

LSTGEE: Longitudinal Analysis of Neuroimaging Data

Yimei, Li^a, Hongtu Zhu^{a, d}, Yasheng Chen^b, Hongyu An^b, John Gilmore^c, Weili Lin^{b, d}, and Dinggang Shen^{b, d}

^aDepartment of Biostatistics, ^bDepartment of Radiology, ^cDepartment of Psychiatry, and ^dBiomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

ABSTRACT

Longitudinal imaging studies are essential to understanding the neural development of neuropsychiatric disorders, substance use disorders, and normal brain. Using appropriate image processing and statistical tools to analyze the imaging, behavioral, and clinical data is critical for optimally exploring and interpreting the findings from those imaging studies. However, the existing imaging processing and statistical methods for analyzing imaging longitudinal measures are primarily developed for cross-sectional neuroimaging studies. The simple use of these cross-sectional tools to longitudinal imaging studies will significantly decrease the statistical power of longitudinal studies in detecting subtle changes of imaging measures and the causal role of time-dependent covariate in disease process.

The main objective of this paper is to develop longitudinal statistics toolbox, called LSTGEE, for the analysis of neuroimaging data from longitudinal studies. We develop generalized estimating equations for jointly modeling imaging measures with behavioral and clinical variables from longitudinal studies. We develop a test procedure based on a score test statistic and a resampling method to test linear hypotheses of unknown parameters, such as associations between brain structure and function and covariates of interest, such as IQ, age, gene, diagnostic groups, and severity of disease. We demonstrate the application of our statistical methods to the detection of the changes of the fractional anisotropy across time in a longitudinal neonate study. Particularly, our results demonstrate that the use of longitudinal statistics can dramatically increase the statistical power in detecting the changes of neuroimaging measures. The proposed approach can be applied to longitudinal data with multiple outcomes and accommodate incomplete and unbalanced data, i.e., subjects with different number of measurements.

Keywords: Covariate, generalized estimating equation, longitudinal, resampling method, score statistic.

1. INTRODUCTION

Longitudinal neuroimaging studies, in which subjects are scanned and measured repeatedly over time, are essential to understanding the neural development of neuropsychiatric disorders, substance use disorders, and normal brain. Many large-scale longitudinal neuroimaging studies, including ADNI (Alzheimer's Disease Neuroimaging Initiative) at <http://www.loni.ucla.edu/ADNI/> and the NIH MRI study of normal brain as in Ref. 1, have been or are being widely conducted to better understand the progress of neuropsychiatric disorders or the normal brain development. Compared to cross-sectional imaging studies, longitudinal imaging studies are able to identify subtle within-subject anatomical and functional changes. However, analysis of longitudinal imaging data has been hindered by the lack of advanced image processing and statistical tools for analyzing complex and correlated imaging data along with behavioral and clinical data. Recently, the cross-sectional image processing and statistical tools may be useful for longitudinal imaging data, but they are in general not optimal in power of detecting subtle changes of correlated imaging measures as discussed in Ref. 2.

The existing statistical analysis of neuroimaging data typically fits a general linear model (LM) or a simple linear mixed model (LMM) to the data from all subjects at each voxel and then generates a statistical parametric map that contains a statistic (or a p-value) at each voxel; see, for example, Refs. 3, 4 and 5. The general linear

Further author information: (Send correspondence to H.T. Zhu)
H.T. Zhu: E-mail: hzhu@bios.unc.edu, Telephone: 1 919 966 7272

Medical Imaging 2009: Image Processing, edited by Josien P. W. Pluim, Benoit M. Dawant,
Proc. of SPIE Vol. 7259, 72590F · © 2009 SPIE
CCC code: 1605-7422/09/\$18 · doi: 10.1117/12.812432

Proc. of SPIE Vol. 7259 72590F-1

model used in the neuroimaging literature was mainly developed for cross-sectional studies, in which neuroimaging measures from different subjects are assumed to be independent in Refs. 5 and 6. For instance, LM usually involves two key assumptions: that the variance of the imaging data is homogeneous across subjects and that the data conform to a Gaussian distribution at each voxel. These two assumptions are important for the valid calculation of parametric distributions in conventional tests (e.g., F test) that assess the statistical significance of parameter estimates in LM. See more discussion in Ref. 7. In particular, the methods of random field theory (RFT) that account for multiple statistical comparisons depend strongly on these assumptions of LM, as well as several additional assumptions (e.g., smoothness of autocorrelation function). Extending RFT to LM without distributional assumption and LMM requires further research. To the best of our knowledge, most existing neuroimaging software platforms including SPM, AFNI, and FSL do not have any valid methods to analyze neuroimaging data from longitudinal studies.

The primary goal of a longitudinal neuroimaging study is to characterize individual change in neuroimaging measurements (e.g., volumetric and morphometric) over time, and the covariates of interest, such as age, diagnostic status, and gender, that influence change. A distinctive feature of longitudinal neuroimaging data is that neuroimaging data have a temporal order; see more discussions in Refs. 8 and 9. Neuroimaging measurements of the same individual usually exhibit positive correlation and the strength of the correlation decreases with the time separation. Moreover, longitudinal data may provide crucial information for a causal role of time-dependent covariate (e.g., exposure) in disease process. Any image processing and statistical tools must account for the temporal correlation among the repeated measures and the spatial correlation among different voxels and regions and failure to account for the temporal and spatial correlation can lead to misleading scientific inferences (see Ref. 8).

We propose two statistical methods for the analysis of neuroimaging data from longitudinal studies. We develop generalized estimating equations for jointly modeling multivariate (or univariate) imaging measures with behavioral and clinical variables from longitudinal studies. We develop a score test statistic to test linear hypotheses of unknown parameters, such as associations between brain structure and function and covariates of interest, such as IQ, age, gene. We also use a resampling method to control for family-wise error rate. We examine the performance of the estimation and test procedures using real imaging data and our results show that the use of longitudinal statistics can dramatically increase statistical power in detecting the subtle changes of brain structure and function.

2. THEORY

We observe imaging, behavioral and clinical data from n subjects in a longitudinal study. Let \mathbf{x}_{ij} be a $q \times 1$ covariate vector, which may contain age, gender, height, and others, for the i -th subject at the j -th time point t_{ij} for $i = 1, \dots, n$ and $j = 1, \dots, m_i$. Here m_i denotes the number of time points for the i -th subject and may differ across subjects. Thus, there are a total $\sum_{i=1}^n m_i = N$ images in the study. For instance, in our longitudinal neonate study, we take the brain images of one cohort of neonates at birth, around one and two years old. Based on observed image data, we compute neuroimaging measures, denoted by $\mathbf{Y}_i = \{\mathbf{y}_{ij}(d) : d \in \mathcal{D}, j = 1, \dots, m_i\}$, across all m_i time points from the i th subject, where d represents a voxel (or a region of interest) on \mathcal{D} , a specific brain region. The imaging measure $\mathbf{y}_{ij}(d)$ at each voxel d can be either univariate or multivariate. For example, the gray matter density and signed Euclidean distance of cortical/subcortical surfaces and the fractional anisotropy are univariate measures, whereas the spherical harmonic shape description (SPHARM) of subcortical surfaces is a three dimensional MRI measure at each point.

2.1 Generalized Estimating Equations

The Generalized Estimating Equations(GEE) approach developed by Liang and Zeger can produce consistent and robust regression estimates for use in analyzing longitudinal or repeated measures (see Refs. 9 and 8). Here, we apply GEE method for jointly modeling multivariate (or univariate) imaging measures with behavioral and clinical variables from longitudinal studies. For notational simplicity, we assume that $\mathbf{y}_{ij}(d)$ are univariate imaging measures and drop d from our notation temporarily. We assume that the model for \mathbf{y}_{ij} at the j -th time point from the i -th subject is

$$E(\mathbf{y}_{ij}) = \mu_{ij} = x_{ij1}\beta_1 + \dots + x_{ijq}\beta_q = \mathbf{x}_{ij}^T\beta \quad \text{for } j = 1, \dots, m_i, \quad (1)$$

where x_{ij1} is usually set at 1, x_{ijk} ($k \geq 2$) can be chosen as time, gender, and others, and β is a $q \times 1$ vector. For instance, in our longitudinal neonate study, we are interested in studying the evolution of growth trajectories for fractional anisotropy in healthy neonatal/pediatric subjects. Specifically, the average evolution of fractional anisotropy in each voxel is described by

$$E[\text{logit}(\mathbf{y}_{ij})] = \beta_1 + \beta_2 \times t_{ij} + \beta_3 \times \text{gender}_i \quad \text{for } j = 1, 2, 3, \quad (2)$$

where $\mathbf{x}_{ij} = (1, t_{ij}, \text{gender}_i)^T$ and \mathbf{y}_{ij} denotes the fractional anisotropy (FA) value. Moreover, because FA belongs to $[0, 1]$, we take a logit-transformation of FA $\text{logit}(t) = \log(t) - \log(1 - t)$ in order to transform the FA value into the whole line. Because neuroimaging measures $(\mathbf{y}_{i1}, \dots, \mathbf{y}_{im_i})^T$ from the same subject are often positive correlated, we assume that $\mathbf{V}_i(\alpha, \sigma) = \text{Var}(\mathbf{Y}_i)$ can be decomposed as $A_i^{1/2}(\xi)R_i(\alpha)A_i^{1/2}(\xi)$, where $A_i = \text{diag}(\sigma_{i1}(\xi), \dots, \sigma_{im_i}(\xi))$, which contains the variances of \mathbf{y}_{ij} , in which ξ is a common parameter vector. In addition, the working correlation matrix $R_i(\alpha)$ represents the correlation among the m_i repeated measurements over time, where α is a vector of parameters. Commonly used working correlation structures include independence structure, exchangeable structure, autoregressive structure, and other structures. We use θ to denote all unknown parameters in the GEE method. Then, we estimate β as the solution to the GEE given by

$$\sum_{i=1}^N \mathbf{D}'_i \mathbf{V}_i(\alpha, \sigma)^{-1} (\mathbf{Y}_i - \mu_i(\beta)) = \mathbf{0}, \quad (3)$$

where $\mathbf{D}_i = \partial \mu_i / \partial \beta$ and $\mu_i = (\mu_{i1}, \dots, \mu_{im_i})^T$.

For fixed β , other parameters in ξ and α can be estimated using either additional estimating equations or moment methods (see Refs. 8 and 9). For instance, we assume the homogeneous variance for all \mathbf{y}_{ij} and then σ^2 can be consistently estimated by using $\hat{\sigma}^2 = \sum_{i=1}^n \sum_{j=1}^{m_i} (\mathbf{y}_{ij} - \mathbf{x}_{ij}^T \beta)^2 / N$. Furthermore, let $S_i(\theta)$ be an $m_i(m_i - 1)/2 \times 1$ vector of $(y_{ij} - \mathbf{x}_{ij}^T \beta)(y_{ij'} - \mathbf{x}_{ij'}^T \beta) / \sigma^2$ for $j < j'$, $j, j' = 1, \dots, m_i$. We assume that the expectation of S_i satisfies

$$E[S_i(\theta)] = C_i(\alpha). \quad (4)$$

Let $\partial_\alpha C_i(\alpha) = \partial C_i(\alpha) / \partial \alpha$. For fixed β and σ , to estimate α , we may either solve

$$U_\alpha(\beta, \alpha, \sigma) = \sum_{i=1}^n [\partial_\alpha C_i(\alpha)]^T [S_i(\theta) - C_i(\alpha)] = \mathbf{0}$$

or minimize

$$\sum_{i=1}^n [S_i(\theta) - C_i(\alpha)]^T [S_i(\theta) - C_i(\alpha)].$$

For instance, let $\alpha \in [0, 1]$ be a correlation coefficient, it is common to assume that

$$E[(y_{ij} - \mathbf{x}_{ij}^T \beta)(y_{ij'} - \mathbf{x}_{ij'}^T \beta) / \sigma^2] = \pm \exp(|t_{ij'} - t_{ij}| \log(\alpha)), \quad (5)$$

in which the sign is determined by the data themselves.

Applying GEE methods has many attractive advantages. First, this model proposed above is very flexible and free of distribution assumption. Second, the GEE estimator is consistent even we misspecified the covariance structure V_i . Third, our inferences using the empirical standard errors are robust even if our knowledge of the covariance structure is imperfect. Fourth, if correlation between \mathbf{y}_{ij} and $\mathbf{y}_{ij'}$ for $j \neq j'$ is correctly specified, the above estimation method can lead to an efficient estimation of β (see Refs. 8 and 9).

Finally, we can develop an iterative algorithm as follows:

In Step 1, we calculate the least squares estimator $\hat{\beta}_{(0)}^T = (\sum_{i=1}^n \sum_{j=1}^{m_i} \mathbf{x}_{ij} \mathbf{x}_{ij}^T)^{-1} \sum_{i=1}^n \sum_{j=1}^{m_i} \mathbf{x}_{ij} \mathbf{y}_{ij}^T$ and then we get $\hat{\sigma}_{(0)} = \sum_{i=1}^n \sum_{j=1}^{m_i} (\mathbf{y}_{ij} - \mathbf{x}_{ij}^T \hat{\beta}_{(0)})^2 / N$. Set $t = 1$.

In Step 2, we solve $\hat{\alpha}_{(t)}$ by minimizing $U_\alpha(\beta_{(t-1)}, \alpha, \sigma_{(t-1)})$ and then we solve the equations

$$\sum_{i=1}^N \mathbf{D}'_i \mathbf{V}_i(\alpha_{(t)}, \sigma_{(t)})^{-1} (\mathbf{Y}_i - \mu_i(\beta)) = \mathbf{0}$$

to obtain $\beta_{(t)}$ and get $\hat{\sigma}_{(t)} = \sum_{i=1}^n \sum_{j=1}^{m_i} (\mathbf{y}_{ij} - \mathbf{x}_{ij}^T \hat{\beta}_{(t)})^2 / N$.

In Step 3, we increase $t = t + 1$ and go back to Step 2. The algorithm stops until the absolute difference between consecutive $\hat{\theta}_{(t)}$ s is smaller than a predefined small number, say 10^{-4} .

2.2 Hypotheses and Test Statistics

In real applications, it is common to test linear hypotheses of β in order to answer various scientific questions involving a comparison of diffusion tensors across two (or more) diagnostic groups or the changes of diffusion tensors across time. See, for example, in Refs. 10, 11 and 12. We can formulate these questions as testing linear hypotheses of θ as follows:

$$H_0 : R\beta = \mathbf{b}_0 \quad \text{vs.} \quad H_1 : R\beta \neq \mathbf{b}_0, \quad (6)$$

where R is an $r \times p$ matrix of full row rank and \mathbf{b}_0 is an $r \times 1$ specified vector. For instance, if we want to study whether the rate of change is statistically significant, then $R = (0, 1, 0)$ and $\mathbf{b}_0 = (0)$.

We test the null hypothesis $H_0 : R\beta = \mathbf{b}_0$ using a score test statistic S_n defined by

$$S_n = \partial_\mu L_n^T \hat{\mathbf{\Gamma}}_{\mu\mu}^{-1} \partial_\mu L_n, \quad (7)$$

where $\mu = R\beta$, $\partial_\mu L_n = n^{-1/2} \sum_{i=1}^n U_{i,\mu}(\tilde{\theta})$ and $\hat{\mathbf{\Gamma}}_{\mu\mu} = n^{-1} \sum_{i=1}^n [U_{i,\mu}(\tilde{\theta}) - \bar{U}_\mu(\tilde{\theta})][U_{i,\mu}(\tilde{\theta}) - \bar{U}_\mu(\tilde{\theta})]^T$, in which $\bar{U}_\mu(\tilde{\theta}) = \sum_{i=1}^n U_{i,\mu}(\tilde{\theta})/n$ and $\tilde{\theta}$ maximizes $\ell_n(\theta)$ under the linear constraints $R\beta = \mathbf{b}_0$. Solving $\tilde{\theta}$ can be done using the Lagrange multiplier method. Exact formula for $\partial_\mu L_n$ can be found in Ref. 13. Particularly, $\hat{\mathbf{\Gamma}}_{\mu\mu}$ is an estimator of the covariance matrix of $\partial_\mu L_n$ and $\hat{\mathbf{\Gamma}}_{\mu\mu}^{-1/2} \partial_\mu L_n$ is approximately a Gaussian random vector having a zero mean vector and an identity covariance matrix under $H_{0,\mu}$. Thus, the statistic S_n is approximately distributed as $\chi^2(r)$, a chi-square distribution with r degrees of freedom (see Ref. 14).

2.3 Test Procedure

In neuroimaging studies, we are often interested in testing the null hypotheses in all voxels d of the region \mathcal{D} under study. In the following, we will introduce voxel d into all mathematical notation, if necessary. To control the family-wise error rate, we consider the maxima of the score test statistics, defined by $S_{\mu,\mathcal{D}} = \max_{d \in \mathcal{D}} S_n(d)$.

To use $S_{\mu,\mathcal{D}}$ as test statistics, we need to know its distributions under the null hypothesis across all voxels of the relevant region. We present a test procedure that is based on the resampling method to approximate the distribution of $S_{\mu,\mathcal{D}}$ (see Refs. 13, 15, 6, and 16). This procedure is essentially a wild bootstrap method for hypothesis test. The test procedure is implemented as follows:

Step 1: At each voxel d , calculate the score test statistic $S_n(d)$ given in (7), and then compute $S_{\mu,\mathcal{D}} = \max_{d \in \mathcal{D}} S_n(d)$;

Step 2: Generate a random sample $\{\eta_i^{(s)} : i = 1, \dots, n\}$ from the distribution F , which is defined by $\eta_i^{(s)} = \pm 1$ with equal probability.

Step 3: At each voxel d of \mathcal{D} , calculate $S_n(d)^{(s)} = \partial_\mu L_n(d)^{(s)T} [\tilde{\mathbf{\Gamma}}_{\mu\mu}(d)]^{-1} \partial_\mu L_n(d)^{(s)}$ and then compute $S_{\mu,\mathcal{D}}^{(s)} = \max_{d \in \mathcal{D}} S_n(d)^{(s)}$, where $\partial_\mu L_n(d)^{(s)} = n^{-1/2} \sum_{i=1}^n U_{i,\mu}(\tilde{\theta}, d) \eta_i^{(s)}$ and $\tilde{\mathbf{\Gamma}}_{\mu\mu}(d) = n^{-1} \sum_{i=1}^n [U_{i,\mu}(\tilde{\theta}, d)]^{\otimes 2}$.

Step 4: Repeat Steps 2-3 S times to obtain $\{S_{\mu,\mathcal{D}}^{(s)} : s = 1, \dots, S\}$. Finally, the p value of $S_{\mu,\mathcal{D}}$ is approximated by $p_{\mu,\mathcal{D}} = S^{-1} \sum_{s=1}^S 1(S_{\mu,\mathcal{D}}^{(s)} \geq S_{\mu,\mathcal{D}})$. If $p_{\mu,\mathcal{D}}$ is smaller than a pre-specified value α , say 0.05, then we reject that the null hypothesis $H_0 : R\beta = \mathbf{b}_0$ holds across all voxels of the region \mathcal{D} .

Step 5: Calculate the p -value of $S_n(d)$ at each voxel d of the region according to $p(d) \approx S^{-1} \sum_{s=1}^S 1(S_n(d)^{(s)} \geq S_n(d))$.

Step 6: Calculate the corrected p -value of $S_n(d)$ at each voxel d of the region using $p_D(d) \approx S^{-1} \sum_{s=1}^S 1(S_{\mu,D}^{(s)} \geq S_n(d))$.

We note several advantages of using the resampling method in the above test procedure. Computationally, the above procedure only requires the computation of $\hat{U}_{i,\mu}(\hat{\theta}, d)$ once and the repeated calculation of $S_n(d)^{(s)}$. Thus, because the permutation method involves refitting the model for each simulated data, the proposed test procedure is computationally much more efficient than the permutation method. Moreover, permutation methods require "complete exchangeability" condition, which is a very strong assumption particularly in the presence of clinical variables as pointed out by Refs. 13 and 17.

3. LONGITUDINAL NEONATE STUDY OF BRAIN DEVELOPMENT

The aim of this study was to quantitatively investigate longitudinal changes in the microstructural integrity of white matter at 0-2 years. White matter as one of the three main solid components of the central nervous system (CNS) consists of bundles of myelinated nerve cell processes (or axons), which connect various gray matter areas (the locations of nerve cell bodies) of the brain to each other, and carry nerve impulses between neurons. The white matter is also important for transferring messages between different areas of gray matter within the CNS.

The response of interest was the fractional anisotropy (FA) images calculated from diffusion tensor images (DTIs) from 16 neonatal/pediatric subjects. Each of these sixteen subjects was longitudinally scanned to acquire T1, T2, DTI and functional connectivity MRI, when they were at about two weeks, 1 year old and 2 years old. A single shot EPI DTI sequence (TR/TE=5400/73 msec) with eddy current compensation was used to acquire DTIs. Diffusion tensor acquisition scheme includes 18 repeated measures of six non-collinear directions, (1,0,1), (-1,0,1), (0,1,1), (0,1,-1), (1,1,0), and (-1,1,0) at a b-value of 1000 s/mm² and a b=0 reference scan. Forty-six contiguous slices with a slice thickness of 2 mm covered a field of view (FOV) of 256 mm² with an isotropic voxel size of 2 × 2 × 2 mm³. High resolution T1 weighted (T1W) images were acquired using a 3D MP-RAGE sequence. A weighted least square estimation method as investigated in Refs. 18 and 19 was used to construct diffusion tensors. Based on the constructed diffusion tensors, we calculated the FA images.

We used a two-step image registration approach to spatially normalize all diffusion tensor images. In Step 1, we chose the T1 weighted image of an arbitrarily chosen forty-one year old female healthy subject as the template, and then we used a B-spline model based on a bi-directional elastic registration method to coregister all other subjects' T1 weighted images to the template. To minimize biases induced by the choice of the template, the symmetry between the image pair to be co-registered was ensured through enforcing the consistency between the forward (from subject X to the template) and backward (from the template to subject X) transformations. In Step 2, we aligned each subject's diffusion tensor images to his/her own anatomical T1 weighted images using a 6 parameter rigid registration tool in FSL 3.2 (Analysis Group, FMRIB, Oxford, UK). In Step 3, we applied the tensor reorientation matrix at each voxel to approximating the transformation matrix (from T1 registration) through a singular value decomposition (SVD) as used in Ref. 20 and 21.

We will use the above FA images to demonstrate the importance of using both longitudinal statistical tools in Section 2. For purpose of comparison, we will particularly select two test statistics (score and Wald test statistics) with/without considering the longitudinal design. By explicitly incorporating longitudinal information, we applied the GEE method (2) with the autoregressive working correlation described above at each voxel. We detected and localized the statistical significance of linear time on the development of FA at all voxels. Here, $R = (0, 1, 0)$ and $\mathbf{b}_0 = (0)$ for the hypothesis on linear time and then we calculated the score and Wald test statistics. Similarly, without incorporating longitudinal information, we treated all observations from the same subject as independent and calculated the score and Wald test statistics under the cross-sectional design. Finally, we have four test statistics and their associated results.

The use of longitudinal statistics can dramatically increase the statistical power in detecting the changes of brain structures and functions. The maxima of the score test values increases from 111.06 for the cross-sectional

design to 980.18 for the longitudinal design (Table 1). The mean and median of all score test values are, respectively, 29.77 and 15.84 for the longitudinal design, whereas those for the cross-sectional design are, respectively, 13.20 and 8.93. We observed an obvious shift towards the right in the empirical cumulative distribution function (CDF) for the longitudinal design as compared with that for the cross-sectional one (Fig. 1 (A)). We also calculated the difference of the score test values for the longitudinal and cross-sectional designs at each voxel of the template. The mean and median of these differences across the whole template are, respectively, 16.56 and 5.86 (Table 1). Most quartiles of these differences are greater than zero and the longitudinal statistics detect more voxels as significant, indicating that incorporating the longitudinal design can dramatically increase the sensitivity in detecting any differences (Fig. 1(B) and (C)). We observed similar scenario for the Wald test statistics.

We calculated the autocorrelation coefficients, defined as $\text{sign} \times \alpha$, across all voxels (Fig. 2 (A)). The values of most autocorrelation coefficients in the white matter are greater than zero with their mean at 0.298 (Fig. 2 (A) and (B)). There is a positive correlation between the autocorrelation coefficients and the differences of the test values for the longitudinal and cross-sectional designs (Fig. 2 (C)).

We applied our test procedure to calculate the p -value $p(d)$ and corrected p -value $p_D(d)$ at each voxel. Color-coded maps of p -values using either the uncorrected $p(d)$ alone or the corrected $p_D(d)$ indicated dramatic time effect within the brain. For instance, all white matter regions are detected as significant (Fig. 3). We observed the significant increase in the number of significant voxels using the longitudinal design compared with the cross-sectional design. Our results reveal the significant change of the water diffusivity in the white matter and gray matter regions in the first three years period since birth (Fig. 3).

Table 1. Comparisons of score and wald test statistics for longitudinal neonate study of brain development. LST denotes longitudinal statistics toolbox, CST denotes cross-sectional statistics toolbox, Score denotes the score test statistics and Wald denotes the wald test statistics.

Statistics	Toolbox	minimum	maximum	mean	median	std
Score	LST	3.7450e-9	980.18	29.77	15.84	42.37
Score	CST	4.61e-10	111.06	13.20	8.93	13.69
Score	LST-CST	-14.66	935.42	16.56	5.86	31.28
Wald	LST	3.7450e-9	1363	30.62	15.83	44.81
Wald	CST	4.6073e-10	612.28	16.96	9.44	23.2060
Wald	LST-CST	-117.88	1274.9	13.65	4.69	27.3896

DISCUSSION

We have developed a set of longitudinal statistics toolbox, called LSTGEE, including GEE, a score test statistic and a resampling method and their association with a set of covariates. The GEE avoids any parametric assumptions regarding responses of interest. We have examined our estimation and test procedures for carrying out statistical inference using real neuroimaging data. We have successfully applied the proposed statistics to our longitudinal neonate data and obtained the growth curves of FA across the whole brain. Our statistical results clearly show significant growth of brain tissues in early brain development. Our results also show that strong positive correlation exists between images at different time from the same subject. We expect a great clinical application of this novel statistical technique. A toolbox for implementing LSTGEE is available upon request and a complete version will be released on <http://www.nitrc.org/> soon.

ACKNOWLEDGMENTS

This work was supported in part by NSF grants SES-06-43663 and BCS-08-26844 and NIH grants UL1-RR025747-01 and R21AG033387 to Dr. Zhu, NIH grants 50 MH064065, R01 MH070890, and RO1 HD053000 to Dr. Gilmore, NIH grants R01NS055754 and R01EB5-34816 to Dr. Lin, and NIH grant 1R01EB006733 to Dr. Shen.

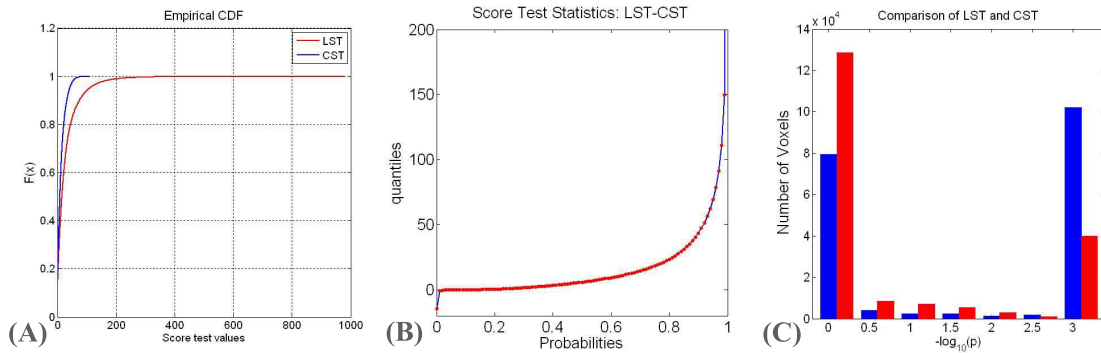


Figure 1. Results from longitudinal neonate study of brain development. Panel (A) presents the empirical cumulative distribution functions (CDFs) for the longitudinal and cross-sectional designs. Panel (B) presents the plot of probabilities against quartiles based on the difference of the score test values for the longitudinal and cross-sectional designs. Panel (C) presents the numbers of voxels, whose $-\log_{10}(p_D(d))$ values are in seven intervals $[k, k + 0.5)$ for $k = 0, 0.5, \dots, 3$, in which $p_D(d)$ are based on the score test values for either the longitudinal design (blue) or the cross-sectional design (red).

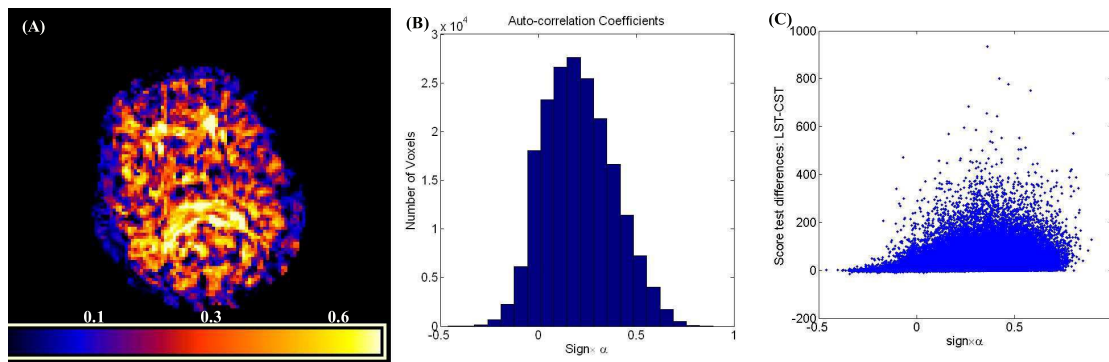


Figure 2. Results from longitudinal neonate study of brain development. Panel (A) presents a color-coded map of autocorrelation coefficient values in a selected slice of the template and the color scale reflects the magnitude of values with black to blue representing smaller values (-0.5-0.1) and red to white representing larger values (0.3-0.8). Panel (B) presents the histogram of the autocorrelation coefficients. Panel (C) presents the plot of the autocorrelation coefficient against the difference between the score test values for the longitudinal and cross-sectional design.

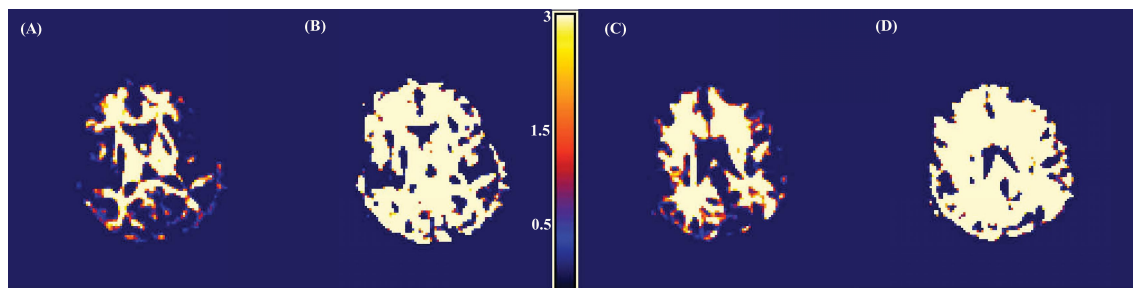


Figure 3. Results from Longitudinal Neonate study of brain development. Panels (A)-(D) present color-coded maps of adjusted $(b) - \log_{10}(p_D(d))$ values in two selected slices of the template and the color scale reflects the magnitude of values of $-\log_{10}(p_D(d))$ with black to blue representing smaller values (0-1) and red to white representing larger values (1.88-3). Panels (A) and (C) are based on the cross-sectional design, whereas panels (B) and (D) are based on the longitudinal design.

REFERENCES

- [1] Almli, C. R., Rivkin, M. J., McKinstry, R. C., and Group., B. D. C., "The nih mri study of normal brain development (objective-2): newborns, infants, toddlers, and preschoolers," *IEEE Transactions on Medical Imaging* **35**, 308–325 (2007).
- [2] Xue, Z., Shen, D., and Davatzikos, C., "Classic: consistent longitudinal alignment and segmentation for serial image computing," *Neuroimage* **30**, 388–399 (2006).
- [3] Friston, K. J., [*Statistical Parametric Mapping: the Analysis of Functional Brain Images*], Academic Press, London (2007).
- [4] Lau, J. C., Lerch, J. P., Sled, J. G., Henkelman, R. M., Evans, A. C., and Bedell, B. J., "Longitudinal neuroanatomical changes determined by deformation-based morphometry in a mouse model of alzheimer's disease," *NeuroImage* **42**, 19–27 (2008).
- [5] Worsley, K. J., Taylor, J., Tomaiuolo, F., and Lerch, J., "Unified univariate and multivariate random field theory," *NeuroImage* **23**, 189–195 (2004).
- [6] Zhu, H. T., Ibrahim, J. G., Tang, N. S., Hao, X. J., Bansal, R., and Peterson, B. G., "A statistical analysis of brain morphology using wild bootstrapping," *IEEE Trans. on Medical Imaging* **26**, 954–966 (2007).
- [7] Nichols, T. and Hayasaka, S., "Controlling the familywise error rate in functional neuroimaging: a comparative review," *Stat Methods Med Res* **12**, 419–426 (2003).
- [8] Diggle, P., Heagerty, P., Liang, K. Y., and Zeger, S., [*Analysis of Longitudinal Data (2nd ed.)*], Oxford University Press, New York (2002).
- [9] Liang, K. and Zeger, S., "Longitudinal data analysis using general linear models," *Biometrika* **73**, 13–22 (1986).
- [10] Lepore, N., Brun, C. A., Chou, Y., Chiang, M., Dutton, R. A., Hayashi, K. M., Luders, E., Lopez, O. L., Aizenstein, H. J., Toga, A. W., Becker, J. T., and Thompson, P. M., "Generalized tensor-based morphometry of hiv/aids using multivariate statistics on deformation tensors," *IEEE Transactions in Medical Imaging* **27**, 129–141 (2008).
- [11] Schwartzman, A., *Random Ellipsoids and False Discovery Rates: Statistics for Diffusion Tensor Imaging Data*, PhD thesis, Stanford University, California (July 2006).
- [12] Whitcher, B., W. J. H. N. and Tuch, D. S., "Statistical group comparison of diffusion tensors via multivariate hypothesis testing," *Magnetic Resonance in Medicine* **57**, 1065–1074 (2007).
- [13] Lin, D. Y., "An efficient monte carlo approach to assessing statistical significance in genomic studies," *Bioinformatics* **6**, 781–787 (2005).
- [14] Lehmann, E. L. and Romano, J. P., [*Testing Statistical Hypotheses*], Springer-Verlag, New York (2006).
- [15] Zhu, H. T., Li, Y. M., Tang, N. S., Bansal, R., Hao, X. J., Weissman, M. M., and Peterson, B. G., "Statistical modelling of brain morphological measures within family pedigrees," *Statistica Sinica* **18**, 1554–1569 (2008).
- [16] Kosorok, M. R., "Bootstraps of sums of independent but not identically distributed stochastic processes," *J. Multivariate Anal.* **84**, 299–318 (2003).
- [17] Dudoit, S., Shaffer, J. P., and Boldrick, J. C., "Multiple hypothesis testing in microarray experiments," *Statist. Sci.* **18**, 71–103 (2003).
- [18] Basser, P. J., Mattiello, J., and LeBihan, D., "Estimation of the effective self- diffusion tensor from the nmr spin echo," *Journal of Magnetic Resonance Ser. B* **103**, 247–254 (1994).
- [19] Zhu, H. T., Zhang, H. P., Ibrahim, J. G., and Peterson, B. G., "Statistical analysis of diffusion tensors in diffusion-weighted magnetic resonance image data (with discussion)," *Journal of the American Statistical Association* **102**, 1085–1102 (2007).
- [20] Alexander, D. C., Pierpaoli, C., Basser, P. J., and Gee, J. C., "Spatial transformations of diffusion tensor magnetic resonance images," *IEEE Transactions on Medical Imaging* **20**, 1131–1139 (2001).
- [21] Xu, D., Mori, S., Shen, D., van Zijl, P., and Davatzikos, C., "Spatial normalization of diffusion tensor fields," *Magnetic Resonance in Medicine* **50**, 175–182 (2003).