Relating Longitudinal and Cross-Sectional Analyses of Atrophy in Alzheimer's: BSI, SIENA and SIENAX

S. Smith[1], A. Rao[2], N. De Stefano[3], M. Jenkinson[1], R. G. Boyes[4], J. M. Schott[4], P. M. Matthews[2], and N. C. Fox[4]

[1]FMRIB, Oxford University, Oxford, United Kingdom, [2]Clinical Imaging Centre, GlaxoSmithKline, London, United Kingdom, [3]Department of Neurological and Behavioural Sciences, University of Siena, Siena, Italy, [4]Dementia Research Centre, Institute of Neurology, UCL, London, United Kingdom

Summary

Brain volume loss measured with structural MRI is accepted as a pathology marker, e.g., for Alzheimer's disease (AD)[3]. Atrophy can be measured with a variety of methods, some of which are longitudinal (measuring atrophy *rate*, for example, % brain volume change (PBVC), using two or more temporally-separated scans) and others being cross-sectional (measuring atrophy *state*, for example, brain volume normalised for head size, with a single scan). Three validated analysis tools for measuring brain atrophy are BSI and SIENA for longitudinal, and SIENAX for cross-sectional. Previous work has shown BSI and SIENA to have similar accuracy - around 0.2% error in PBVC, but no papers have compared the methods directly. Here we present a comparison using data from AD and controls. We focus on comparing BSI and SIENA longitudinal atrophy estimation, and on comparing SIENA longitudinal atrophy with SIENAX cross-sectional atrophy estimation (using the earliest scan taken from each subject). **We find excellent correspondence between BSI and SIENA, and strong correspondence between SIENA longitudinal and SIENAX cross-sectional atrophy.**

Data

46 AD, 22 age-matched controls (both groups 70±7y). Each subject scanned several times over 2 years. At some sessions, two separate scans were taken, to allow for test-retest error characterisation. T1-weighted scans: 1.5T GE Signa, IRSPGR, TE/TI/TR=6.4/650/3000ms; 1x1x1.5mm voxels. Previous BSI analyses have investigated many aspects of the dataset, including different sources of variability (cross-session, cross-subject, etc.), the dependence of estimation sensitivity on factors such as inter-scan interval, and use of ventricular size as a disease marker[3].

Methods

BSI refers to "Boundary Shift Integral", the measure of brain boundary shift used[1,2]. BSI is largely automated, utilising manual intervention at the brain-extraction stage (>98% reproducibility). The repeat scan is registered to the baseline using using affine registration. The scans are then differentially bias corrected before calculating the BSI, which integrates the inter-scan intensity difference near the brain boundary. The absolute scaling of the BSI (to produce a PBVC estimate) is calibrated using manual measurements of brain volume on each scan.

SIENA gives a fully automated analysis of whole-brain atrophy[4]; it is available as part of FSL[5]. SIENA extracts brain and skull; the brain images are then aligned to each other, using the skull images to constrain the scaling; both brain images are resampled into the space halfway between the two, avoiding bias due to asymmetric interpolation blurring. Tissue segmentation finds brain edge points, then edge displacement (between the two timepoints) is estimated to sub-voxel accuracy, by aligning the peaks of the spatial derivatives of the intensity profiles of the two images. Finally, the mean edge displacement is converted into PBVC using self-calibration based on automated image re-scaling and re-estimation of mean displacement.

SIENAX gives a fully automated analysis of normalised (for head size) brain volume (NBV)[4]. SIENAX starts by extracting brain and skull images. The brain is then affine-registered to MNI152 (using the skull to determine scaling). Next, tissue-type segmentation is applied, to give volume of brain tissue (including separate estimates of volumes of GM, WM, peripheral GM and ventricular CSF).

Results

Correlation of SIENA and BSI. We estimated PBVC for each consecutive pair of timepoints, as well as the maximum interval, e.g., t_0-t_1, t_1-t_2 .... t_4-t_5 and the full interval, t_0-t_5. All PBVC values, unnormalised for interval, were compared between SIENA and BSI. The two methods are in excellent agreement; r=0.87, p<0.0001, median-absolute-difference=0.25% (see figure). Not only are the two measures in close correlation; even the absolute scaling of the two measures is close. Scaling one set of measures to best match the other gives a relative scaling ratio of 1.2. Note that any such scaling has no effect on a method's pathology discrimination power.

Comparison of scan-rescan comparisons. Each subject had up to 3 pairs of within-visit scan-rescan images, to allow estimation of test-retest error. The [mean, mean-abs, median-abs] error values for SIENA (scaling-ratio-adjusted to make comparison valid) were [0.00, 0.27, 0.16] and BSI [0.11, 0.22, 0.17]. Hence there is no bias (nonzero mean) with SIENA, and very slight bias in BSI. On the other hand, BSI has slightly better mean absolute error, and very similar median absolute error. As with the above, this is consistent with published estimations of ~0.2% error rates for both methods.

Comparison of incremental atrophy summation with first-last difference. We summed atrophy estimation across the set of incremental timepoint pairings and compared with the direct first-last timepoint atrophy estimation. A plot of t_0-t_{final} against the sum of the incremental differences is a useful validation of many of the possible sources of estimation error. Both methods gave good (and similar) results; the [mean, median] absolute difference is SIENA=[0.20%, 0.16%] BSI=[0.20%, 0.18%] - in both cases again agreeing with 0.2% overall error rate.

AD-control t-test: sensitivity to disease effect. A useful test of the sensitivity of a method relative to its error is via a t-test between two groups that we believe should show a difference. We took the t_0-t_{final} PBVC p.a. and carried out a two-group t-test. Values are not ratio-adjusted, as this makes no difference to the t-test. The mean±sd (sd is across subjects) PBVC p.a. for SIENA was controls=0.53±0.45%, AD=2.43±1.34%. BSI gave controls=0.38±0.35%, AD=1.80±0.95%. This gives SIENA t_{66}=8.6 and BSI t_{66}=8.9, both p<10^{-10}. Hence both methods give a very significant group difference between AD and controls, with the two methods having extremely similar discrimination power.

SIENAX Normalised Brain Volume. If atrophy is occurring during the study, one would expect that to some extent it has already been taking place before the study. Hence a single-timepoint measure of brain volume should correlate with atrophy rate. One would expect such a marker of disease to be less sensitive than longitudinal atrophy, primarily because cross-subject variability in brain size is a source of error in the use of NBV as a cross-subject discriminant, and this is not a large confound in longitudinal estimation. However, it would be of great interest to be able to discriminate immediately, upon just the first scan. SIENAX was run on the first timepoint of all subjects; it estimated all volumes to be reduced in AD, except ventricular CSF (↑50%). There is nearly twice as much reduction of GM (↓14%) than WM (↓7%) in AD. Of the various volumetric measures estimated by SIENAX, the most sensitive group discriminant is NBV, with a t_{66}=7.04, i.e., very strong discrimination power - less than that found using longitudinal atrophy (SIENA t_{66}=8.6), but not very much less. The correlation of SIENAX NBV with SIENA PBVC is strong: r=0.71, p<0.0001 (see figure).
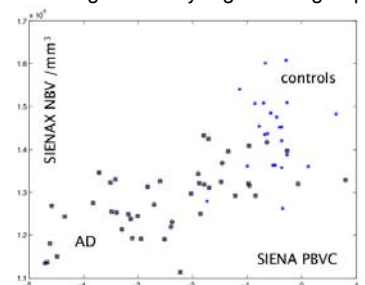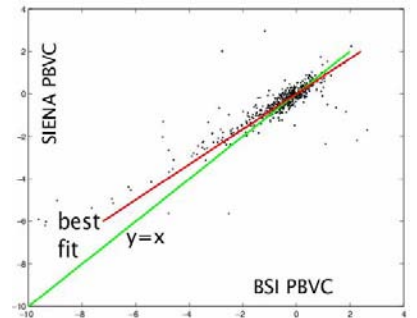
References [1]Fox, JMRI 1997. [2]Freeborough, IEEE-TMI 1997. [3]Frost, Stats. in Med. 2004. [4]Smith, NeuroImage 2002. [5]Smith, NeuroImage 2004.