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

Randy L. Buckner,[a,b,c,d,*] Denise Head,[a,b] Jamie Parker,[a,b] Anthony F. Fotenos,[e] Daniel Marcus,[a,b] John C. Morris,[f] and Abraham Z. Snyder[c,f]

[a] Department of Psychology, Washington University, USA
[b] Howard Hughes Medical Institute, USA
[c] Mallinckrodt Institute of Radiology, Washington University School of Medicine, USA
[d] Department of Anatomy and Neurobiology, Washington University School of Medicine, USA
[e] Division of Biology and Biomedical Sciences, Washington University School of Medicine, USA
[f] Department of Neurology, Washington University School of Medicine, USA

Atlas normalization, as commonly used by functional data analysis, provides an automated solution to the widely encountered problem of correcting for head size variation in regional and whole-brain morphometric analyses, so long as an age- and population-appropriate target atlas is used. In the present article, we develop and validate an atlas normalization procedure for head size correction using manual total intracranial volume (TIV) measurement as a reference. The target image used for atlas transformation consisted of a merged young and old-adult template specifically created for cross age-span normalization. Automated atlas transformation generated the Atlas Scaling Factor (ASF) defined as the volume-scaling factor required to match each individual to the atlas target. Because atlas normalization equates head size, the ASF should be proportional to TIV. A validation analysis was performed on 147 subjects to evaluate ASF as a proxy for manual TIV measurement. In addition, 19 subjects were imaged on multiple days to assess test–retest reliability. Results indicated that the ASF was (1) equivalent to manual TIV normalization ($r = 0.93$), (2) reliable across multiple imaging sessions ($r = 1.00$; mean absolute percentage of difference = 0.51%), (3) able to correct between-gender head size differences, and (4) minimally biased in demented older adults with marked atrophy. Hippocampal volume differences between nondemented ($n = 49$) and demented ($n = 50$) older adults (measured manually) were equivalent whether corrected using manual TIV or automated ASF (effect sizes of 1.29 and 1.46, respectively). To provide normative values, ASF was used to automatically derive estimated TIV (eTIV) in 335 subjects aged 15–96 including both clinically characterized nondemented ($n = 77$) and demented ($n = 90$) older adults. Differences in eTIV between nondemented and demented groups were negligible, thus failing to support the hypothesis that large premorbid brain size moderates Alzheimer's disease. Gender was the only robust factor that influenced eTIV. Men showed an approximately ~12% larger eTIV than women. These results demonstrate that atlas normalization using appropriate template images provides a robust, automated method for head size correction that is equivalent to manual TIV correction in studies of aging and dementia. Thus, atlas normalization provides a common framework for both morphometric and functional data analysis.
© 2004 Elsevier Inc. All rights reserved.

Keywords: Alzheimer's disease; Aging; Morphometry; Cognitive reserve; Brain volume; ICV; TIV; MCI; fMRI; PET; Hippocampus

Introduction

Morphometric analysis of brain structure confronts the challenge that head size markedly differs among individuals—a challenge also present in functional data analysis. Measurement of regional volumes, rates of atrophy, and estimated premorbid whole-brain volume require some form of procedure to measure and account for head size variation. Quantities that have been used for this purpose in morphometric analysis include height (Bartzokis et al., 2001; Raz et al., 1997), head circumference (Graves et al., 1996; Schofield et al., 1997), and MRI-based measurement of the total intracranial volume (TIV; also sometimes called TICV and ICV) (Blatter et al., 1995; Edland et al., 2002; Jenkins et al., 2000; Mathalon et al., 1993). The usual assumption, supported by postmortem findings (Davis and Wright, 1977; Epstein and Epstein, 1978), is that these measures serve as proxy variables for the premorbid (unatrophied) brain volume. Normalization by head size increases the robustness of expected mophometric effects, for example, gray-matter

* Corresponding author. Department of Psychology, HHMI at Washington University, Campus Box 1125, One Brookings Drive, St. Louis, MO 63130. Fax: +1-314-935-4711.
E-mail address: rbuckner@artsci.wustl.edu (R.L. Buckner).

volume differences in aging (Mathalon et al., 1993). Moreover, as clinical practice moves to routine use of morphometric variables, automated head size measurement and correction will be required to compare individual patients against normative values.

Here we present a fast, automated procedure for head size correction that is validated against manual TIV measurement. The method uses the volume-scaling factor derived by registration of each individual to an atlas template to automatically generate a good estimate of TIV. Atlas normalization, of this form, is also commonly used in functional data analysis. By using atlas-based head size normalization, the same validated procedures derive robust morphometric quantities for structural comparisons as well as normalized images for between-subject and between-group functional data analysis.

Methods

Overview

We evaluated automated atlas-based registration as a means of measuring and correcting for variation in head size. Each subject was registered to an atlas-representative template. The template consisted of a merged, averaged image representing the atlas of Talairach and Tournoux (1988). This template was generated from data acquired in 12 young and 12 healthy old adults. The Atlas Scaling Factor (ASF) was computed as the determinant of the affine transform connecting each individual to the atlas-representative template. The ASF represents the whole-brain volume expansion (or contraction) required to register each individual to the template. To determine ASF reliability, data from the same subjects imaged on two separate days were independently processed and compared. Two procedures were employed to determine validity. First, total intracranial volume in atlas space (TIV_{atl}) and its conversion to native space (TIV_{nat}) were manually measured using T1- and T2-weighted images, thereby providing a gold standard measure of head size; the relation between ASF and TIV_{nat} was explored. A concern regarding automated atlas-based registration in older adults is that atrophy in aging and dementia may bias ASF, in particular, by overexpanding brains with greater atrophy. The relation of ASF to atrophy was investigated by computing normalized whole-brain volume (nWBV) using a previously validated tissue-segmentation procedure (Zhang et al.,

2001). We also measured ASF longitudinally over 4 years in an individual with unusually accelerated atrophy. Second, to examine criterion validity, hippocampal volumes were manually measured and normalized using both automated ASF and manual TIV_{nat}. Volume differences between age-matched demented and nondemented groups were compared using both normalization procedures. Finally, after establishing the reliability and validity of atlas-based normalization, estimated total intracranial volume (eTIV) was automatically derived from ASF in an expanded group of 335 subjects aged 15–96 including both nondemented older adults and older adults with dementia of the Alzheimer type (DAT) (Table 1).

Subjects

Three hundred thirty-five subjects aged 15–96 (194 females, 141 males; mean age 53.2 years) were consented and participated in accordance with guidelines of the Washington University Human Studies Committee. Of these, 147 contributed to a core validation data set that included automated and manual measures (Table 2). The expanded group of 335 individuals was used to derive normative data (Table 3). One additional individual, followed longitudinally, was not included in other analyses but was included as a case study. Subjects were recruited from the general Washington University community and also from the longitudinal sample of the Washington University Alzheimer Disease Research Center (ADRC). All subjects were screened for neurological impairment (other than dementia in the demented sample), current depression, and psychoactive medications. Adults 60 and over were clinically screened for dementia and classified based on the Clinical Dementia Rating (CDR) (Morris, 1993). Exclusion criteria included dementia other than DAT (e.g., vascular dementia, primary progressive aphasia). Of the individuals classified with the CDR, 77 were CDR 0 (no dementia), 59 were CDR 0.5 (very mild dementia), 29 were CDR 1 (mild dementia), and 2 were CDR 2 (moderate dementia).

Image acquisition

All imaging was conducted at 1.5 T (Siemens Vision scanner, Erlangen Germany). Head movement was minimized by cushioning and a thermoplastic face mask. Headphones were provided for communication. A vitamin E capsule was placed over the left forehead. Positioning was low in the head coil (toward the feet)

Table 1
Measures, abbreviations, and their relations

TIV_{atl}	Total intracranial volume (cm^3) head size adjusted by use of atlas space. Manually measured estimate of the cranial cavity based on T1- and T2-weighted sagittal images.
TIV_{nat}	Total intracranial volume (cm^3) in native space. The above manual estimate scaled to reflect the native acquisition space.
ASF	Atlas Scaling Factor (unitless). Computed scaling factor that transforms native-space brain and skull to the atlas target (i.e., the determinant of the transform matrix).
eTIV	Estimated total intracranial volume (cm^3). Automated estimate of total intracranial volume in native space derived from the ASF.
HCV_{atl}	Hippocampal volume (mm^3) head size adjusted by use of atlas space. Manually measured estimate of hippocampal volume.
HCV_{nat}	Hippocampal volume (mm^3) in native space. The above manual estimate scaled to reflect the native acquisition space.
HCV_{adj}	Hippocampal volume (mm^3) normalized via covariance correction for head size.
nWBV	Normalized whole-brain volume (%) Automated tissue segmentation based estimate of brain volume (gray-plus white-matter). Normalized to percentage based on the atlas target mask.

The relations between the measures above can be summarized as follows. (1) TIV_{atl} divided by ASF yields TIV_{nat}. (2) Thus, TIV_{atl} and TIV_{nat} are manual measurements of total intracranial volume with and without transformation into atlas space, respectively. (3) eTIV is a fully automated estimate of TIV_{nat}. (4) HCV_{atl} divided by ASF yields HCV_{nat}.

Table 2
Clinical and demographic characteristics for the validation sample

	Young (n = 48)	Nondemented old (n = 49)	Demented old (DAT) (n = 50)
Gender (F/M)	23/25	36/13	31/19
Age (years)	22.5 (3.0)	76.8 (6.9)	77.3 (5.5)
CDR (0/0.5/1/2)	n/a	49/0/0/0	0/33/16/1
Education (years)	n/a	14.8 (2.9)	13.3 (2.7)
MMSE	n/a	28.9 (1.3)	24.0 (4.2)

Note. Sample includes 147 individuals. DAT = dementia of the Alzheimer type. F = female; M = male. Young adults are aged 18–30; old adults are 65–93. CDR = Clinical Dementia Rating (CDR scores of 0, 0.5, 1, and 2 indicate no, very mild, mild, and moderate dementia, respectively; Morris, 1993). MMSE = Mini-Mental State Examination (score range = 0–30, with 30 being high functioning; Folstein et al., 1975). Age, Education, and MMSE are presented as means with SD in parentheses. n/a = not available.

to optimize imaging of the cerebral cortex. Three or four T1-weighted MP-RAGE (Mugler and Brookeman, 1990) scans were acquired in each subject. MP-RAGE parameters were empirically optimized for gray–white contrast [TR = 9.7 ms, TE = 4 ms, flip angle = 10°, inversion time (TI) = 20 ms, delay time (TD) = 200 ms, 256 × 256 (1 × 1 mm) in-plane resolution, one hundred twenty-eight 1.25-mm slices without gaps]. The MP-RAGE data were averaged offline (with correction for head movement) to increase the contrast to noise ratio in all procedures involving manual tracing, segmentation, and measurement of nWBV. For the subset of subjects for whom total intracranial volume was manually measured, T2-weighted images were also acquired using a turbo spin-echo (TSE) sequence [TR = 6150 ms, TE = 15 ms, 256 × 256 (1 × 1 mm) in-plane resolution, sixty-four 2-mm slices with 2-mm gaps]. TSE images were collected twice shifting by 2 mm between acquisitions to provide contiguous coverage.

Construction of the merged young- and old-adult atlas target

The basis of atlas normalization is the target to which individual subjects are registered. The atlas-representative target used here was constructed from 24 adults (12 young, 12 nondemented old) and made to conform to the Talairach and Tournoux (1988) atlas using the method of Lancaster et al. (1995). Details concerning creation of the combined (young plus old) target are given in Appendix A. An important feature of this target is the inclusion of anatomic variation representative of the population under study. As can be seen in Fig. 1, the template exhibits blurring of the ventricular boundaries, as expected from the inclusion of young (unatrophied) adults as well as older adults with marked atrophy. The template is somewhat larger than the average brain, a property that reflects the Talairach and Tournoux (1988) atlas. An atlas mask was constructed by manually segmenting the brain out of the young-adult averaged template, blurring 5 mm full-width at half-maximum (FWHM) in each direction, and then thresholding at 25%. This mask, which roughly corresponds to the intracranial volume and measures 1755 cm^3, was used as the reference for atlas-normalized total intracranial volume [see Eq. (3)].

Registration of individuals to the atlas target and MP-RAGE averaging

Several registration steps were used to generate structural images resampled to 1 mm^3 voxels in the atlas space of Talairach and Tournoux (1988) (Head et al., 2004). The basic strategy involved transforming an individual's brain, through a single affine transformation, to match the atlas-representative target (Ojemann et al., 1997). For each individual, only the first acquired MP-RAGE was directly registered to the atlas-representative target using a 12-parameter affine transform. Registration was driven by the brain as a loose-fitting mask was applied to exclude skull and extracranial features. The remaining MP-RAGE images were registered to the first (rigid body plus cardinal axis stretch) and transformed to the atlas-representative target through a single, composite interpolation. The TSE images were similarly registered to the atlas-representative target through within-subject registration to the first MP-RAGE.

Table 3
Estimated total intracranial volume (eTIV) by gender and diagnostic category

Subjects	n	Age (SD)	eTIV (SD)	eTIV quantiles		
				25th	Median	75th
Young/middle						
Male	80	29.1 (11.9)	1592.4 (115.1)	1505.3	1590.0	1676.8
Female	88	29.1 (11.6)	1427.8 (133.6)	1337.0	1438.4	1524.3
Nondemented (CDR 0)						
Male	20	76.8 (8.4)	1560.7 (163.7)	1418.2	1556.2	1697.9
Female	57	76.7 (9.0)	1379.0 (117.6)	1312.4	1370.6	1474.4
Demented (DAT, CDR 0.5/1/2)						
Male	41	77.7 (6.8)	1563.3 (156.3)	1467.8	1556.4	1642.2
Female	49	78.6 (7.0)	1424.8 (135.9)	1335.2	1423.1	1476.7

Note. Sample includes 335 individuals. Young/Middle-aged adults are between 15 and 59. Older adults (both nondemented and demented) are 60 and over. DAT = dementia of the Alzheimer type. CDR = Clinical Dementia Rating (see text). Age is in years. eTIV is in cm^3.

MERGED
YOUNG   YOUNG/OLD   OLD
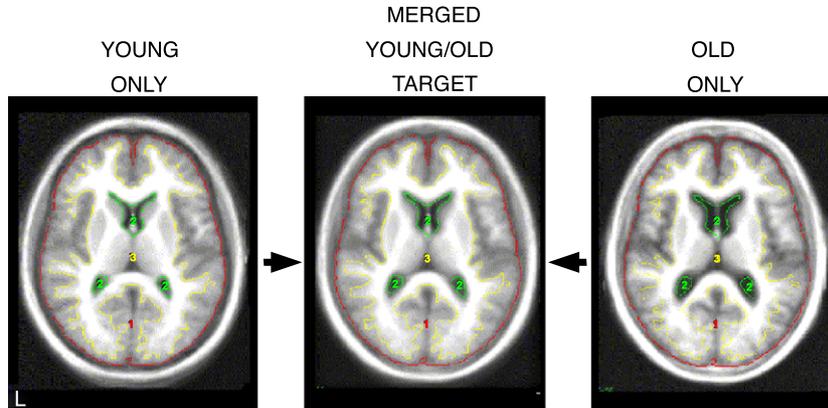ONLY   TARGET   ONLY

Fig. 1. Axial sections illustrate the merged young and old-adult atlas template. Averages of the atlas-forming 12 young adults (left), 12 old adults (right), and merged young and old adults (center) are displayed using sections 4 mm above the anterior-commissure posterior-commissure (AC–PC) plane (based on the atlas of Talairach and Tournoux, 1988). Colored lines appear on each image as reference to the young-adult contours. Atlas normalization in this paper uses the merged atlas for all individuals.

Registration of the first MP-RAGE image to the atlas-representative target generated the key metric of this report, the Atlas Scaling Factor (ASF), computed as the reciprocal of the atlas transform determinant. *The ASF represents the volume expansion (ASF > 1) or contraction (ASF < 1) required to register each individual to the atlas template.* This metric is similar to the volume-scaling factor described in Smith et al. (2000).

The final product of MP-RAGE preprocessing is illustrated in Fig. 2 for three representative individuals. High contrast-to-noise (attributable to the sequence parameters and averaging) is evident. The figure also demonstrates that local structural variation (including atrophy) remains after affine transformation.

Manual total intracranial volume measurement

Manual total intracranial volume measurement used a procedure modified from Eritaia et al. (2000). Coregistered T1- and T2-weighted images in atlas space (see Fig. 3) were simultaneously viewed using Analyze software (Analyze Version 4.0, Mayo Clinic, Rochester, MN). Two operators (JP and DH), both blind to subject's exact age, gender, and cognitive status, each manually outlined TIV on half of the sample. Inter-rater reliability was assessed on 10 randomly selected brains (mean age = 58.7 years; three young, three nondemented old, and four old DAT). The reliability coefficient, using intraclass correlations presuming ran-

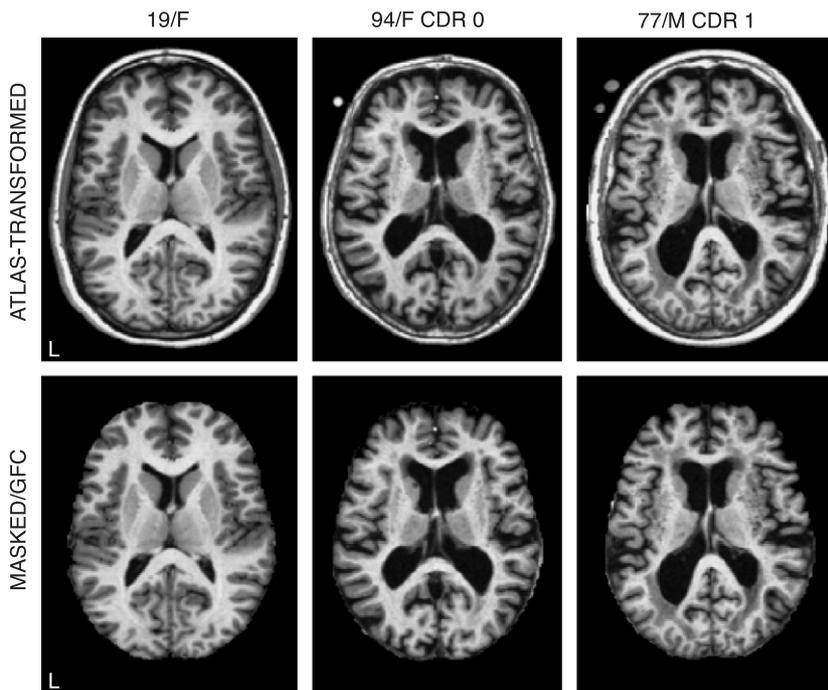19/F   94/F CDR 0   77/M CDR 1

Fig. 2. Representative atlas-normalized images are shown for three individuals including a 19-year-old female (left), a 94-year-old nondemented female (middle), and a 77-year-old male with dementia of the Alzheimer type (DAT) (right). The top row displays images following atlas normalization and within-subject averaging, but before any additional processing. The left forehead reference marker can be seen in the two older adults. The bottom row displays corresponding gain-field-corrected (GFC) images after the atlas-based mask has been applied.
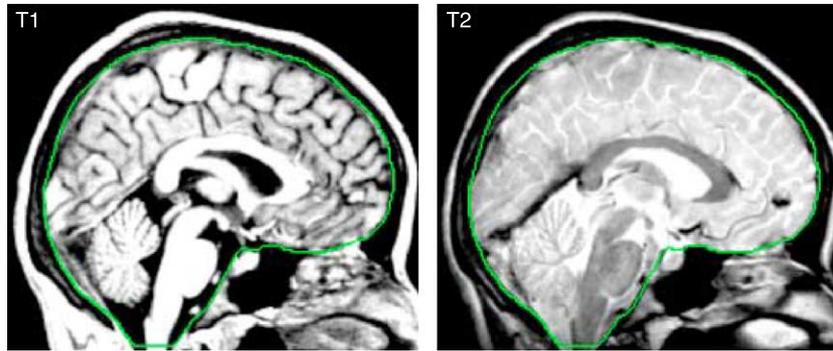
Fig. 3. Example total manual intracranial volume (TIV$_{atl}$) tracings in atlas space are shown for a 20-year-old female in sagittal section. The T1-weighted image is displayed to the left and the T2-weighted image to the right. Registered T1- and T2-weighted images were simultaneously viewed during tracing. The green contour is the actual tracing of TIV$_{atl}$ on this section. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

dom selection of raters (ICC[2]; Shrout and Fleiss, 1979), was 0.99. Images were visualized on an 18-in. interactive LCD display with an accompanying grip pen (Cintiq 15X LCD Tablet, Wacom, Vancouver, WA). The cranial cavity was manually outlined on sections beginning at the midsagittal plane and then on every 10th slice (1-cm intervals) outward in both directions. The midsagittal section was determined primarily by the clarity of the cerebral aqueduct, using the septum pellucidum as a second landmark when the cerebral aqueduct was less clearly visualized. Tracings were made on the supratentorial dural margins. As in Eritaia et al. (2000), when the dura was not visible the cerebral contour was outlined. Other landmarks were the undersurfaces of the frontal lobe, dorsum sellae, clivus, and at the craniovetebral junction, the attachment of the dura to the posterior foramen magnum cutting across to the anterior arch of C1. Volume calculation (here and for the hippocampus) used the Cavalieri procedure (Rosen and Harry, 1990). Atlas-normalized measures (TIV$_{atl}$) were converted to native-space measures (TIV$_{nat}$) by dividing by the ASF. In this manner, (1) native-space TIV$_{nat}$ and (2) atlas-normalized TIV$_{atl}$ were tabulated for each subject.

Normalized whole-brain volume measurement

Normalized whole-brain volume (nWBV) was computed by first intensity normalizing the MP-RAGE data based on Styner et al. (2000) and then using a validated segmentation tool to classify brain tissue as cerebral spinal fluid (CSF), gray, or white matter (Smith, 2002; Zhang et al., 2001). First, correction of intensity inhomogeneity was accomplished by an automated procedure to minimize intensity variation within contiguous regions. Contiguous region boundaries were determined (without brain masking) based on intensity limits and contour (intensity gradient) detection. The bias field was modeled as a general, second-order polynomial in three dimensions (10 free parameters) (Styner et al., 2000). Next, segmentation began with an initial estimation step to obtain and classify tissue parameters. A three-step, expectation-maximization algorithm then updated class labels and tissue parameters to iterate toward the maximum likelihood estimates of a hidden Markov random field model. The model used spatial proximity to constrain the probability with which voxels of a given intensity are assigned to each tissue class. Finally, brain volume was estimated as the sum of white and gray-matter voxels within the atlas-based brain mask and expressed as the percentage of the mask.

Manual hippocampal volume measurement

Hippocampal measurement was derived from Head et al. (submitted for publication) using procedures similar to Jack et al. (1992a,b, 1997) and Killiany et al. (1993). One operator (DH), blind to participant age, gender, and cognitive status, performed the measurements. Reliability was assessed on 10 randomly selected brains measured on two occasions separated by 2 weeks (mean age = 56.6 years; four young, four nondemented old, and two old DAT). The reliability coefficient, using ICC[2], was 0.98. Tracings were made on atlas-transformed MP-RAGE data recut to align the long axis of the hippocampus perpendicular to the coronal plane (31° tilted about the *X*-axis relative to standard AC-PC orientation). Alignment to the long axis of the hippocampus provides consistent visualization over subjects (Jack et al., 1995). The data were resampled to 0.5 mm³ voxels (by trilinear interpolation) on loading into Analyze to enable precise tracing. Measurements were made on 23–30 coronal images separated by 1.5 mm, that is, every third slice. The most anterior slice on which the hippocampus emerged inferior to the amygdala formed the rostral border. The caudal border was the slice on which the fornices rise after leaving the fimbria of the hippocampus. Measurement included the hippocampal formation, dentate gyrus, alveus, fimbria, and portions of the subiculum. Measured atlas space volumes (HCV$_{atl}$) were converted to native space (HCV$_{nat}$) by division by ASF. Both HCV$_{atl}$ and HCV$_{nat}$ include volumes summed over both hemispheres.

Manual head size correction of hippocampal volume

For comparative purposes, head size correction of hippocampal volumes was derived using a procedure common to the literature. Normalized hippocampal volumes were computed using a *covariance* approach (Free et al., 1995; Jack et al., 1989; see Mathalon et al., 1993, for discussion), which estimates the hippocampal volume by removing the influence of head size. These head size corrected hippocampal volumes were generated using only the manual measurement of TIV$_{nat}$. Specifically, head size adjusted hippocampal volume was estimated using the formula:

$$HCV_{adj} = HCV_{nat} - b(TIV_{nat} - \text{Mean } TIV_{nat}) \tag{1}$$

where HCV$_{adj}$ is the covariance-adjusted (corrected) volume, HCV$_{nat}$ is the hippocampal volume in native space, *b* is the slope

of the volume regression on TIV_{nat}. TIV_{nat} is as described previously and Mean TIV_{nat} is the sample mean of the TIV_{nat}. These hippocampal volumes, which were estimated based on manual measures typical of the literature (HCV_{nat}) and also on a covariance correction procedure (HCV_{adj}), were directly compared to the hippocampal volumes as measured in atlas space (HCV_{atl}).

Results

Atlas normalization reduces head size variance and equates head size across the age span

The first set of results concerns the effect of atlas normalization on head size as measured by manual total intracranial volume. Fig. 4 plots the relation between age and total intracranial volume measured in atlas space (TIV_{atl}) and also in native space (TIV_{nat}). The sample consists of the 147 individuals described in Table 2. Several features of atlas normalization are apparent.

First, TIV_{nat} shows a slight age-related volume decrease manifesting as a small, but significant, negative correlation [slope = -1.05 cm^3/year; $r = 0.19$; $F(1,145) = 5.20$, $P < 0.05$], perhaps reflecting an age-associated cohort effect. In contrast, TIV_{atl} is flat across the age span with no detectable correlation [-0.09 cm^3/year; $r = 0.04$; $F(1,145) < 1$]. Thus, atlas transformation eliminates systematic young-adult vs. old-adult head size differences.

Second, the sample variance for TIV_{atl} is less than that for TIV_{nat}. Quantitatively, the standard deviations are 67.8 and 147.5 cm^3, respectively. This effect is reflected in Fig. 4 as a reduced 95% confidence interval (CI) for TIV_{atl} relative to TIV_{nat} at all

ages. The corresponding coefficients of variation (standard deviation divided by the mean) are 3.9% vs. 10.2%.

Finally, TIV_{atl} is systematically larger than TIV_{nat} by about 15%. This difference reflects the 1988 atlas of Talairach and Tournoux and has no biological significance. However, a similar inflation must affect all measures performed in atlas space. In principle, it is possible to scale atlas-normalized measured volumes to eliminate the systematic inflation (using a constant scaling factor across all subjects and volumes). With the exception of the normative values presented in the Estimated total intracranial volume (eTIV) in 335 subjects section [see Eq. (3)], we do not perform such an adjustment so as to present the data as actually measured in Talairach and Tournoux (1988) atlas space.

Atlas normalization is highly reliable across independent imaging sessions

The second set of results quantifies the reliability of atlas normalization. To estimate ASF reliability, 19 subjects were imaged on two separate days (within a 60-day interval; mean interval = 22 days; mean age = 31.9, range = 18–87) and the image sets were processed independently. In this manner, the influences of head positioning, day-to-day scanner variability, and other sources of uncontrolled ASF variance were estimated. Across the 19 subjects, the mean ASF ranged from 1.00 to 1.44. The correlation between ASF estimates from the 2 days was nearly perfect ($r = 1.00$). Moreover, the mean absolute percentage difference (measured as the absolute difference between the two daily measures divided by their mean, expressed in percentage) was small (mean = 0.51%; range = 0.04–1.45% across the 19 individuals).
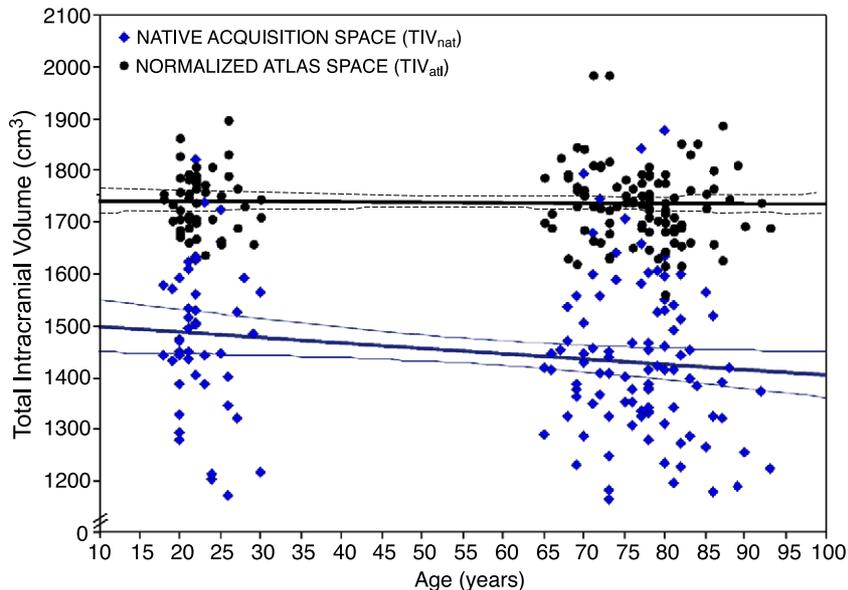

Fig. 4. Atlas normalization reduces variance and equates manual measures of total intracranial volume across the age span. Manually, measures are plotted as a function of age for the 147 subjects listed in Table 2. The blue diamonds and lines plot native acquisition space measures without any head size correction (denoted as TIV_{nat}). The darker line is the regression line and the lighter curved line represents the 95% confidence interval (CI). The black circles and lines plot measures corrected by atlas transformation (denoted as TIV_{atl}). Note the reduced variance about the TIV_{atl} and its flat slope. The increase in absolute volume for TIV_{atl} is due to the specific, arbitrary size of the atlas target and should not be interpreted (see text). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Atlas normalization corrects for gender-based head size differences

Head size differs between men and women. An effective head size normalization should correct this difference. Using the 147 individuals from Table 2, Fig. 5 shows that atlas normalization corrects for gender-based head size differences (TIV_{nat} and TIV_{atl} are the dependent variables). Men were 12.4% larger than women before correction (TIV_{nat}) and 1.3% smaller following correction (TIV_{atl}). It is unclear whether this residual error reflects variance or a small, systematic gender bias. The baseline difference of 12.4% between men and women is similar, but slightly smaller, than that recently measured in several other reports. Blatter et al. (1995) noted a 13.8–16.5% larger male total intracranial volume across five different samples that varied in age (mean = 15.0%). Jenkins et al. (2000) estimated 14.9% and 16.2% larger male total intracranial volume in their control and AD patients, respectively. Edland et al. (2002) recently reported similar estimates of 14.0% and 16.1%. Whitwell et al. (2001) reported a 12.4% difference nearly identical to that reported here. The issue of gender difference in baseline head size is revisited in the estimated total intracranial volume (eTIV) in 335 subjects section.

Atlas normalization based on ASF is proportional to manual measurement of TIV

The relation between atlas normalization using ASF and normalization using manual measurements of TIV is central to the potential use of ASF as a routine metric for head size correction, and also to the validation of atlas-based normalization as an appropriate step in between-group functional data analysis. Fig. 6 (top) plots ASF against TIV_{nat} for the 147 individuals from Table 3. The two measures are highly correlated ($r = 0.93$). Thus, the size of the head in native acquisition space is nearly fully reflected in the automated derivation of ASF. Moreover, across the three groups (young, nondemented old, and demented old), the relation is quite similar. These results suggest that it is feasible to use

Fig. 6. ASF is proportionate to TIV and is minimally influenced by whole-brain volume (atrophy). The top panel plots the automatically derived ASF vs. manually measured TIV_{nat} for each of the 147 subjects in Table 2 ($r = 0.93$). The points represent the young adults (black), the nondemented old adults (green), and demented old adults (red). The regression line for each group is plotted separately; all three are highly similar indicating ASF appropriately captures variance in TIV_{nat}. The bottom panel plots ASF vs. normalized whole-brain volume (nWBV) as measured by tissue segmentation (see text). Note that the nondemented old adults show markedly reduced brain volume compared to young adults, and that demented old adults are reduced further. Nonetheless, ASF is unaffected ($r = 0.04$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

automated measurement of ASF as a proxy for manual measurement of TIV.

Atlas normalization is minimally biased by atrophy

An intuitive concern regarding atlas-based automated head size correction is that it may be biased by atrophy. This might result if smaller, atrophied brains are overexpanded to match the atlas target. We observed this effect in the past when registering old-adult brains to a young-adult-only atlas template (Buckner RL, Snyder AZ, unpublished observations). The merged (young plus old) adult template was developed to solve this problem (see Appendix A). The similar relations between ASF and TIV_{nat} in both young- and old-adult groups (Fig. 6, top) indicate substantial bias is unlikely. Nonetheless, because this is a significant issue and
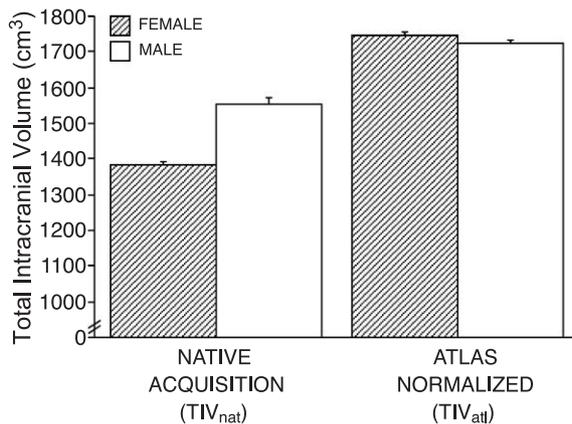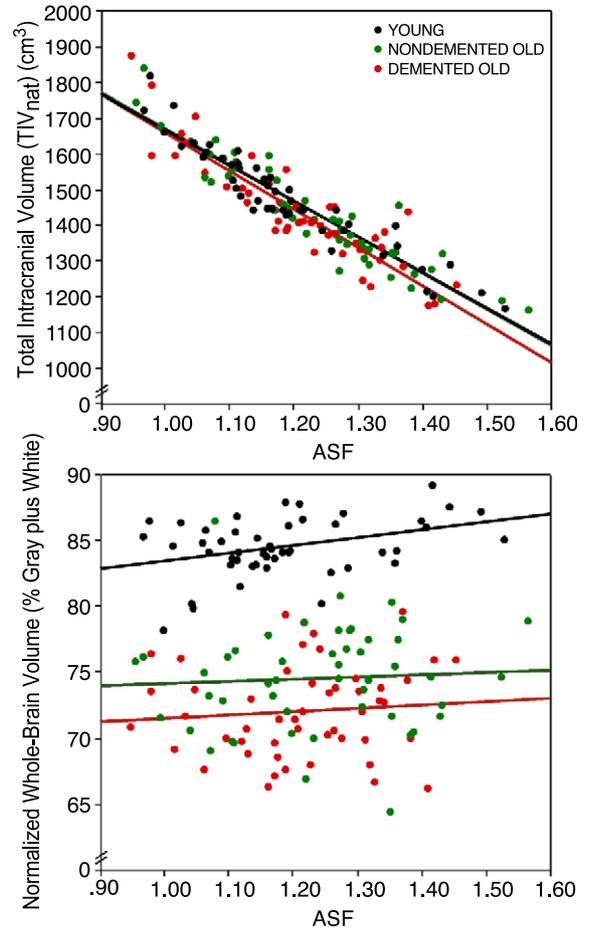

Fig. 5. ASF can be used to correct for gender-based head size differences. Mean manually measured total intracranial volumes are plotted for native acquisition (no head size correction; TIV_{nat}) and for atlas correction (TIV_{atl}). Error bars indicate standard error of the mean. Note the near equivalence of mean TIV_{atl} in males and females following atlas normalization.

atrophy often correlates with the groups being compared, we further explored the influence of atrophy using two separate procedures—one based on cross-sectional measurement of whole-brain volume differences and a second based on longitudinaly measured atrophy.

First, normalized whole-brain volume (nWBV) was measured as an estimate of atrophy and plotted against ASF for young, nondemented old, and demented old groups (Fig. 6, bottom). There was minimal, if any, relation between ASF and nWBV ($r = 0.04$). Atrophy bias would be expected to manifest as larger ASF for atrophied brains. Moreover, across the three groups, the lack of relation is similarly apparent despite a large baseline group effect on nWBV. Nondemented old adults showed significantly reduced nWBV as compared to young adults [unpaired $t(95) = 15.41$, $P < 0.001$], and demented old adults showed significant reduction as compared to nondemented old adults [unpaired $t(95) = 3.40$, $P < 0.001$]. Thus, automated measurement of ASF was not influenced by marked brain volume differences among separate age and dementia groups.

Second, the relation of ASF and atrophy was examined in a longitudinal case study of an individual with marked, accelerated atrophy (see Buckner et al., 2002). This male subject with atypical DAT (enrolled at age 66) was chosen to represent a worst-case scenario: his brain was asymmetric and he showed substantial whole-brain atrophy over a period of 1233 days. Fig. 7 plots the results. Clear atrophy is present while ASF remains stable. Quan-

titatively, the mean absolute percentage of difference between his first measurement and the four others was 0.84% placing the variance of his longitudinal ASF measures in a range similar to that noted for the test–retest reliability analysis. The difference between his two most discrepant ASF measures (between days 0 and 970) represents a 1.11% absolute percentage difference. These results suggest that, in addition to being minimally influenced by atrophy, atlas normalization shows longitudinal stability over multiple years (see similar analysis of manual TIV in Whitwell et al., 2001).

Atlas normalization can be used to detect dementia-associated hippocampal volume differences

As a test of criterion validity, atlas-based hippocampal volumes were explored in the context of detecting dementia effects. For this analysis, the separation of hippocampal volume between nondemented and DAT subjects was used as the gold standard (Convit et al., 1993; Csernansky et al., 2000; DeCarli et al., 1995; Head et al., submitted for publication; Jack et al., 1992a,b; Killiany et al., 1993; Laakso et al., 1995; Soininen et al., 1994). The question asked was how atlas-based volumes (fully automated head size correction) compared to normalized volumes based on the manual measures of TIV_{nat}. Fig. 8 shows the results. For this analysis, the hippocampal volumes were held constant; the independent variable was whether the volume was normalized using TIV_{nat}, as described
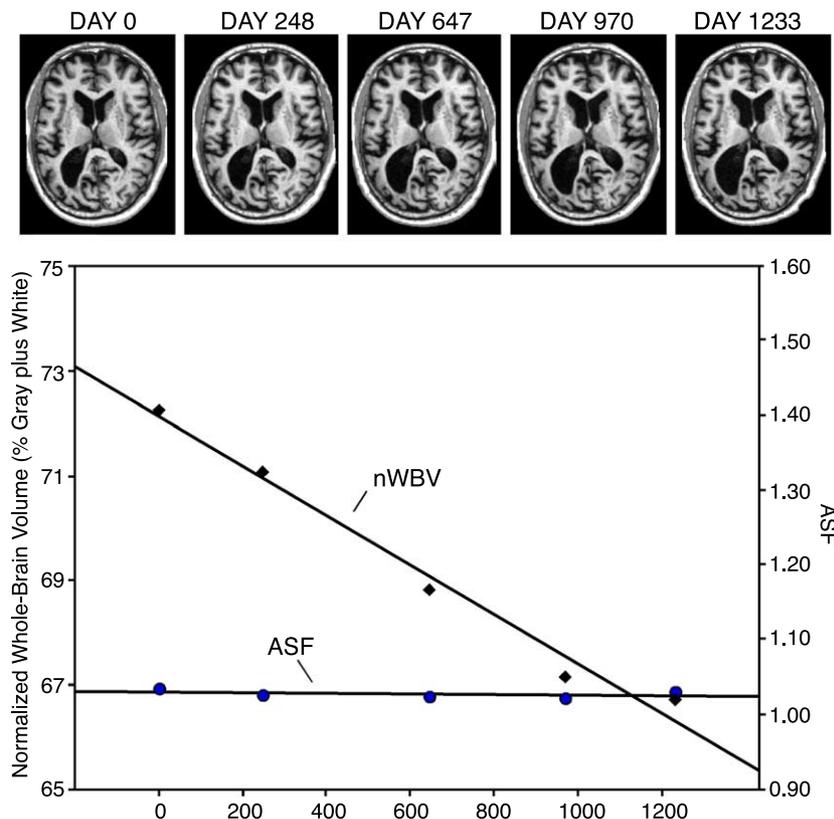

Fig. 7. ASF is longitudinally stable in the presence of marked progressive atrophy. The top axial images show sequential atlas-normalized images from five separate sessions with their relative acquisition time (referenced to the first session) noted in days. Atrophy is observed visually as the enlarging of the ventricles. The bottom panel simultaneously plots ASF (left axis) and normalized whole-brain volume (nWBV) (right axis). Note that ASF remains stable over time, while whole-brain volume markedly decreases consistent with the visually apparent atrophy.
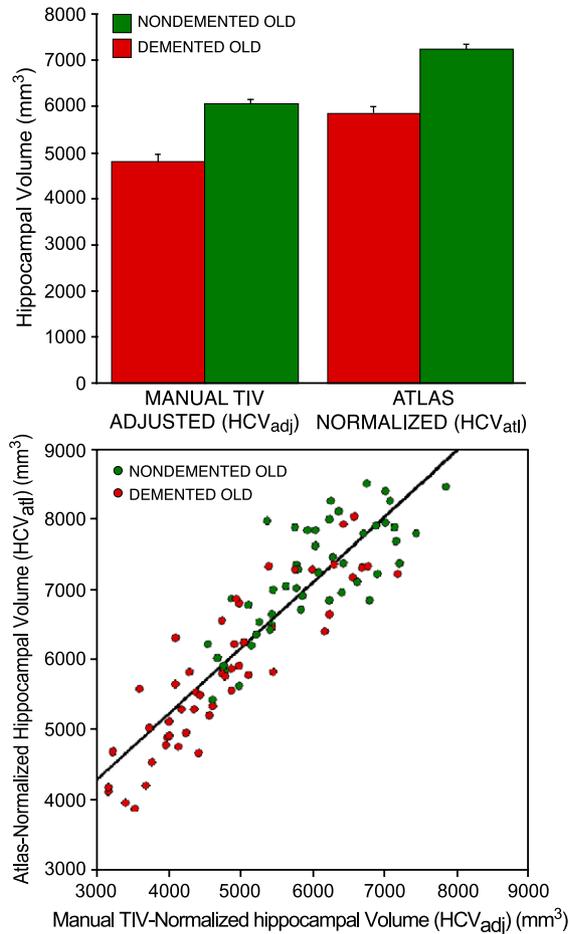
Fig. 8. Normalized hippocampal volumes using manually measured procedures typical of the literature (HCV$_{adj}$) and those exclusively using automated atlas transformation (HCV$_{atl}$) are directly compared. The top panel plots mean hippocampal volumes for HCV$_{adj}$ and HCV$_{atl}$. Error bars indicate standard error of the mean. Note that similar differences between nondemented (red bar) and demented (green bar) hippocampal volumes are noted for both forms of correction. The bottom panel plots individual values again color-coding them based on dementia status. Note the strong relation between the two methods of correction and the clear separation of the nondemented (red circles) and demented (green circles) hippocampal volumes. The increase in overall hippocampal volumes for HCV$_{atl}$ as compared to HCV$_{adj}$ is due to the specific, arbitrary size of the atlas target and should not be interpreted (see text). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in Eq. (1) (yielding HCV$_{adj}$), or whether the volume was simply that measured in atlas space (HCV$_{atl}$). Both HCV$_{adj}$ and HCV$_{atl}$ showed robust effects of dementia status, with significant volume reduction in demented individuals [unpaired $t(97) = 7.23$ and 6.40 for HCV$_{atl}$ and HCV$_{adj}$ volumes, respectively; both $P < 0.001$].

To compare their relative power, the effect sizes of the hippocampal volume measures were computed using Cohen's (1988) d:

$$d = (HCV_{ND} - HCV_D)/Sqrt((SD_{ND}^2 + SD_D^2)/2) \qquad (2)$$

where HCV$_{ND}$ and HCV$_D$ were the mean hippocampal volumes for the nondemented and demented older adults and SD$_{ND}$ and

SD$_D$ were the respective standard deviations. Using this metric, the effect size for the atlas-based hippocampal volumes (HCV$_{atl}$) was, if anything, greater than that for the manually normalized volumes (HCV$_{adj}$) (1.46 vs. 1.29). Both effect sizes were greater than those measured without any head size normalization (HCV$_{nat}$) (1.11). In practice, these values mean that atlas-based hippocampal volumes (HCV$_{atl}$) show 30% overlap between demented and nondemented individuals, compared to 40% without normalization (HCV$_{nat}$) (Cohen, 1988). The relation between atlas-based (HCV$_{atl}$) and manually normalized (HCV$_{adj}$) volumes is further plotted in Fig. 8 (bottom) for each individual. A strong correlation for hippocampal volume between atlas-based (HCV$_{atl}$) and manual normalization (HCV$_{adj}$) was observed ($r = 0.90$).

Estimated total intracranial volume (eTIV) in 335 subjects

Having established the reliability and validity of atlas normalization, we next provide normative values based on an expanded sample of 335 subjects. ASF can be converted to an estimate of TIV$_{nat}$ by scaling the ASF metric by the atlas mask volume. For our target atlas:

$$eTIV = 1755cm^3/ASF \qquad (3)$$

where eTIV is the estimated TIV$_{nat}$ and 1755 cm^3 is a constant representing the atlas mask volume. (If the constant is empirically derived to equate eTIV with manually measured TIV$_{nat}$, based on the 147 subjects in Table 2, it is similarly estimated to be 1738 mm^3. All eTIV values reported in this paper can thus be multiplied by 0.9903 if an empirically derived constant for eTIV estimation, based on manual TIV$_{nat}$, is desired.)

Fig. 9 and Tables 3 and 4 provide normative data for eTIV. eTIV showed strong gender effects in each of the young or middle, nondemented old, and demented old groups [$t(166) = 8.51$, $t(75) = 5.35$, $t(88) = 4.50$, respectively; all $P < 0.001$] and a trend for a very small age effect [$r = -0.13$; $F(1,192) = 3.17$, $P = 0.07$ for females and $r = -0.08$; $F(1,139) < 1$ for males]. Direct eTIV comparison of demented and nondemented groups failed to reach significance for either the males [unpaired $t(59)=0.06$, $P = 0.95$] or
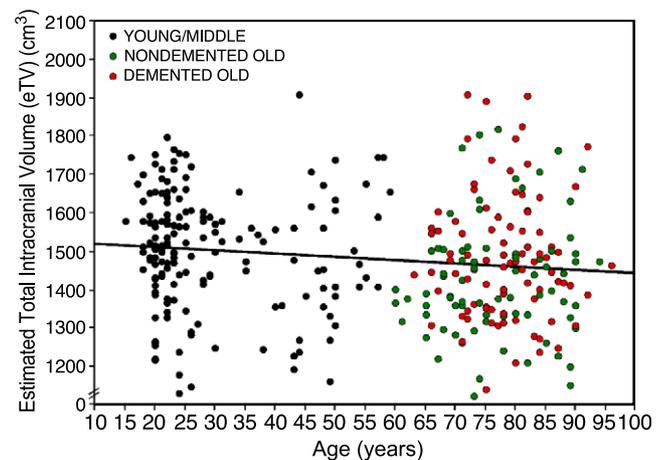

Fig. 9. eTIV, reflecting a fully automated estimate of total intracranial volume in native acquisition space, is plotted for 335 individuals. A regression line is plotted for the entire cohort (aged 15–96) (black line).

Table 4
Estimated total intracranial volume (eTIV) by Blatter et al. (1995) decade groupings

Age decade	n	Age (SD)	eTIV (SD)	Values taken from Blatter et al. (1995)		
				n	Age (SD)	TIV (SD)
Female						
16−25	53	21.3 (2.1)	1457.2 (136.5)	20	21.3 (2.5)	1400.3 (93.1)
26−35	10	28.4 (2.6)	1385.2 (133.9)	24	30.7 (3.1)	1329.7 (112.4)
36−45	12	40.7 (3.0)	1392.1 (138.0)	22	40.7 (3.0)	1351.6 (110.8)
46−55	12	50.8 (2.5)	1370.6 (92.5)	24	50.4 (2.7)	1343.2 (146.4)
56−65	8	62.0 (3.0)	1367.2 (53.8)	15	59.8 (2.2)	1335.5 (83.0)
66−75	38	71.1 (2.7)	1386.1 (119.7)	0	n/a	n/a
76−85	41	80.2 (2.5)	1409.1 (131.3)	0	n/a	n/a
86−95	19	89.2 (2.1)	1419.8 (157.1)	0	n/a	n/a
Male						
16−25	46	21.4 (2.1)	1598.4 (105.2)	24	23.2 (1.9)	1594.1 (101.4)
26−35	14	28.9 (2.2)	1556.5 (99.0)	19	30.9 (3.3)	1549.4 (81.9)
36−45	5	40.2 (4.8)	1526.6 (243.8)	16	41.1 (2.7)	1545.9 (103.9)
46−55	10	48.8 (2.7)	1612.4 (99.2)	15	51.1 (2.7)	1532.9 (89.3)
56−65	6	59.7 (3.2)	1578.7 (172.3)	15	60.6 (3.0)	1548.9 (106.1)
66−75	25	71.1 (3.1)	1573.2 (163.4)	0	n/a	n/a
76−85	24	80.8 (2.7)	1600.8 (136.9)	0	n/a	n/a
86−95	10	87.8 (1.8)	1481.9 (162.1)	0	n/a	n/a

Note. Sample includes 333 individuals (all individuals between 16 and 95). Age is in years. eTIV and TIV are in cm^3. Age bins are divided by decade to be similar to Blatter et al. (1995) and therefore are of unequal sample size. Estimates from samples with fewer than 10 subjects should be considered cautiously.

females $[t(104) = 1.86, P = 0.07]$. Table 4 plots eTIV in relation to age-span decades as previously reported for normative data by Blatter et al. (1995).

Discussion

Automated measures of head size using atlas-based transformation are proportional to manual total intracranial volume normalization ($r = 0.93$) and reliable from 1 day to the next ($r = 1.00$). Potential bias due to atrophy minimally influenced head size measurement indicating that atlas-based normalization is appropriate for comparisons between populations in studies that seek to quantify global and regional volume differences. The normative values generated by our atlas-based procedure are discussed first followed by a discussion of the relevance of the procedure to morphometric and functional imaging studies.

Estimated total intracranial volume (eTIV) and relevance to the cerebral reserve hypothesis

Normative data for eTIV are presented for 335 subjects in Tables 3 and 4, and Fig. 9. Summarizing the results, (1) gender effects were robust in all groups explored, (2) age-associated differences were minimal (although a slight trend cannot be ruled out), and (3) differences between demented and nondemented individuals were minimal.

Gender was the only robust influence on eTIV. Our data suggest a roughly 12% increase in male over female eTIV. As discussed earlier, others have generally estimated a 14–16% difference. While the origin of this slight difference is unclear, it should be noted that similar gender differences were also observed for our manual TIV$_{nat}$ measures. Thus, the difference is unlikely to arise from the atlas normalization procedure but rather to characteristics of the sample. Several examples of gender-difference values similar to those obtained here have been reported. Reiss et al.

(1996) noted an approximately 10% cerebral volume difference between young boys and girls (aged 5–17). Courchesne et al. (2000) report an 11.1% difference in a sample of individuals aged 2–80. Whitwell et al. (2001) noted a 12.4% male greater than female TIV in a sample similar to that presented here.

Normative eTIV values are relevant to the cerebral reserve hypothesis. Several reports suggest that large premorbid brain volume is protective against developing clinically symptomatic dementia (Graves et al., 1996; Katzman et al., 1988; Schofield et al., 1995, 1997; see also Cummings et al., 1998), while others fail to find support for a reserve hypothesis (Edland et al., 2002; Jenkins et al., 2000). Consistent with Jenkins et al. (2000) and Edland et al. (2002), our data provide minimal support for this hypothesis, at least insofar as TIV is the relevant measure of reserve. We do not interpret the nonsignificant trend in females (which goes in opposite direction to the reserve hypothesis) in light of the nearly identical eTIVs in males.

Relevance to morphometric analysis

Head size normalization, based on direct or indirect measurement of the cranial cavity, is commonly used in regional and whole-brain volume analysis. Postmortem studies suggest that the cranial cavity is a good proxy for premorbid brain volume. Using an elegant procedure, Davis and Wright (1977) measured 100 postmortem brain volumes and then compared the values to the cranial cavity volumes measured by inflating a water-filled balloon within the skull. Their results indicated that brain volume variance is nearly fully accounted for by a linear relation to cranial cavity volume in young and middle-aged adults. Atrophied brains deviate from this relation because the skull changes minimally with age (see Fig. 5 of Davis and Wright, 1977 for a particularly clear demonstration of this relation). Thus, the cranial cavity volume represents an estimate of premorbid brain size before global and local atrophy (see Whitwell et al., 2001, for further discussion). Because between-subject head size variance is so

great, baseline normalization, or inclusion of a head size measure as a covariate, is important to morphometric analysis. Studies of criterion validity note that gray-matter volume differences in aging are better detected after head size normalization (Mathalon et al., 1993), a finding replicated in our data for hippocampal volume difference in dementia. In practice, head size correction, most often using manually measured TIV, is commonly employed in research and clinical studies of aging and dementia (e.g., Jernigan et al., 1990, and Blatter et al., 1995; Hesselink, 1990; Jack et al., 1992a,b; Killiany et al., 1993; Raz et al., 1997; Resnick et al., 2000; Salat et al., 1999; Sullivan et al., 1995). What is novel in the present study is the demonstration that automated atlas-based measurement is as effective as manual measurement in head size normalization.

Atlas-based head size measurement and correction has two immediate uses in morphometric analysis. The first use is as an efficient, robust estimate of TIV. We refer to atlas-based estimated TIV as eTIV to differentiate its derivation from that of direct measurement (Table 1). Atlas-based eTIV is proportional to TIV and highly reliable. Thus, it can serve as a covariate or normalization factor in morphometric analysis of regional and whole-brain volume, much in the same manner of manually measured TIV (or alternative automated measures such as estimation of the skull boundary; e.g., see Smith et al., 2001, 2002).

Second, atlas-based normalization provides a direct means of head normalization for morphometric analysis: measurements can be made directly on the atlas-transformed images (see Goldszal et al., 1998, for an interesting, related approach; see applications of their procedure in Resnick et al., 2000, 2003). Head et al. (2004) have recently applied an atlas-based volume normalization to regional measures of diffusion tensor imaging (DTI) data. Atlas transformation, in one step, places individuals in a consistent orientation, adjusts for between-subject head size variation, and allows for between-modality measurements to be made on registered images (such as was done here for TIV$_{atl}$ using T1- and T2-weighted images; Fig. 3). Atlas transformation does require interpolation. Most in-use manual tracing procedures already induce a single interpolation to align axes and/or adjust for tilted head position. However, some methods optimize *acquisition* to non-orthogonal planes and thus avoid interpolation (e.g., Jack et al., 1995).

A further subtle feature of using atlas-transformed images for morphometric measures relates to how the head size correction is implemented. Regional volumes measured in transformed images are inherently *proportion* corrected, as contrasted to *head size-residual* scores or volumes corrected with a covariance approach {Eq. (1); Free et al., 1995; Jack et al., 1989; Mathalon et al., 1993}. Some correction procedures may be better than others, depending on the analysis goals[1]. Of importance, atlas-based normalization does not obligate any specific statistical correction procedure (see also Albrecht et al., 1993), but additional steps may be required to obtain the desired correction. For example, the covariance method can be applied to correct for head size following atlas-based

normalization using the ASF to derive the native space volume and then using eTIV as the covariate. Alternatively, the volumes in atlas space can be retained which is analogous to proportion correction.

Relevance to functional brain imaging

Group-based functional imaging studies often use, or begin with, affine transformation to an atlas target (e.g., Bäckman et al., 1999; Bookheimer et al., 2000; Cabeza et al., 1997; Grady et al., 1995; Logan et al., 2002; Lustig et al., 2003; Madden et al., 1999). *The present results suggest that if an appropriate atlas target is employed, such as the merged young and old-adult template in* Fig. 1, *head size normalization in functional imaging comparisons will be proportional to TIV.* Atrophy (whole-brain volume differences) as well as residual local differences will be retained. The magnitude of the residual whole-brain volume differences can be seen in Fig. 6 (for comparisons between young, nondemented, and demented groups). Local distortion is a separate issue and not formally addressed in the present manuscript. Nonlinear transformation of local distortion may be one solution to this problem, when anatomic registration is the goal. We are in the process of quantifying residual local anatomic distortion in aging and dementia following affine atlas-based transformation (O'Brien et al., 2003). Depending on the goals, preserving or removing local anatomic variance may be desirable. For example, in group-based functional data comparisons, removal of anatomic variance is likely optimal. In volume or shape-based morphometric analysis, the residual local variance is the measure of interest.

Atlas normalization enables a unified approach to morphometric and functional data analysis

For many applications, the present head size normalization procedure will be interchangeable with other procedures, such as that developed by Smith et al. (2002) for cross-sectional analysis, so long as appropriate target atlases are used that accommodate the sample. As discussed above, the present methods provide a robust within-subject head size reference (eTIV) that can be used as a covariate in morphometric analysis. However, for some purposes the present approach may have additional utility. In particular, one by-product of the present atlas-based approach is that a general framework emerges for conducting structural, functional, and molecular data analysis in the same, common data space. The present method and validation suggest that this common data space, in terms of its global scaling properties, is proportionate to head size normalization as used in traditional morphometric investigations, even when individuals with atrophy are included.

Atlas-based approaches have been used in functional data analysis for nearly two decades (e.g., Fox et al., 1985). While there are limits of atlas-based normalization because of residual interindividual variability, the general approach has proven useful for contrasting and pooling functional data across subjects, and also for constructing probabilistic atlases (Mazziotta et al., 1995). By employing an atlas-based approach to head size normalization in morphometric analysis, between-subject and between-modality comparability is gained. Variance between subjects as well as commonalities can be explored. For example, within our data environment, we commonly compare structural, diffusion tensor,

---

[1] Atchley et al. (1976) provide empirical and theoretical analysis of the hazards of using ratio variables as scaling factors, particularly in the context of falsely inducing (or increasing) correlations among measures being scaled. Mathalon et al. (1993) provide a thoughtful discussion on varied correction procedures in regional measures.

and functional images within the same reference frame within and between subjects. Molecular images, such as those that target amyloid (Klunk et al., 2004), can also be registered. Alignment to a common atlas facilitates such comparisons. Measures made in one modality in one subject can transfer or be compared to other modalities and to other subjects, within limits of between-modality and between-subject registration algorithms. Voxel-based morphometry (VBM; Ashburner and Friston, 2000), the "regional analysis of volumes examined in normalized space" (RAVENS) approach (Goldszal et al., 1998), and our own analysis of DTI data (Head et al., 2004) are some examples of structural/morphometric methods that take advantage of atlas-based normalization.

Another way of conceptualizing this additional utility is in its flexibility. Manual and even automated measures of region volumes and total intracranial volume can be made in atlas-normalized images much in the same manner as they are made in native (or more commonly, rotated) image space. Thus, atlas-based transformations are largely compatible with existing approaches and analyses. Atlas-normalized images can, for example, be input into existing automated tools (e.g., Smith et al., 2001, 2002) with the penalty of an additional interpolation. In addition, measured volumes in atlas space can be directly compared between subjects. At its most basic level, this uniformity facilitates visualization. For example, in normalized space, manually traced hippocampal volumes from many subjects can be called up simultaneously and easily referenced to one another. In its most elaborate form, common registration of morphometric and other data allows probabilistic and modal properties of different data types to be compared. For example, it will be interesting to observe how regional amyloid patterns relate to structural atrophy patterns and function as measured by FDG-PET and MRI. Atlas-based normalization allows spatial patterns to be visualized and quantitatively compared across data modalities in an efficient manner.

Limitations of an atlas-based approach

Automated atlas-based head size normalization has limitations. One limitation is the limits of its accuracy. The present study yielded highly reliable results with a mean absolute percentage difference (MAPD) of about one half of a percent. While this error term is small, some applications benefit from even more precise estimates, such as longitudinal estimates of brain atrophy (Fox and Freeborough, 1997; Smith et al., 2001, 2002). In this regard, the present methods, which are based on independent (cross-sectional) registration of each image to a common target, could likely be improved for longitudinal data analysis by first computing the serial registration of the multiple within-subject images.

A further limitation derives from the construction and use of the target atlas. In our initial attempts to derive an efficient and robust target atlas, a young-adult template was used that failed for many older adults. When the template was constructed to include the anatomic features and image contrast properties common to older adults (Fig. 1, middle panel), atlas-registration almost always succeeded yielding the results presented in this paper. However, failure may still occur when anatomic variation or contrast properties outside the range of the target atlas exists. In the present data set, this kind of failure likely minimally influenced estimates as patients with marked insult-induced brain changes were not included and correlation with a gold-standard measure (manual TIV_{nat}) was strong.

Only one instance of clear bias due to atypical anatomy was identified. It is possible that more subtle cases of bias exist and contribute to the error terms and presently reported variance estimates. In this case, there was an 11.7% difference between manual TIV_{nat} and eTIV. The patient was an 80-year-old woman with DAT. Review of the imaging data showed ventricular enlargement typical of that seen in normal-pressure hydrocephalus. The diagnosis was not made because of the absence of neurologic symptoms (apart from dementia). On one hand, the infrequency of failure and rough correspondence between TIV_{nat} and eTIV in this individual are evidence of the robustness of the method. On the other hand, this case is a reminder that automated procedures need not, and should not, be unsupervised. For example, the atlas-normalized images with superimposed target atlas reference marks can be included in reports or simply visualized as a quality control check. Furthermore, future refinements of atlas-based head size corrections should explore targets containing a wider range of variation, particularly for population-based and clinical studies that recruit a wider range of patients.

An interesting, if speculative, possibility is that atlas-based transformation, and its goodness of fit to multiple target atlases will, in itself, be an informative measure. For example, in the case discussed above, the individual might better register to a hydrocephalus atlas providing a computer-aided diagnostic that the candidate is at risk for the disease.

Appendix A

A.1. Preparation of the combined (young + old) atlas-representative target image

Older individuals systematically differ from young adults, particularly in the degree of cerebral atrophy. Thus, atlas transformation of older adults may be affected by bias if the target image represents exclusively young-adult anatomy. Our initial use of an atlas-representative target image designed for fMRI of young adults led to systematic overexpansion of older brains (Buckner RL, Snyder AZ, unpublished observations). The combined target was created to address this problem. (Atlas transformation using separate targets for each age group was avoided because observed group differences could then be attributed to systematic differences in the respective targets. Moreover, complexities arise in the case of intermediate aged individuals.)

To derive the combined target, atlas-representative standard images were separately prepared for young ($n = 12$; $F = 6$; mean age = 23.4) and old adults ($n = 12$; $F = 9$; mean age = 75.2; all CDR = 0) using the algorithm described below. Each subject used as an atlas prior contributed four MP-RAGE scans (Mugler and Brookeman,

1991; see Buckner et al., 2000) acquired identically to the present experimental data; these images were corrected for interscan head movement, averaged and bias field corrected (see Methods) before submission to the new standard-generating algorithm. As a final step, the newly generated young and old standards were combined by simple voxel-wise averaging to generate the merged target (see Fig. 1).

The strategy for atlas generation involves an iterative scheme that, for each age group, yields an average image (*new standard*) matched to a preexisting atlas-representative target (*original standard*). The approach has been used previously to create an atlas-representative target image for nonhuman primates (Black et al., 2001) but is here described in greater detail. All transforms are 12-parameter affines (4 × 4 matrices). The original standard atlas-representative target image that began the process represented the Talairach and Tournoux (1988) atlas according to the formulation of Lancaster et al. (1995). This original standard image has been used within the laboratory for many studies (e.g., Wheeler et al., 2000).

Before iteration, an approximate new standard was generated by averaging and atlas transformation computed based on the original standard:

Step 1. Compute individual MP-RAGE → new standard transforms.

Step 2. Generate a new average by application of the most recently computed transforms to the individual images.

Step 3. Compute the generated average → preexisting standard transform.

Step 4. Perform transform composition (right matrix multiply results of step 1 by the result of step 4) to obtain updated individual MP-RAGE → new standard transforms.

Step 5. Regenerate the average as in step 2.

Iteration of steps 1–5 generates a new standard average matched to the preexisting standard while simultaneously minimizing mismatch between individuals. Five iterations were sufficient to achieve convergence of all transforms (up to four decimal places).

Additional algorithmic details include the following: In listitem1step 1, the individual MP-RAGE images were preblurred (6.9 mm FWHM 3D Gaussian profile). The objective function was simple intensity correlation, or equivalently, minimization of difference image variance (Snyder, 1995). listitem3Step 3 included no preblur and used a cross-modal registration method based on alignment of intensity gradients related to the method of Andersson et al. (1995). This cross-modal registration algorithm takes into account intensity gradient relative orientation by inclusion in the objective function ($\langle|\nabla I_1||\nabla I_2|\cos^2\theta\rangle$) of an angular factor (Pluim et al., 2000) that is maximal when gradients are either aligned or anti-aligned. It is reasonable to assume that listitem3step 3 was driven, in large part, by alignment of gradients at the gray–white boundary, a prominent feature of the present MP-RAGE data. Formulation of the objective function in terms of intensity gradients also renders listitem3step 3 relatively insensitive to variations in MRI contrast, for example, due to variation in sequence parameters such as flip angle and TD. This is an important feature of the present strategy as the MP-RAGE sequence used to make the new standard atlas-representative image (and to acquire the present experimental data) has considerably better gray–white contrast than that used to create the original standard.

References

Albrecht, G.H., Gelvin, B.R., Hartman, S.E., 1993. Ratios as a size adjustment in morphometrics. Am. J. Phys. Anthropol. 91, 441–468.

Andersson, J.L., Sundin, A., Valind, S., 1995. A method for coregistration of PET and MR brain images. J. Nucl. Med. 36, 1307–1315.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry—The methods. NeuroImage 11, 805–821.

Atchley, W.R., Gaskins, C.T., Anderson, D., 1976. Statistical properties of ratios. I. Empirical results. Syst. Zool. 25, 137–148.

Bäckman, L., Andersson, J.L., Nyberg, L., Winblad, B., Nordberg, A., Almkvist, O., 1999. Brain regions associated with episodic retrieval in normal aging and Alzheimer's disease. Neurology 52, 1861–1870.

Bartzokis, G., Beckson, M., Lu, P.H., nuechterlein, K.H., Edwards, N., Mintz, J., 2001. Age-related changes in frontal and temporal lobe volumes in men. Arch. Gen. Psychiatry 58, 461–465.

Black, K.J., Snyder, A.Z., Koller, J.M., Gado, M.H., Perlmutter, J.S., 2001. Template images for nonhuman primate neuroimaging: 1. Baboon. NeuroImage 14, 736–743.

Blatter, D.D., Bigler, E.D., Gale, S.D., Johnson, S.C., Anderson, C.V., Burnett, B.M., Parker, N., Kurth, S., Horn, S.D., 1995. Quantitative volumetric analysis of brain MR: normative database spanning 5 decades of life. Am. J. Neuroradiol. 16, 241–251.

Bookheimer, S.Y., Strojwas, M.H., Cohen, M.S., Saunders, A.M., Pericak-Vance, M.A., Mazziotta, J.C., Small, G.W., 2000. Patterns of brain activation in people at risk for Alzheimer's disease. N. Engl. J. Med. 343, 450–456.

Buckner, R.L., Snyder, A.Z., Sanders, A.L., Raichle, M.E., Morris, J.C., 2000. Functional brain imaging of young, nondemented, and demented older adults. J. Cogn. Neurosci. 12 (Suppl. 2), 24–34.

Buckner, R.L., Snyder, A.Z., Williams, L., Gold, B.T., Sergent-Marshall, S., Balota, D.A., Morris, J.C., 2002. Automated assessment of atrophy progression in a case of progressive aphasia. 8th International Conference on Alzheimer's Disease and Related Disorders.

Cabeza, R., Grady, C.L., Nyberg, L., McIntosh, A.R., Tulving, E., Kapur, S., Jennings, J.M., Houle, S., Craik, F.I., 1997. Age-related differences in neural activity during memory encoding and retrieval: a positron emission tomography study. J. Neurosci. 17, 391–400.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Erlbaum, Hillsdale, NJ.

Convit, A., de Leon, M.J., Golomb, J., George, A.E., Tarshish, C.Y., Bobinski, M., Tsui, W., De Santi, S., Wegiel, J., Wisniewski, H., 1993. Hippocampal atrophy in early Alzheimer's disease: anatomic specificity and validation. Psychiatr. Q. 64, 371–387.

Courchesne, E., Chisum, H.J., Townsend, J., Cowles, A., Covington, J., Egaas, B., Harwood, M., Hinds, S., Press, G.A., 2000. Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. Radiology 216, 672–682.

Csernansky, J.G., Wang, L., Joshi, S., Miller, J.P., Grado, M., Kido, D., McKeel, D., Morris, J.C., Miller, M.I., 2000. Early DAT is distinguished from aging by high-dimensional mapping of the hippocampus. Dementia of the Alzheimer type. Neurology 55, 1636–1643.

Cummings, J.L., Vinters, H.V., Cole, G.M., Khachaturian, Z.S., 1998. Alzheimer's disease: etiologies, pathophysiology, cognitive reserve, and treatment opportunities. Neurology 51, S2–S17.

Davis, P.J.M., Wright, E.A., 1977. A new method for measuring cranial cavity volume and its application to the assessment of cerebral atrophy at autopsy. Neuropathology and Applied Neurobiology 3, 341–358.

DeCarli, C., Murphy, D.G., McIntosh, A.R., Teichberg, D., Schapiro, M.B., Horwitz, B., 1995. Discriminant analysis of MRI measures as a method to determine the presence of dementia of the Alzheimer type. Psychiatry Res. 57, 119–130.

Edland, S.D., Xu, Y., Plevak, M., O'Brien, P., Tangalos, E.G., Petersen, R.C., Jack, C.R.J., 2002. Total intracranial volume: normative values and lack of association with Alzheimer's disease. Neurology 59, 272–274.

Epstein, H.T., Epstein, E.B., 1978. The relationship between brain weight

and head circumference from birth to age 18 years. Am. J. Phys. Anthropol. 48, 471–473.

Eritaia, J., Wood, S.J., Stuart, G.W., Bridle, N., Dudgeon, P., Maruff, P., Velakoulis, D., Pantelis, C., 2000. An optimized method for estimating intracranial volume from magnetic resonance images. Magn. Reson. Med. 44, 973–977.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198.

Fox, N.C., Freeborough, P.A., 1997. Brain atrophy progression measured from registered serial MRI: validation and application to Alzheimer's disease. J. Magn. Reson. Imaging 7, 1069–1075.

Free, S.L., Bergin, P.S., Fish, D.R., Cook, M.J., Shorvon, S.D., Stevens, J.M., 1995. Methods for normalization of hippocampal volumes measured with MR. Am. J. Neuroradiol. 16, 637–643.

Goldszal, A.F., Davatzikos, C., Pham, D.L., Yan, M.X., Bryan, R.N., Resnick, S.M., 1998. An image-processing system for qualitative and quantitative volumetric analysis of brain images. J. Comput. Assist. Tomogr. 22, 827–837.

Grady, C.L., McIntosh, A.R., Horwitz, B., Maisog, J.M., Ungerleider, L.G., Mentis, M.J., Pietrini, P., Schapiro, M.B., Haxby, J.V., 1995. Age-related reductions in human recognition memory due to impaired encoding. Science 269, 218–221.

Graves, A.B., Mortimer, J.A., Larson, E.B., Wenzlow, A., Bowen, J.D., McCormick, W.C., 1996. Head circumference as a measure of cognitive reserve. Association with severity of impairment in Alzheimer's disease. Br. J. Psychiatry 169, 86–92.

Head, D., Buckner, R.L., Shimony, J.S., Williams, L.E., Akbudak, E., Conturo, T.E., McAvoy, M., Morris, J.C., Snyder, A.Z., 2004. Differential vulnerability of anterior white matter in nondemented aging with minimal acceleration in dementia of the Alzheimer type: evidence from diffusion tensor imaging. Cereb. Cortex 14, 410–423.

Head, D., Snyder, A.Z., Girton, L.E., Morris, J.C., Buckner, R.L., submitted for publication. Neuroanatomical double dissociation between normal aging and Alzheimer's disease. Cereb. Cortex.

Jack, C.R.J., Twomey, C.K., Zinsmeister, A.R., Sharbrough, F.W., Petersen, R.C., Cascino, G.D., 1989. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. Radiology 172, 549–554.

Jack, C.R.J., Petersen, R.C., O'Brien, P.C., Tangalos, E.G., 1992a. MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. Neurology 42, 183–188.

Jack, C.R.J., Sharbrough, F.W., Cascino, G.D., Hirschorn, K.A., O'Brien, P.C., Marsh, W.R., 1992b. Magnetic resonance image-based hippocampal volumetry: correlation with outcome after temporal lobectomy. Ann. Neurol. 31, 138–146.

Jack, C.R.J., Theodore, W.H., Cook, M., McCarthy, G., 1995. MRI-based hippocampal volumetrics: data acquisition, normal ranges, and optimal protocol. Magn. Reson. Imaging 13, 1057–1064.

Jack, C.R.J., Petersen, R.C., Xu, Y.C., Waring, S.C., O'Brien, P.C., Tangalos, E.G., Smith, G.E., Ivnik, R.J., Kokmen, E., 1997. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology 49, 786–794.

Jenkins, R., Fox, N.C., Rossor, A.M., Harvey, R.J., Rossor, M.N., 2000. Intracranial volume and Alzheimer disease: evidence against the cerebral reserve hypothesis. Arch. Neurol. 57, 220–224.

Jernigan, T.L., Press, G.A., Hesselink, J.R., 1990. Methods for measuring morphologic features on magnetic resonance images. Validation and normal aging. Arch. Neurol. 47, 27–32.

Katzman, R., Terry, R., DeTeresa, R., Brown, T., Davies, P., Fuld, P., Renbing, X., Peck, A., 1988. Clinical, pathological, and neurochemical changes in dementia: a subgroup with preserved mental status and numerous neocortical plaques. Ann. Neurol. 23, 138–144.

Killiany, R.J., Moss, M.B., Albert, M.S., Sandor, T., Tieman, J., Jolesz, F., 1993. Temporal lobe regions on magnetic resonance imaging identify patients with early Alzheimer's disease. Arch. Neurol. 50, 949–954.

Klunk, W.E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., et al.,

2004. Imaging brain amyloid in Alzhemier's disease with Pittsburgh Compound-B. Ann. Neurol. 55, 306–319.

Laakso, M.P., Soininen, H., Partanen, K., Helkala, E.L., Hartikainen, P., Vainio, P., Hallikainen, M., Hanninen, T. , Riekkinen Sr., P.J. 1995. Volumes of hippocampus, amygdala and frontal lobes in the MRI-based diagnosis of early Alzheimer's disease: correlation with memory functions. J. Neural Trasm.: Parkinson's Dis. Dementia Sect. 9, 73–86.

Lancaster, J.L., Glass, T.G., Lankipalli, B.R., Downs, H., Mayberg, H., Fox, P.T., 1995. A modality-independent approach to spatial normalization of tomographic images of the human brain. Hum. Brain Mapp. 3, 209–223.

Logan, J.M., Sanders, A.L., Snyder, A.Z., Morris, J.C., Buckner, R.L., 2002. Under-recruitment and nonselective recruitment: dissociable neural mechanisms associated with aging. Neuron 33, 827–840.

Lustig, C., Snyder, A.Z., Bhakta, M., O'Brien, K.C., McAvoy, M., Raichle, M.E., Morris, J.C., Buckner, R.L., 2003. Functional deactivations: change with age and dementia of the Alzheimer type. Proc. Natl. Acad. Sci. U. S. A. 100, 14504–14509.

Madden, D.J., Turkington, T.G., Provenzale, J.M., Denny, L.L., Hawk, T.C., Gottlob, L.R., Coleman, R.E., 1999. Adult age differences in the functional neuroanatomy of verbal recognition memory. Hum. Brain Mapp. 7, 115–135.

Mathalon, D.H., Sullivan, E.V., Rawles, J.M., Pfefferbaum, A., 1993. Correction for head size in brain-imaging measurements. Psychiatry Res. 50, 121–139.

Mazziotta, J.C., Toga, A.W., Evans, A., Fox, P., Lancaster, J., 1995. A probabilistic atlas of the human brain: theory and rationale for its development. NeuroImage 2, 89–101.

Morris, J.C., 1993. The clinical dementia rating (CDR): Current version and scoring rules. Neurology 43, 2412–2414.

Mugler III, J.P., Brookeman, J.R., 1990. Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP-RAGE). Magn. Reson. Med. 15, 152–157.

Mugler III, J.P., Brookeman, J.R., 1991. Rapid three-dimensional T1-weighted MR imaging with the MP-RAGE sequence. J. Magn. Reson. Imaging 1, 561–567.

O'Brien, K.C., Head, D., Williams, L.E., Marcus, D., Snyder, A.Z., Morris, J.C., Buckner, R.L., 2003. Evaluation of spatial normalization in non-demented and demented older adult populations. Soc. Neurosci. Abstr.

Ojemann, J.G., Akbudak, E., Snyder, A.Z., McKinstry, R.C., Raichle, M.E., Conturo, T.E., 1997. Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. NeuroImage 6, 156–167.

Pluim, J.P., Maintz, J.B., Viergever, M.A., 2000. Image registration by maximization of combined mutual information and gradient information. IEEE Trans. Med. Imag. 19, 809–814.

Raz, N., Gunning, F.M., Head, D., Dupuis, J.H., McQuain, J., Briggs, S.D., Loken, W.J., Thornton, A.E., Acker, J.D., 1997. Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. Cereb. Cortex 7, 268–282.

Reiss, A.L., Abrams, M.T., Singer, H.S., Ross, J.L., Denckla, M.B., 1996. Brain development, gender, and IQ in children. A volumetric imaging study. Brain 119, 1763–1774.

Resnick, S.M., Goldszal, A.F., Davatzikos, C., Golski, S., Kraut, M.A., Metter, E.J., Bryan, R.N., Zonderman, A.B., 2000. One-year age changes in MRI brain volumes in older adults. Cereb. Cortex 10, 464–472.

Resnick, S.M., Pham, D.L., Kraut, M.A., Zonderman, A.B., Davatzikos, C., 2003. Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. J. Neurosci. 23, 3295–3301.

Rosen, G.D., Harry, J.D., 1990. Brain volume estimation from serial section measurements: a comparison of methodologies. J. Neurosci. Methods 35, 115–124.

Salat, D.H., Kaye, J.A., Janowsky, J.S., 1999. Prefrontal gray and white matter volumes in healthy aging and Alzheimer disease. Arch. Neurol. 56, 338–344.

Schofield, P.W., Mosesson, R.E., Stern, Y., Mayeux, R., 1995. The age at

onset of Alzheimer's disease and an intracranial area measurement. A relationship. Arch. Neurol. 52, 95–98.

Schofield, P.W., Logroscino, G., Andrews, H.F., Albert, S., Stern, Y., 1997. An association between head circumference and Alzheimer's disease in a population-based study of aging and dementia. Neurology 49, 30–37.

Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 86, 420–428.

Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155.

Smith, S.M., De Stefano, N., Jenkinson, M., Matthews, P.M., 2001. Normalized accurate measurement of longitudinal brain change. J. Comput. Assist. Tomogr. 25, 466–475.

Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P.M., Federico, A., De Stefano, N., 2002. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. NeuroImage 17, 479–489.

Snyder, A.Z., 1995. Difference image vs. ratio image error function forms in PET–PET realignment. In: Myer, R., Cunningham, V.J., Bailey, D.L., Jones, T. (Eds.), Quantification of Brain Function Using PET. Academic Press, San Diego, pp. 131–137.

Soininen, H.S., Partanen, K., Pitkanen, A., Vainio, P., Hanninen, T., Halli-kainen, M., Koivisto, K., Riekkinen Sr., P.J., 1994. Volumetric MRI analysis of the amygdala and the hippocampus in subjects with age-associated memory impairment: correlation to visual and verbal memory. Neurology 44, 1660–1668.

Styner, M., Brechbuhler, C., Szekely, G., Gerig, G., 2000. Parametric estimate of intensity inhomogeneities applied to MRI. IEEE Trans. Med. Imag. 19, 153–165.

Sullivan, E.V., Marsh, L., Mathalon, D.H., Lim, K.O., Pfefferbaum, A., 1995. Age-related decline in MRI volumes of temporal-lobe gray-matter but not hippocampus. Neurobiol. Aging 16, 591–606.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme, New York.

Wheeler, M.E., Petersen, S.E., Buckner, R.L., 2000. Memory's echo: vivid remembering reactivates sensory-specific cortex. Proc. Natl. Acad. Sci. U. S. A. 97, 11125–11129.

Whitwell, J.L., Crum, W.R., Watt, H.C., Fox, N.C., 2001. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. Am. J. Neuroradiol. 22, 1483–1489.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imag. 20, 45–57.