Age-related differences in regional brain volumes: A comparison of optimized voxel-based morphometry to manual volumetry

Kristen M. Kennedy[a], **Kirk I. Erickson**[b], **Karen M. Rodrigue**[a], **Michelle W. Voss**[b], **Stan J. Colcombe**[c], **Arthur F. Kramer**[b], **James D. Acker**[d], and **Naftali Raz**[a,*]

[a] Institute of Gerontology, Department of Psychology, Wayne State University, Detroit, MI

[b] Beckman Institute for Advanced Sciences and Technology and Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL

[c] School of Psychology, University of Wales, Bangor, United Kingdom

[d] Mid-South Diagnostic Imaging/Therapeutics, Memphis, TN

Abstract

Regional manual volumetry is the gold standard of *in vivo* neuroanatomy, but is labor-intensive, can be imperfectly reliable, and allows for measuring limited number of regions. Voxel-based morphometry (VBM) has perfect repeatability and assesses local structure across the whole brain. However, its anatomic validity is unclear, and with its increasing popularity, a systematic comparison of VBM to manual volumetry is necessary. The few existing comparison studies are limited by small samples, qualitative comparisons, and limited selection and modest reliability of manual measures. Our goal was to overcome those limitations by *quantitatively* comparing optimized VBM findings with highly reliable multiple regional measures in a large sample (N = 200) across a wide agespan (18–81). We report a complex pattern of similarities and differences. Peak values of VBM volume estimates (modulated density) produced stronger age differences and a different spatial distribution from manual measures. However, when we aggregated VBM-derived information across voxels contained in specific anatomically defined regions (masks), the patterns of age differences became more similar, although important discrepancies emerged. Notably, VBM revealed stronger age differences in the regions bordering CSF and white matter areas prone to leukoaraiosis, and VBM was more likely to report nonlinearities in age-volume relationships. In the white matter regions, manual measures showed stronger negative associations with age than the corresponding VBM-based masks. We conclude that VBM provides realistic estimates of age differences in the regional gray matter only when applied to anatomically defined regions, but overestimates effects when individual peaks are interpreted. It may be beneficial to use VBM as a first-pass strategy, followed by manual measurement of anatomically-defined regions.

*Corresponding author. Institute of Gerontology, Wayne State University, 87 E Ferry, 226 Knapp Bldg, Detroit, MI 48202. Tel. (313) 577-2297; Fax: (508) 256-5689; electronic mail: nraz@wayne.edu.

Keywords

Voxel-Based Morphometry; manual morphometry; aging; gray matter; white matter; regional brain volume; MRI; nonlinear age trends; method comparison; ICV correction

1. Introduction

A complex pattern of behavioral and cognitive changes characterizes adult development and aging. Understanding of the biological roots of these changes is predicated on understanding age-related transformations of the brain. The advent of magnetic resonance imaging (MRI) facilitated assessment of the effects of age on brain structure and made it routine (for reviews see Raz, 2000; Raz & Rodrigue, 2006). During the first decade of wide-spread use of MRI, the traditional neuroanatomical and anthropometric methods were applied to quantification of human brain structures observed *in vivo.* In manual methods of brain demarcation and measurement, the MR image, though originated in digital data, is treated as a faithful analog of brain anatomy and its regions and structures are traced by human operators directly on the images. Manual volumetry has considerable face validity but is time-consuming, and requires substantial training aimed at minimizing operator-induced error. Moreover, estimation of regional volumes through manual morphometry (manual volumetry) can be performed on only a limited number of brain regions at a time, thus requiring *a priori* hypotheses that allow narrowing down the number of such pre-selected target regions..

In the past decade, significant progress has been made towards ameliorating the abovementioned shortcomings of manual volumetry by developing semi-automated methods of evaluating regional differences in brain structure. In contrast to manual volumetry, which targets specific regions of interest (ROIs), automated methods use voxel-level information to estimate differences in local gray matter content in a standard stereotaxic space (Ashburner & Friston, 2000) and can be these voxel-based methods of estimating local density and volume of the gray matter can be grouped under the general rubric of Voxel-Based Morphometry (VBM). As the name implies, VBM takes advantage of digital information available for each volume element (voxel) of the three-dimensional image. They hold a promise of perfectly reproducible hypotheses-free assessment of the whole brain that can be routinely conducted by minimally trained personnel with only cursory knowledge of neuroanatomy. Because of the high degree of automation, VBM permits high-throughput and efficient workflow in evaluation of large samples of brains in a reasonable time-frame. It is hardly surprising, then that since its introduction, VBM has gained significant popularity. For example, typing "Voxel-Based Morphometry" in a PubMed database on October 27, 2007 yielded 670 references, with a wide thematic range from idiopathic headache (May et al., 1999) to putative neuroanatomical peculiarities of professional mathematicians (Aydin et al., 2007).

As the popularity of VBM is increasing rapidly and publications are amassing, there is a growing need for systematic comparison of the automated methods with manual methods of neuromorphometry that, due to their high anatomical validity, are held as the current gold standard. Despite concerns about the potentially problematic aspects of VBM that were expressed at its introduction (Bookstein, 2001), only a few limited-scope comparisons between the methods have been conducted to date.

Although some significant differences among the VBM methods are apparent at the procedural and algorithmic levels, the basic stages of image processing are similar. VBM methods use a skull-stripped segmented brain image as their starting point. Segmentation algorithms divide the imaged brain tissue into gray matter, white matter and cerebrospinal fluid (CSF) with subsequent assigning of each voxel the probability of representing one of the three

compartments, generally on the basis of differences in signal intensity. Local values of signal intensity can be studied in relation to some specified individual or group property (e.g., age) on a voxel-by-voxel basis. Because individual brains differ substantially in their geometric properties, comparison of individual's MR-based brain images requires conversion of the data into a canonical space. In VBM such normalization of the data to a standard space necessitates several corrective steps involving application of filters that, while improving signal-to-noise ratio, reduce the effective resolution of the initial high-resolution MR images. Resolution loss is often as much as 1000-fold: from $1 \times 1 \times 1$ mm pixel on an acquired image to $10 \times 10 \times 10$ mm pixel on a filtered image. Therefore, segmentation, registration, and normalization steps can introduce distortions and are a main concern in evaluating VBM performance. The difficulties of interpreting VBM findings are especially significant when small brain structures are concerned (Bookstein, 2001; Crum et al., 2003; Davatzikos, 2004; Eckert et al., 2006; Senjem et al., 2005).

Although several attempts to evaluate VBM against manual volumetric methods have been made, many aspects of these comparative studies need improvement and/or updating. First, in all extant studies, comparisons between VBM and manual tracing methods were limited to either total brain volume or a very small number of manually traced regions of interest (ROIs) (Allen et al, 2005; Cardenas et al., 2003; Dorion et al, 2001; Douaud et al, 2006; Giuliani et al, 2005; Gong et al, 2005; Good et al, 2002; Kubicki et al., 2002; Senjem et al., 2005; Tapp et al, 2006; Testa et al., 2004; Tisserand et al., 2002). Second, reliability of manual measures employed in some of the comparative studies is either unspecified or measured by suboptimal reliability indices such as Pearson correlations or coefficients of variation (Douaud et al., 2006; Good et al., 2002; Tapp et al., 2006). Third, some comparative studies employed non-optimized VBM methods (Giuliani et al., 2005; Tisserand et al., 2002) that produce density rather than volume estimates. Fourth, the extant comparative studies relied on relatively small sample sizes ranging between 10 and 75 subjects, thus limiting the statistical power of the comparisons.

Most of the extant studies are limited to qualitative comparisons between the methods. In some studies, categorical measures, such as success in detecting diagnostic group differences were used as a measure of method differences (Cardenas et al., 2003; Testa et al., 2004). In one study, limited to frontal regions (Tisserand et al., 2002), the variance explained by manual and VBM-based measures was used as the means of a direct comparison of the manual and semiautomated techniques. In another study, a comparison between manual stereology of the prefrontal cortex and optimized VBM in young and middle-aged healthy adults revealed lack of convergence in gauging age differences in the frontal volumes, despite both methods finding an association between the medial prefrontal volume and fluid intelligence (Gong et al., 2005). Because of limited overlap of VBM findings with the results of manual volumetry the authors concluded that despite its speed, reliability and high degree of automation, VBM cannot replace manual measurements. Several authors recommended that the two methods should be used in tandem as complementary measures (Giuliani et al., 2005; Kubicki et al., 2002; Testa et al., 2004; Tisserand et al., 2002) without specifying the exact roles for either method, and recommended "studying correlation between techniques [to] bring quantitative information of their degree of variation" (Dorion et al, 2001). However such comparisons require sufficient statistical power. For example, a small-sample study aimed at predicting conversion to dementia found no difference among automated, semi-automated, and manual methods (Cardenas et al., 2003). Finally, although comparisons of VBM and manual measures are certainly sparse, even fewer of the comparison studies (Tapp et al., 2006 in canine model; Tisserand et al., 2002) focused on normal aging.

The purpose of the present study was, therefore, to address the limitations identified in the current literature and to *quantitatively* compare the manual measures of multiple ROIs obtained

from unmanipulated MR images to the full-brain regional volume estimates derived from optimized VBM procedures. In addition to correspondence between the methods in measuring regional volumes, we evaluated the methods in their ability to detect age-related differences in the brains of healthy adults.

For this comparison study, we used a cross-sectional sample of 200 largely healthy non-demented adults that included a small subsample of persons with controlled hypertension. The age of the participants covered most of the adult lifespan (18 – 80 years). Manual measures of thirteen ROIs (eleven in gray and two in white matter) and their correlations with age in this sample were published in Raz et al (2004). To expand the scope of regional measures and to provide additional opportunities to compare performance of two methods in the white matter, we included two additional white matter regions to this comparison study. Thus we had 15 manually defined ROIs available for a more extensive comparison than ever reported. Further, the regions selected as the standard of comparison are highly reliable. The reliability index employed in this study (intraclass correlation, ICC(2), Shrout & Fleiss, 1979) is more stringent and comprehensive than frequently used product-moment correlations or coefficient of variation. Intraclass correlation takes into account information about the order of observations, the distance among them, and the mean differences in the observations produced by two raters. Moreover, in the specific ICC formula, ICC(2), the error variance is computed under an assumption of random observers (Shrout & Fleiss, 1979). This assumption makes it more difficult to attain high reliability, but ensures generalizability of the measures. Lastly, we performed quantitative comparisons between the two techniques in addition to the qualitative comparisons that have been reported in previous comparison studies.

2. Materials and methods

2.1 Participants

The data for this study came from a published sample (Raz et al., 2004) that consisted of 200 healthy right-handed, native English speaking participants (119 women, 81 men) across a wide age span (18 – 81 yrs, mean 46.93) with at least high school education (mean number of years of formal schooling was 15.76 yrs). Sample composition by decades is provided in Table 1.

The participants reported no history of cardiovascular disease (except treated hypertension), diabetes, thyroid problems, drug/alcohol abuse, neurological or psychiatric conditions, or head trauma with loss of consciousness for more than 5 min. They denied using anti-seizure, anxiolytics, antidepressants, or recreational drugs or a habit of consuming more than three alcoholic drinks per day. All participants were screened for dementia and depression using a modified Blessed Information-Memory-Concentration Test (Blessed, Tomlinson, & Roth, 1968) with a cut-off of 30 and Geriatric Depression Questionnaire (Radloff, 1977) with a cut-off of 15. An experienced neuroradiologist (JDA) examined the MR scans for space occupying lesions and neurodegenerative disease. Participants were strongly right-handed, indicated by a score of 75% and higher on the Edinburgh Handedness Questionnaire (Oldfield, 1971).

2.2 MRI protocol

The images were obtained on a 1.5T GE Signa scanner (General Electric Co., Milwaukee, WI). Volumes were measured on the full-brain images acquired with T1-weighted 3-D spoiled gradient recalled (SPGR) sequence with 124 contiguous transaxial slices, echo time (TE) = 5 ms, repetition time (TR) = 24 ms, field of view (FOV) = 22 × 22 cm, acquisition matrix 256 × 192, slice thickness = 1.3 mm, and flip angle = 30°. Thus, the acquired voxel size was .86 × .86 × 1.30 mm. The imaging site was fully staffed by medical physicists and MRI technologists, and the scanners were routinely calibrated using a standard GE phantom.

2.2.1 Manual volumetric morphometry—Most of the volumes of the individual regions of interest (ROIs) were taken from our previously published study (two white matter ROIs, pre- and post-central white matter, were added for this comparison) and detailed descriptions of demarcation and tracing rules are available therein (Raz et al., 2004). Briefly, we repartitioned the SPGR data set into $.86 \times .86 \times 1.5$-mm-thick slices and manually aligned each participant's 3D volume using BrainImage v2.3.3 (Reiss et al., 1995) to correct for tilt (interhemispheric fissure), pitch (anterior commissure – posterior commissure), and rotation (orbits size equality) of the head (see Figure 1A). Hippocampus was traced on the scans after alignment along its long axis. The ROIs were manually traced on each 1.5 mm slice on each brain using NIH Image v1.62 software (http://rsb.info.nih.gov/nih-image/) by reliable expert raters with conservative ICC(2) coefficients (Shrout & Fleiss, 1979) of .90 –.97 (see Table 2 for the list of regional reliabilities). The volumes (cm^3) were computed and adjusted for body size differences via regression/ANCOVA equations using manually traced intracranial volume (ICV). We examined the 15 ROIs in each hemisphere (Raz et al., 2004), as listed in Table 2 and illustrated in Figure 2, including the two additional white matter regions adjacent to pre- and post-central gyri.

2.2.2 Voxel-based morphometry—To process the SPGR scans in the Voxel-Based Morphometry (VBM) framework (Ashburner & Friston, 2000; 2001; Mechelli et al., 2005) we used the FMRIB Software Library (FSL 3.2) package (http://www.fmrib.ox.ac.uk/fsl/; Smith et al., 2004). Our decision was based on considerations of convenience and does not imply endorsement of that specific software or assertion of its advantages over other popular VBM software (e.g., SPM, Statistical Parametric Mapping, http://www.fil.ion.ucl.ac.uk/spm/). The goal of this study was to compare the VBM approach with manual volumetry and not to evaluate relative merits of specific software packages.

We applied the optimized VBM method, which has been reported to significantly reduce registration error and to reconstruct voxel volumes from density values via Jacobian-based correction (Ashburner & Friston, 2001; Good et al., 2001). We chose to apply the most typical, robust and well-researched VBM technique to ensure the generalizability of our comparison analyses. We have used this method in several other studies (Colcombe et al., 2003; Colcombe et al., 2005; Erickson et al., 2005; Erickson et al., 2007). This method consists of a "two-level" approach, where the first level refers to the initial processing of the images, followed by a "second-level" of processing which refers to re-running these similar steps on the study-specific template.

To begin, all brains were stripped of extracranial tissue (e.g., skull, para-orbital tissue, dura) using an accurate and robust deformable model algorithm (BET; Smith, 2002). Each brain was then visually inspected and any excess extracranial matter was manually removed. It is crucial that image histograms of voxel intensity values represent only brain matter as the presence of skull or fat in an image submitted to segmentation will alter the range of intensity values classified as white matter and result in increased tissue misclassifications.

The skull-stripped images were segmented into 3D gray, white and CSF probability density maps via a semi-automated algorithm (FAST; Zhang et al., 2001) that uses a Hidden Markov Random Fields model which, in addition to mixture histogram, takes into account spatial information and allows an improved estimate of tissue distribution in each voxel. The segmented probability maps in stereotaxic space were used as tissue priors to seed the tissue histogram distribution for the second-level (study-specific) segmentation. Skull-stripped brains were registered to MNI standard stereotaxic space using a robust 12-parameter affine transform based on a correlation ratio cost function with sinc interpolation (FLIRT; Jenkinson et al., 2002). The transformation matrices resulting from registration were then applied to the segmented gray and white matter images, which produced segmented tissue images in standard

space. We then created a study-specific template by averaging these images into a composite image that represents the average brain of our sample and then smoothed with a 4 mm HWHM 3D Gaussian kernel (~ 8 mm FWHM). The smoothed image became the study-specific template used for later registrations.

To complete the optimized VBM process, the "second-level" of processing, the original skull-stripped brains (in native space) were registered to the study-specific template and segmented using the study-specific gray, white, and CSF probability maps (priors) (Good et al., 2001). These partial volume maps were inspected and corrected for misclassification of tissue segmentation at edge voxels at the gray/white and white/CSF borders. These maps were then registered to the study-specific template via a 12-parameter affine transform and spatially smoothed with a 4-mm HWHM kernel. Because significant stretching may occur during spatial transformation from native to study-specific space we computed the determinant of this second-level transformation Jacobian matrix and multiplied the gray, white, and CSF probability maps of the individual brains by the Jacobian determinant. This step attempts to preserve information about changes in voxel volume that were induced by spatial registration. The resulting modulated probability maps (illustrated in Figure 1B and hereafter referred to as the modulated data) are used throughout this report. One participant was not included in the VBM analysis due to an unrecoverable artifact in the scan that negatively affected the segmentation algorithm. Thus, the final comparative sample consisted of 199 brains.

The gray matter, white matter, and CSF maps for each participant yield a partial volume estimate (PVE) of each brain tissue type for each voxel in stereotaxic space. Those PVEs were used as the primary dependent measures in statistical analyses. Specifically, age, sex, and hypertensive status were entered into a general linear model as independent variables of interest, with the gray and white matter PVE map voxel-wise values as the dependent variables of interest. Thus, a separate model was run for each voxel using FEAT (FMRI Expert Analysis Tool; Smith et al., 2004), with manually traced ICV as a covariate. A set of linear contrasts was then used to test the regression parameters, from which a z statistic was generated for each voxel. The resulting analysis yielded statistical parametric maps (SPMs) for regressed variables for each tissue type. Statistical parametric maps were thresholded at an uncorrected voxel-wise threshold of $p < .001$ ($z = 3.10$).

2.2.3 Manually drawn ROI masks for VBM—To directly compare the manually measured volumes with the volumes estimated from VBM, we used hand-drawn binary ROI masks. The masks were drawn using the FSLview module on T1-weighted coronal images of a skull-stripped brain of a representative participant. For the purpose of mask generation, the representative subject was defined as the one with the least root mean squared difference (RMSD) from the mean sample-specific template image. We created the masks in native space of the subject with the lowest RMSD (98.1, range 98–153) and then registered each mask into study-specific standard space. The list of mask regions (see Figure 3 for an example) was identical to that of the manual volumetry study (Raz et al., 2004), with an addition of two peri-Rolandic white matter ROIs. Note that, NIH Image software used for original measures did not allow us to save ROI objects, hence the need to create the masks in FSLview for this comparison study.

The correlation between the volumes estimated from the manually demarcated ROIs and the volume (number of voxels $\times$ 8 mm^3) encompassed by the ROI masks was $r = .98$. Thus, we demonstrated excellent reliability correspondence between the manually drawn ROI masks and original manual tracing. Each ROI mask was applied to the group (age variable Z-statistic map) and individual (partial volume estimation, PVE) image data, yielding two separate data sets. One contained average Z-score values for each region/mask and the values for each factor in the VBM model and the other consisted of regional data for each subject. Thus the peak Z-

scores obtained from the VBM analysis could be identified within the manually traced ROIs for both the group and individual data. Finally, we computed summary statistics (min, max, and mean intensity, # voxels, and volume) within the masked areas to directly compare with manual volumes.

3. Results

3.1 Manual volumetric morphometry

First, we summarize the results of manual measures analyses as reported in our previous publication (Raz et al., 2004). To examine the differential effects of age on the manually traced regional brain volumes, we used the General Linear Model, in which regional volume adjusted for ICV was a dependent variable, age was continuous predictor, sex was a categorical predictor, and hemisphere was a repeated measures factor. Although men still had larger regional volumes (with the exception of IPL), a Sex × ROI interaction further indicated the magnitude of sex effects differed across the brain regions from moderate (d = .66 for VC, .51 for PHG, and .48 for ACG, .41 for HC, .38 for FW) to low (d < .3 for SSC, MS, IT, OFC) to negligible (d = .17 for DLPFC and .16 for parietal white). The observed pattern of sex effects was not consistent with that observed in our previous sample drawn from the same population (Raz et al., 1997). However, a Sex × Age interaction was observed in HC, FG, and Frontal white matter, where sex differences were observed only among the young participants.

We found neither significant effect of diagnosis of hypertension or hormone replacement therapy nor interactions among them. Correlations between age and regional volumes adjusted for intracranial volume (ICV) varied throughout the brain ranging widely from $r = .04$ in the inferior parietal lobule to $r = -.63$ in the dorsolateral prefrontal cortex (see Table 2 for details).

3.2 Voxel-based morphometry

The optimized VBM analysis yielded widespread effect of age on brain modulated density (see Figure 4 for visual inspection of the age effect magnitude and Table 3 for peak coordinates). Inspection of peak coordinates and visual inspection of the gray matter statistical parametric maps (SPM; see Figure 4) revealed significant age-related reduction in modulated density in all lobes. However, the strongest effects were limited to the regions of parenchyma-CSF interface such as Sylvian and interhemispheric fissures, especially the superior temporal gyrus (see Table 3). Age differences in the white matter modulated density were considerably less widespread and concentrated in the anterior and middle corpus callosum, periventricular areas and cerebellar peduncles (see Figure 5 and Table 3). Note that in the whole-brain VBM analysis, the age effect on gray matter was stronger than on white matter. Specifically, as illustrated in Figures 4 and 5, significant gray matter voxels associated with age occupied greater space than white matter voxels. Furthermore, the peak Z statistics were higher in the gray matter compared to white matter (max Z = 10.5 vs. max Z = 7.2).

We observed several circumscribed regional differences between hypertensive participants (n = 19) and their normotensive peers as well as a significant Hypertension × Age interaction. Hypertension exacerbated the age-related differences. However, the extent of spatial distribution of those effects was limited. Specifically, the main effects of hypertension (reduced modulated density) were limited to gray matter in the left medial frontal and right and left inferior parietal regions. Peaks for the Age × Hypertension effect were limited to white matter in the left and right medial frontal, right superior temporal gyrus, right inferior parietal, and left posterior parietal regions. A right medial frontal cluster was also observed in gray matter, but contained only three voxels (max z = 3.16), and overlapped with the right medial frontal white matter cluster. Specific coordinates, cluster size, and max Z scores (peaks) for the Age × Hypertension effect can be found in Table 4. It is notable that overlapping peaks in gray and

white matter SPM's oftentimes occurred near the gray and white matter borders. As mentioned in the Method section, for VBM analyses, probability maps are the dependent variables of interest. Thus even voxels with low probability of being gray or white matter (e.g., .2 to .5) are still continuous variables that may show statistical relations, albeit within a restricted range of intensities.

In several small clusters of voxels, we noted sex differences across the age span such that greater modulated density was observed in women compared to men. Specific regions of sex differences in gray matter included the right superior temporal gyrus, right cingulate gyrus, left medial frontal, and the left inferior temporal regions. Regions of sex differences in the white matter included the left pre-central gyrus, right post-central gyrus, right inferior parietal, and right medial temporal regions. In addition, there was an Age × Sex interaction: men showed increasingly reduced modulated white matter density with age in the cingulate gyrus, right post-central gyrus, left and right pre-central gyrus, right superior temporal gyrus, and the right medial frontal region. Specific coordinates, cluster size, and peak Z scores for the sex effects are presented in Table 5.

Note that the sex and hypertension effects were very small in both spatial extent and magnitude. Moreover, they disappeared after a cluster-wise threshold of $p < .01$ and voxel-wise threshold of $z = 2.33$ were applied. The sex differences that were observed when total brain voxels served as a covariate (as seems to be a convention in many VBM studies) were substantially attenuated when an extra-cerebral index, ICV, was used as a covariate instead. For example, without ICV as a covariate there were 4138 significant ($z > 3.1$) gray matter voxels and 4096 significant white matter voxels. With ICV entered in the model, however, the size of sexually dimorphic clusters dropped to 253 significant gray matter voxels and 289 significant white matter voxels.

3.3 Comparison of VBM to manual volumetry

Age differences in multiple cortical and white matter regions were apparent on VBM significance maps (Figures 4 and 5). In addition to the prefrontal regions that evidenced the highest correlations with age in manual volumetry, the VBM peak analysis placed the insula and superior temporal gyrus at the top of the most age-sensitive regions list. Notably, the highest peak Z values for age effects were observed in the regions bordering the major fissures: Sylvian and interhemispheric. In the white matter compartment, age differences were limited almost exclusively to the periventricular regions and cerebellar peduncles. The top six gray and white regional peaks ranked by the age-effect Z values are presented in Table 3.

Although visual inspection of age-related differences (Z scores) may be informative, the results obtained by the two methods need to be compared quantitatively. With that goal in mind we conducted several comparisons. First, we extracted the partial volume estimate (PVE, or proportion of gray or white matter) for each peak voxel from the list of the highest age-associated Z-scores for each subject (see Table 3 for coordinates and statistics). Those peak PVEs were correlated with age (see correlations in Table 3). The correlations between age and VBM-derived peak PVE ranged from $r = -.26$ for the right postcentral gyrus white matter to $r = -.79$ for right superior temporal gyrus/insula gray matter. By comparison, the correlations between manually measured volumes and age ranged from $r = .04$ for inferior parietal lobule gray matter to $r = -.63$ for dorsolateral prefrontal cortex gray matter (see Tables 2 and 3 for all regions). However, voxel-based measures obtained from the peaks that already showed the largest age-related differences would be biased towards finding large age effects. Therefore the second step was to examine the peak values within the anatomically defined ROIs comparable to those from which regional volumes were obtained manually.

The maximum voxel-wise age differences for the ROIs that correspond to manually traced regions are reported in Table 6. This comparison shows that every gray matter ROI and all but

three white matter ROIs contained at least some voxels with significant age-differences. However, a single-voxel peak located in a specific region does not have a clear anatomical meaning and cannot serve as a representative of the region. For a quantitative comparison of the manual and VBM results, we conducted the third analysis, in which VBM-derived density estimates were aggregated across the ROIs demarcated according to the manual volumetry rules (Raz et al., 2004).

3.3.1 VBM regional mask analyses—We re-ran the same FSL model as before, but restricted it to the voxels within the binary masked ROIs. For each ith subject, within each of jth ROIs and each kth voxel, we computed the PVE_{ijk}, and compared ROI-specific age-related Z values and partial volume estimates to manual volume and age effects. The descriptive statistics for the age differences in aggregated regional PVEs are presented in Table 6, whereas the PVE sample means for all ROIs are listed in Table 7.

Correlations between the ICV-adjusted partial volume estimates from the masked regions and the volume from the manually traced ROIs were quite low, ranging from $r = .56$ for the dorsolateral prefrontal area to $r = -.01$ for the inferior parietal white matter (see Table 7 for all region's correlations between the two methods). Age effects in all cortical gray matter ROIs were negative (see Table 7), indicating reduced local modulated density in older brains. Although in eight out of 11 cortical regions the magnitude of age differences was greater for VBM-based estimates, the proportion was not significantly different from chance: sign test $p = .23$. For three out of four white matter ROIs similar, though smaller, negative age differences were observed with the manual measures, but not with the VBM estimates. No significant differences were noted in IPw. For all white matter ROIs, manual measures produced larger age effects. Overall, there was a weak association between the magnitude of age differences in white matter obtained with manual and VBM-derived measures: Spearman $\rho = .40$, $p < .05$, one-tailed. However, when the most discrepant region, inferior parietal lobule, was removed from the set, the association strengthened to $\rho = .58$, $p < .05$. Notably, across all ROIs, greater PVE variability (as indexed by the coefficient of variation, CV) was associated with greater age-related differences (Spearman $\rho = .59$, $p < .05$) and greater negative correlations between PVE and age: Spearman $\rho = -.69$, $p < .01$. We used Steiger's Z* statistic, which takes into account the dependence between the members of bivariate correlations (Steiger, 1980), to quantitatively compare the correlations of age and regional volumes obtained by the two methods for significant departure. These differences are listed in Table 7.

We conducted a general linear model analysis with regional PVEs as multivariate dependent variables, age (re-centered at its sample mean) as a continuous predictor, and sex as a categorical independent variable. For the gray matter regions, the ROI and the Hemisphere served as two repeated measure factors with 11 and 2 levels, respectively. The analyses revealed the main effect of age: $F(1, 196) = 306.67$, $p < .001$, $\eta^2 = .61$. However, as indicated by a significant Age × ROI interaction ($F(10, 196) = 19.56$, $p < .001$, $\eta^2 = .01$), the magnitude of age differences varied across ROIs. The correlations between age and PVE ranged between $r = -.40$ for the IT to $r = -.80$ for the MC. Within ROIs, correlations with age were sometimes larger for the left hemisphere and sometimes for the right. In all cases these differences in correlation magnitude were trivial and were confined to the second digit after the floating point. Nonetheless, because of strong dependence between the hemispheric PVEs, a small but significant Age × ROI × Hemisphere effect ensued: $F(10, 196) = 2.38$, $p < .05$, $\eta^2 = .01$.

In a similar model for the white matter, ROI (four levels), and Hemisphere (two levels) were repeated measure factors. This analyses revealed a significant effect of Hemisphere: $F(1, 196) = 255.57$, $p < .001$, $\eta^2 = .39$. Neither Age nor Sex showed significant main effects (both $F < 1$). However, a significant ROI × Age interaction ($F[3, 588] = 32.26$, $p < .001$, $\eta^2 = .06$) indicated that some ROIs showed age differences. Indeed, the correlation with age was

significant only for the prefrontal and primary motor (precentral gyrus) area white matter, but not for the inferior parietal or post-central gyrus. Furthermore, the significant triple interaction ROI × Age × Hemisphere (F[3, 588] = 10.56, $p < .01$, $\eta^2 = .03$) indicated that age effects varied not only across ROIs but also within ROI. Whereas left IPw and right SSw showed no age differences, among the ROIs that exhibited smaller volumes with age, the left hemisphere effects were larger than those on the right.

To illustrate the differences in the observed age effects, we provide side-by-side scatterplots depicting the regression of each ROI (adjusted for intracranial volume) on age for the two methods (see Figures 6 through 9). Visual inspection of the plots suggests several method-related differences in the distributions of volume estimates and age. For some regions (e.g. DLPFC, OFC and PHG), the VBM-based estimates produced isolated outliers that appear to enhance the association with age. Removal of the outliers, however, had no effect on the correlations between ROI PVEs beyond the third decimal place, and therefore, all the observations were retained in the analyses.

For some ROIs, we observed nonlinear effects, i.e., age-related differences in PVE became greater with age (Figures 6–9). We formally tested each of the regressions of PVE or volume on age for a quadratic fit (in no instance was a cubic polynomial a better fit). The results (tabulated in Table 8) indicated that the magnitude of nonlinear effects found by PVE and by manual findings correlated, although the agreement was far from perfect: $\rho = .55$, $p < .05$. The PVE index found nonlinearity more frequently (in six ROIs) than manual measures (only three), sign test: $p = .02$.

4. Discussion

In this study we have compared VBM to the current gold standard of volumetry, manual tracing, in their ability to detect age differences. We improved upon the limitations of the existing comparisons by significantly increasing the sample size and the number of regions that were traced manually with a higher level of reliability, and by providing a quantitative means of evaluating these two techniques via binary masks based on the manual ROIs. The analysis of voxel-aggregated ROIs is atypical for a VBM approach to analysis of structural brain properties. In the extant literature, assessment of neuroanatomically plausible ROIs is downplayed in favor of peak analysis of statistically defined but not biologically plausible clusters of pixels. Manual volumetry, in contrast, is based on anatomically defined gross anatomical units, such as gyri and nuclei. Thus, creation of equivalent ROI-based objects in VBM was necessary for a valid quantitative comparison to manual methods. It is worth noting that this approach not only did not detract from VBM usefulness, but in fact, complemented the semi-automated approach by revealing regional non-linear effects of age. Such effects would not have been detected otherwise.

In a qualitative comparison, we found that despite a substantial regional overlap in the pattern of age differences in regional brain structure, discrepancies between the methods were apparent. VBM peak analysis yielded a more widespread and less differential pattern of age differences than manual measures. According to manual volumetry (Raz et al., 2004), dorsolateral prefrontal cortex exhibited the greatest age-related differences, whereas significantly weaker associations were observed in the PFC white matter, sensory-motor, and visual association regions. Also, according to the manual measures, the primary visual, anterior cingulate, the inferior parietal cortices, and the parietal white matter volumes were not adversely affected by age. In contrast, in optimized VBM analyses of peak Z distribution, the strongest negative associations with age were found in the insula, superior temporal gyrus, and precentral gyrus. Although we had not assessed it for this sample (Raz et al., 2004), we have since manually measured volume of the insula in our laboratory in another sample of 115

healthy adults (19–83 yrs old) and found it correlated with age $r = -.27$ (Raz et al, unpublished data). This is a much weaker association with age than that of prefrontal cortex. In fact, nine other measured regions had a stronger age effect than the insula. Thus, the strongest VBM effects were limited to the regions of parenchyma-CSF interface such as Sylvian and interhemispheric fissures, especially the superior temporal gyrus/insula areas. It is plausible that this discrepancy reflects the fact that due to smoothing, there may appear to be reduced gray matter in a gyrus, simply because the neighboring sulcus has widened with age, either due to gray or white tissue loss elsewhere. Or it could be due to age-related changes in MRI contrast, discussed further below. It is likely, however a product of misregistration that can occur in the VBM method, or the combined effect of all of these problems.

Regarding the white matter, VBM peak analyses revealed age differences mainly in the periventricular regions. Thus, significant VBM peaks of age differences in density were more likely to appear in the regions that either bordered large pools of cerebrospinal fluid (CSF) or were likely to contain regions of CSF that on T2-weighted images appear as white-matter hyperintensities. The latter is hardly surprising, as in published VBM analyses of gray matter volumes, ventriculomegaly is associated with artifactual findings of disease-related reduction in local density (Duran et al, 2006; Senjem et al., 2005). In the past, the designers and major practitioners of the VBM methods (Ashburner & Friston, 2001; Good et al., 2001) have cautioned about a common problem with VBM: voxels on the edge of white matter and ventricles often appearing as gray matter. Consequently, registration works the least well in instances where there are large shape/size differences between groups, especially around the ventricles. Unfortunately, ventriculomegaly is one of the most common neuroanatomical features of aging and of many psychiatric and neurodegenerative diseases, the very populations that VBM users are most interested in studying. A systematic investigation of the effects of spatial normalization is needed to address concerns that the degree of deformation of individual brains fitted to a common template may depend on age (e.g., Bookstein, 2001).

The observed elevated sensitivity of VBM to age differences in regions highly prone to partial volume artifacts and white matter heterogeneity suggests that the patterns of VBM results should not be taken on face value and the abundance of positive findings should not necessarily be viewed as a sign of methodological superiority. For example, the insula, (a structure that lies in the depth of the Sylvian fissure in the circular gyrus and near the claustrum) is frequently reported as a location of significant age- and disease-related differences (e.g., Guiliani et al., 2005; Kubicki et al., 2002 and 238 studies in Pubmed bibliographical database accessed on October 23, 2007). When the image is spatially smoothed, and its resolution reduced sometimes more than 1000-fold, small structures are averaged and small differences are amplified (see also Allen et al., 2005). Those amplified differences include relatively small imperfections of registration, thus confounding the subsequent voxel-based analyses (Bookstein, 2001). When a region that is commonly found as the focus of significant differences by VBM analyses is also the region that would suffer the most from misregistration, extra caution in interpretation of results is advisable.

Another potential source of error in the highly-automated VBM method is incomplete removal of the non-brain matter from the images in the skull-stripping phase. Separation of brain and non-brain tissue requires expertise and thus cannot be totally automated, adding a necessary semi-automatic step of removing discrepant tissue by hand, as viewed by the expert eye vs and automated algorithm before proceeding with processing. In the current study we visually examined each brain in detail after every step of the VBM process including skull-stripping where we checked and manually removed any excess extracranial tissue not removed via the automated algorithm.

Although not the main focus of this study, sex differences were examined in both manual and VBM analyses. Manual volumetry revealed several sexually dimorphic regions, with men having significantly larger volumes in the visual cortex, parahippocampal gyrus, anterior cingulate gyrus, the hippocampus, and prefrontal white matter. That pattern of sex differences was not entirely consistent with that observed in our previous sample drawn from the same population (Raz et al., 1997) which reported significant sex differences only in the hippocampal and primary visual cortex volumes, highlighting the unstable nature of sex effects. Lack of clear replication of sex differences across samples has been reported in other studies as well (e.g., Cowell et al., 2007, Chen et al., 2006; DeCarli et al., 2005; Nopoulos et al., 2000). Several sex differences in modulated peak gray matter density were observed in the VBM analyses when we applied the typically used correction for total brain volume. They all favored women. However, when manually measured intracranial volume was used as a covariate, the sex differences all but disappeared. This finding reinforces the recommendation to use an extra-cerebral index such as manually or semi-automatically traced intracranial vault volume as a covariate for head size correction (Walhovd et al., 2005).

In the context of an aging brain study, the possibility of sexually dimorphic age trajectories warrants attention. We examined the interaction between sex and age-related volume differences and found no agreement between manual volumetry and VBM. With manual volumetry, sex differences in hippocampus, fusiform gyrus and prefrontal white matter disappeared with age. VBM revealed age-dependent sex differences in a non-overlapping set of regions: peri-rolandic gray, peri-Sylvian white, and pericallosal gray and white. Notably, all VBM-detected interactive effects fell along the banks of the two major fissures (Sylvian and Rolandic) and along the major white matter structure – the corpus callosum. A finding that may reflect age-related reduction in gray-white-CSF contrast. Moreover, previous reports indicated disproportional expansion of sulcal space in men (Coffey et al., 1998; Gur et al., 1991), increased cortical thickness in women (Sowell et al., 2007) and young adults (Luders et al., 2006) potential sources of bias in VBM findings.

The VBM analyses (but not the manual measures) found small circumscribed areas of age-related modulated density loss that accelerated with hypertension. Most of those findings, however, were of questionable robustness and did not survive application of more stringent statistical inference criteria. Nonetheless, those that remained significant at the adjusted p-levels corresponded to the hypertension-related differences observed in other cross-sectional samples drawn from this same population as well as in a longitudinal follow-up of a portion of this sample (Head et al., 2002; Raz, et al., 2003; Raz et al., 2005; Raz et al., 2007a; Raz et al., 2007b). Thus, in the case of hypertension, VBM was more likely to detect what appears to be a replicable difference in regional brain structure.

When analyses of statistically-derived isolated peaks were replaced by biologically more plausible analyses of homogenous anatomically defined regions via aggregation of PVE values (in masks), the patterns of findings obtained by either method became more similar. However, VBM failed to detect age differences in white matter regions, showing slight increases in modulated density for some of them, which may be due to errors in segmentation and/or registration. Notably, the magnitude of age differences detected by the regional PVE approach was greater in the regions of elevated pixel PVE variability. Thus, in future applications of VBM to the study of individual and age-related differences, it may be advisable to pay at least as much attention to within-region variability as to between-region differences in parenchymal properties.

The shapes of estimated trajectories of regional brain aging differed between the methods. In some regions, such as the hippocampus and the prefrontal white matter, nonlinearity (age-related acceleration of estimated declines) was detected by both methods. In other regions,

inter-method discrepancies were observed. Except for one region (fusiform gyrus), VBM analyses were significantly more likely to detect nonlinearity than did the manual measures. In this study, it is impossible to declare which of these effects are "true" and which are spurious. We infer, however, that the observed white matter nonlinearities reflect the reality of brain aging as they are in accord with multiple converging findings (see Raz and Rodrigue, 2006 for a review). Thus, VBM may be more sensitive to nonlinear trends in the white matter volumes. However, nonlinearities in at least some of the VBM gray matter regions may be artifactual. They may reflect disproportionately significant distortion in the regions of high individual variability and high proneness to MRI artifacts, such as increased flow in the vicinity of the parahippocampal gyrus, or as a result of smoothing in these small areas.

Dependence on high image quality and homogeneity is an inherent limitation of the VBM methods. In VBM, tissue classification is based on calculated voxel intensities, which changes as a decelerating quadratic function of age (Cho et al., 1997). The result of these T1-weighted changes is age-related alterations of image contrast. Thus, in the brains of older participants, age-related reduction of gray matter T1 may result in gray matter pixels appearing more similar to the white ones than on the images of younger brains. In addition, flow and motion artifacts create contrast situations that are problematic for intensity-based approaches and can be resolved only by imposing anatomical constraints. The human eye is likely better than an algorithm in taking these contrast differences into account, and correctly assessing the likelihood of the white-gray distinction. Trained human operators guided by the "top-down" influence of their knowledge of anatomy have several advantages over digital algorithmic approaches. The attempts to incorporate "training-like" information in computer decision making (e.g., Fischl et al., 2002) seem to work better for simple rather than anatomically complex structures.

Conclusions

The results of this study reveal significant discrepancy between manual volumetry and VBM analysis of age differences when the latter is confined to identification of statistically significant peak Z values. However, when compared on a common ground, (i.e. with regards to anatomically definable regions), semi-automated and manual methods of evaluating regional brain structure yield similar pattern of results, some notable exceptions notwithstanding. The discrepancies tend to occur in the regions of increased inter-voxel variability. However, when highly circumscribed isolated differences in density (the peaks) are taken as indicators of structural integrity, the correspondence between the two approaches is low. Because isolated clusters of voxels have no clear biological meaning and because VBM algorithms are sensitive to localized artifactual differences in intensity, voxel-by-voxel analyses of the whole brain can be interpreted only as an exploratory hypotheses-generating preamble to a careful analysis of potential artifacts and examination of the structural properties in anatomically definable regions. There is a danger of amassing VBM studies to collectively reify potentially spurious findings. For example, when different VBM studies produce the same results (i.e., the insula), this might be more likely to be taken as "validation" of a finding rather than as perpetuation of an artifact.[1]

Variability within given regions may be informative of age-related processes compromising brain integrity and should be examined with no lesser attention than the between-region mean differences. A similar point was recently made in regards to interpretation of fMRI experiments on purely statistical grounds: there is no evidence that "red" pixels are statistically different from "orange" pixels nearby, yet the former are accepted as more meaningful findings than the latter (Jernigan et al., 2003). Unfortunately, that important argument somehow did not exert

---

[1]We are grateful to an anonymous reviewer for pointing out this important implication.

due influence on the current research practices in the field of neuroimaging. Lastly, VBM methods may be better suited for detection of nonlinearities in age trends but this matter needs careful investigation considering a possibility that nonlinearities may reflect registration artifacts.

In sum, for an optimal evaluation of age-related differences in brain structure, VBM as it is utilized now should be used judicially as a hypothesis-generating device. The maps of group or individual differences generated by VBM should be used in conjunction with regional volumetry aimed at testing specific hypotheses about plausible neuroanatomical units. In this way, each technique will accentuate the advantages and ameliorate the limitations of the other.

Acknowledgments

References

Allen JS, Bruss J, Brown CK, Damasio H. Methods for studying the aging brain: Volumetric analyses versus VBM [Commentary]. Neurobiol Aging 2005;26:1275–1278.

Ashburner JA, Friston KJ. Voxel-based morphometry—the methods. NeuroImage 2000;11:805–821. [PubMed: 10860804]

Ashburner JA, Friston KJ. Why voxel-based morphometry should be used. Neuroimage 2001;14:1238–1241. [PubMed: 11707080]

Aydin K, Ucar A, Oguz KK, Okur OO, Agayev A, Unal Z, Yilmaz S, Ozturk C. Increased gray matter density in the parietal cortex of mathematicians: A Voxel-Based Morphometry study. Am J Neuroradiol. 2007 Oct 5;[Epub ahead of print]

Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and senile change in the cerebral gray matter of elderly subjects. Br J Psychiatry 1968;114:797–811. [PubMed: 5662937]

Bookstein FL. Voxel-based morphometry should not be used with imperfectly registered images. NeuroImage 2001;14:1454–1462. [PubMed: 11707101]

Cardenas VA, Du AT, Hardin D, Ezekiel F, Weber P, Jagust WF, Chui HC, Schuff N, Weiner MW. Comparison of methods for measuring longitudinal brain change in cognitive impairment and dementia. Neurobiol Aging 2003;24:537–544. [PubMed: 12714110]

Chen X, Wen W, Anstey KJ, Sachdev PS. Effects of cerebrovascular risk factors on gray matter volume in adults aged 60–64 years: a voxel-based morphometric study. Psychiatry Res 2006;147:105–114. [PubMed: 16962291]

Cho S, Jones D, Reddick WE, Ogg RJ, Steen RG. Establishing norms for age-related changes in proton T1 of human brain tissue in vivo. Magn Reson Imaging 1997;15:1133–1143. [PubMed: 9408134]

Coffey CE, Lucke JF, Saxton JA, Ratcliff G, Unitas LJ, Billig B, Bryan RN. Sex differences in brain aging: a quantitative magnetic resonance imaging study. Arch Neurol 1998;55:169–179. [PubMed: 9482358]

Colcombe SJ, Erickson KI, Raz N, Webb AG, Cohen NJ, McAuley E, Kramer AF. Aerobic fitness reduces brain tissue loss in aging humans. J Gerontol Biol Sci Med Sci 2003;58:176–180.

Colcombe SJ, Kramer AF, Erickson KI, Scalf P. The implications of cortical recruitment and brain morphology for individual differences in inhibitory function in aging humans. Psychol Aging 2005;20:363–375. [PubMed: 16248697]

Cowell PE, Sluming VA, Wilkinson ID, Cezayirli E, Romanowski CA, Webb JA, Keller SS, Mayes A, Roberts N. Effects of sex and age on regional prefrontal brain volume in two human cohorts. Eur J Neurosci 2007;25:307–318. [PubMed: 17241292]

Crum WR, Griffin LD, Hill DLG, Hawkes DJ. Zen and the art of medical image registration: correspondence, homology, and quality. NeuroImage 2003;20:1425–1437. [PubMed: 14642457]

Davatzikos C. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. NeuroImage 2004;23:17–20. [PubMed: 15325347]

DeCarli C, Massaro J, Harvey D, Hald J, Tullberg M, Au R, Beiser A, D'Agostino R, Wolf PA. Measures of brain morphology and infarction in the framingham heart study: establishing what is normal. Neurobiol Aging 2005;26:491–510. [PubMed: 15653178]

Dorion AA, Salomon O, Zanca M, Duyme M, Capron C. Magnetic resonance imaging and cerebral volumic evaluation: comparison of four post-processing techniques. Comput Biol Med 2001;31:215–227. [PubMed: 11334632]

Douaud G, Gaura V, Ribeiro MJ, Lethimonnier F, Maroy R, Verny C, Krystkowiak P, Damier P, Bachoud-Levi AC, Hantraye P, Remy P. Distribution of gray matter in Huntington's disease patients: A combined ROI-based and voxel-based morphometric study. NeuroImage 2006;32:1562–1575. [PubMed: 16875847]

Duran, FSL.; Valente, AA.; Miguel, EC.; Castro, CC.; Busatto, GF. Risk of artifacts due to enlarged ventricles using voxel-based morphometry studies. Poster presented at Organization for Human Brain Mapping conference; Florence, IT: Neuroimage; Jun. 2006 p. 31

Eckert MA, Tenforde A, Galaburda AM, Bellugi U, Korenberg JR, Mills D, Reiss AL. To modulate or not to modulate: Differing results in uniquely shaped Williams syndrome brains. NeuroImage 2006;32:1001–1007. [PubMed: 16806978]

Erickson KI, Colcombe SJ, Raz N, Korol DL, Scalf P, Webb A, Cohen NJ, McAuley E, Kramer AF. Selective sparing of brain tissue in postmenopausal women receiving hormone replacement therapy. Neurobiol Aging 2005;26:1205–1213. [PubMed: 15917105]

Erickson KI, Colcombe SJ, Elavsky S, McAuley E, Korol DL, Scalf PE, Kramer AF. Interactive effects of fitness and hormone treatment on brain health in postmenopausal women. Neurobiol Aging 2007;28:179–185. [PubMed: 16406152]

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33:41–55.

Giuliani NR, Calhoun VD, Pearlson GD, Francis A, Buchanan RW. Voxel-based morphometry versus region of interest: a comparison of two methods for analyzing gray matter differences in schizophrenia. Schizophre Res 2005;74:135–147.

Gong QY, Sluming V, Mayes A, Keller S, Barrick T, Cezayirli E, Roberts N. Voxel-based morphometry and stereology provide convergent evidence of the importance of medial prefrontal cortex for fluid intelligence in healthy adults. NeuroImage 2005;25:1175–1186. [PubMed: 15850735]

Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RSJ. Cerebral asymmetry and the effects of sex and handedness on brain structure: A voxel-based morphometric analysis of 465 normal adult human brains. NeuroImage 2001a;14:685–700. [PubMed: 11506541]

Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RSJ. A voxel-based morphometric study of ageing in 465 normal adult human brains. NeuroImage 2001b;14:21–36. [PubMed: 11525331]

Good CD, Scahill RI, Fox NC, Ashburner J, Friston KJ, Chan D, Crum WR, Rossor MN, Frackowiak RSJ. Automatic differentiation of anatomical patterns in the human brain: Validation with studies of degenerative dementias. NeuroImage 2002;17:29–46. [PubMed: 12482066]

Gur RC, Mozley PD, Resnick SM, Gottlieb GL, Kohn M, Zimmerman R, Herman G, Atlas S, Grossman R, Berretta D, et al. Gender differences in age effect on brain atrophy measured by magnetic resonance imaging. Proc Natl Acad Sci USA 1991;88:2845–2849. [PubMed: 2011592]

Gur RC, Gunning-Dixon F, Bilker WB, Gur RE. Sex differences in temporo-limbic and frontal brain volumes of healthy adults. Cereb Cortex 2002;12:998–1003. [PubMed: 12183399]

Head D, Raz N, Gunning-Dixon F, Williamson A, Acker JD. Age-related differences in the course of cognitive skill acquisition: the role of regional cortical shrinkage and cognitive resources. Psychol Aging 2002;17:72–84. [PubMed: 11931289]

Honea R, Crow TJ, Passingham D, Mackay CE. Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. Am J Psychiatry 2005;162:2233–2245. [PubMed: 16330585]

Jenkinson M, Smith SM. A global optimization method for robust affine registration of brain images. Med Image Anal 2001;5:143–156. [PubMed: 11516708]

Jernigan TL, Gamst AC, Fennema-Notestine C, Ostergaard AL. More "mapping" in brain mapping: statistical comparison of effects. Hum Brain Mapp 2003;19:90–95. [PubMed: 12768533]

Kubicki M, Shenton ME, Salisbury DF, Hirayasu Y, Kasai K, Kikinis R, Jolesz FA, McCarley RW. Voxel-based morphometric analysis of gray matter in first episode schizophrenia. NeuroImage 2002;17:1711–1719. [PubMed: 12498745]

Luders E, Narr KL, Thompson PM, Rex DE, Woods RP, Deluca H, Jancke L, Toga AW. Gender effects on cortical thickness and the influence of scaling. Hum Brain Mapp 2006;27:314–324. [PubMed: 16124013]

May A, Ashburner J, Buchel C, McGonigle DJ, Friston KJ, Frackowiak RS, Goadsby PJ. Correlation between structural and functional changes in brain in an idiopathic headache syndrome. Nat Med 1999;57:836–838. [PubMed: 10395332]

Mechelli A, Price CJ, Friston KJ, Ashburner J. Voxel-based morphometry of the human brain: methods and applications. Curr Med Imaging Rev 2005;11:1–9.

Mehta S, Grabowski TJ, Trivedi Y, Damasio H. Evaluation of voxel-based morphometry for focal lesion detection in individuals. NeuroImage 2003;20:1438–1454. [PubMed: 14642458]

Morrell MJ, Twigg G. Neural consequences of sleep disordered breathing: the role of intermittent hypoxia. Adv Exp Med Biol 2006;588:75–88. [PubMed: 17089881]

Nopoulos P, Flaum M, O'Leary D, Andreasen NC. Sexual dimorphism in the human brain: evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. Psychiatry Res 2000;98:1–13. [PubMed: 10708922]

Oldfield RC. The assessment and analysis of handedness. Neuropsychologia 1971;9:97–113. [PubMed: 5146491]

Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psych Measure 1977;1:385–401.

Raz N, Gunning FM, Head D, Dupuis JH, McQuain JD, Briggs SD, Loken WJ, Thornton AE, Acker JD. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. Cereb Cortex 1997;7:268–282. [PubMed: 9143446]

Raz N, Gunning-Dixon F, Head D, Rodrigue KM, Williamson A, Acker JD. Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: replicability of regional differences in volume. Neurobiol Aging 2004;25:377–396. [PubMed: 15123343]

Raz N, Rodrigue KM, Head D, Kennedy KM, Acker JD. Differential changes in the medial temporal structures: a five-year longitudinal study. Neurology 2004;62:433–438. [PubMed: 14872026]

Raz N, Rodrigue KM. Differential aging of the brain: patterns, cognitive correlates and modifiers. Neurosci Biobehav Rev 2006;30:730–748. [PubMed: 16919333]

Raz N, Rodrigue K, Acker J. Hypertension and the brain: vulnerability of the prefrontal regions and executive functions. Beh Neurosci 2003;17:1169–1180.

Raz N, Rodrigue KM, Kennedy KM, Acker JD. Vascular health and longitudinal changes in brain and cognition in middle-aged and older adults. Neuropsychol 2007a;21:149–157.

Raz N, Lindenberger U, Ghisletta P, Rodrigue KM, Kennedy KM, Acker JD. Neuroanatomical correlates of fluid intelligence in healthy adults and persons with vascular risk factors. Cereb Cortex. 200710.1093/cercor/bhm108

Reiss, A.; Hennessey, J.; Rubin, M.; Beach, L.; Subramaniam, B. BrainImage, v. 2.3.3. Behavioral Neurogenetics and Neuroimaging Research Center, Division of the Kennedy Krieger Institute; Baltimore, MD: 1995.

Salmond CH, Ashburner J, Vargha-Khadem F, Connelly A, Gadian DG, Friston KJ. Distributional assumptions in voxel-based morphometry. Neuroimage 2002;17:1027–1030. [PubMed: 12377176]

Senjem ML, Gunter JL, Shiung MM, Petersen RC, Jack CR. Comparison of different methodological implementations of voxel-based morphometry in neurodegenerative disease. NeuroImage 2005;26:600–608. [PubMed: 15907317]

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing raters reliability. Psychol Bull 1979;86:420–428. [PubMed: 18839484]

Smith S. Fast robust automated brain extraction. Hum Brain Mapp 2002;17:143–155. [PubMed: 12391568]

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy R, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 2004;23(S1):208–219.

Sowell ER, Peterson BS, Kan E, Woods RP, Yoshii J, Bansal R, Xu D, Zhu H, Thompson PM, Toga AW. Sex differences in cortical thickness mapped in 176 healthy individuals between 7 and 87 years of age. Cereb Cortex 2007;17:1550–1560. [PubMed: 16945978]

Steiger JH. Tests for comparing elements of a correlation matrix. Psychol Bull 1980;87:245–251.

Tapp PD, Head K, Head E, Milgram NW, Muggenburg BA, Su MY. Application of an automated voxel-based morphometry technique to assess regional gray and white matter brain atrophy in a canine model of aging. NeuroImage 2006;29:234–244. [PubMed: 16275022]

Testa C, Laakso MP, Sabattoli F, Rossi R, Beltramello A, Soininen H, Frisoni GB. A comparison between the accuracy of voxel-based morphometry and hippocampal volumetry in Alzheimer's disease. JMRI 2004;19:274–282. [PubMed: 14994294]

Tisserand DJ, Pruessner JC, Arigita EJS, van Boxtel MJ, Evans AC, Jolles J, Uylings HBM. Regional frontal cortical volumes decrease differentially in aging: an MRI study to compare volumetric approaches and voxel-based morphometry. NeuroImage 2002;17:657–669. [PubMed: 12377141]

Walhovd KB, Fjell AM, Reinvang I, Lundervold A, Dale A, Quinn BT, Salat D, Makris N, Fischl B. Neuroanatomical aging: universal but not uniform. Neurobiol Aging 2005;26:1279–1282.

Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm. IEEE T Med Imaging 2001;20:45–57.
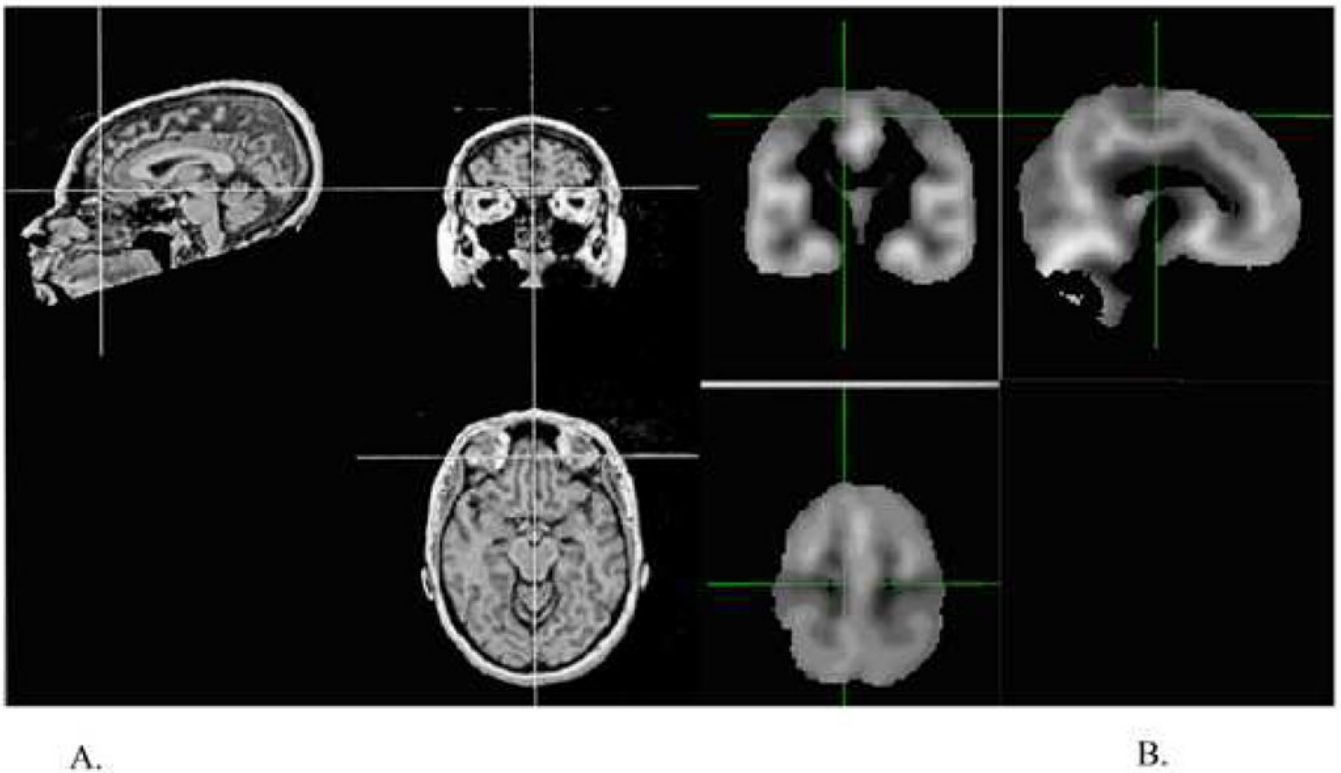
Figure 1.
Examples of MRI images prepared for manual (A) and VBM (B) processing. Manual alignment involving rotation of the brain volume to correct for head tilt, pitch, and rotation was performed on each subject's high-resolution scan (A). For VBM (B), segmentation map produced after skull stripping, stretching to standard space, and smoothing. Note the observable difference in the resulting image resolution quality between the two methods.
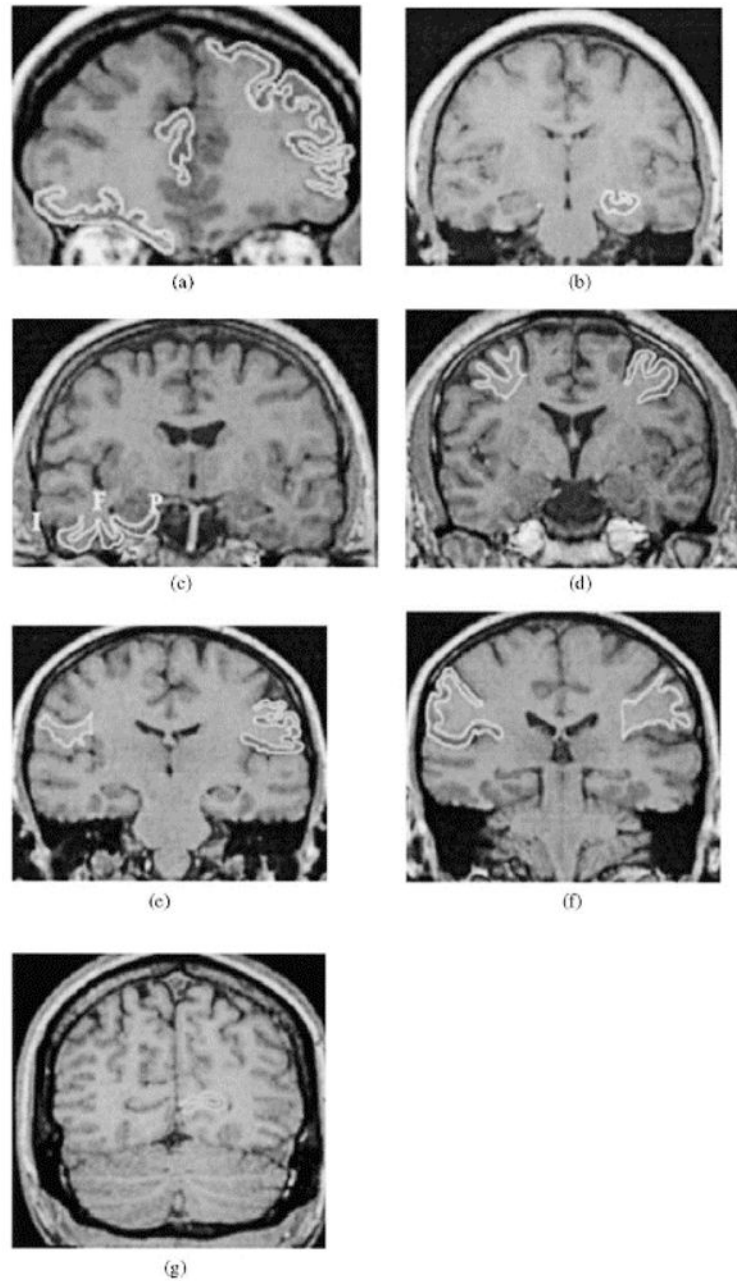
Figure 2.
Examples of manually traced Regions of Interest (ROI): (a) dorsolateral prefrontal cortex (left), orbito-frontal cortex (right), and anterior cingulate gyrus (middle); (b) hippocampus; (c) clockwise: parahippocampal gyrus (P), fusiform (F) and inferior temporal (I) cortices; (d) pre-central gyrus: cortex (left), pre-central white matter (right); (e) post-central gyrus: cortex (left), post-central white matter (right); (f) inferior parietal lobule; cortex (right), white matter (left); (g) visual (calcarine) cortex.
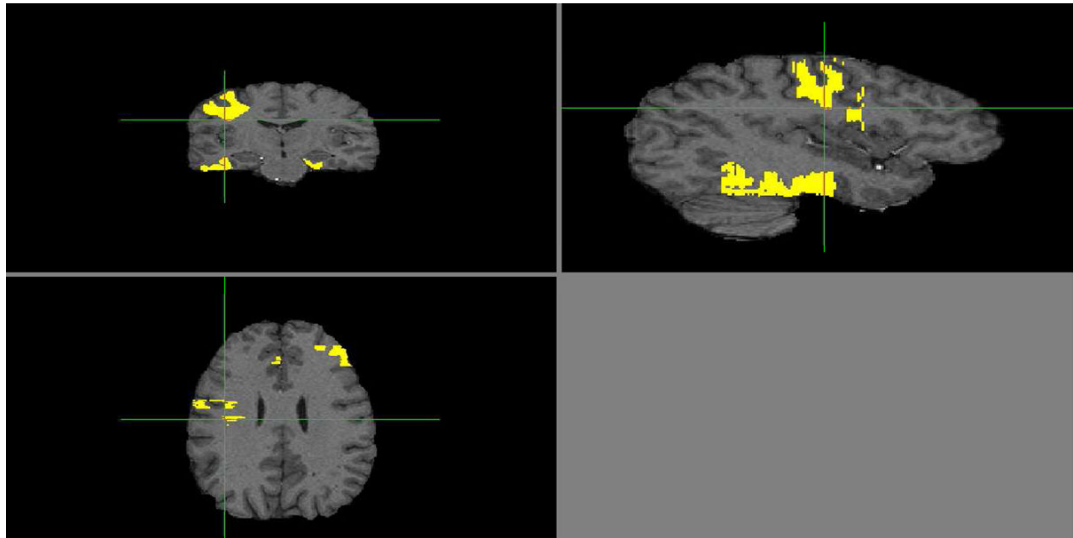
Figure 3.
Example of manually drawn ROI masks for VBM analysis. The same ROIs were manually drawn on a representative subject in native space and then registered to standard space to create the masks for a quantitative comparison of VBM and manual volumetry.
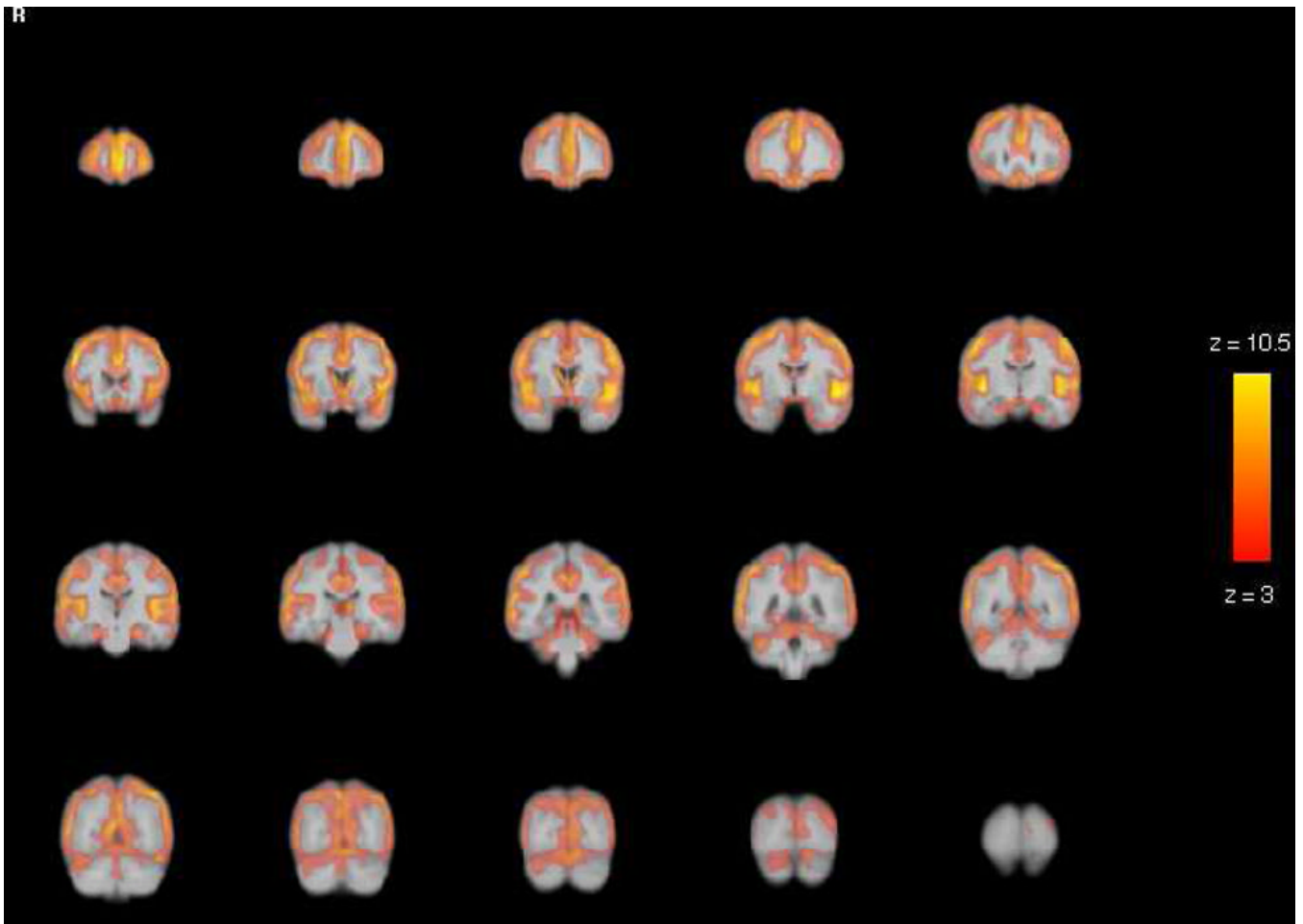
Figure 4.
VBM Gray Matter Results: Statistical parametric maps for age differences in gray matter revealed significant age-related reduction in modulated density in all lobes. The strongest effects were limited to the regions of parenchyma-CSF interface such as Sylvian and interhemispheric fissures, especially the superior temporal gyrus. The SPM is a map of significant Z-statistic values that correspond in color to the magnitude legend to the right, with the yellow voxels having an effect of up to 10.5. Note that because these are Z values, they can be directly compared to the magnitude of effects for white matter in Figure 5.
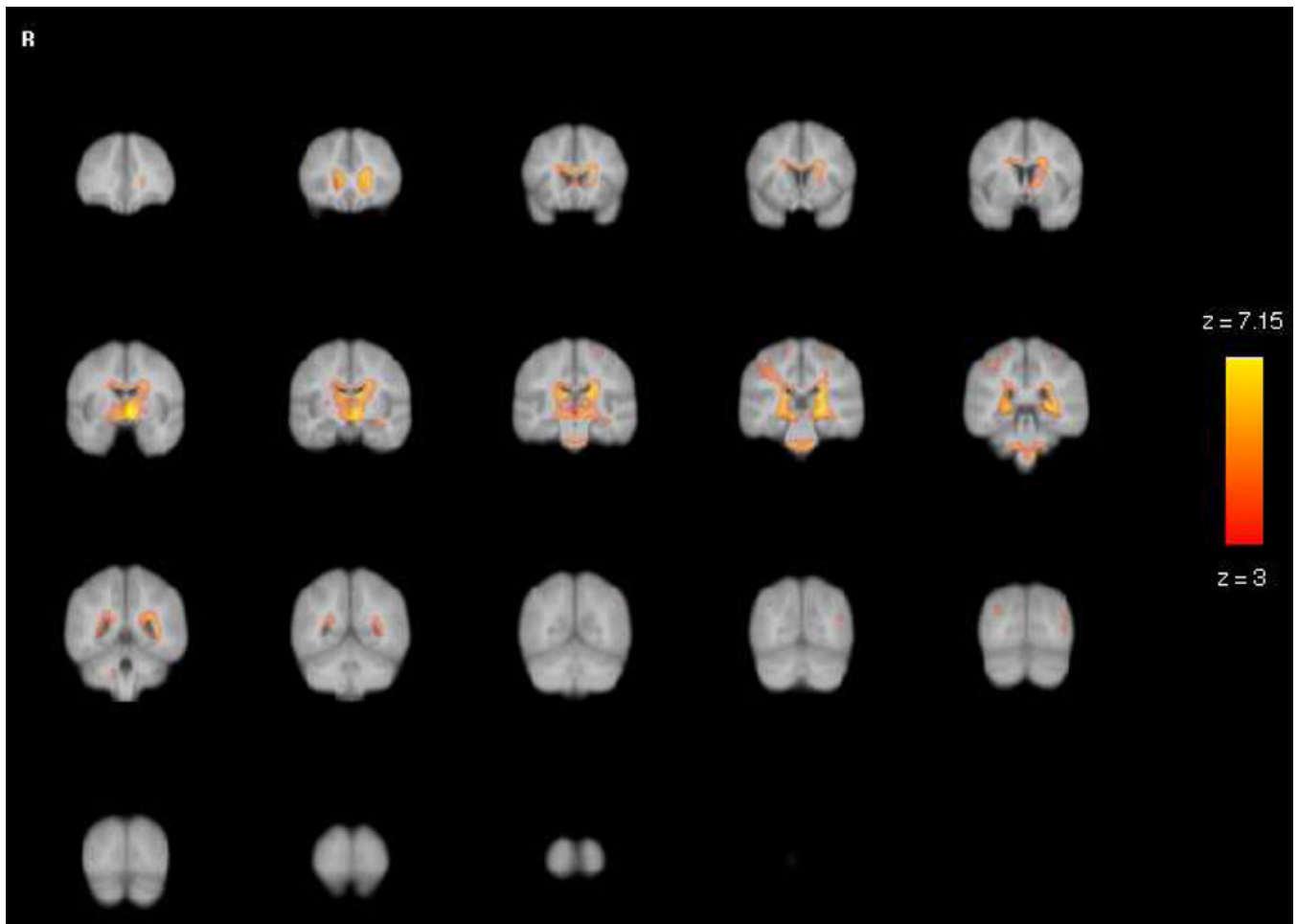
Figure 5.
VBM White Matter Results: Statistical parametric maps for age differences in white matter revealed age-related differences in modulated density were considerably less widespread than in gray matter and concentrated in the anterior and middle corpus callosum, periventricular areas and cerebellar peduncles. Note that the largest effects for white matter were in the yellow voxels, with values of up to 7.15, substantially smaller than the gray matter effects (of up to 10.5) shown in Figure 4.
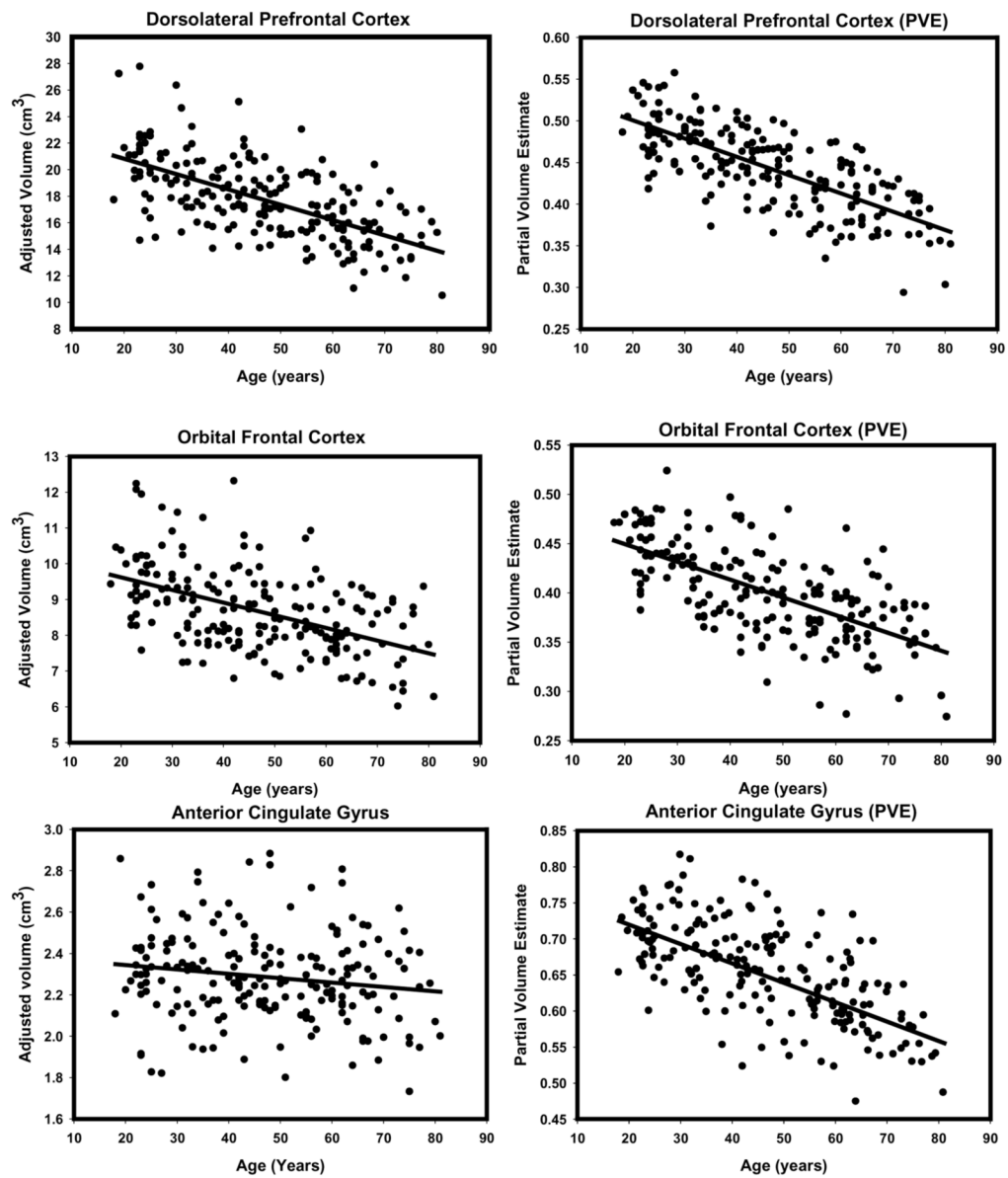
Figure 6.
Scatterplots of bivariate relations between age manual volumes or VBM-derived partial volume estimates for prefrontal gray matter regions (dorsolateral prefrontal cortex, orbitofrontal cortex, anterior cingulate gyrus). All regressions were conducted using intracranial volume adjustment.
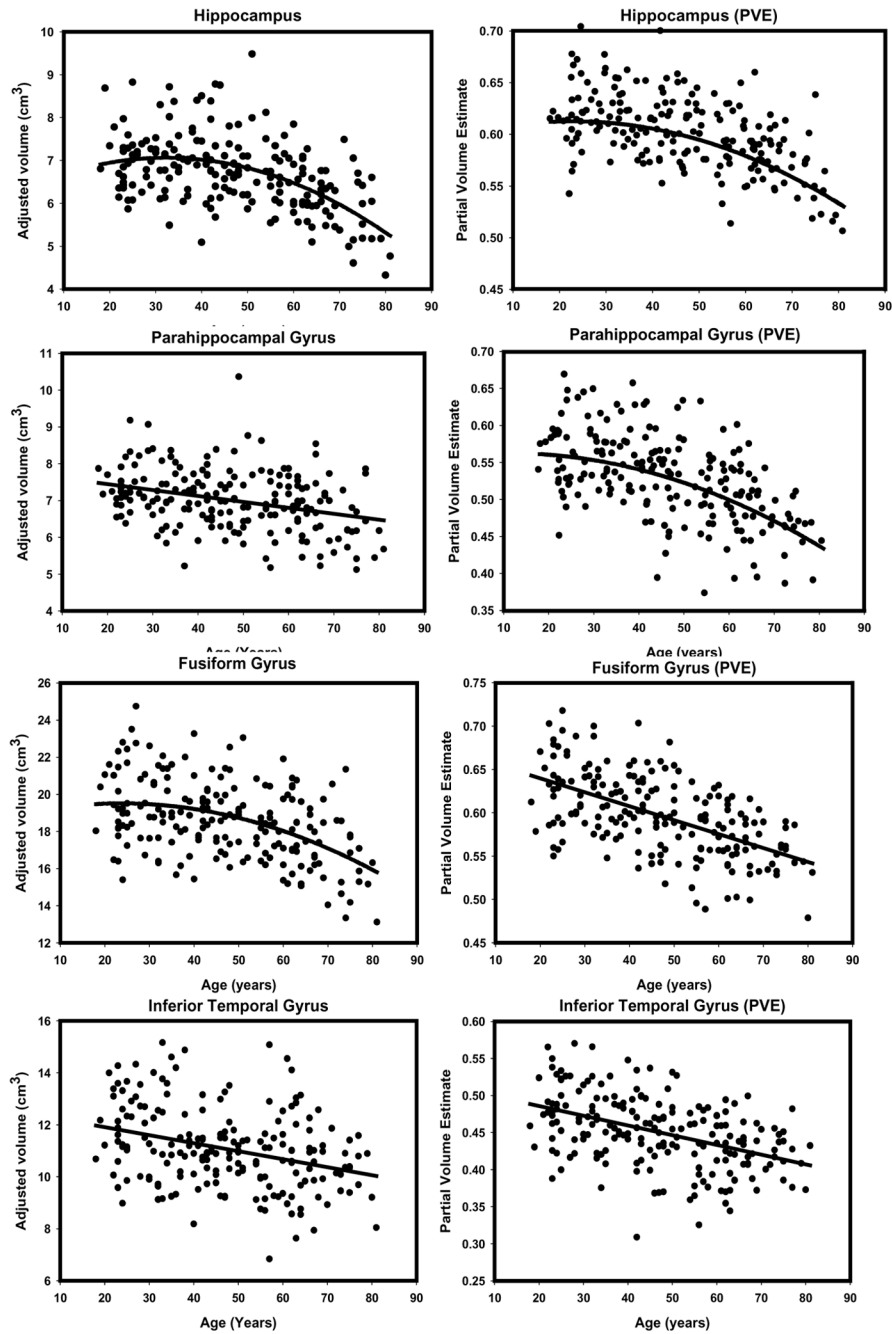
Figure 7.
Scatterplots of bivariate relations between age manual volumes or VBM-derived partial volume estimates for temporal lobe gray matter regions (hippocampus, parahippocampal gyrus, fusiform gyrus, inferior temporal gyrus). All regressions were conducted using intracranial volume adjustment.
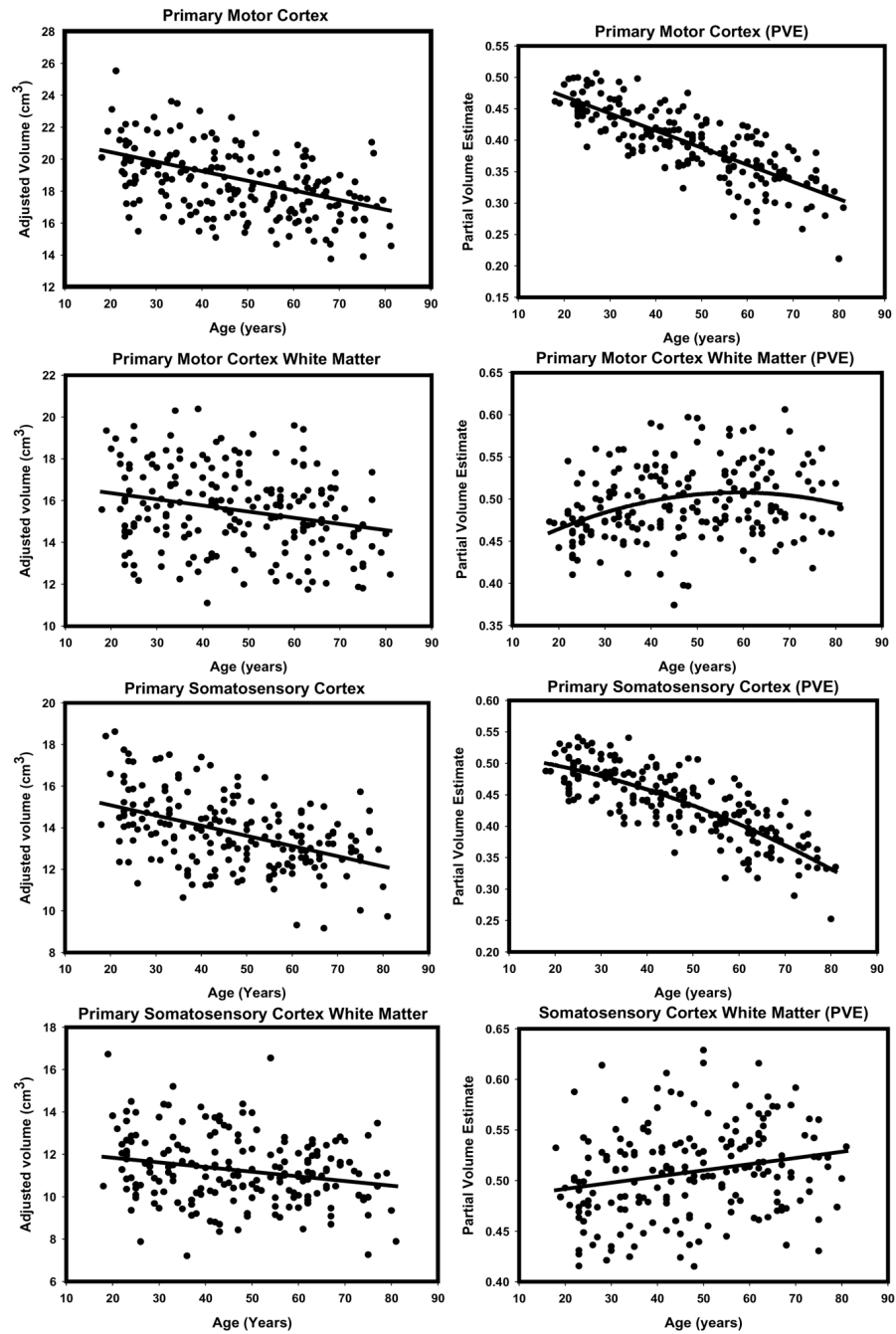
Figure 8.
Scatterplots of bivariate relations between age manual volumes or VBM-derived partial volume estimates for primary motor and somatosensory cortices (gray and white matter). All regressions were conducted using intracranial volume adjustment.
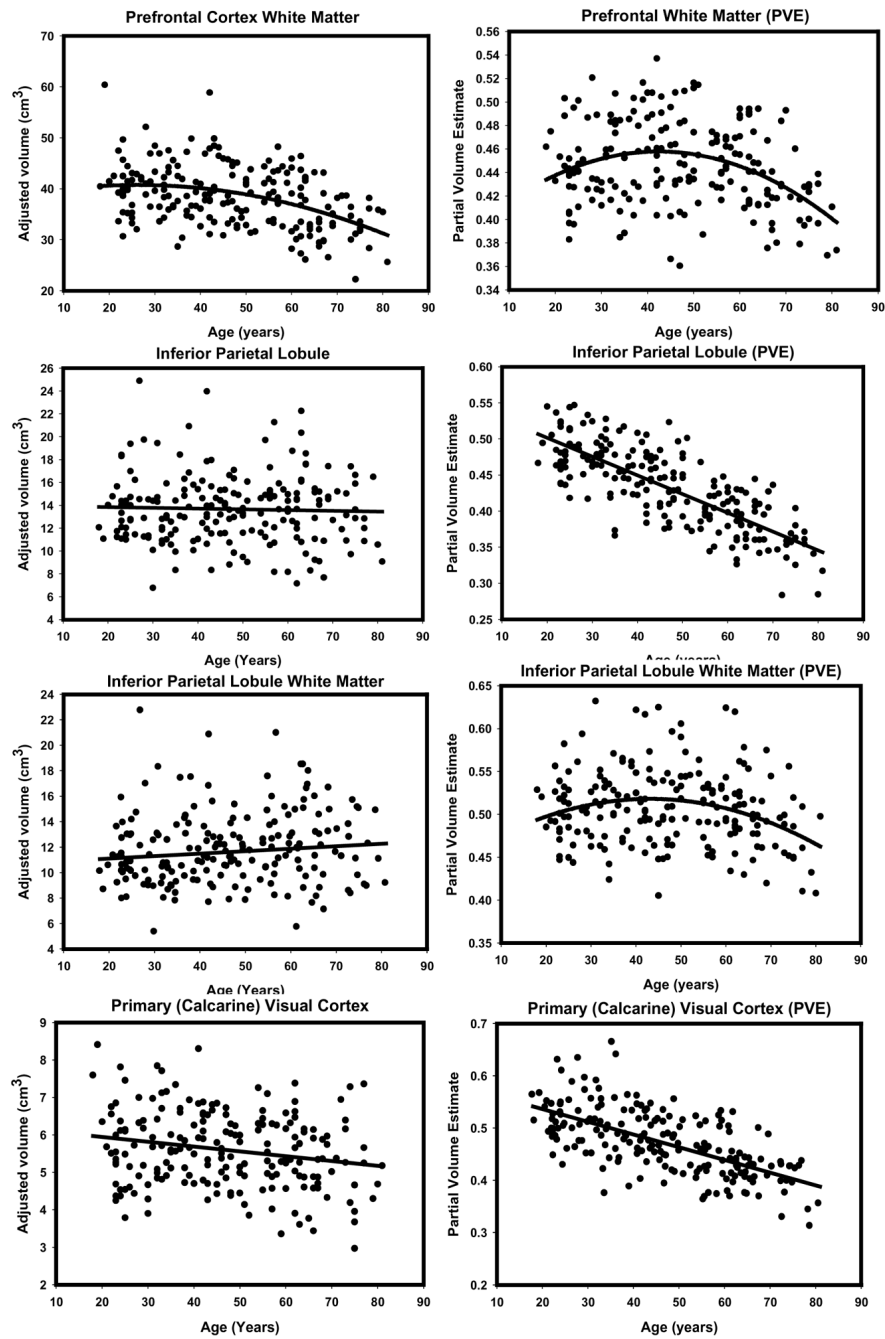
Figure 9.
Scatterplots of bivariate relations between age manual volumes or VBM-derived partial volume estimates for prefrontal cortex white matter, inferior parietal lobule (gray and white matter), primary visual (calcarine) cortex. All regressions were conducted using intracranial volume adjustment.

Table 1

Sample description by age decades.

Age Decades (years)	N	N Men	N Women	N Hypertensives	Education, years (mean ± SD)
18–30	41	18	23	0	16.37±1.71
31–40	35	14	21	3	16.34±2.70
41–50	42	18	24	5	16.09±2.31
51–60	30	8	22	2	15.17±2.89
61–70	35	15	20	6	14.83±2.29
71–80	17	8	9	4	15.23±2.68

Note. N – number; SD – standard deviation.

Table 2

Manual volumetry: Regional volumes, reliabilities, and correlations with age

Region Of Interest	Volume (cm^3)		r Age	ICC (2)
	L	**R**		
Anterior cingulate gyrus (ACG)	1.14	1.14	−.14	.90
Dorsolateral prefrontal cortex (DLPFC)	8.81	8.88	−.63	.96
Fusiform gyrus (FG)	9.16	9.43	−.44	.93
Prefrontal white matter (FW)	19.17	19.26	−.39	.95
Hippocampus (HC)	3.36	3.33	−.46	.97
Inferior parietal lobule (IP)	6.72	6.94	.04	.96
Inferior parietal white matter (IPW)	5.81	5.80	.11	.96
Inferior temporal gyrus (IT)	5.63	5.45	−.33	.95
Pre-central gyrus (MC)	9.40	9.44	−.47	.90
Pre-central gyrus white (MCW)	7.77	7.79	−.24	.96
Orbitofrontal cortex (OFC)	4.35	4.31	−.49	.93
Parahippocampal gyrus (PHG)	3.61	3.39	−.33	.97
Post-central gyrus (SSC)	6.68	7.08	−.47	.96
Post-central gyrus white (SSCW)	5.47	5.78	−.22	.96
Visual (pericalcarine) cortex (VC)	2.78	2.81	−.21	.94

Note: L – volume (cm^3) of left ROI after adjustment for intracranial volume; R – volume (cm^3) of right ROI after adjustment for ICV; r Age – correlation of brain volume with age after adjustment for intracranial volume; ICC – intraclass correlation coefficient, a measure of reliability of tracing.

Table 3

Location and coordinates of voxels demonstrating the highest age-association (adjusted for intracranial volume) for both gray and white matter and mean and standard deviation partial volume estimates (PVEs) from those peak voxels and relation to age.

Peak location	MNI x,y,z	max Z	PVE Mean ±(SD)	PVE r_{age}	PVE-Age R^2
Gray Matter					
L Superior Temporal Gyrus/Insula	−46, −4, −8	10.5	.58 (.080)	−.77	0.59
R Superior Temporal Gyrus	46, −10, −2	10.2	.57 (.080)	−.70	0.49
L Superior Frontal Gyrus	−2, 62, 0	9.88	.59 (.088)	−.78	0.60
R Medial Frontal Gyrus	6, 62, 6	9.76	.59 (.069)	−.74	0.55
R Superior Temporal Gyrus/Insula	54, 0, 2	9.35	.64 (.073)	−.79	0.62
L PreCentral/Inferior Frontal Gyrus	−48, 8, 30	9.33	.60 (.088)	−.73	0.53
White Matter					
L Midbrain	−2, −4, −2	7.15	.52 (.104)	−.66	0.43
Lateral Ventricle	−24, −36, 8	6.67	.63 (.114)	−.46	0.21
L Sublobar Caudate Tail	−16, −24, 20	6.53	.63 (.120)	−.55	0.31
L Thalamus	−8, −16, 18	6.45	.59 (.120)	−.51	0.26
L Dorsal midbrain	−4, −12, 16	6.27	.77 (.091)	−.50	0.25
L Corpus Callosum, anterior	−14, 30, 6	6.17	.62 (.108)	−.56	0.33
R Post Central Gyrus	38, −30, 44	4.11	.58 (.080)	−.26	0.07

Note. All reported results use modulated data. L – left hemisphere; R – right hemisphere; MNI x y z – coordinates in Montreal Neurological Institute space; r Age – correlation of PVE with age after adjustment for intracranial volume; PVE-Age R^2 – squared multiple correlation, i.e. the portion of variance accounted for by age.

Table 4

Location and coordinates of voxels demonstrating the strongest association with diagnosis of hypertension.

Peak location	Gray MNI x,y,z	#voxels	max Z	White MNI x,y,z	#voxels	max Z
Main Effect: Hypertension						
R Inferior Parietal	56,−38,26	69	4.41	—	—	—
L Frontal Medial	−10, 26, 44	35	4.09	—	—	—
L Frontal Medial	−10, 40, 36	24	3.81	—	—	—
L Frontal Medial	−12, 18, 58	15	3.52	—	—	—
L Inferior Parietal	−54, −38, 26	7	3.47	—	—	—
Interaction: Age × Hypertension						
L Frontal Medial	—	—	—	−12, 24, 44	44	4.13
R Superior Temporal Gyrus	—	—	—	48, −10, 8	32	4.11
R Frontal Medial	—	—	—	18, 12, 56	85	3.93
R Inferior Parietal	—	—	—	58, −38, 26	23	3.97
L Posterior Parietal	—	—	—	−22, −64,56	36	3.91
L Frontal Medial	—	—	—	−6, 50, 28	83	3.72
L Frontal Medial	—	—	—	−8, 42, 36	83	3.65
L Frontal Medial	—	—	—	−10, 60, −2	29	3.60

Note. All reported results are from the modulated data. #voxels is the number of contiguous voxels (cluster size) surrounding the peaks. L – left hemisphere, R – right hemisphere.

Table 5

Location and coordinates of voxels demonstrating the strongest sex differences.

Peak location	Gray MNI x,y,z	#voxels	max Z	White MNI x,y,z	#voxels	max Z
Main Effect of Sex (F > M)						
L PreCentral Gyrus	—	—	—	−24, −22, 56	107	4.36
R Superior Temporal Gyrus	42, −20, −2	81	3.74	—	—	—
R PostCentral Gyrus/Inferior	—	—	—	56, −14, 26	50	3.72
R PostCentral Gyrus	—	—	—	46, −26, 46	52	3.71
R Cingulate Gyrus	14, 16, 24	34	3.68	—	—	—
R Cingulate Gyrus	14, 30, 22	34	3.65	—	—	—
L Frontal Medial	−14, 32, 32	12	3.65	—	—	—
L Inferior Temporal	−34, −4, −46	30	3.62	—	—	—
R Cingulate Gyrus	14, 12, 26	34	3.60	—	—	—
R Superior Temporal Gyrus	40, 0, −22	67	3.80	—	—	—
R Inferior Parietal	—	—	—	46, −22, 28	50	3.32
R Medial Temporal	—	—	—	28, −30, −4	39	3.27
Age × Sex Interaction						
Cingulate Gyrus/Posterior	0, −28, 38	29	3.64	—	—	—
R PostCentral Gyrus	42, −26, 44	78	3.62	—	—	—
R Frontal Medial	4, 34, 26	52	3.61	—	—	—
L PreCentral Gyrus	−24, −22, 54	16	3.53	—	—	—
R PreCentral Gyrus	42, −8, 38	27	3.41	—	—	—
R Superior Temporal Gyrus	—	—	—	44, 2, −18	67	3.41
R Cingulate Gyrus	—	—	—	12, 22, 22	26	3.4

Note. All reported results use modulated data. #voxels is a number of contiguous voxels (cluster size) surrounding peaks. L – left hemisphere, R – right hemisphere.

Table 6

Age Differences in ICV-Adjusted Manually Measured Volumes of Mask ROIs: Gray Matter

ROI	Max z-score	Peak coordinates			# of voxels	Volume (mm³)	Mean	SD
		x	y	z				
ACGL	8.89	−4	36	22	363	2904	5.51	2.30
ACGR	8.49	4	38	16	449	3592	4.97	2.10
DLPFCL	8.93	−6	56	28	2966	23728	4.36	1.83
DLPFCR	7.28	30	44	32	3063	24504	4.11	1.49
FGL	7.14	−46	−60	−22	3105	24840	2.42	1.50
FGR	6.20	40	−48	−28	3381	27048	2.56	1.55
HCL	6.37	−24	−24	−18	1143	9144	1.89	2.70
HCR	5.59	18	−12	−22	1181	9448	1.73	2.36
IPL	8.65	−40	−24	16	1743	13944	4.36	1.65
IPR	8.02	52	−16	48	2033	16264	4.09	2.30
ITL	6.58	−46	−54	−26	1721	13768	2.43	1.34
ITR	7.98	64	−32	−4	1905	15240	2.52	1.66
MCL	10.03	−46	−4	−2	3195	25560	5.31	1.92
MCR	9.35	54	0	2	3313	26504	4.99	2.12
OFCL	7.62	−34	30	−14	1420	11360	4.17	1.80
OFCR	7.64	28	32	−16	1375	11000	4.24	1.53
PHGL	6.07	−14	−40	−18	1424	11392	3.30	1.57
PHGR	5.50	20	−22	20	1587	12696	2.87	1.33
SSCL	9.78	−46	−6	0	2451	19608	5.13	2.07
SSCR	8.77	48	−4	2	2038	16304	4.68	2.10
VCL	6.30	−8	−76	12	1236	9888	4.41	1.43
VCR	7.55	8	−60	12	1187	9496	3.92	1.58

Table 6a. Age Differences in ICV-Adjusted Manually Measured Volumes of Mask ROIs: White Matter

ROI	Max z-score	Peak coordinates			# of voxels	Volume (mm³)	Mean	SD
		x	y	z				
FWL	5.97	−14	32	4	4190	33520	0.34	1.46

Table 6a. Age Differences in ICV-Adjusted Manually Measured Volumes of Mask ROIs: White Matter

ROI	Max z-score	Peak coordinates			# of voxels	Volume (mm^3)	Mean	SD
FWR	4.84	16	32	4	4198	33584	0.00	1.38
IPWL	3.15	−54	−18	36	1320	10560	0.03	1.33
IPWR	4.15	38	−30	42	1567	12536	0.65	1.68
MCWL	2.01*	−24	2	32	2360	18880	−1.29	1.27
MCWR	3.26	24	−6	30	2662	21296	−0.55	1.28
SSCWL	2.33*	−58	−16	32	1882	15056	−1.23	1.45
SSCWR	3.10*	60	−16	24	1523	12184	−0.45	1.19

*
Note. = max peak z-score did not reach significance, indicating no significant effect of age within this ROI. All results reported use modulated data. L – left hemisphere; R – right hemisphere; Volume = # of voxels multiplied by 8mm^3. See Table 2 for ROI abbreviations.

Table 7

Summary statistics of VBM partial volume estimates (PVEs) from manually drawn ROI masks and their age effects compared to manual volume, all adjusted for ICV ($N = 199$).

| ROI | Partial Volume Estimates | | | | | r_{Age} | | | r_{PVE_Manual} | Z^* for Δr |
	Min	Max	Mean	SD	C.V.	PVE	Manual			
ACG	0.47	0.81	0.65	0.06	0.10	−0.66	−0.16		0.15	6.47 ***
DLPFC	0.29	0.56	0.44	0.04	0.11	−0.73	−0.63		0.56	2.27 *
FG	0.48	0.72	0.60	0.06	0.08	−0.57	−0.43		0.29	2.03 *
FW	0.36	0.54	0.45	−0.01	0.08	−0.21	−0.39		0.35	2.37 *
HC	0.50	0.69	0.59	0.14	0.06	−0.56	−0.46		0.42	1.6
IP	0.28	0.55	0.43	0.42	0.13	−0.77	−0.04		0.06	9.61 ***
IPW	0.41	0.63	0.51	0.21	0.09	−0.14	0.11		−0.01	2.48 *
IT	0.31	0.57	0.45	0.40	0.11	−0.45	−0.33		0.14	1.45
MC	0.21	0.51	0.40	0.26	0.14	−0.81	−0.47		0.42	6.90 ***
MCW	0.37	0.61	0.50	0.43	0.09	0.24	−0.25		0.21	5.57 ***
OFC	0.27	0.53	0.40	0.22	0.11	−0.65	−0.49		0.39	2.70 **
PHG	0.37	0.66	0.52	0.32	0.11	−0.55	−0.33		0.26	2.98 **
SSC	0.25	0.54	0.44	0.06	0.13	−0.79	−0.47		0.43	6.34 ***
SSCW	0.42	0.63	0.51	0.04	0.09	0.23	−0.22		0.22	5.13 ***
VC	0.31	0.66	0.47	0.06	0.13	−0.64	−0.21		0.32	6.16 ***

Note. ICV = intracranial volume; PVE = partial volume estimate (the proportion of gray matter within a voxel across voxels included in the mask); Z^* = Steiger's Z for the change in correlation. Modulated data were used for all indices;

*
p < .05,

**
p < .01,

***
p < .001. See Table 2 for ROI abbreviations.

Table 8

Test of nonlinearity of effects of age on regional ROIs across the age span for manual ROIs and PVEs from binary masks.

| | Slope | | | | R^2 | | | | $F_{(1, 196)}$ | |
| | Linear (cm³/year) | | Quadratic (cm³/year²) | | Linear | | Quadratic | | | |
ROI	Manual	PVE	Man	PVE	Man	PVE	Man	PVE	Manual	PVE
ACG	−0.002	−0.003	−0.00005	−0.00003	0.026	0.438	0.030	0.448	<1	3.55
DLPFC	−0.115	−0.002	0.00014	−0.000002	0.400	0.527	0.401	0.527	<1	<1
FG	−0.055	−0.002	−0.00113	−0.00001	0.186	0.331	0.206	0.332	4.77[*]	<1
FW	−0.143	−0.00045	−0.00327	−0.00004	0.153	0.042	0.173	0.127	4.67[*]	19.04[**]
HC	−0.024	−0.001	−0.00177	−0.00002	0.216	0.311	0.270	0.341	14.59[**]	8.99[**]
IP	−0.007	−0.003	−0.00037	−0.00001	0.001	0.599	0.002	0.603	<1	2.03
IPW	0.019	−0.00037	−0.00068	−0.00004	0.013	0.019	0.016	0.071	<1	10.83[**]
IT	−0.031	−0.001	0.00021	0.00001	0.108	0.201	0.109	0.203	<1	<1
MC	−0.060	−0.003	0.00039	0.000002	0.222	0.652	0.225	0.652	<1	<1
MCW	−0.030	0.001	−0.00091	−0.00003	0.060	0.056	0.074	0.084	2.96	6.00[*]
OFC	−0.035	−0.002	0.00035	0.00001	0.243	0.418	0.249	0.424	1.52	1.81
PHG	−0.016	−0.002	−0.00028	−0.00003	0.107	0.305	0.114	0.319	1.71	3.98[*]
SSC	−0.049	−0.003	0.00081	−0.00002	0.218	0.625	0.232	0.634	3.57	4.84[*]
SSCW	−0.022	0.001	0.00019	−0.00002	0.050	0.052	0.051	0.069	<1	3.70
VC	−0.013	−0.002	−0.00020	−0.00002	0.043	0.415	0.046	0.420	<1	1.69

Note. Man = manual, PVE = partial volume estimate. Linear slope indicates estimate of rate of change in volume (cm³/year); quadratic slope is estimate of acceleration in volume (cm³/year²).

[*] At (1, 196) df, F = 3.89 for p = 0.05;

[**] F = 6.76, p = 0.01. See Table 2 for ROI abbreviations.