



Published in final edited form as:

*Proc Soc Photo Opt Instrum Eng.* 2013 March 13; 8669: . doi:10.1117/12.2006727.

## Effects of T2-Weighted MRI Based Cranial Volume Measurements on Studies of the Aging Brain

Phong Vuong, David Drucker, Chris Schwarz, Evan Fletcher, Charles DeCarli, and Owen Carmichael for the Alzheimer's Disease Neuroimaging Initiative

### Abstract

Many brain aging studies use total intracranial volume (TIV) as a proxy measure of premorbid brain size that is unaffected by neurodegeneration. T1-weighted Magnetic Resonance Imaging (MRI) sequences are commonly used to measure TIV, but T2-weighted MRI sequences provide superior contrast between the cerebrospinal fluid (CSF) bounding the premorbid brain space and surrounding dura mater. In this study, we compared T1-based and T2-based TIV measurements to assess the practical impact of this superior contrast on studies of brain aging. 810 Alzheimer's Disease Neuroimaging Initiative (ADNI) participants, including healthy elders and those with mild cognitive impairment (MCI) and Alzheimer's Disease (AD), received T1-weighted and T2-weighted MRI at their baseline evaluation. TIV was automatically estimated from T1-weighted images using FreeSurfer version 4.3 (T1TIV), and an automated active contour method was used to estimate TIV from T2-weighted images (T2TIV). The correlation between T1TIV and T2TIV was high (.93), and disagreement was greater on larger heads. However, correcting a FreeSurfer-based measure of total parenchymal volume by dividing it by T2TIV led to stronger expected associations with a standardized measure of cognitive dysfunction (MMSE) in Poisson regression models among individuals with AD ( $z=1.73$  vs. 1.09) and MCI ( $z=3.15$  vs. 2.79) than a corresponding parenchymal volume measure divided by T1TIV. This effect was enhanced when the analysis was restricted to the cases where T1TIV and T2TIV disagreed the most. These findings suggest that T2-based TIV measurements may be higher fidelity than T1-based TIV measurements, thus leading to greater sensitivity to detect biologically plausible brain-behavior associations.

### Keywords

total intracranial volume; segmentation; FreeSurfer; MRI; image processing

### 1. PURPOSE

A large body of brain aging studies has determined that neurodegenerative diseases are associated with reductions in the volume of brain tissue in a variety of anatomical regions<sup>1,2,3</sup>. These studies commonly correct for normal variations in head size by adjusting the raw absolute head size measurements between subjects. One method of head size correction is to divide the brain region over the total intracranial volume (TIV) –also known as the region-to-ICV ratio or proportional method<sup>4</sup>. The rationale for this correction is that in the context of degenerative disease, it is desirable for brain region volume measurements to capture not only the current amount of brain tissue but also disease-related volume reductions from peak values attained during young adulthood; however, inter-individual variability in peak regional volumes is substantial<sup>5</sup>, making current regional volumes questionable proxy measures of disease-related reductions. Because TIV is an index of overall head size that is unaffected by disease<sup>6,7</sup>, it is a viable proxy for peak brain volume attainment; however, the ability to obtain accurate and reliable TIV measurements is difficult. Manual tracing of the cranial vault is impractical for large-scale studies, leading to

a variety of automated approaches; these have been shown to possess varying degrees of measurement noise<sup>8</sup>, which in turn introduce measurement noise into TIV-corrected brain region volumes. Therefore, an ongoing research goal is to develop high-fidelity TIV estimation methods that should in turn enhance the fidelity of TIV-corrected MRI measures, thus leading to an enhanced ability to identify associations between MRI measures and various factors related to neurodegenerative diseases such as Alzheimer's disease (AD).

While other automated TIV estimation methods such as SPM and FreeSurfer take T1-weighted MRI scans as input<sup>9,10</sup>, T1-weighted MRI complicates TIV estimation due to a lack of clear tissue contrast. Contrast between the cranial cerebrospinal fluid (CSF) space that bounds the healthy young brain and surrounding dura mater is minimal under T1 weighting (see Figure 2), and contrasts between dura, bone, fat, and skin can vary widely depending on numerous acquisition parameters. Alternatively, T2 weighting provides a clear contrast between the cranial CSF space and surrounding tissue, and thus a clear target for TIV delineation. Therefore, we developed a TIV estimation method based on T2-weighted MRI and compared it to a popular existing T1-weighted MRI based method from the publicly available FreeSurfer software package<sup>11</sup>. FreeSurfer estimates TIV by dividing a known atlas TIV over the Atlas Scaling Factor (ASF). This ASF is computed by taking the determinant of the affine transformation matrix used to align the subject brain to a common atlas space<sup>9</sup>. We hypothesized that the superior contrast between CSF and surrounding cranial tissues in T2-weighted MRI sequences would enable higher-fidelity TIV segmentations, which would in turn result in stronger expected associations between TIV-corrected brain tissue volumes and cognitive function in elderly individuals with AD and its prodromal condition, mild cognitive impairment (MCI).

## 2. METHODS

We implemented a brain volume segmentation pipeline that delineates TIV given a T2-weighted MRI scan. We tested our algorithm on 810 Alzheimer's Disease Neuroimaging Initiative (ADNI) subjects and visually assessed TIV quality in each scan (Figure 1). Finally, we compared T2-weighted MRI based TIV volume (T2TIV) against T1TIV in terms of their ability to give rise to biologically-plausible associations between brain tissue variables and clinically-relevant markers of cognitive function.

### 2.1 T2TIV Segmentation Pipeline

Our processing pipeline required as input corresponding T1-weighted and T2-weighted MRI scans from the same scan session. First, we linearly co-registered the T2-weighted image space to the T1-weighted image space. Next we ran our own implementation of the Brain Extraction Tool<sup>12</sup> (BET) on the T1-weighted image to generate a binary brain surface mask that included the parenchyma, and ventricular and sulcal CSF. Then an expert tracer manually corrected segmentation errors generated by the BET software. We developed a total intracranial segmentation algorithm that took as input the cleaned binary brain surface mask and the T2-weighted image and returned as output a TIV binary mask. We defined the output of the T2-weighted TIV (T2TIV) segmentation algorithm to include the parenchyma, ventricular and sulcal CSF and the CSF between the parenchyma and surrounding dura mater. We visually graded the T2TIV segmentation and manually corrected the segmentation errors only if the T2TIV received a grade of 1 or less (Figure 1).

**2.1.1 T2TIV Segmentation Algorithm**—We developed a T2TIV segmentation algorithm using a mesh growing active contour method. We wrote the algorithm in C++ and incorporated ITK and VTK libraries for processing MRI sequences. Our algorithm extended BET's segmentation pipeline and modified its mesh inflation step by redefining the mesh's boundary stopping condition. We speculated that by initializing our T2TIV segmentation

algorithm with a reasonable three-dimensional surface mesh of the brain we could reliably inflate the mesh points outward until the mesh surface contacted the CSF and dura mater boundary. We previously mentioned that boundary detection of the inner cranial vault using a T1-weighted MRI sequence is difficult because the intensity contrast between CSF and dura mater tissue both appear dark. However, we observed that in a T2-weighted sequence, the CSF appears bright against the dark dura mater. We used this sharp contrast, which clearly defines the inner cranial vault, to form the basis for our T2TIV segmentation algorithm. Our algorithm took as input a T2-weighted MRI sequences and a binary mask of the brain. It then constructed a reasonable brain mesh model using the binary brain mask. Next we inflated the mesh such that the surface of the mesh lines up with the edge of the inner cranial vault, using the voxel intensities from the T2-weighted MRI sequence as reference.

Our T2TIV segmentation algorithm began by initializing a spherical polygonal mesh with 2563 points and 5120 triangular faces. The center of the mesh sphere was placed at the center of mass of the BET brain tissue mask after manual cleanup. The mesh points were inflated outwards along the direction of the point's surface normal until it reached the boundary of the binary brain mask. We found that this process was a simpler and more robust way to generate a smooth and numerically well-behaved brain surface mesh compared to directly fitting a mesh to the brain tissue mask. We then estimated the T2TIV by iteratively expanding this initial mesh, using the T2-weighted image for guidance.

For each point on the mesh, the algorithm sampled ten contiguous voxel intensities along the point's outward and inward normal vector –generating a set of twenty voxel intensities. The intensities in this set were then grouped into three approximate categories corresponding to CSF (brightest), parenchyma (medium brightness), and dura mater (darkest). We used a rudimentary tissue classification method by which intensities were assigned a tissue type based on its intensity difference from the local maximum (CSF), median (parenchyma), or minimum (dura mater) intensities of that set. Mesh points that have CSF or parenchyma tissues types along the outward facing surface normal vector were pushed one voxel unit outward, while points that have dura mater tissue type in the inward facing surface normal were pushed one voxel unit inward. Points meeting neither of these criteria remained stationary.

To maintain smooth morphology of the mesh points, we employed BET's vector displacement method, which determined the direction for each point's perturbation step. For each point, the BET algorithm computed and summed together three individual vectors to form a displacement vector, (1) planar smoothing (move points such that neighboring points were of equal distance along those point's surface plane), (2) orthogonal smoothing (move points such that their orthogonal distance from their corresponding surface plane was within a distance), (3) direction of perturbation (control whether points were moved inwards or outwards along the surface normal). This mesh expansion process was repeated for 5000 iterations. Finally the T2TIV mesh was converted into a binary mask format for validation and clean up.

**2.1.2 T2TIV Segmentation Grading**—We visually rated all 810 T2TIV segmentations on a scale of 0 to 3, representing gross and large-scale errors (0), multiple small and localized errors (1), a single small and localized error (2) and a lack of noteworthy errors (3). We found that of the 810 subjects, 165 T2-based TIV segmentations scored a 3, 510 scored a 2 and 135 scored a 1 (see Table 1). We manually cleaned all T2-based TIV segmentations that scored a 1, boosting their scores to 3. Our automated T2TIV segmentation pipeline returned an unsatisfactory TIV with a rate of 16%, which was caused by (1) an invalid initial binary brain mask input that was not properly cleaned of

segmentation errors and/or (2) an invalid registration of the T2-weighted image to the T1-weighted image space.

All 810 T2TIV segmentations were scored using our own 2D image slice-viewing program (see Figure 3). T2TIV segmentations that received a grade of 1 or less were manually fixed or “cleaned” using our own paint tool program, which allowed us to remove or add voxels that were incorrectly classified as brain tissue.

## 2.2 Statistical Analysis

The two goals of statistical analysis were (1) to quantify differences between our T2-weighted MRI based TIV (T2TIV) and a T1-weighted MRI based TIV (T1TIV) output by FreeSurfer version 4.3, and (2) to determine whether expected, biologically-plausible associations between brain MRI variables and clinically relevant cognitive markers were stronger when brain MRI variables are corrected by T2TIV as opposed to T1TIV. The FreeSurfer TIV method was described previously<sup>9</sup>. We used linear regression with T2TIV as the dependent variable and T1TIV as the independent variable to assess the relation between the two TIV variables. We used the intercept and slope of the regression equation to identify additive and multiplicative biases between the two methods, and used graphical plots to determine whether agreement between the two methods varied by head size. We then used a brain volume (BV) and a hippocampus volume (HV) variables provided by FreeSurfer, both representing the total volume of gray and white matter in the brain region, to construct four variables that represent BV corrected by head size and HV corrected by head size (BV/T1TIV and BV/T2TIV, and HV/T1TIV and HV/T2TIV respectively). We then used two Poisson regression models to assess the strengths of association between the four TIV variables and the integer-valued 30-point Mini Mental State Exam (MMSE) score<sup>13</sup>. Each Poisson regression model had BV and HV divided by one of the two TIV variables as the independent variable and 30 minus MMSE as the dependent variable. Individual models were estimated among baseline-MCI and baseline-AD groups, recognizing systematic group differences in the MMSE distribution that could bias pooled models, as well as the limited measurement sensitivity of MMSE among cognitively normal individuals<sup>14</sup>. Because a large body of literature supports the view that diminished cognitive function as measured by MMSE is associated with reduced brain volume in MCI and AD<sup>15</sup>, stronger associations between corrected BV and 30-MMSE, and corrected HV and 30-MMSE are considered more biologically plausible.

In a secondary analysis, the Poisson regression models were re-estimated on a restricted data set, the 10% of scans for which the absolute difference between T1TIV and T2TIV was maximal. This gave us a means of adjudicating such discrepant cases in terms of whether one or the other measure provided more biologically plausible associations with cognition.

## 3. RESULTS

An example brain surface and T2TIV boundary are shown overlaid onto corresponding T1-weighted and T2-weighted images in Figure 2. A scatter plot of the regression of T1TIV against T2TIV is shown in Figure 4. The correlation between methods was high (Pearson's  $r=.93$ ), and there was little evidence of systematic additive or multiplicative biases between methods (intercept= $1.6e+04$ , slope= $.97$ , see Figure 4). However, visual plots of regression residuals across the range of T2TIV values (Figure 4) suggested that the two methods disagreed more strongly on TIV of larger heads.

In Poisson regression models, brain volume divided by T2TIV and hippocampus volume divided by T2TIV was more strongly associated with MMSE than brain volume divided by T1TIV and hippocampus volume divided by T1TIV in both MCI and AD groups (Table 2).

This effect was enhanced by reducing the analysis to discrepant cases; in this secondary analysis, the difference in the Poisson model test statistic was greater between the corresponding BV/T1TIV and BV/T2TIV and corresponding Hip/T2TIV and Hip/T2TIV models despite the inclusion of dramatically fewer scans.

#### 4. DISCUSSION

The key finding of this study is that brain volume is more strongly associated with a standardized measure of cognitive function in individuals with MCI and AD when it is corrected by T2TIV compared to T1TIV. The data are consistent with the hypothesis that T2TIV provides a higher-fidelity estimate of TIV due to superior contrast between the CSF space in the cranium and surrounding tissues; that is, dividing by T1TIV may have diluted the strength of association between brain volume and cognition because T1TIV contains greater measurement noise. The fact that this effect is enhanced in the sub-group of cases where T1TIV and T2TIV disagreed most suggests that these extreme discrepancies may reflect errors on the part of T1TIV. The strength of the study is the direct comparison of T1-based and T2-based TIV measurements on a large group (N=810) of MRI scans of elderly individuals; the key limitation is the lack of ground-truth measurement of the size of the cranial vault to determine the absolute accuracy of the two approaches.

#### 5. CONCLUSIONS

We developed an automated T2-weighted MRI based TIV segmentation pipeline and compared it to FreeSurfer's T1-weighted MRI based TIV method on 810 ADNI subjects. We hypothesized that T2-weighted TIV has higher fidelity over T1-weighted TIV due to the clearer contrast between the CSF space and surrounding tissue along the inner cranial boundary in T2-weighted images, which would in turn result in stronger expected associations between TIV-corrected brain tissue volumes and cognitive function in elderly individuals with AD and its prodromal condition, mild cognitive impairment (MCI). We compared head size correction using T2TIV and T1TIV and our results suggested that head size correction using T2-weighted TIV measurements does strengthen associations between brain volume and cognitive function in individuals with MCI and AD over head size correction using FreeSurfer's T1-based TIV. Future neuroimaging studies of the aging brain should carefully consider whether acquisition of T2-weighted MRI would enhance the study of brain-behavior relationships. This work was supported by NIH grants P30 AG010129, K01 AG030514, and U01 AG024904.

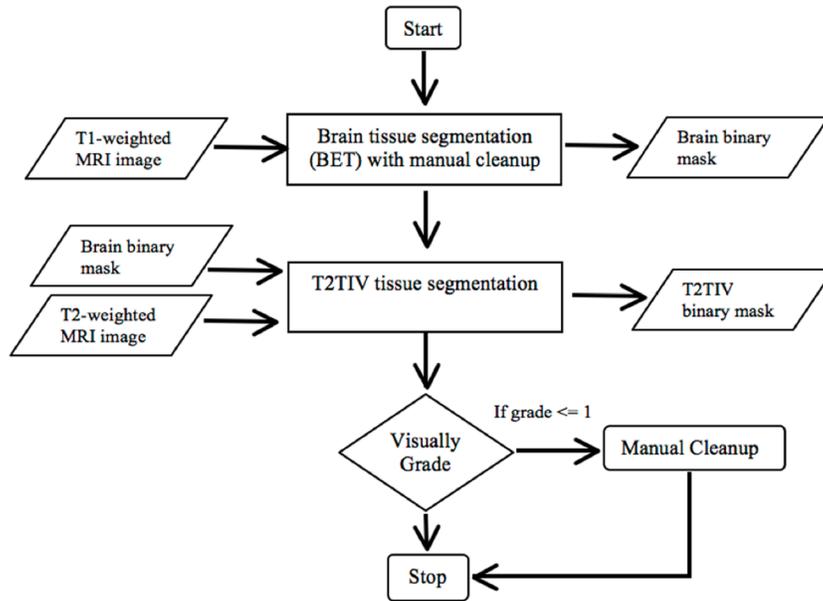
#### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

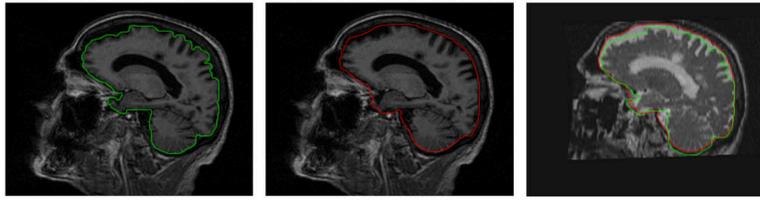
#### References

1. Jack CR, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, Petersen RC, Trojanowski JQ. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol.* 2010; 9(1):119–128. [PubMed: 20083042]
2. Schuff N, Woerner N, Boreta L, Kornfield T, Shaw LM, Trojanowski JQ, Thompson PM, Jack CR, Weiner MW. the Alzheimer's; Disease Neuroimaging Initiative. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain.* 2008; 132(4): 1067–1077. [PubMed: 19251758]
3. Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermersch P, Kuiper M, Steining M, Wolters EC, Valk J. Atrophy of medial temporal lobes on MRI in 'probable' Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *Journal of Neurology, Neurosurgery & Psychiatry.* 1992; 55(10):967–972.

4. Sanfilipo MP, Benedict RHB, Zivadinov R, Bakshi R. Correction for intracranial volume in analysis of whole brain atrophy in multiple sclerosis: the proportion vs. residual method. *Neuroimage*. 2004; 22(4):1732–1743. [PubMed: 15275929]
5. DeCarli C, Massaro J, Harvey D, Hald J. Measures of brain morphology and infarction in the Framingham heart study: establishing what is normal. *Neurobiology of Aging*. 2005; 26(4):491–510. [PubMed: 15653178]
6. Jenkins R, Fox NC, Rossor AM, Harvey RJ, Rossor MN. Intracranial Volume and Alzheimer Disease: Evidence Against the Cerebral Reserve Hypothesis. *Arch Neurol*. 2000; 57(2):220–224. [PubMed: 10681081]
7. Pfefferbaum A, Mathalon DH, Sullivan EV, Rawles JM, Zipursky RB, Lim KO. A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Arch Neurol*. 1994; 51(9):874. [PubMed: 8080387]
8. Ridgway G, Barnes J, Pepple T. Estimation of total intracranial volume; a comparison of methods. *Alzheimer's and Dementia*. 2011
9. Buckner RL, Head D, Parker J, Fotenos AF, Marcus D, Morris JC, Snyder AZ. A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *Neuroimage*. 2004; 23(2):724–738. [PubMed: 15488422]
10. Pengas G, Pereira JMS, Williams GB, Nestor PJ. Comparative reliability of total intracranial volume estimation methods and the influence of atrophy in a longitudinal semantic dementia cohort. *J Neuroimaging*. 2009; 19(1):37–46. [PubMed: 18494772]
11. Fischl B. FreeSurfer. *Neuroimage*. 2012; 62(2):774–781. [PubMed: 22248573]
12. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002; 17(3):143–155. [PubMed: 12391568]
13. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975; 12(3):189–198. [PubMed: 1202204]
14. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005; 53(4):695–699. [PubMed: 15817019]
15. Chan D, Janssen J, Whitwell J, Watt H. Change in rates of cerebral atrophy over time in early-onset Alzheimer's disease: longitudinal MRI study. *The Lancet*. 2003; 362(9390):1121–1122.

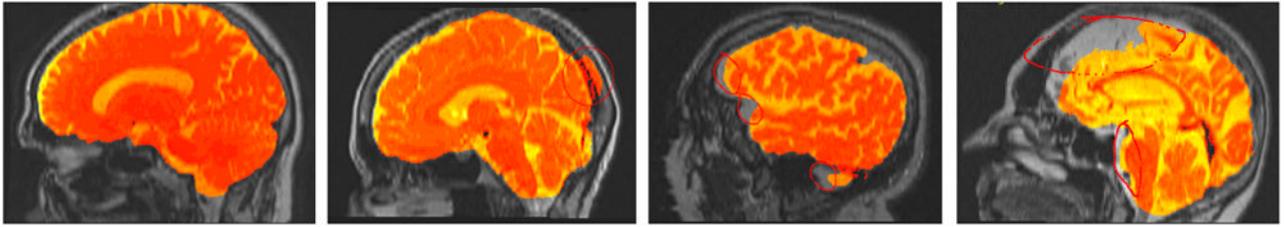


**Figure 1.** We present a flow chart that shows the TIV segmentation pipeline starting from the T1-weighted BET segmentation and ending at either the *visually grade T2TIV* step or the *manual cleanup* step. TIV segmentations that were rated with a grade of 1, meaning a TIV with multiple segmentation errors in large regions, are manually fixed and turned into a grade of 3.



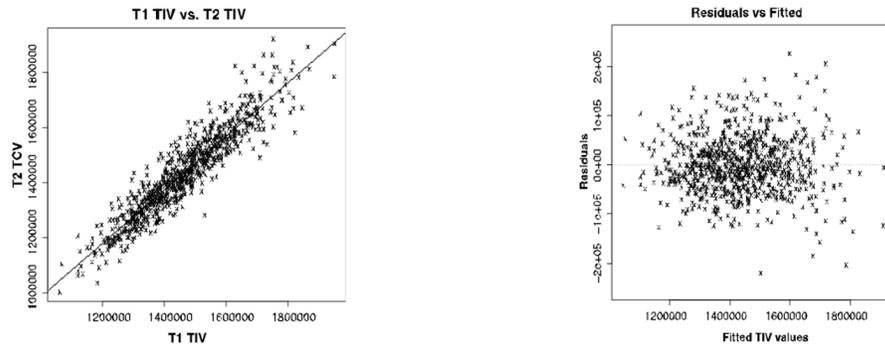
**Figure 2.**

(Left) Sagittal slice of a T1-weighted MRI with the brain surface segmentation outlined in green. (Middle) T1-weighted MRI with the T2-based TIV segmentation outlined in red. Note that in the superior prefrontal cortex, the correct position of the TIV surface is non-obvious based on T1-weighted image intensities. (Right) Co-registered T2-weighted MRI with the brain surface and TIV segmentations from the left and middle images overlaid in green and red.



**Figure 3.**

Reading from left and to right, we present a series of T2TIV tissue segmentation overlaid on top of its corresponding T2-weighted MRI sequence in order of descending segmentation quality. The left most image depicts a TIV that has received a visual grade quality of 3 (excellent brain tissue segmentation with little to no errors), follow by a TIV that has received a visual grade of 2 (good brain tissue segmentation with no more than one region of segmentation error), follow by a TIV that has received a visual grade of 1 (fair brain tissue segmentation with few regions of segmentation errors), and finally on the far right shows a TIV that has received a visual grade of 0 (unsatisfactory brain tissue segmentation with multiple segmentation errors).



**Figure 4.** (Left) Plot of T2TCV ( $m^3$ ) vs. T1TIV ( $m^3$ ). (Right) Plot of Residuals ( $m^3$ ) vs. Fitted ( $m^3$ )

We present a tally of T2TIV segmentation grades after running the T2TIV segmentation and before manual cleanup, and the overall percentage that each grade received out of 810 ADNI scans.

**Table 1**

Grade	0	1	2	3
Number of TIV	0	135	510	165
Percentage	0	16%	63%	20%

**Table 2**

Strength of association between brain volume corrected by T1-based (T1TIV) and T2-based (T2TIV) total intracranial volume in Poisson regression models including all ADNI participants diagnosed with MCI or AD at baseline (All Cases) and the 10% of cases for which T1TIV and T2TCV disagreed to the greatest degree (Discrepant Cases). The number of individuals in each regression model (N), the test statistic for significance of the brain volume predictor in the Poisson regression model ( $z$ ), and the significance of that test statistic ( $p$ ) are shown. Lesser values for  $z$  and  $p$  indicate stronger associations between the brain volume predictor and the outcome variable (MMSE).

	All Cases				Discrepant Cases				
	Brain/T1TIV	Brain/T2TIV	HV/T1TIV	HV/T2TIV	Brain/T1TIV	Brain/T2TIV	HV/T1TIV	HV/T2TIV	
N	387				190				
MCI	$z$	-2.79	-3.15	-3.84	-4.07	-0.39	-1.13	-1.75	-1.97
	$p$	0.005	0.002	1.3e-4	4.7e-05	0.70	0.26	0.08	0.049
	N	87				40			
AD	$z$	-1.16	-1.73	-2.337	-2.632	-0.53	-1.21	-0.172	-0.714
	$p$	0.25	0.08	0.0194	0.008	0.59	0.23	0.863	0.475