{\mathbf{y}}_i) = \left(\tilde{\mathbf{y}}_{i1}^\top, \dots, \tilde{\mathbf{y}}_{iJ_i}^\top \right)^\top$, can be rewritten as $\text{vec}(\tilde{\mathbf{y}}_i) = \mathbf{B}_i \boldsymbol{\omega}_i$, where $\mathbf{B}_i = (\mathbf{B}_i^X; \mathbf{B}_i^W)$, $\mathbf{B}_i^X = \mathbf{1}_{J_i} \otimes \hat{\Phi}_X^0 + \mathbf{T}_i \otimes \hat{\Phi}_X^1$, $\mathbf{B}_i^W = \mathbf{I}_{J_i} \otimes \hat{\Phi}_W$, where $\mathbf{T}_i = (T_{i1}, \dots, T_{iJ_i})^\top$, $\boldsymbol{\omega}_i = \left(\boldsymbol{\xi}_i^\top, \boldsymbol{\zeta}_i^\top \right)^\top$, the subject level principal scores $\boldsymbol{\zeta}_i = \left(\zeta_{i1}^\top, \dots, \zeta_{iJ_i}^\top \right)^\top$, and $\mathbf{1}_{J_i}$ is a $J_i \times 1$ vector of ones. Then the scores can be estimated as $\hat{\boldsymbol{\omega}}_i = \left(\hat{\mathbf{B}}_i^\top \hat{\mathbf{B}}_i \right)^{-1} \hat{\mathbf{B}}_i^\top \text{vec}(\tilde{\mathbf{y}}_i)$. Due to linearity the estimated scores are the same in both models (2.6) and (2.7). Details of the matrix calculation and additional theoretical results of HD-LFPCA can be found in [Zipunnikov and others \(2012\)](#).

3. SIMULATION

In this section, we present a simulation study to test the performance of LFPCA in RAVENS-VBM analysis. We investigate if LFPCA can identify subject-specific signals from noise, particularly registration errors, which often dominate signals in VBM analyses. Also, we identify cross-sectional and longitudinal variation when they exist.

3.1 Simulation Setup

We design a simulation study to mimic longitudinal analysis of RAVENS images. For the purpose of illustration, we use 2D images with $200 \times 200 = 40,000$ pixels. We generate images from 50 subjects ($N = 50$) with three follow-ups. To replicate RAVENS processing routine, we assume that all images are registered to a template space. Figure 2 displays simulated RAVENS images from 5 randomly chosen subjects. Each column represents four longitudinally collected images of the same subject.

Each image mimics four canonical brain structures: background (B), white matter (W), ventricles (V), and gray matter (G). Those four components are simplified and shown as a background, a big square, a small square inside the big square, and a rectangle at the bottom, respectively. Registration errors are introduced via random rigid shifts of simulated structures as described below.

The shapes of G, V and W are modeled as follows: the side lengths of W and V were generated from $2 \cdot \lfloor N(90, 2) \rfloor$ and $2 \cdot \lfloor N(30, 0.5) \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. The width of G is the same as that of W in the same image and the height is generated from $\lfloor N(10, 1) \rfloor$. For visit-to-visit registration errors, all images are shifted to both x- and y-directions. The amount of each shift is generated from $\lfloor N(0, 1) \rfloor$ for W and $\lfloor N(0, 0.25) \rfloor$ for V. In Figure 2, the sizes of G, V and W are different across subjects and their locations vary at different visits.

The intensities of W, V and G are generated from $N(55, 5)$, $N(15, 5)$ and $N(20, 5)$, respectively. For the longitudinal variation, we imposed an increasing trend (expansion) for V with the random slope generated from $N(2, 4)$, and atrophy for G with the random slopes generated from $N(-2, 4)$. In Figure 2, the images from the first subject, which are displayed at the first column, show the longitudinal patterns. In the images, the color of V changes from darker gray to brighter gray, which represents longitudinal enlargements of V. Similarly, the colors of W and G changed to darker colors, which represent longitudinal atrophy.

3.2 Simulation Results

Figure 3 shows the first five pairs of subject-specific components (Φ_X). The baseline components (Φ_X^0) are displayed in the top row and their corresponding longitudinal components (Φ_X^1) are displayed in the second row. Each image is colored with a blue(negative)-gray(0)-red(positive) color scheme. The first subject-specific component (Figure 3 (a) first column) represents cross-sectional variations of the intensities of the W. The second component captures subject-specific registration errors, which only depend on cross-sectional variation. The third and fourth components represent the size of V and G. The fifth component shows longitudinal patterns of V and G. For a subject with positive score, the area V enlarges over the time and that of G shrinks, matching the truth used in simulation.

One useful feature of LFPCA is that contributions of the longitudinal and baseline components within each subject-specific component can be quantified on a $[0, 1]$ scale. A subject-specific eigen-

vector is the stacked vector of baseline and longitudinal components: $\Phi_{X,k} = \left\{ \Phi_{X,k}^0{}^\top, \Phi_{X,k}^1{}^\top \right\}^\top$, such that $\|\Phi_{X,k}\|^2 = \|\Phi_{X,k}^0\|^2 + \|\Phi_{X,k}^1\|^2 = 1$. For each component, the variation or the contribution of the longitudinal component can be calculated as $\frac{\|\Phi_{X,k}^1\|^2}{\|\Phi_{X,k}^0\|^2 + \|\Phi_{X,k}^1\|^2}$. Combined with the contribution of each subject-specific component to the total variation, Figure 3 (b) displays variations explained by the first 10 subject-specific components with the proportion of the longitudinal components within each subject. Each bar plot intensity represents the amount of variation explained by each subject-specific component and is comprised of variations explained by the longitudinal component (magenta) and the baseline component (cyan). The top of each bar displays numerical values of the variation explained by the subject-specific component with the variation explained by the longitudinal component within the subject-specific component in parenthesis. Note that the fifth principal component has the highest longitudinal-baseline ratio among all ten components. This provides a strong indication that the fifth component should be essentially treated as a longitudinal component. Using both visual and quantitative methods, we can conclude that the first four components represent baseline variation and registration error and the fifth component reveals longitudinal variation. In the data set, the longitudinal variation and baseline variations are independent, which agrees with the simulation setting.

Figure 3 (c) displays the boxplots of first 10 subject-specific LFPC scores. A subject-specific LFPC score represents every subject-specific LFPC in terms of voxel intensities as a distance from the mean image along the appropriate subject-specific LFPC axis. This score is a weighted sum of the RAVENS image intensities, with weights proportional to the strength of the relationship between the LFPCs and RAVENS map. For the first four components, the scores represent baseline sizes/shapes. For example, a subject with a positive value of the first LFPC score has larger W. A subject with a positive value of the fifth LFPC score has longitudinal enlargement of V and longitudinal atrophy of G. Principal scores can be efficiently utilized as explanatory variables representing particular brain regions in regression analysis relating them to covariates of interest such as a disease status.

An advantage of LFPCA is its ability to couple baseline and longitudinal variation. The longitudinal component is added to the baseline with the time used as a multiplicative weight. Figure 4 illustrates the temporal trajectory of principal component loadings. We display only the first component, which does not appear to change over time. This pattern is replicated in the first four components. This indicates that the first four components mostly represent baseline variation. The fifth component loading does appear to change over time, while the baseline loading has relatively lower intensities compared to the longitudinal loading.

Figure 5 (a) displays the first ten components of the imaging deviation. All components appear to be registration errors. Figure 5 (b) displays the barplots of the variation explained by the first ten visit-to-visit components. Roughly 31% of the total variation is explained by visit-to-visit components. No patterns can be detected for the scores of the imaging deviation.

To summarize, our simulation studies convincingly demonstrate the power and flexibility of LFPCA to address some of key challenges of brain imaging. In particular, LFPCA managed to estimate and separate longitudinal and cross-sectional variation in a complex imaging simulation design with registration errors. The main part of the analysis can be automated and performed robustly with no operator input.

4. RAVENS IMAGE ANALYSIS

In this section, we analyze structural images obtained from MS patients using both linear mixed models with random intercepts and slopes and LFPCA.

High resolution 3D magnetization-prepared rapid acquisition of gradient echoes (MPRAGE; acquired resolution: $1.1\text{mm} \times 1.1\text{mm} \times 1.1\text{mm}$; TR: $\sim 10\text{ms}$; TE: 6ms ; TI= 835ms ; flip angle: 8deg ; SENSE factor: 2 ; averages: 1) were acquired on a 3.0 T MRI scanner (Intera, Philips Medical Systems). Forty eight MS patients (aged 42 ± 12 years at baseline) were enrolled in a longitudinal study of brain volume change. The study population included 33 female and 16 male patients; 28 patients with relapsing-remitting MS (RRMS), 13 patients with secondary progressive MS (SPMS), 5 patients with primary progressive MS (PPMS) and 2 patients with clinically isolated syndrome (CIS). One hundred forty eight T1 images have been acquired, with three images per subject for 44 subjects and 4 images per subject for 3 subjects. The average time interval between scans was 368 days (± 27). All images were spatially normalized via registration of T1 maps into the mean template, generated using ANTS (*Avants and others, 2010, 2011*) from 30 randomly chosen MS patients among those with more than three visits.

All T1 images were segmented into GM, WM, VN, and lesions with Lesion-TOADS (*Shiee and others, 2010*). Lesions were then relabeled as WM for all subsequent processing. After segmentation, the final tissue maps of GM, WM, and VN were normalized using HAMMER-SUITE (*Shen and Davatzikos, 2002*) to generate RAVENS images. Finally, the RAVENS maps were separately smoothed with 4 mm FWHM using SPM8.

4.1 Classical VBM Analysis using Linear Mixed Model

First, we applied traditional VBM analysis using a linear mixed model to find a longitudinal trend. Many previous longitudinal studies have applied pairwise comparisons between two time points (*Driemeyer and others, 2008*). This study attempts to discover constant longitudinal trends over the time, i.e., focusing on the atrophy or enlargement rates. This may elucidate disease progression patterns of the patients. For the i th subject j th visit, the RAVENS image at voxel v follows the model:

$$y_{ij}(v) = \beta_i^0(v) + \beta_i^1(v)t_{ij} + \epsilon_{ij}(v), \quad (4.8)$$

where $\beta_i^0(v) \sim N(\beta_0(v), \sigma_0^2(v))$, $\beta_i^1(v) \sim N(\beta_1(v), \sigma_1^2(v))$, $\epsilon_{ij}(v) \sim N(0, \sigma_\epsilon^2(v))$.

The time variable T_{ij} represents the number of days since the first (baseline) visit. This analysis focuses on the population mean of the longitudinal trend $\beta_1^1(v)$. After an FDR correction (*Benjamini and Yekutieli, 2001*) combined with cluster level thresholding, there are significant clusters with spatial extent more than 20 voxels. Table 1 shows information about the significant clusters, including cluster size, maximum or minimum t-values within each cluster and its location, and center of gravity of the clusters. We do not include an image of VBM results, since the clusters are very sparse having small cluster sizes. The results show that the longitudinal patterns do not appear to be significant for the most of brain regions. We suspect that it is because of the large variation within and between images for this very heterogeneous disease. We apply HD-LFPCA to uncover more subtle underlying variation.

4.2 HD-LFPCA results

We present the LFPCA results for ventricular RAVENS images in Figure 6. Figure 6 (a) and (b) display the amount of variation explained by subject-specific components and subject and

visit-specific components to the total variation, respectively. Figure 6 (a) displays variation explained by the first 10 subject-specific components along the proportion of longitudinal variation represented within each subject-specific component. Each bar’s height represents the percent of variation explained by each subject-specific component. It is color coded by the proportion of the variation explained by the longitudinal component and the baseline component. The top of each bar displays the variation explained by the subject-specific component; the fraction of that variation that is explained by the longitudinal component is given in parentheses.

The first subject-specific LFPC explains 45% of the overall variation, almost completely due to the cross-sectional part. The longitudinal part explains 81% of the variation within the second subject-specific LFPC. Figure 6 (b) displays variation explained by the first 10 subject-visit-specific components to the total variation. The remaining LFPCs explain less than 1% of the total variation.

Most of the subject-specific LFPCs are driven by cross-sectional variation, which possibly include registration errors. The longitudinal changes are mainly captured by the second LFPC, which explains about 8% of the total variation. This provides an explanation as to why traditional VBM using linear mixed models did not find meaningful longitudinal patterns.

Figures 7–8 show the first two pairs of LFPCs of ventricles. Figure 7 (a) and (b) show the baseline and longitudinal components of the first LFPC. The LFPC loadings are color-coded with a red-yellow color scheme for positive values and blue-cyan for negative values. The first components reveal baseline ventricular morphometric variation, while the longitudinal component has relatively lower intensities. To investigate the characteristics of the first component, we fit the linear regression with covariates of interest and volumes of 6 ROIs obtained by the Lesion-TOADS segmentation algorithm. Figure 7 (c) displays scatter plots of the LFPC scores with covariates, baseline age, baseline Expanded Disability Status Scale (EDSS) score, and volumes of 6 ROIs (thalamus, ventricle, cortical gray matter volume, caudate, sulcal CSF, putamen). The dashed lines overlaid on the scatter plots are the linear regression lines and are colored as red when the linear trend is significant and as green otherwise.

The significant correlation between the first subject-specific score and baseline VN volume (first row, fourth column) confirms that the first component represents baseline variation (R^2 : 0.9684), i.e., a subject with a positive score has larger ventricles at the baseline. The scores are significantly correlated to the subject’s baseline age (R^2 : 0.1402) and three gray matter ROIs (thalamus, caudate, and putamen).

Figure 8 (a) and (b) display the baseline and longitudinal components of the second subject-specific component, respectively. A subject with a positive score tends to have a larger regional volume at the baseline (yellow-red colored voxels in Figure 8 (a)) and longitudinal enlargement. The second subject-specific scores have significant relations with the baseline age, EDSS, thalamus volume, VN volumes, sulcal CSF volume and putamen volume. The second component scores have higher correlation with baseline age (R^2 : 0.2371) and EDSS (R^2 : 0.2053) than the first component scores. This indicates that the longitudinal enlargement of ventricles is related to disease progression.

5. DISCUSSION

In this manuscript, we described and evaluated a coherent methodology for the study of longitudinal RAVENS - or other methodological - volumetric imaging studies. Our simulation studies demonstrate that LFPCA tightly links the analysis methodology with the morphometric image processing stream. We demonstrated that LFPCA can uncover interesting, yet subtle, directions

of longitudinal variation in a case where independent voxel-level investigations fail. Of note, this study represents the first application of the high dimensional variation of LFPCA to RAVENS images. Related work includes [Zipunnikov *and others* \(2012\)](#), who investigated DTI imaging data and [Zipunnikov *and others* \(2011\)](#) and [Zipunnikov *and others* \(2011\)](#), who studied RAVENS images with cross-sectional and clustered (but not longitudinal) settings. Moreover, this manuscript represents the first application of LFPCA to morphometric imaging in multiple sclerosis.

A key insight from the simulation studies is the ability of LFPCA to uncover interesting directions of variation in the presence of errors from registration to a template. Previously, registration errors were handled via either extremely aggressive smoothing during post-registration processing or by improved normalization algorithms. While improved algorithms are certainly a desirable goal, all normalization algorithms must be tuned and suffer from trade offs (such as bias/variance). Our results suggest the possibility of employing less aggressive normalization.

In our study of MS, we found that the majority of variation is focused in cross-sectional components. This will likely be true in any study of adults, as variation in head size, brain size and intracranial volume will vary far more substantially than longitudinal decline, not unlike if one were to study adult cross-sectional and longitudinal trends in heights. It would be of interest to apply LFPCA to developmental populations or populations with severe progressive brain disorders significantly after disease onset.

We have applied similar analysis to gray matter and white matter RAVENS images. Figure 9 shows variation explained by first 10 subject-specific LFPCs in gray matter and white matter RAVENS images. In gray matter, around 20% variation comes from the longitudinal part across all subject-specific LFPCs. Lower proportions of variation, around 15%, are explained by the longitudinal part in white matter. Unlike the ventricles, any subject-specific component of gray and white matter is not dominated by the longitudinal part. We speculate it is due to spatial heterogeneity of longitudinal brain atrophy, which may depend on subject-specific disease progression.

The correlation between subject-specific LFPC scores of ventricles and EDSS indicates that EDSS is better associated with longitudinal ventricular enlargement than baseline ventricular size. This implies ventricular enlargement is a sensitive measurement of disease progression. Some cross-sectional MS patient studies have reported that brain atrophy is related to irreversible clinical disability in (MS) and ventricular enlargement may be a sensitive marker of this tissue loss that is seen at all stages of MS ([Turner *and others*, 2001](#); [Benedict *and others*, 2005](#); [Hildebrandt *and others*, 2006](#); [Tekok-Kilic *and others*, 2007](#)). In existing longitudinal studies, longitudinal ventricular enlargement and gray matter atrophy have been detected in both ROI and VBM with paired t-test or factor models ([Simon *and others*, 1999](#); [Dalton *and others*, 2002, 2006](#); [Kalkers *and others*, 2002](#); [Sepulcre *and others*, 2006](#); [Lukas *and others*, 2010](#); [Bendfeldt *and others*, 2009](#)), which agrees with our finding. Unlike other methods, LFPCA is able to show spatially heterogeneous patterns of longitudinal enlargement, which ROI based methods cannot provide.

ACKNOWLEDGMENTS

The project described was supported by grants numbered R01EB012547 and P41 EB015909 from the National Institute of Biomedical Imaging and Bioengineering and R01NS060910 and R01NS070906 from the National Institute of Neurological Disorders and Stroke. Daniel S. Reich is supported by the Intramural Research Program of the National Institute of Neurological Disorders and Stroke. *Conflict of Interest:* None declared.

REFERENCES

- ASHBURNER, J. AND FRISTON, K.J. (2000). Voxel-Based Morphometry—the Methods. *NeuroImage* **11**(6), 805–821.
- ASHBURNER, J. AND FRISTON, K.J. (2001). Why Voxel-Based Morphometry Would be Used. *NeuroImage* **14**(6), 1238–1243.
- AVANTS, B.B., TUSTISON, N.J., SONG, G., COOK, P.A., KLEIN, A. AND GEE, J.C. (2011). A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration. *NeuroImage* **54**(3), 2033–2044.
- AVANTS, B.B., YUSHKEVICH, P., PLUTA, J., MINKOFF, D., KORCZYKOWSKI, M., DETRE, J. AND GEE, J.C. (2010). The Optimal Template Effect in Hippocampus Studies of Diseased Populations. *NeuroImage* **49**(3), 2457–2466.
- BARTZOKIS, G., BECKSON, M., LU, P.H., NUECHTERLEIN, K.H., EDWARDS, N. AND MINTZ, J. (2001). Age-Related Changes in Frontal and Temporal Lobe Volumes in Men: a Magnetic Resonance Imaging Study. *Archives of General Psychiatry* **58**(5), 461.
- BENDFELDT, K., HOFSTETTER, L., KUSTER, P., TRAUD, S., MUELLER-LENKE, N., NAEGELIN, Y., KAPPOS, L., GASS, A., NICHOLS, T.E., BARKHOF, F. *and others.* (2011). Longitudinal Gray Matter Changes in Multiple Sclerosis: Differential Scanner and Overall Disease-Related Effects. *Human Brain Mapping* **33**, 1225–1245.
- BENDFELDT, K., KUSTER, P., TRAUD, S., EGGER, H., WINKLHOFER, S., MUELLER-LENKE, N., NAEGELIN, Y., GASS, A., KAPPOS, L., MATTHEWS, P.M. *and others.* (2009). Association of Regional Gray Matter Volume Loss and Progression of White Matter Lesions in Multiple Sclerosis—A Longitudinal Voxel-Based Morphometry Study. *NeuroImage* **45**(1), 60–67.
- BENEDICT, R.H.B., ZIVADINOV, R., CARONE, D.A., WEINSTOCK-GUTTMAN, B., GAINES, J., MAGGIORE, C., SHARMA, J., TOMASSI, M.A. AND BAKSHI, R. (2005). Regional Lobar Atrophy Predicts Memory Impairment in Multiple Sclerosis. *American Journal of Neuroradiology* **26**(7), 1824–1831.
- BENJAMINI, Y. AND YEKUTIELI, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 1165–1188.
- BERMEL, R.A., INNUS, M.D., TJOA, C.W. AND BAKSHI, R. (2003). Selective Caudate Atrophy in Multiple Sclerosis: A 3D MRI Parcellation Study. *Neuroreport* **14**(3), 335–339.
- CHEN, S. AND DUBOIS BOWMAN, F. (2011). A novel support vector classifier for longitudinal high-dimensional data and its application to neuroimaging data. *Statistical Analysis and Data Mining* **4**(6), 604–611.
- DALTON, CM, BREX, PA, JENKINS, R., FOX, NC, MISZKIEL, KA, CRUM, WR, O’RIORDAN, JI, PLANT, GT, THOMPSON, AJ AND MILLER, DH. (2002). Progressive Ventricular Enlargement in Patients with Clinically Isolated Syndromes is Associated with the Early Development of Multiple Sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* **73**(2), 141–147.
- DALTON, CM, MISZKIEL, KA, O’CONNOR, PW, PLANT, GT, RICE, GPA AND MILLER, DH. (2006). Ventricular enlargement in ms. *Neurology* **66**(5), 693–698.
- DAVATZIKOS, C. (1998). Mapping Image Data to Stereotaxic Spaces: Applications to Brain Mapping. *Human Brain Mapping* **6**(5-6), 334–338.
- DAVATZIKOS, C., GENC, A., XU, D. AND RESNICK, S.M. (2001). Voxel-Based Morphometry Using the RAVENS Maps: Methods and Validation Using Simulated Longitudinal Atrophy. *NeuroImage* **14**(6), 1361–1369.
- DERADO, G., BOWMAN, F.D.B. AND ZHANG, L. (2012). Predicting brain activity using a bayesian spatial model. *Statistical Methods in Medical Research*.
- DI, C.Z., CRAINCICANU, C.M., CAFFO, B.S. AND PUNJABI, N.M. (2009). Multilevel Functional Principal Component Analysis. *Annals of Applied Statistics* **3**(1), 458–488.
- DRIEMEYER, J., BOYKE, J., GASER, C., BÜCHEL, C. AND MAY, A. (2008). Changes in Gray Matter Induced by Learning—Revisited. *PLoS One* **3**(7), E2669.
- GOLDSZAL, A.F., DAVATZIKOS, C., PHAM, D.L., YAN, M.X.H., BRYAN, R.N. AND RESNICK, S.M. (1998). An Image-Processing System for Qualitative and Quantitative Volumetric Analysis of Brain Images. *Journal of Computer Assisted Tomography* **22**(5), 827–837.
- GREVEN, S., CRAINCICANU, C., CAFFO, B. AND REICH, D. (2011). Longitudinal Functional Principal Component Analysis. *Recent Advances In Functional Data Analysis and Related Topics*, 149–154.

- HILDEBRANDT, H., HAHN, HK, KRAUS, JA, SCHULTE-HERBRÜGGEN, A., SCHWARZE, B. AND SCHWENDEMANN, G. (2006). Memory Performance in Multiple Sclerosis Patients Correlates with Central Brain Atrophy. *Multiple Sclerosis* **12**(4), 428–436.
- KALKERS, N.F., AMEZIANE, N., BOT, J.C.J., MINNEBOO, A., POLMAN, C.H. AND BARKHOF, F. (2002). Longitudinal Brain Volume Measurement in Multiple Sclerosis: Rate of Brain Atrophy Is Independent of the Disease Subtype. *Archives of Neurology* **59**(10), 1572.
- KARHUNEN, KARI. (1947). Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.* **1947**(37), 79.
- LOÈVE, M. (1948). *Fonctions Aléatoires du Second Ordre.*
- LUKAS, C., MINNEBOO, A., DE GROOT, V., MORAAAL, B., KNOL, D.L., POLMAN, C.H., BARKHOF, F. AND VRENKEN, H. (2010). Early central atrophy rate predicts 5 year clinical outcome in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* **81**(12), 1351–1356.
- SEPULCRE, J., SASTRE-GARRIGA, J., CERCIGNANI, M., INGLE, G.T., MILLER, D.H. AND THOMPSON, A.J. (2006). Regional Gray Matter Atrophy in Early Primary Progressive Multiple Sclerosis: A Voxel-Based Morphometry Study. *Archives of Neurology* **63**(8), 1175.
- SHEN, D. AND DAVATZIKOS, C. (2002). HAMMER: Hierarchical Attribute Matching Mechanism for Elastic Registration. *Medical Imaging, IEEE Transactions On* **21**(11), 1421–1439.
- SHEN, D. AND DAVATZIKOS, C. (2003). Very High-Resolution Morphometry Using Mass-Preserving Deformations and HAMMER Elastic Registration. *NeuroImage* **18**(1), 28–41.
- SHIEE, N., BAZIN, P.L., OZTURK, A., REICH, D.S., CALABRESI, P.A. AND PHAM, D.L. (2010). A Topology-Preserving Approach to the Segmentation of Brain Images with Multiple Sclerosis Lesions. *NeuroImage* **49**(2), 1524–1535.
- SIMON, JH, JACOBS, LD, CAMPION, MK, RUDICK, RA, COOKFAIR, DL, HERNDON, RM, RICHERT, JR, SALAZAR, AM, FISCHER, JS, GOODKIN, DE *and others.* (1999). A Longitudinal Study of Brain Atrophy in Relapsing Multiple Sclerosis. *Neurology* **53**(1), 139–139.
- SMITH, S.M., JENKINSON, M., WOOLRICH, M.W., BECKMANN, C.F., BEHRENS, T.E.J., JOHANSEN-BERG, H., BANNISTER, P.R., DE LUCA, M., DROBNJAK, I., FLITNEY, D.E. *and others.* (2004). Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage* **23**, S208–S219.
- TEKOK-KILIC, A., BENEDICT, R.H.B., WEINSTOCK-GUTTMAN, B., DWYER, M.G., CARONE, D., SRINIVASARAGHAVAN, B., YELLA, V., ABDELRAHMAN, N. AND OTHERS. (2007). Independent Contributions of Cortical Gray Matter Atrophy and Ventricle Enlargement for Predicting Neuropsychological Impairment in Multiple Sclerosis. *NeuroImage* **36**(4), 1294–1300.
- TURNER, B., RAMLI, N., BLUMHARDT, LD AND JASPER, T. (2001). Ventricular enlargement in multiple sclerosis: A comparison of three-dimensional and linear mri estimates. *Neuroradiology* **43**(8), 608–614.
- WORSLEY, KJ, TAYLOR, JE, CARBONELL, F., CHUNG, MK, DUERDEN, E., BERNHARDT, B., LYTTTELTON, O., BOUCHER, M. AND EVANS, AC. (2009). Surfstat: A matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. *NeuroImage* **47**, S102.
- ZIPUNNIKOV, V., CAFFO, B., YOUSEM, D.M., DAVATZIKOS, C., SCHWARTZ, B.S. AND CRAINICEANU, C. (2011a). Functional Principal Components Model for High-Dimensional Brain Imaging. *NeuroImage* **58**, 772–784.
- ZIPUNNIKOV, V., CAFFO, B., YOUSEM, D.M., DAVATZIKOS, C., SCHWARTZ, B.S. AND CRAINICEANU, C. (2011b). Multilevel functional principal component analysis for high-dimensional data. *Journal of Computational and Graphical Statistics* **20**(4), 852–873.
- ZIPUNNIKOV, V., GREVEN, S., CAFFO, B., REICH, D.S. AND CRAINICEANU, C. (2012). Longitudinal High-Dimensional Data Analysis. *Johns Hopkins University, Dept. of Biostatistics Working Papers*, 234.