



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Psychiatry Research: Neuroimaging 131 (2004) 169–176

PSYCHIATRY
RESEARCH
NEUROIMAGING

www.elsevier.com/locate/psychresns

Controlling for premorbid brain size in imaging studies: T1-derived cranium scaling factor vs. T2-derived intracranial vault volume

George Fein^{a,*}, Victoria Di Sclafani^a, Colin Taylor^a, Kirk Moon^b, Jerome Barakos^b,
Hoang Tran^a, Bennett Landman^a, Robert Shumway^c

^aNeurobehavioral Research, Inc., 201 Tamal Vista Boulevard, Corte Madera, CA 94925, USA

^bDepartment of Radiology, California Pacific Medical Center, San Francisco, CA, USA

^cDepartment of Statistics, University of California, Davis, CA, USA

Received 29 May 2003; received in revised form 13 October 2003; accepted 27 October 2003

Abstract

Intracranial vault (ICV) volume, obtained from T2-weighted magnetic resonance imaging (MRI), is generally used to estimate premorbid brain size in imaging studies. T1-weighted sequences lack the signal characteristics for ICV measurements [they have poor contrast at the outer boundary of sulcal cranium scaling factor (CSF)] but are valuable in imaging studies due to their excellent gray vs. white matter contrast. Smith et al. [NeuroImage 17 (2002) 479] suggested a T1-derived cranium scaling factor as an alternative control variable for premorbid brain size in cross-sectional studies. This index, which is computed using the SIENAX software, is a scaling factor comparing an individual's skull to a template skull derived from the Montreal Neurological Institute (MNI) average of 152 T1 studies (the MNI152). SIENAX computes coarsely defined estimates for the individual and MNI skulls rather than well-defined volumes. To test how well this approach would work as a control variable for premorbid brain size in cross-sectional studies, we compared the T1-derived cranium scaling factor to T2-derived ICV measurements in a sample of 92 individuals: 39 white males, 22 white females, and 31 African-American males, with an age range of 26–78 years. The correlation between T1- and T2-derived variables was 0.94 and did not differ across subject groups. The T1-derived cranium scaling factor accounted for a statistically significant portion (87%) of the variance of the T2-derived ICV measure and thus is a good surrogate for ICV measurement of premorbid brain size as a reference measure in MRI atrophy studies. Furthermore, neither race, sex, nor age accounted for any additional variance in ICV, indicating that neither race-, gender-, nor age-associated cranial bone thickness effects were present in this data set.

© 2004 Elsevier Ireland Ltd. All rights reserved.

Keywords: Brain atrophy; Premorbid brain size; Intracranial vault; SIENAX

1. Introduction

The in vivo quantification of the volume of brain structures is benefiting from both advances in magnetic

resonance imaging (MRI) sequences and advances in image analysis. An index of the size of the cranial vault must be used to control for normal variation in premorbid brain size in imaging studies comparing regional or structural brain volumes between clinical samples. Brain atrophy can only be assessed in a cross-sectional study if current brain volume is examined relative to premorbid brain size. Finally, premor-

* Corresponding author. Tel.: +1-415-927-7676; fax: +1-415-924-2903.

E-mail address: george@nbresearch.com (G. Fein).

bid brain volume is important in its own right, as it is associated with the brain's ability to compensate for changes associated with aging and disease (often called 'compensatory' or 'functional' reserve) (Katzman et al., 1988; Schofield et al., 1995; Graves et al., 1996; Mori et al., 1997; Di Sclafani et al., 1998).

Since the growth of the cranium (those bones that house the brain, but not the rest of the skull) is driven by brain growth, the intracranial vault (ICV) volume is the generally accepted measure of premorbid brain size. The ICV can be measured most easily on T2-weighted MRIs, where the boundary of the ICV is defined by the outer boundary of the sulcal cranium scaling factor (CSF). Newer T1-weighted volumetric sequences are now the norm for tissue atrophy studies because they have higher resolution in a given imaging time and much better tissue perspicuity than T2 sequences (especially in gray matter vs. white matter contrast). Unfortunately, the ICV cannot be volumed on T1 images using existing software because of a lack of contrast at the outer boundary of sulcal CSF.

In addition to the measurement of sulcal CSF, T2-weighted sequences had been used for identification of white matter signal hyperintensities (WMSH, often a marker of white matter morbidity). Fluid-attenuated inversion recovery (FLAIR) sequences have become the protocol of choice for the assessment of WMSH, leaving the delineation of sulcal CSF (and its attendant ICV boundary) as the primary reason to retain T2-weighted sequences in structural imaging protocols.

Recently, Smith et al. (2002) suggested a T1-derived cranium scaling factor as an alternative to the T2-derived ICV volume as an index of premorbid brain size for imaging studies. We assessed the validity of this T1-derived cranium scaling factor as a premorbid brain size index by comparing it to T2-derived ICV volumes in a large sample. Our sample was constructed to assess biases associated with age, gender, or racial differences in cranial bone thickness.

2. Methods

2.1. Subjects

The MRI data were taken from subjects and controls in a study of cocaine abuse and a study of the interaction of alcohol dependence and aging.

This sample of 92 individuals included 39 white males, 22 white females, and 31 African-American males (we had difficulty recruiting African-American females). Subjects ranged in age from 26 to 78 years (Table 1).

2.2. MRI acquisition

All studies were performed on a 1.5-T MR system (Siemens Vision, Erlangen, Germany) using a head coil with quadrature detection. The imaging protocol included proton density and T2-weighted spin-echo axial images (2500/20, 80 ms/1 NEX, 3-mm contiguous slices, and in-plane resolution of $0.94 \times 0.94 \text{ mm}^2$) and T1-weighted 3D MP-RAGE coronal images (10/4/1 NEX, $1 \times 1 \text{ mm}^2$ in-plane resolution, and 1.4-mm-continuous slices).

2.2.1. Image analysis

The outer boundary of the sulcal CSF was delineated on the T2 images using an algorithm developed by Itti et al. (1997). This algorithm relies on the identification of the outer boundary of the signal of sulcal CSF (bright in T2-weighting). This method often requires editing by a human operator (especially in the area of the connection of the eyes to the brain by the optic nerves).

The T1 images were processed using the FSL software package (SIENAX, <http://www.fmrib.ox.ac.uk/analysis/research/siena>) from the University of Oxford (Smith et al., 2002). SIENAX utilizes the fact that the scalp has a bright signal to delineate the outer

Table 1
Subject demographics

Group	Number	Age (years)	T2-derived ICV (cc^3)	T1-derived cranium scaling factor
African-American males	31	40.4 (6.7 S.D.) (31–53)	1446 (113 S.D.)	0.76 (0.058 S.D.)
White males	39	54.7 (14.1 S.D.) (26–78)	1531 (138 S.D.)	0.82 (0.078 S.D.)
White females	22	56.6 (14.3 S.D.) (29–76)	1367 (88 S.D.)	0.73 (0.045 S.D.)
All subjects	92	50.3 (14.0 S.D.) (26–78)	1463 (135 S.D.)	0.78 (0.074 S.D.)

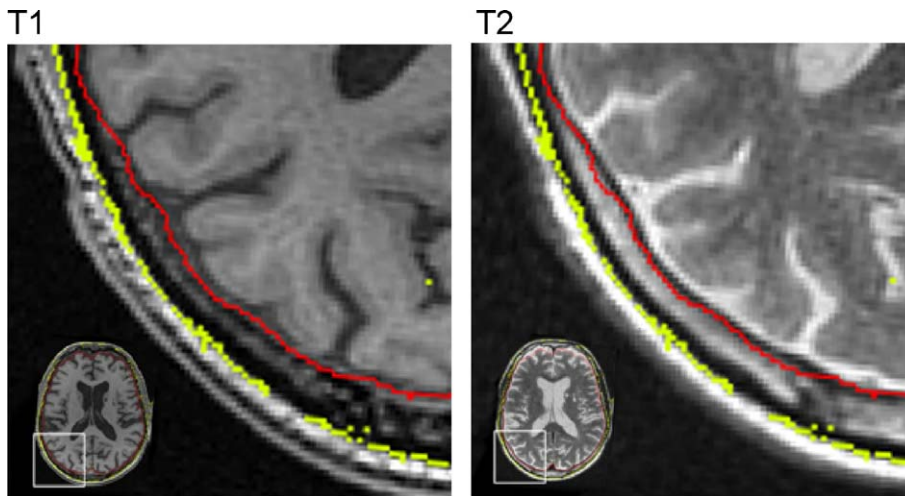


Fig. 1. T1- and T2-weighted sections from a subject with atrophy. The inlaid transverse sections highlight the enlarged subsection. The BET contours (solid red lines) are the outer boundary of the brain that was extracted from the T1-weighted image. The SIENAX outer-skull estimates (discontinuous yellow dots) show the coarsely defined outer boundary of the cranium from the T1-weighted image. These points form a discontinuous boundary that approximates the cranium. SIENAX utilizes both the brain and cranium estimates in a joint registration method to align the cranium boundary (yellow dots) to the MNI152 template. The reciprocal of the determinant of the transformation matrix is the T1-derived cranium scaling factor.

boundary of the cranium. First, SIENAX uses the brain extraction tool (BET, <http://www.fmrib.ox.ac.uk/analysis/research/bet>) to produce a model of the brain surface, including the cerebellum and the upper part of the brain stem (Fig. 1). Next, each point on the surface of the brain is taken as the start of a search outwards for the exterior boundary of the skull. Once the most distant point of low intensity (before the bright scalp) is found, the first peak in the gradient beyond this point is defined as the location on the skull's exterior. Using T1-weighted images to find the outer boundary of the skull appears to work quite well on visual inspection, even in regions of overlying muscle (dark signal) or in areas of visible bone marrow (bright signal). The outer boundary of the skull created by BET is not a well-defined surface, but rather a collection of discontinuous points. While not a complete surface, the points are reliable enough for SIENAX to compute a scaling factor between the subject skull and a template skull. We note that the template is also sparsely defined along the lower boundary of the brain, so the cranium location principally determines subject alignment since the algorithms can only compare skull estimates.

In SIENAX, the extracted brain is first co-registered to the MNI152 brain image using FSL's cor-

egistration tool (FLIRT, <http://www.fmrib.ox.ac.uk/analysis/research/flirt>) [MNI152 is an average T1 brain image constructed from 152 normal subjects at Montreal Neurological Institute (MNI)]. The rotation and translation parameters from this coregistration are then used in a constrained affine registration between the individual skull and the MNI152 skull template (generated from the MNI152 T1 image in the fashion described above); thus, only scaling and skew are allowed to vary. Finally, SIENAX computes the determinant of this transformation matrix and denotes it as the 'v-scaling' parameter. This parameter represents the relative scaling of the individual skull to the MNI152 skull. The reciprocal of v-scaling is what we refer to as the T1-derived cranium scaling factor.

We note that FSL never actually computes the MNI152 cranium volume or the individual subject cranium volume. In T1-weighted images, the cranium volumes are ill-defined due to the low contrast at the skull-scalp boundary and the resulting coarseness of that boundary. As described above, SIENAX uses a coarse skull outline and the brain outline for a joint registration process that avoids defining the cranium as a closed surface. It is especially important to reiterate

that v-scaling is a derived parameter from this joint registration rather than a directly observed quantity. Furthermore, while multiplication by a constant could give v-scaling physical units (e.g., cc^3), we believe that would be misleading and that v-scaling should be used as a covariate (rather than a predictor). Again, note that FSL uses v-scaling as a scaling variable to control for intersubject differences in premorbid brain volume. In a similar vein, we used the reciprocal of v-scaling in predicting premorbid differences in brain volume as measured by ICV. SIENAX, BET, and FLIRT are described in detail on FSL's Web site: <http://www.fmrib.ox.ac.uk/fsl/>.

2.3. Statistical analysis

The data were analyzed using the GLM procedure of the Statistical Analysis System (SAS) (SAS Institute, 1990) to perform analysis of covariance (ANCOVA). The standard T2-derived ICV measure was the dependent variable, and the T1-derived cranium scaling factor was the covariate. We examined the

following variables: (1) the strength of association between the dependent variable and the covariate (i.e., how good a predictor of ICV is the T1-based measure?); (2) whether the association between the dependent variable and the covariate differed across subject groups (i.e., African American males, white males, and white females); (3) whether any group differences in the dependent variable remained after controlling for the covariate (i.e., must we account for any cranial bone thickness differences between groups when using the cranium-based estimate of ICV?); (4) whether the association between the dependent variable and the covariate changed as a function of age; and (5) whether any age effects on the dependent variable remained after controlling for the covariate (Fig. 2).

3. Results

The correlation between the T1-derived cranium scaling factor and ICV across all 92 subjects was

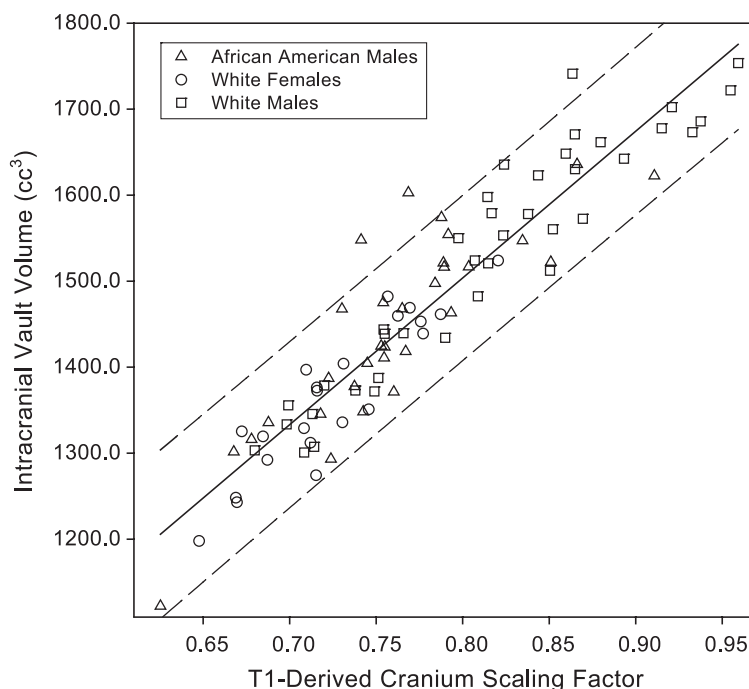


Fig. 2. Scatter plot showing the strong association between the ICV and the SIENAX T1-derived cranium scaling factor. Subjects from the three groups are indicated separately. The regression line and 95% prediction intervals are presented—only 3 of 92 subjects fall outside the 95% prediction interval.

0.939. The T1-derived cranium scaling factor accounted for 87% of the variance in the T2-derived ICV across all subjects ($P < 0.001$). The three group (white males, white females, and African-American males) by volume index interaction was nonsignificant ($P > 0.90$), indicating that the association between the T1-derived cranium scaling factor and ICV was comparable across groups. (The correlation between the T1-derived cranium scaling factor and ICV was 0.83 in African-American males, 0.89 in white females, and 0.95 in white males.) The residual effect of group on ICV (after removing variance associated with the T1-derived cranium scaling factor) was also nonsignificant ($P > 0.40$). This implies that there were no mean differences in the T1-derived cranium volume index after accounting for the variance associated with the covariate. Age did not modulate the association between the T1-derived cranium scaling factor and ICV ($P > 0.18$), and the residual effect of age on ICV (after removal of variance associated with the cranium scaling factor) was also nonsignificant ($P > 0.24$). Thus, the relationship between ICV and the T1-derived cranium scaling factor was not affected by the subject's age.

4. Discussion

An index of premorbid brain size is essential in imaging studies of brain atrophy since atrophy can only be estimated by examining how current brain volumes differ from premorbid volumes. It is also of value in controlling for normal variation in brain size in studies comparing regional or structure-specific brain volumes between clinical samples (e.g., studies comparing the volume of the dorsolateral prefrontal cortex or the thalamus between schizophrenic and normal individuals). Finally, it is of value as an index of brain 'compensatory reserve' (Katzman et al., 1988; Schofield et al., 1995; Graves et al., 1996; Mori et al., 1997; Di Sclafani et al., 1998). The current study establishes the validity of the T1-derived cranium scaling factor obtained from T1-weighted sequences by the SIENAX software (Smith et al., 2002) as an index of premorbid brain size. Validity was established using the generally accepted standard of ICV volume derived from T2-weighted sequences.

The T1-derived cranium scaling factor ranged from about 0.65 to 0.95 for the 92 subjects in our study. We found this surprising having expected the T1-derived cranium scaling factor to range from below 1.0 to above 1.0. In response to questions about this (both from us and other users of SIENAX), Smith (personal communication) noted that the MNI152 template is approximately 10% larger than the true population mean. In FSL's database of over 443 subjects, the T1-derived cranium scaling factor was below 1.0 for all subjects. This is an interesting anomaly, but it neither affects the use of SIENAX nor the use of the MNI152 template, which is primarily used as a spatial rather than as a size template. We further note that we have a subject (not included in this study) with a cranium scaling factor above 1.0 at this time.

Neither group (race and sex) nor age modulated the correlation between the T1-derived cranium scaling factor and ICV. We note that the cranium scaling factor is based on the outer boundary of the cranium while the ICV falls strictly within the inner cranial boundary. Thus, population variation in skull thickness could have a systematic effect on the T1-derived cranium scaling factor. Recent work by Lynnerup (2001) found no significant correlations between cranial thickness and sex, age, height, weight, or history of chronic drug and alcohol abuse. Previous work in adults (Adeloye and Kattan, 1975) has shown only localized variations in cranium thickness associated with race (African Americans and Caucasians) and sex. Since the reported cranium thickness variations are relatively small (0–3 mm) and spatially isolated, our finding that the T1-derived cranium scaling factor is not confounded with group membership is not surprising.

T1-weighted images are the standard in volumetric brain studies because of their excellent gray/white/CSF contrast. Their limitation for such studies is that their poor contrast at the sulcal CSF/cranium boundary makes accurate estimation of sulcal CSF and intracranial vault volumes impossible. Several recent studies have addressed both differentiating brain from nonbrain components and tissue compartment segmentation of T1 images (Smith et al., 2002; Kitagaki et al., 1998; Hata, 2000). While these methods have been successful at extracting brain structures, they have not demonstrated an ability to

unambiguously define the boundaries of the cranial vault volume because of the lack of contrast at the outer boundary of the sulcal CSF. The current finding overcomes the ICV estimation limitation. Although accurate sulcal CSF volumes still cannot be estimated from T1-weighted sequences, it is unclear whether such measures have additional value in morphometric studies where brain tissue volumes can be accurately measured.

Acknowledgements

This work was supported by NIAAA Grants AA11311 and AA11569, and NIDA Grant DA09543 (Principal Investigator for all grants, George Fein, PhD).

Appendix A

SIENAX uses several FSL software tools including BET for skull removal, FLIRT for linear image registration, and FAST (<http://www.fmrib.ox.ac.uk/analysis/research/fast>) for tissue segmentation. SIENAX accepts a command line argument denoted by '-f', which sets the fractional intensity threshold for BET processing. The allowable value ranges from 0

to 1 with a default value of 0.5. Small values correspond to less skull removal (e.g., a larger brain outline). To assess the effects of this threshold on the inverse v-scaling factor, we selected three subjects that spanned the range of observed inverse v-scaling values. Fig. A1 presents the result of varying BET thresholds from 0.05 to 0.95 on inverse v-scaling for subjects LOW ($1/v\text{-scaling}$)=0.62, AVG ($1/v\text{-scaling}$)=0.75, and HIGH ($1/v\text{-scaling}$)=1.03).

The data show that the inverse v-scaling factor remains relatively unchanged over BET thresholds ranging from 0.1 to 0.6. The reason inverse v-scaling values remain relatively invariant in this range of thresholds becomes apparent after examining the BET results. The first row of Fig. A2 shows slice 62 from subject AVG with the BET threshold at 0.1 (A) and 0.6 (B). There are obvious differences in the location of the brain–skull boundary, but for the most part, the brains are very close in size.

Outside this linear inverse v-scaling range, two scenarios exist. The second row of Fig. A2 shows the effects of BET when the threshold is too high; (C) shows slice 62 from AVG with the BET threshold at 0.7 and (D) shows it at 0.8. It is visually apparent that using a threshold of 0.8 results in a much smaller brain volume than the one using 0.7. The difference in the inverse v-scaling factor when the BET threshold goes from 0.7 to 0.8 is about 60%. Because so much

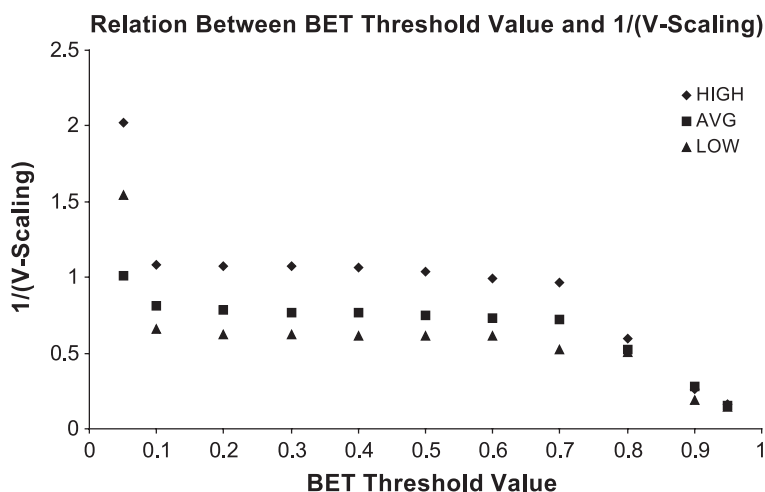


Fig. A1. Inverse v-scaling vs. BET threshold values for subjects LOW ($1/v\text{-scaling}$)=0.62, AVG ($1/v\text{-scaling}$)=0.75, and HIGH ($1/v\text{-scaling}$)=1.03).

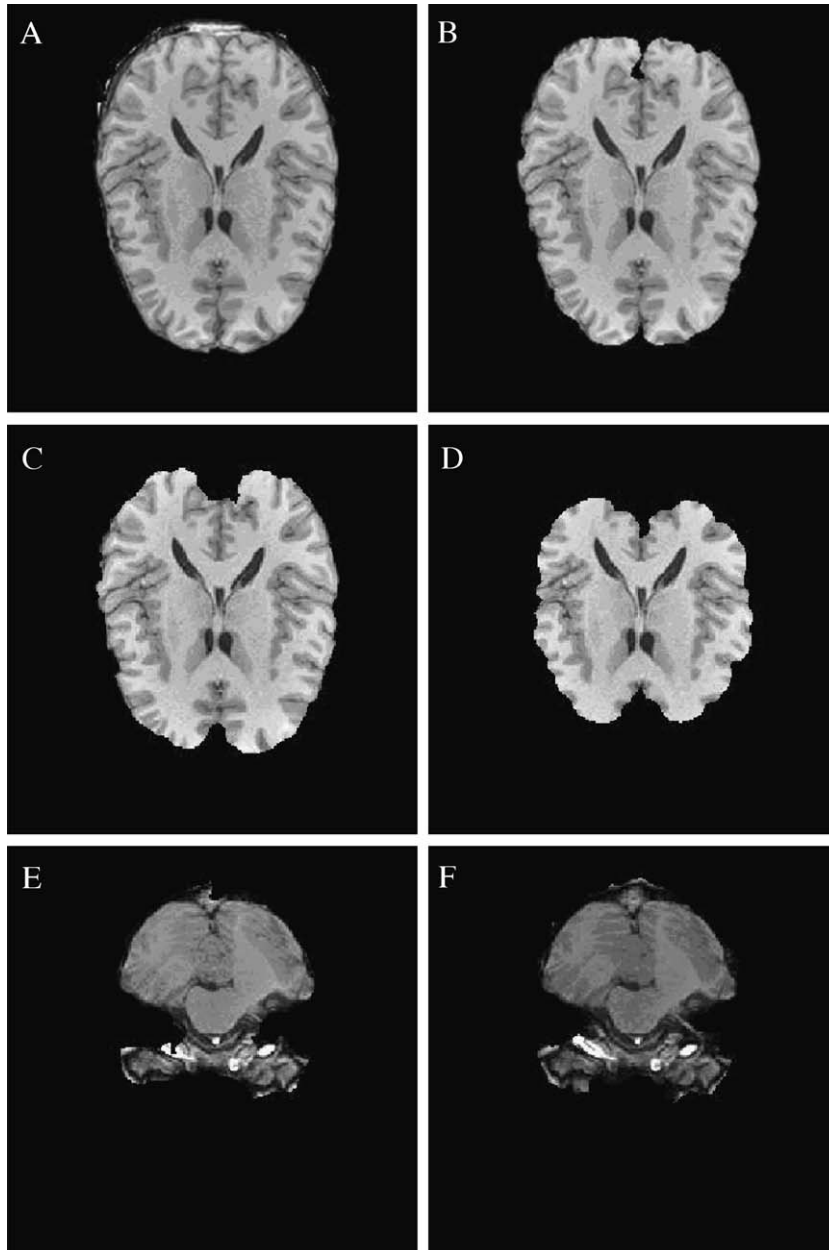


Fig. A2. Slice 62 from subject AVG with BET threshold at 0.1 (A), 0.6 (B), 0.7 (C), and 0.8 (D). Slice 15 from subject AVG with BET threshold at 0.1 (E) and 0.05 (F).

of the brain is cut off, the $1/v$ -scaling factor is not a reliable estimator when using high threshold values.

The inverse v -scaling value also changes significantly when the BET threshold is too small. In this case, more of the skull is included as part of the brain.

In the limiting case where the threshold is 0, the entire head image is 'extracted' as the brain. This becomes problematic for SIENAX because it attempts to align the stripped skull results with stripped skull templates. Failure to remove the skull interferes with this regis-

tration step. The third row of Fig. A2 shows slice 15 from AVG with BET thresholds of 0.1 (E) and 0.05 (F). As we would expect, the figure on the right shows that more nonbrain tissue is classified as brain when the $1/v$ -scaling factor is below 0.1.

The inverse v -scaling factor derived from SIENAX is relatively invariant across a range of BET threshold values. This attribute allows the inverse v -scaling factor to be reliably used as a covariate to normalize for premorbid brain size. The BET threshold parameter that should be used for generating the inverse v -scaling factor should be within the range from 0.1 to 0.6. Ideally, this threshold value should preserve as much of the brain as possible without including non-brain tissues. We currently use a threshold value of 0.5 (the default) for all of our SIENAX computations, which achieves a reasonable measure for $1/v$ -scaling. For our other brain imaging work, we manually edit the brain mask produced by BET to generate more accurate brain–skull boundaries.

References

- Adeloye, A., Kattan, K.R., Silverman, F.N., 1975. Thickness of the normal skull in the American Blacks and Whites. *American Journal of Physical Anthropology* 43, 23–30.
- Di Sclafani, V., Clark, H.W., Tolou-Shams, M., Bloomer, C.W., Salas, G.A., Norman, D., Fein, G.H., 1998. Premorbid brain size as a determinant of functional reserve in recently abstinent crack-cocaine and crack-cocaine/alcohol dependent subjects. *Journal of the International Neuropsychological Society* 4, 559–565.
- 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. *British Journal of Psychiatry* 169, 86–92.
- Hata, Y.K.S., Hirano, S., Kitagaki, H., Mori, E., 2000. Automated segmentation of human brain MR images aided by fuzzy information granulation and fuzzy inference. *Systems, Man and Cybernetics, Part C, IEEE Transactions on*, 30, 381–395.
- Itti, L., Chang, L., Mangin, J.F., Darcourt, J., Ernst, T., 1997. Robust multimodality registration for brain mapping. *Human Brain Mapping* 5, 3–17.
- 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 neurocortical plaques. *Annals of Neurology* 23, 138–144.
- Kitagaki, H., Mori, E., Yamaji, S., Ishii, K., Hirono, N., Kobashi, S., Hata, Y., 1998. Frontotemporal dementia and Alzheimer disease: evaluation of cortical atrophy with automated hemispheric surface display generated with MR images. *Radiology* 208, 431–439.
- Lynnerup, N., 2001. Cranial thickness in relation to age, sex and general body build in a Danish forensic sample. *Forensic Science International* 117, 45–51.
- Mori, E., Hirono, N., Yamashita, H., Imamura, T., Ikejiri, Y., Ikeda, M., Kitagaki, H., Shimomura, T., Yoneda, Y., 1997. Premorbid brain size as a determinant of reserve capacity against intellectual decline in Alzheimer's disease. *American Journal of Psychiatry* 154, 18–24.
- SAS Institute, 1990. *SAS Procedures Guide: Version 6*. SAS Institute, Cary, NC.
- Schofield, P.W., Mosesson, R.E., Stern, Y., Mayeux, R., Schofield, P., 1995. The age at onset of Alzheimer's disease and an intracranial area measurement. *Archives of Neurology* 52, 95–98.
- Smith, S.M., 2002. Fast robust automated brain extraction. *Human Brain Mapping* 17, 143–155.
- 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.