

# Asymmetric Bias in User Guided Segmentations of Brain Structures

Martin Styner<sup>a,b</sup>, Rachel G Smith<sup>b</sup>, Michael M Graves<sup>b</sup>, Matthew W Mosconi<sup>b</sup>, Sarah Peterson<sup>b</sup>, Scott White<sup>b</sup>, Joe Blocher<sup>b</sup>, Mohammed El-Sayed<sup>c</sup> Heather C Hazlett<sup>b</sup>

<sup>a</sup>Department of Psychiatry, University of North Carolina, Chapel Hill NC, USA;

<sup>b</sup>Department of Computer Science, University of North Carolina, Chapel Hill NC, USA

<sup>c</sup>Mansoura University, Mansoura, Egypt

## ABSTRACT

Brain morphometric studies often incorporate comparative asymmetry analyses of left and right hemispheric brain structures. In this work we show evidence that common methods of user guided structural segmentation exhibit strong left-right asymmetric biases and thus fundamentally influence any left-right asymmetry analyses. We studied several structural segmentation methods with varying degree of user interaction from pure manual outlining to nearly fully automatic procedures. The methods were applied to MR images and their corresponding left-right mirrored images from an adult and a pediatric study. Several expert raters performed the segmentations of all structures. The asymmetric segmentation bias is assessed by comparing the left-right volumetric asymmetry in the original and mirrored datasets, as well as by testing each sides volumetric differences to a zero mean standard t-tests.

The structural segmentations of caudate, putamen, globus pallidus, amygdala and hippocampus showed a highly significant asymmetric bias using methods with considerable manual outlining or landmark placement. Only the lateral ventricle segmentation revealed no asymmetric bias due to the high degree of automation and a high intensity contrast on its boundary. Our segmentation methods have been adapted in that they are applied to only one of the hemispheres in an image and its left-right mirrored image. Our work suggests that existing studies of hemispheric asymmetry without similar precautions should be interpreted in a new, skeptical light.

Evidence of an asymmetric segmentation bias is novel and unknown to the imaging community. This result seems less surprising to the visual perception community and its likely cause is differences in perception of oppositely curved 3D structures.

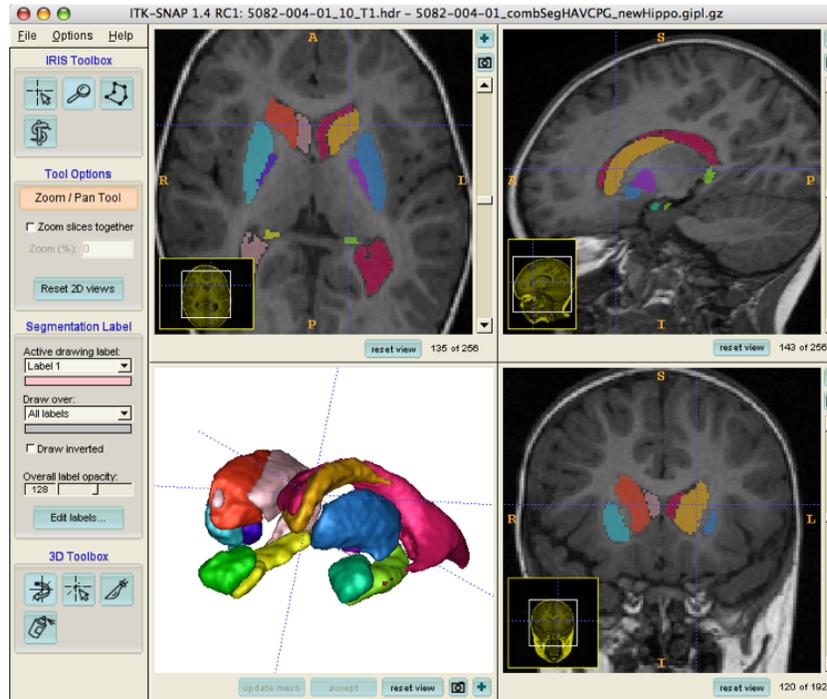
## 1. INTRODUCTION

Quantitative assessment of pathological and normal brain morphometry difference in brain structures is commonly determined via a segmentation processes applied to MRI or CT imagery. Such morphometric studies often incorporate comparative asymmetry analyses of left and right hemispheric brain structures, such as the hippocampus, thalamus, lateral ventricles or cortical parcellations.<sup>1-8</sup> Most of these studies are based on structural segmentations with varying degree of user interaction ranging from pure manual traces to initialization with a few landmarks. In this manuscript we investigate the influence of a possible user induced asymmetric bias in such studies. Few methods are fully automatic<sup>9,10</sup> and these methods are not studied here.

In this work we show evidence that common methods of user guided structural segmentation exhibit strong left-right asymmetric biases and thus fundamentally influence such left-right morphometric asymmetry analyses. In the next section we will discuss the variety of segmentation methods studied for determining volumetry of brain structures, as well as the testing methodology. This is followed by our results in two separate asymmetry studies and discussion section.

## 2. METHODS

In this section we describe segmentation methods with varying degree of user interaction for the segmentation of deep brain structures, as well our testing methodology.



**Figure 1.** The ITK SNAP segmentation tool is one of the segmentation tools employed in our asymmetric bias study. Semi-automatic caudate segmentation, as well as purely manual segmentations of the putamen and globus pallidus were performed using this tool. Here a visualization of the segmentations of all brain structures studied in this manuscript from a selected pediatric dataset is shown.

## 2.1. Segmentation Methods

We studied several structural fully 3D methods with varying degree of user interaction for the segmentation of the lateral ventricles, caudate, putamen, globus pallidus, hippocampus and amygdala. The most time demanding and user interactive method is pure manual outlining by expert raters. This method was chosen for the segmentation of the amygdala, putamen and globus pallidus. The putamen and the globus pallidus are closely situated next to each other and often no clear separation is evident for the full boundary of the structures in the grayscale image. Similarly, the amygdala has no clear grayscale boundaries both anteriorly and posteriorly. Human expert decisions of the structural boundaries thus influence the segmentation quite strongly. These segmentations are performed using the ITK SNAP segmentation \* tool by outlining the structures in all three orthogonal slice directions using the polygon drawing tool (see Figure 1). The segmentations are checked for consistency across the three orthogonal slice directions, as well as in the visualization of the 3D surfaces.

Limited user interaction is employed for the caudate segmentation. Here, an initial approximation of the segmentation is obtained via the deformable contour tool in ITK SNAP<sup>11</sup>. The contouring tool is based on a narrow band levelset implementation of a geodesic active snake segmentation. This initial segmentation is adapted only in the regions abutting anteriorly at the nucleus accumbens via manual outlining, as there is no clear separation between the caudate and nucleus accumbens visible in the grayscale image. As for the case of the purely manually segmented structures, the caudate segmentations are checked for consistency across the three orthogonal slice directions, as well as in the visualization of the 3D surfaces.

The user interaction is also limited in the hippocampus segmentation where manual landmark placement is necessary as an initialization. The hippocampal landmarks initializes and constrains the deformation field

Email: martin.styner@ieee.org, WWW: www.ia.unc.edu

\*ITK SNAP can be downloaded at <http://www.itksnap.org>

of a diffeomorphic, fluid deformable registration that maps a prior template atlas with known hippocampus surface<sup>7,12,13</sup> to the image. The computed deformation map is then applied to the atlas hippocampus surface and the user reviews the mapped surface in the 3D orthogonal slice views. The user then has the choice to adjust the landmarks and recompute the deformation process. This landmark adjustment is performed iteratively until the segmentation receives the user’s approval, which normally takes just a few iterations.

The segmentation of the lateral ventricle is nearly fully automatic. First an automatic, atlas based tissue segmentation procedure<sup>14,15</sup> is applied to the grayscale image and classifies it into cerebro-spinal fluid (CSF), white matter and gray matter. The ventricle segmentation is then performed on the probabilistic CSF map rather than the original grayscale image. As for the caudate, the segmentation is computed using the deformable contour tool in ITK SNAP. The user interaction is limited to providing a seed point and stopping the deformation upon convergence. Only in if the deformable contour leaks into the third ventricle, small user interaction is additionally necessary to cutoff the third ventricle at its small connection.

At <http://www.psychiatry.unc.edu/autismresearch/mri/roiprotocols.htm>, all segmentation protocols are described in detail with the exact steps necessary in the expert decisions for manual boundaries. All segmentation protocols were performed by trained expert raters in our studies, who showed reliability in a separate testing datasets via intra-class correlation evaluation.

## 2.2. Testing Methodology

In order to study the possible asymmetric segmentation bias, the segmentation methods were applied to a set of MR images and their corresponding left-right mirrored images. The MR images were selected from different studies and were segmented if possible by multiple raters. The asymmetric segmentation bias is then assessed by the following comparisons:

- Asymmetry  $A$  measured by relative left-right volumetric difference:  $A = 2 \cdot \frac{L-R}{L+R}$ , where  $L$  is the volume of the left hemispheric structure and  $R$  is the volume of the right hemispheric structure. The asymmetry value should be consistent for the original and mirrored dataset and can be assessed using a standard paired t-test.
- Each sides volumetric difference between the segmentations (original-vs-mirrored difference): If there is no asymmetric bias present then the volumetric difference of one side’s segmentation minus its mirrored counterpart should be Gaussian distributed with zero mean. An asymmetric bias can thus be easily detected via standard Students t-test of these differences. The result of the left and right hemispheric volumetric measurements can be pooled for this analysis. In our plots in this manuscript, we show the relative volumetric difference  $= 2 \cdot \frac{V_{orig} - V_{mirr}}{V_{orig} + V_{mirr}}$ .

For the analysis presented below, the volumetric measurements in each brain hemisphere are labeled in regard to their original correct left/right label. E.g., the left hippocampus volume in a left-right mirrored dataset is the volume of the structure segmented by the user as the right hippocampus.

## 3. RESULTS

### 3.1. Testing Datasets

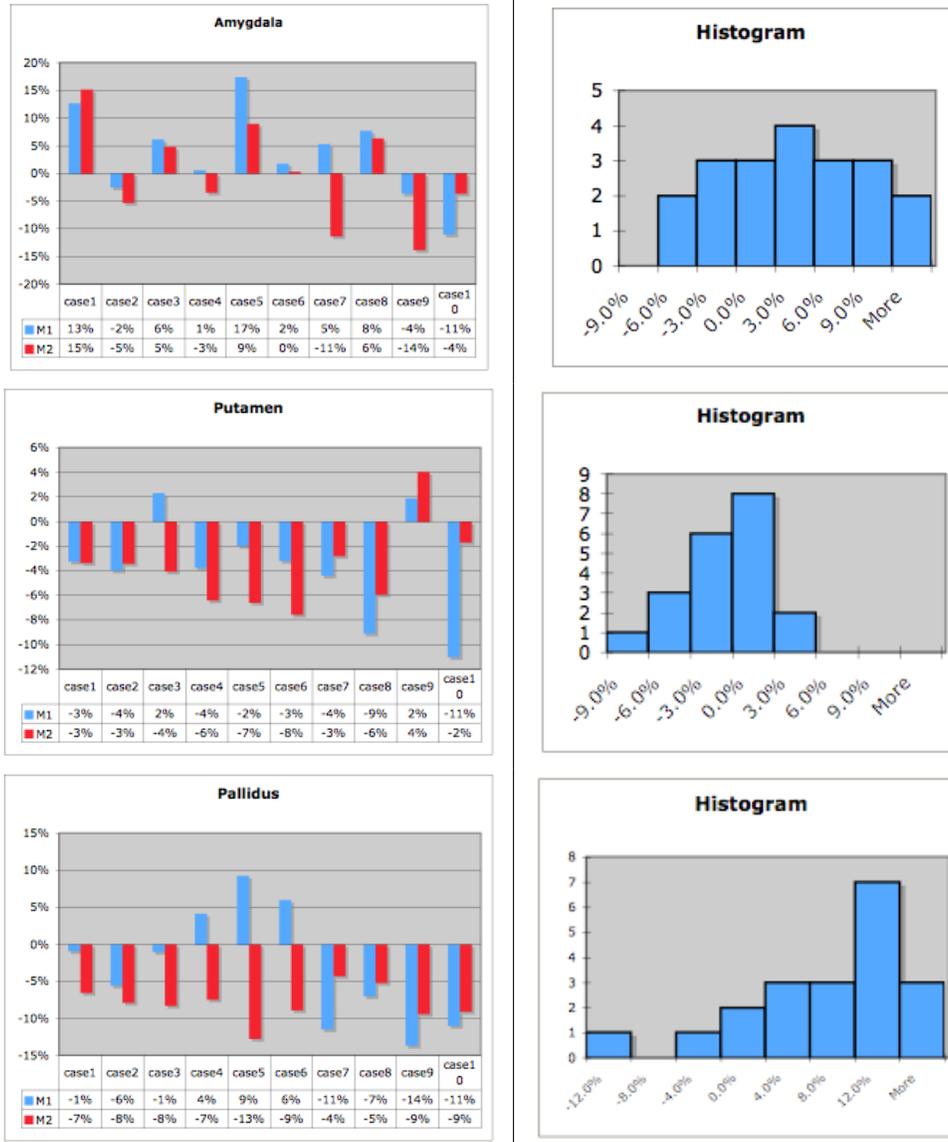
MR images were randomly selected from both an adult (5 images, all healthy controls) and a pediatric study (10 images, healthy controls and autistic subjects between ages 2 and 4). The images were mirrored at their left-right axis and all datasets were segmented in random order.

5 different expert raters (JB, SW, MES, MM, SP) performed the hippocampal segmentations, whereas the other structures were segmented each by a single expert (lateral ventricle: RS, amygdala: SP, caudate, putamen, globus pallidus: MG). All raters are right handed and have been trained with their reliability tested (JB, SW, MES for adult data, MM, SP, RS, MG for pediatric data) using intra-class correlation coefficients: caudate  $> 0.96$ , hippocampus  $> 0.99$ , lateral ventricles  $> 0.99$ , putamen  $> 0.93$ , globus pallidus  $> 0.83$ , amygdala  $> 0.89$ .

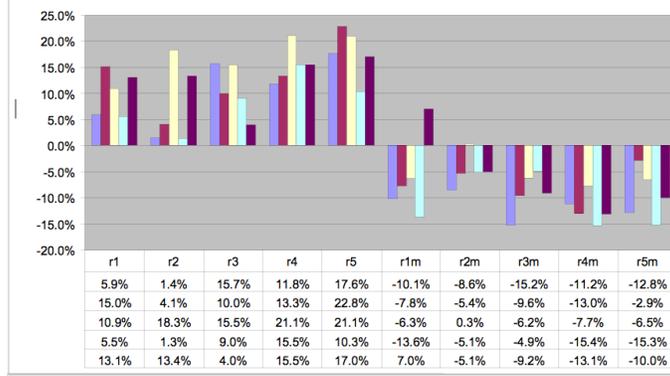
The hippocampus segmentation was tested on both the adult and pediatric study, whereas the other structures were only tested on the pediatric data, as the expert raters were only trained on such datasets.

### 3.2. Bias in Pure Manual Segmentation

For the purely manually segmented structures, we found evidence for an asymmetric segmentation bias. Specifically, we found a clear asymmetric bias using the original-vs-mirrored difference (amygdala  $p = 0.05$ , putamen  $p < 0.0001$ , globus pallidus  $p = 0.0002$ ). For the asymmetry value the putamen and amygdala showed a non-significant trend ( $p = 0.13$  and  $p = 0.12$ ), where as the globus pallidus did not show any differences ( $p > 0.5$ ). Figure 2 shows the results in the respective plots. The measurements further show that the segmentation variability is quite high for these manual segmentations compared to those of more automated methods described further below.



**Figure 2.** Asymmetric bias analysis in pure manual segmentations. Left column: volumetric left-right asymmetry measurements of the amygdala, putamen and globus pallidus in the pediatric study (original images = M1, left-right mirrored images = M2). Right column: The histogram plot of the original-vs-mirrored volumetric difference is non-symmetric at the origin (for putamen and globus pallidus) and/or its mean (mean amygdala = 1.61%, mean putamen = -3.07%, mean globus pallidus = 1.39) different from zero.



**Figure 3.** Volumetric left-right asymmetry measurements  $A = 2 \cdot \frac{L-R}{L+R}$  of the hippocampus by 5 expert raters on 5 randomly selected cases (r1-r5) in an adult study and on their left-right mirrored images (r1m-r5m). A clear left-right asymmetry ( $R_L L$ ) is inverted when mirroring the dataset ( $L_L R$ ).

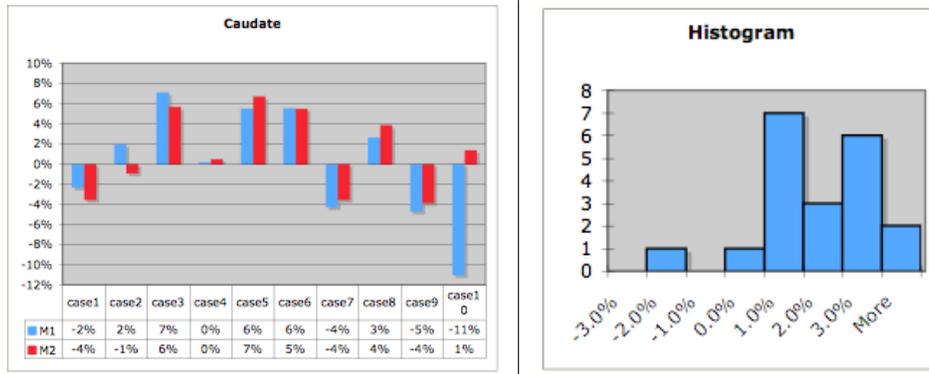
### 3.3. Bias in Landmark Initialized Hippocampus Segmentation

The strongest asymmetric segmentation bias was found in the hippocampus segmentation (see Figure 3 for a plot of the adult data). The relatively strong left-right volumetric asymmetry is inverted when mirroring the dataset. The paired t-test that test for the presence of an asymmetric bias results in p-values of  $p = 10^{-7}$  for the adult study and  $p = 10^{-8}$  for the pediatric study. Due to clear asymmetric bias we did not test additionally the original-vs-mirrored difference.

We assessed whether the asymmetric segmentation bias was potentially introduced by the template deformation methodology for the hippocampus segmentation. When given the same landmarks to the original and mirrored datasets (the corresponding landmarks were also mirrored), the segmentation results were virtually the same without any detectable bias. Further, we could not find any bias in regard to the template selection as we confirmed the results with a separate hippocampal template. The deformation methodology itself proved thus to be very stable to left-right mirroring the datasets, but not its user interaction.

### 3.4. Bias in Semi-automatic Caudate Segmentations

In the caudate segmentations, we also found a clear asymmetric bias using the original-vs-mirrored difference ( $p < 0.0001$ ), but not for the asymmetry value ( $p > 0.4$ ). Figure 4 shows the results in the respective plots.



**Figure 4.** Left: volumetric left-right asymmetry measurements of the caudate in the pediatric study (original images = M1, left-right mirrored images = M2). Right: The histogram plot of the original-vs-mirrored volumetric difference is non-symmetric at the origin and its mean (mean = 1.39%) different from zero.

### 3.5. Bias in Semi-automatic Lateral Ventricle Segmentations

The lateral ventricle segmentation did not show any asymmetric bias in our studies, both when testing the asymmetry measurement ( $p > 0.2$ ) and the original-vs-mirrored difference ( $p > 0.99$ ). In Figure 5, the left-right asymmetry is not consistently across the cases, as was the case for the hippocampus. Also, no clear difference between the results of the original and left-right mirrored images is visible.



**Figure 5.** Left: volumetric left-right asymmetry measurements of the lateral ventricles in the pediatric study (original images = M1, left-right mirrored images = M2). No left-right asymmetry is consistently visible across the cases, nor is there a clear difference between the results of the original and left-right mirrored images. Right: The histogram plot of the original-vs-mirrored volumetric difference seems symmetric at the origin and its mean (mean = 0.01%) is close to zero.

## 4. DISCUSSION

The absence of an asymmetric bias for the lateral ventricle is due to the high degree of segmentation automation. It is thus surprising to note that the strongest asymmetric bias is visible in the hippocampus segmentation, which also has limited user interaction. The differences in local contrast of the two structures, the different type of user-interaction as well as the difference in operating on a probability map compared to a grayscale MR image may be the main reasons for this discrepancy. The lateral ventricles shows perfect contrast in the probability maps, whereas the hippocampus contrast is quite low at the anterior, superior and posterior parts of its boundary.

While both hippocampus and lateral ventricle segmentations show small intra and inter-rater variability, all manually segmented structures have larger variability. This segmentation variability reduces the sensitivity to detect the asymmetric segmentation bias, which was nevertheless very clearly detected in our studies.

In our studies, the original-vs-mirrored volumetric differences was a more sensitive measurement for the presence of an asymmetric bias than the asymmetry value itself, probably due to the pooling of the left and right hemispheric measurements and the resulting gain in sample size by a factor of two.

The evidence of an asymmetric segmentation bias in our studies is novel and unknown to the imaging community. Based on our communications with experts in the visual perception community, this result seems less surprising to them and its likely cause is differences in perception of oppositely curved 3D structures. To our knowledge no specific studies have been performed on this topic and so currently no definite answer on the origin of the asymmetric segmentation bias can be given. In our next steps in this matter, we plan to analyze our current data using local shape analysis in order to localize any consistent asymmetric biases. This should result in initial evidence what perception bias is likely to cause the results reported here.

We ourselves stumbled upon this asymmetric bias when we analyzed a hippocampus dataset in a schizophrenia study with the same deformable registration method used in this manuscript. During the pre-processing steps the grayscale images were inadvertently left-right mirrored. This mirroring was detected at later stage and

corrected. We realized though that the segmentation of the mirrored images and those of the original images showed inconsistent, even inverted asymmetry. This led to the initial adult hippocampus study and subsequently to the pediatric study of the other brain structures.

In general, segmentation methodology can be adapted to be independent of this bias using two possible solutions: a) slice by slice segmentation can be applied on sagittal slices with a medial-to-lateral presentation; b) any segmentation method can be applied to only one of the hemispheres in both the original image and in its left-right mirrored image; a simple reverse mirroring joins both segmentations in the original image. In our studies, we are employing the latter method. Limiting user interaction can also reduce this bias, as our lateral ventricle segmentation results have shown.

## 5. CONCLUSION

We have presented in this paper a first set of studies investigating an asymmetric segmentation bias for left and right hemispheric brain structures. Through simple changes in the segmentation practices, the asymmetric bias can and should be avoided. The results of our studies further suggest that existing studies of left-right hemispheric asymmetry without such precautions have to be interpreted in a new, skeptical light.

## 6. ACKNOWLEDGMENT

We would like to thank Dale Purves, Duke University, and Donald Mershon, North Carolina State University, for insightful discussions about the origin of this bias. This research has been/is supported by the following grants: UNC Neurodevelopmental Disorders Research Center HD 03110, the NIH Conte Center MH064065, and NIH RO1 MH61696 and NIMH MH64580.

## REFERENCES

1. K. L. Narr, R. M. Bilder, E. Luders, P. M. Thompson, R. P. Woods, D. Robinson, P. R. Szeszko, T. Dimtcheva, M. Gurbani, and A. W. Toga, "Asymmetries of cortical shape: Effects of handedness, sex and schizophrenia," *NeuroImage* **34**, pp. 939–48, Feb. 2007.
2. G. Bersani, A. Quartini, O. Piperopoulos, A. Iannitelli, M. Paolemili, D. Pucci, C. D. Biasi, G. Gualdi, and P. Pancheri, "Brain abnormalities in schizophrenia. a qualitative comparative study of schizophrenic patients and control individuals assessed by magnetic resonance imaging," *Journal of neuroradiology. Journal de neuroradiologie* **33**, pp. 152–7, June 2006.
3. B. K. Puri, A. J. Richardson, A. Oatridge, J. V. Hajnal, and N. Saeed, "Cerebral ventricular asymmetry in schizophrenia: a high resolution 3d magnetic resonance imaging study," *International journal of psychophysiology : official journal of the International Organization of Psychophysiology* **34**, pp. 207–11, Dec. 1999.
4. R. F. Deicken, Y. Eliaz, L. Chosiad, R. Feiwell, and L. Rogers, "Magnetic resonance imaging of the thalamus in male patients with schizophrenia," *Schizophrenia research* **58**, pp. 135–44, Dec. 2002.
5. J. Barnes, J. L. Whitwell, C. Frost, K. A. Josephs, M. Rossor, and N. C. Fox, "Measurements of the amygdala and hippocampus in pathologically confirmed alzheimer disease and frontotemporal lobar degeneration," *Archives of neurology* **63**, pp. 1434–9, Oct. 2006.
6. L. Bonilha, E. Kobayashi, F. Cendes, and L. M. Li, "Protocol for volumetric segmentation of medial temporal structures using high-resolution 3-d magnetic resonance imaging," *Human brain mapping* **22**, pp. 145–54, June 2004.
7. L. Wang, S. C. Joshi, M. I. Miller, and J. G. Csernansky, "Statistical analysis of hippocampal asymmetry in schizophrenia," *NeuroImage* **14**, pp. 531–45, Sept. 2001.
8. A. Kumar, W. Bilker, H. Lavretsky, and G. Gottlieb, "Volumetric asymmetries in late-onset mood disorders: an attenuation of frontal asymmetry with depression severity," *Psychiatry research* **100**, pp. 41–7, Nov. 2000.
9. N. Makris, J. Kaiser, C. Haselgrove, L. J. Seidman, J. Biederman, D. Boriel, E. M. Valera, G. M. Papadimitriou, B. Fischl, V. S. C. Jr., and D. N. Kennedy, "Human cerebral cortex: A system for the integration of volume- and surface-based representations," *NeuroImage* **33**, pp. 139 – 153, 2006.

10. I. S. Sigalovsky, B. Fischl, and J. R. Melcher, "Mapping an intrinsic mr property of gray matter in auditory cortex of living humans: a possible marker for primary cortex and hemispheric differences," *NeuroImage* **32**, pp. 1524–37, Oct. 2006.
11. P. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. Gee, and G. Gerig, "User-guided 3d active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *NeuroImage* **31**, pp. 1116 – 1128, 2006.
12. J. Csernansky, S. Joshi, L. Wang, J. Haller, M. Gado, J. Miller, U. Grenander, and M. Miller, "Hippocampal morphometry in schizophrenia via high dimensional brain mapping," *Proc. Natl. Acad. Sci. USA* **95**, pp. 11406–11411, September 1998.
13. M. I. Miller, S. C. Joshi, and G. E. Christensen, "Large deformation fluid diffeomorphisms for landmark and image matching," in *Brain Warping*, A. W. Toga, ed., pp. 115–131, Academic Press, 1999.
14. M. Prastawa, J. H. Gilmore, and W. L. and Guido Gerig, "Automatic segmentation of mr images of the developing newborn brain," *Medical Image Analysis* **9**, pp. 457–466, October 2005.
15. M. Styner, H. C. Charles, J. Park, J. Lieberman, and G. Gerig, "Multi-site validation of image analysis methods - assessing intra and inter-site variability," in *SPIE Medical Imaging*, **4684**, pp. 278–286, 2002.