NeuroImage 47 (2009) 1659–1665

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Automated quantification of caudate atrophy by local registration of serial MRI: Evaluation and application in Huntington's disease

Nicola Z. Hobbs [a,e,*], Susie M.D. Henley [a], Edward J. Wild [a], Kelvin K. Leung [a,f], Chris Frost [a,b], Roger A. Barker [c], Rachael I. Scahill [e], Josephine Barnes [a], Sarah J. Tabrizi [d,e,1], Nick C. Fox [a,d,1]

[a] Dementia Research Centre, UCL Institute of Neurology, University College London, UK
[b] Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, UK
[c] Brain Repair Centre, Department of Clinical Neurosciences, Addenbrooke's Hospital, Cambridge, UK
[d] Department of Clinical Neurology, National Hospital for Neurology and Neurosurgery, Queen Square, London, UK
[e] Department of Neurodegenerative Disease, UCL Institute of Neurology, University College London, UK
[f] Centre for Medical Image Computing (CMIC), Department of Medical Physics and Bioengineering, University College London, UK

ARTICLE INFO

Article history:
Received 19 February 2009
Revised 27 May 2009
Accepted 1 June 2009
Available online 10 June 2009

ABSTRACT

Objective: Caudate atrophy rate measured from serial MRI is proposed as a biomarker of HD progression that may be of use in assessing putative disease-modifying agents. Manual measurement techniques are the most widely applied but are time-consuming. We describe and evaluate an automated technique based on a local registration and boundary shift integral (BSI) approach at the caudate–CSF and caudate–white matter boundaries; caudate boundary shift integral (CBSI). Methods: Two-year caudate volume change was measured in controls, premanifest HD and early HD using the CBSI and compared with a detailed manual measure in terms of 1) raw caudate volume change, 2) group differentiation, 3) associations with clinical variables and 4) rater requirements. CBSI additivity was assessed by comparing measurements over a single scan pair (baseline → 2 years), with the sum of measurements from two scan pairs (baseline → 1 year → 2 years). Results: Techniques produced comparable caudate volume change measurements, although CBSI under-reported by 0.04 ml relative to manual. Both techniques distinguished controls, premanifest and early HD with a stepwise increase in rates across groups. Higher rates (CBSI and manual) were associated with increased proximity to estimated disease onset but not clinical change scores. CBSI reduced rater requirements by 2/3 (2 h per subject) relative to manual for this three time-point investigation. CBSI measurements over one scan pair showed good agreement with the sum of measurements from two scan pairs. Conclusions: CBSI results were comparable to a manual measure but with reduced rater requirements. CBSI may be of use in large-scale studies of HD.

Introduction

Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder caused by a CAG triplet repeat expansion in the gene encoding the protein *huntingtin* (Huntington Study Group, 1996). The average age of onset is approximately 40 years, and disease progression is relentless, causing abnormalities of cognition, movement and behaviour.

The earliest and most striking pathological changes in HD occur in the striatum (Gutekunst et al., 2002; Halliday et al., 1998). Cross-sectional MRI studies have demonstrated atrophy of the caudate nucleus in manifest HD and in premanifest HD up to approximately 15 years prior to predicted disease onset (Aylward et al., 2004; Paulsen et al., 2008). Caudate atrophy has been shown to progress with advancing disease (Aylward et al., 1997; Aylward et al., 2004), and therefore has been proposed as a biomarker of disease progression which may be of use in assessing the efficacy of putative disease-modifying agents in clinical trials (Aylward, 2007).

The most widely applied technique for measuring caudate atrophy by volumetry is manual segmentation. This involves an expert rater tracing around the outline of the caudate on every 'slice' of the MR image following a validated protocol (for example Aylward et al., 2000). In a longitudinal study this is carried out at all time-points and volumes measured at each time-point subtracted to estimate the amount of atrophy over the scanning interval. This approach is time-consuming and subject to high inter- and intra-rater measurement variability. Automation of measurement would be particularly valuable for large trials using caudate atrophy rate as an outcome measure. Several computer-assisted alternatives to manual segmentation have already been developed including probabilistic models based on manually-labelled training sets (Fischl et al., 2002), elastic registration with anatomical atlases (Iosifescu et al., 1997), deformable surface

* Corresponding author. Dementia Research Centre, Box 16, National Hospital for Neurology and Neurosurgery, Queen Square, London WC1N 3BG, UK. Fax: +44 20 7676 2066.
E-mail address: hobbs@drc.ion.ucl.ac.uk (N.Z. Hobbs).
[1] Denotes equal senior author.

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.06.003

models (Wallius et al., 2008), and knowledge-driven algorithms (Xia et al., 2007). However, for many reasons including confidence in reliability and anatomical accuracy of labelling, manual segmentation is still the most widely applied technique.

One approach for longitudinal volumetry is to measure volume change directly from the difference images of registered scan pairs. The brain boundary shift integral (BSI) is a semi-automated technique for measuring cerebral atrophy from affine-registered serial MRI using changes in voxel intensity between scan pairs at the brain boundary (Freeborough and Fox, 1997a). The BSI has proved successful for measuring whole-brain atrophy from serial MRI in a range of neurological disorders including Alzheimer's disease (AD) (Schott et al., 2005), multiple sclerosis (Anderson et al., 2007), HD (Henley et al., 2006), and progressive supranuclear palsy (Paviour et al., 2006).

In this study we describe a novel method for measuring the volume of caudate atrophy from serial MRI using a boundary shift integral approach at both the caudate–CSF and caudate–white matter boundaries. We refer to the technique as the caudate boundary shift integral (CBSI). The utility of the CBSI is evaluated by comparison with a detailed manual measure in three groups of individuals; controls, premanifest HD and early manifest HD. We examine 1) measurement agreement in caudate volume loss over 2 years 2) group separation on the basis of caudate atrophy rates, 3) associations between caudate atrophy rate and a range of standard clinical measures and 4) rater and processing requirements. We also assess the additivity of the CBSI by comparing CBSI measurements over a single scan pair (baseline→2 years), with the sum of CBSI measurements obtained from two scan pairs (baseline→1 year→2 years).

Methods

Summary

Caudate volume change was measured over 2 years using both manual delineation and the automated CBSI in 13 controls, 17 premanifest HD and 26 early HD subjects. Techniques were compared in terms of agreement in raw caudate volume change measurements, ability to distinguish between subject groups on the basis of caudate atrophy rates (%/year), associations with clinical variables, and processing and rater requirements.

12 controls, 16 premanifest HD and 26 early HD had serial scans at baseline, 1 year and 2 years (one control and one premanifest subject were unable to attend the 1-year scan). To assess the consistency of the CBSI over variable scanning intervals we compared the net volume change over two serial scan pairs, CBSI A→B→C, with that over one scan pair spanning the same interval, CBSI A→C. This assessment tests the additivity of the CBSI measures.

Subjects

HD subjects were recruited from the multidisciplinary HD clinic at the National Hospital for Neurology and Neurosurgery, London and the Huntington's Disease Clinic at Addenbrooke's Hospital, Cambridge. Premanifest HD was defined as a diagnostic confidence score of less than four on the Unified Huntington's Disease Rating Scale (UHDRS) (Huntington Study Group, 1996); early HD as a diagnostic confidence score of four with a UHDRS total functional capacity score of seven or more (Shoulson and Fahn, 1979). Controls were partners or spouses of mutation carriers, or individuals previously at risk of HD with a negative genetic test for the mutation. Subjects gave written informed consent in accordance with the Declaration of Helsinki, and the study had local research ethics committee and Hospital Trust approval. No subjects were taking any medication known or suspected to influence brain volume or atrophy rate (see supplementary table summarizing drug usage, S1), and subjects with concomitant neurological illnesses were excluded.

Image acquisition

Acquisition was consistent between subjects and time-points. Subjects underwent T1-weighted volumetric imaging on a 1.5-T scanner unit using an inversion-recovery prepared FAST spoiled GRASS sequence with a 24 cm×75% field of view and 256×256 matrix reconstructed to provide 124 contiguous 1.5-mm thick coronal slices. In-plane pixel dimensions: 0.9375×0.9375 mm. Acquisition parameters: repetition time = 12 ms; echo time = 5.2 ms; flip angle = 13°; inversion time = 650 ms; receiver bandwidth = 16 kHz, NEX = 1. Acquisition time was ~9 min.

Image processing and analysis

All scans were corrected for intensity inhomogeneities using the N3 algorithm (Sled et al., 1998). Image processing was carried out using the software package MIDAS (Medical Image Display and Analysis Software) (Freeborough et al., 1997b).

Manual caudate segmentation

Brains were spatially-normalized into MNI 305 atlas space (Mazziotta et al., 1995) with a six degrees-of-freedom (dof) registration (3 translations, 3 rotations, no scaling). This positioned the images in a similar space, important for the segmentation protocol, whilst ensuring that each image retained its original size. Scan-pairs in MNI space were subsequently co-registered to baseline using an affine (12 dof) (3 translations, 3 rotations, 3 scalings and 3 shears) brain–brain registration, to correct for any voxel-size changes due to slight variation in the scanner magnetic field. To reduce the potential for rater bias to influence results all segmentations were carried out blind to:

1. scan laterality; images were flipped across the mid-sagittal plane to create two mirror-image scans and the caudate was always delineated on the right-hand side of the image,
2. diagnosis; scans were de-identified,
3. time-point; baseline and registered-repeat scans for the same subject were presented on the monitor simultaneously in random order. This also enabled the rater to apply consistent arbitrary cut-offs at all time-points, thus maximizing the uniformity of the technique.

Caudate segmentation included the head and body of the caudate, with the medial border defined by the lateral ventricle, and the lateral border by the internal capsule. Caudate nuclei were segmented in the axial plane with the sagittal and coronal planes available for cross-reference.

The segmentation protocol included a semi-automated initialisation step based on percentages of the mean brain intensity (MBI). The rater manually positioned a 'seed' (marker) inside the caudate, and intensity-threshold constraints at 69% MBI and 110% MBI formed an initial outline of the caudate. This was followed by extensive and careful manual editing, including over-riding of the initial outline where necessary. To improve reproducibility the nucleus accumbens was removed on coronal view; and inferiorally, segmentation stopped one slice below the last slice in which the ventricle was visible in the axial view. The slender caudate tail, a posterior extension of the caudate running parallel to the lateral ventricle, was excluded in the sagittal view whilst ensuring that the caudate remained continuous in the axial view. We refer to this method as "manual delineation" or "manual segmentation" throughout this manuscript.

Caudate volume was computed by multiplying the ROI area by the slice thickness. Caudate volume loss was calculated by subtracting the

follow-up caudate volume from the baseline caudate volume. All segmentations were carried out by two raters whose intra- and inter-rater reproducibility (intraclass correlation coefficients) was >0.99 (mean absolute difference ~4% of mean caudate volume).

Caudate boundary shift integral (CBSI)

Baseline caudate regions were manually segmented as described above. For each subject the left and right baseline caudate regions were dilated by two voxels and used to drive a local 6 dof (3 translations, 3 rotations) caudate–caudate registration (Woods et al., 1998), see Fig. 1-A. Only the voxels inside the dilated region in the baseline image were included when computing the cost function of the linear registration. This resulted in a good rigid alignment of caudates between the baseline and the repeat images.

A region of interest containing the shift in the caudate boundary between the locally-registered scan pairs was generated. This was the XOR of the baseline caudate region eroded by one voxel and the baseline caudate region dilated by one voxel (see Fig. 1-B for an example CBSI region). The volume of caudate loss (atrophy) was estimated by computing the three-dimensional integral of all the boundary shifts within this region of interest, using the BSI algorithm, with boundary shift being determined directly from voxel intensities (Freeborough and Fox, 1997a). A double intensity window was included for the CBSI calculation in order to capture boundary shift at both the caudate–CSF border, and the caudate–WM border (Fig. 2).

The optimal parameters for the CBSI calculation were chosen using an automatic intensity window selection method. In order to capture most of the tissue-type change between the caudate and CSF (or caudate and WM), it was desirable to ignore changes within the same tissue type, and to maximize changes between different tissue types. Therefore, the lower intensity window for capturing CSF and caudate change was chosen to be ($I_{\text{CSF mean}} + I_{\text{CSF std}}$, $I_{\text{caudate mean}} - I_{\text{caudate std}}$), and the higher intensity window for capturing caudate and WM change was chosen to be ($I_{\text{caudate mean}} + I_{\text{caudate std}}$, $I_{\text{WM mean}} - I_{\text{WM std}}$) where $I_{\text{CSF mean}}$, $I_{\text{CSF std}}$, $I_{\text{caudate mean}}$, $I_{\text{caudate std}}$, $I_{\text{WM mean}}$ and $I_{\text{WM std}}$ are the mean and standard deviation of CSF-, caudate- and WM-intensities respectively, see Fig. 2. $I_{\text{caudate mean}}$ and $I_{\text{caudate std}}$ were estimated over the baseline caudate region, and $I_{\text{CSF mean}}$, $I_{\text{CSF std}}$, $I_{\text{WM mean}}$ and $I_{\text{WM std}}$ were estimated using a *k*-means clustering method over a dilated (by three voxels) caudate region, with the caudate masked. The optimal intensity windows were taken to be the median of all the subject-specific optima. This method resulted in a lower intensity window of 0.39 to 0.82 for measuring volume change at the caudate–CSF boundary, and a higher intensity window of 0.99 to 1.05 for measuring volume change at the caudate–WM boundary. The CBSI was the sum of the lower and higher intensity window BSIs.

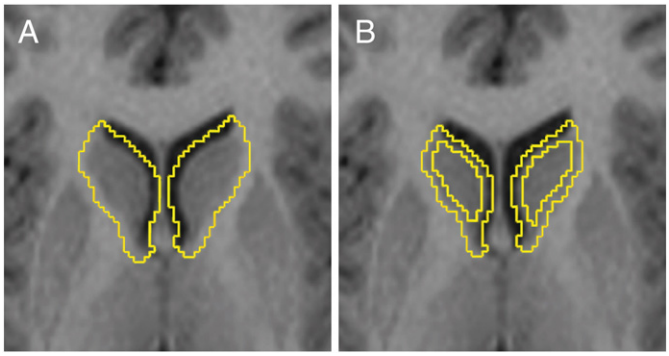

Fig. 1. (A) Example region used to drive the local rigid caudate–caudate registration. (B) CBSI region of interest encompassing the shift in caudate boundary between locally-registered scan pairs.

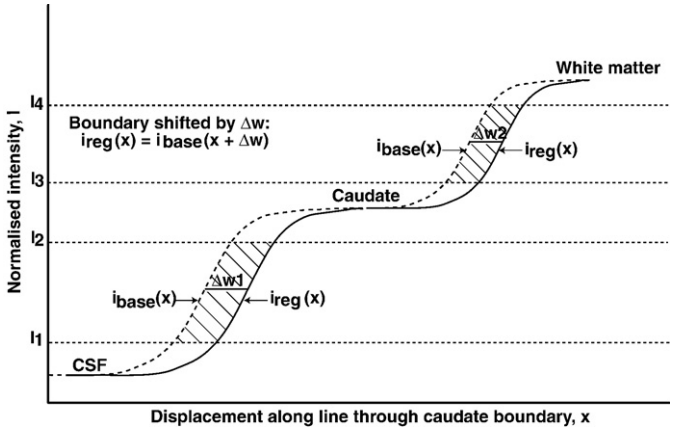

Fig. 2. Idealised one-dimensional representation of the change in intensity associated with a shift in caudate boundary between a baseline scan and a locally-registered repeat scan. The dotted line represents the intensity profile (boundary or edge) for the baseline scan; the solid line represents the same edge on the repeat scan; the shift in edge, Δw, is due to caudate atrophy. I_1 to I_2 is the intensity window for estimating change at the caudate–CSF boundary, where $I_1 = I_{\text{CSF mean}} + I_{\text{CSF std}}$, $I_2 = I_{\text{caudate mean}} - I_{\text{caudate std}}$. I_3 to I_4 is the intensity window for estimating change at the caudate–WM boundary, where $I_3 = I_{\text{caudate mean}} + I_{\text{caudate std}}$, $I_4 = I_{\text{WM mean}} - I_{\text{WM std}}$.

Quality control

Significant motion artefacts are common in HD and may reduce the reliability of both techniques. To control for this, locally-registered scan pairs were visually assessed for artefacts and/or mis-registration by one investigator blinded to subject identity. The region over which the CBSI was calculated was assessed to ensure that it was sufficiently large to incorporate the shifts in the caudate boundary, whilst at the same time not distorting the measure by including boundaries of structures other than the caudate (e.g. the putamen).

Statistical analysis

Data were analysed using STATA v10.0 (Stata Corporation, College Station, TX).

Agreement between measurement techniques

Only the 0-to-2 year data were used to assess agreement (to increase SNR). This was assessed using a Bland–Altman plot (Bland and Altman, 1986). Pitman's variance ratio test assessed whether there was a difference in within-group variability between techniques. Paired *t*-tests were used to test for differences in mean caudate volume loss between techniques.

Group separation on the basis of caudate atrophy rate

Caudate volume loss was converted to a percentage of baseline caudate volume and annualized using a logarithmic scale. Analysis on a logarithmic scale has the advantage that doublings and halving are treated as effects of equal magnitude. Mean atrophy rates for each group (%/year) were calculated by back transformation with standard deviations (SDs) calculated from variance transformation formulae. Between-group differences in atrophy rate were investigated using a linear regression model with robust standard errors to allow for heterogeneity in variance between groups, and adjusting for age and gender. Pair-wise between-group differences were divided by the residual standard deviation to estimate effect sizes. To simultaneously allow for both the imprecision in estimated between-group differences and the residual standard deviation, bias-corrected and accelerated bootstrap confidence intervals for effect sizes were calculated using 2000 bootstrap replications (Efron and Tibshirani, 1993). In order to compare effect sizes for the two techniques, the two

Table 1
Demographic data.

	Control	Premanifest HD	Early HD
N	12	16	18 (13 St.1, 5 St.2)[a]
Gender, M:F	6:6	7:9	6:12
Age at baseline assessment, years	43.6 (9.5)	38.6 (8.7)	45.9 (9.3)
CAG repeat length	–	42 (1.6)	43.8 (2.3)
Predicted years to onset[b]	–	17.0 (6.2) range 9–30	–
Disease duration, years	–	–	4.9 (2.8) range 1.0–10.5
Scanning interval; baseline to 2 years, years	2.2 (0.1)	2.3 (0.1)	2.3 (0.1)
Scanning interval; baseline to 1 year, years	1.0 (0.05)	1.0 (0.1)	1.0 (0.1)

Data are presented as mean (SD).
[a] St. 1 is Stage 1, St. 2 is Stage 2 (Shoulson and Fahn, 1979).
[b] Calculated from age at baseline assessment using the equation from Langbehn et al. (2004), with onset defined as 60% chance of showing signs.

models were fitted simultaneously using generalised least squares and 95% bootstrap confidence intervals for the differences in effect sizes calculated.

Associations with clinical variables

Associations between caudate atrophy rate and rate of change in the UHDRS total functional capacity (TFC), independence score (IS) and motor score were investigated using linear regression models with only the early HD subjects included in the analysis, and adjusting for age and gender. Similar models were used to investigate the relationship between caudate atrophy rate and estimated 5-year onset probability in premanifest subjects (Langbehn et al., 2004), and CAG repeat length in premanifest and early HD subjects separately.

CBSI additivity

A Bland–Altman plot assessed agreement between CBSI estimates measured over two serial scan pairs, A→B→C, and one serial scan pair, A→C, both spanning 2 years (Bland and Altman, 1986). Pitman's variance ratio test assessed whether there was a difference in variability between measurements. Paired *t*-tests were used to test for differences in mean caudate volume loss between estimates.

Results

Quality control

Visual assessment showed all local registrations to be of acceptable quality. CBSI regions were all sufficiently large to accommodate the shift in caudate boundary between time-points. 8/56 scan pairs were rejected due to motion (1/13 control, 1/17 premanifest, and 6/26 early HD). An additional 2/56 scan pairs were rejected due to intensity artefact within the caudate region (2/26 early HD).

Demographics

Demographic data for included subjects are shown in Table 1. As expected the premanifest group was significantly younger than the early HD group ($p=0.03$). There was no significant difference between premanifest and controls ($p=0.2$), or early HD and controls ($p=0.5$).

Agreement between measurement techniques

There was reasonably good agreement between the CBSI and the manual measurements, see Figs. 3-A and B. The CBSI had a tendency to underestimate change relative to the manual measure with a mean difference in volume loss estimations between techniques of 0.05 ml (95% CI 0.02 ml to 0.08 ml). Underestimation was consistent in direction in all three subgroups although this was only statistically significant in the premanifest HD group (Table 2). Importantly, the Bland–Altman plot demonstrated no evidence of the under-reporting increasing as the measure of interest (caudate volume loss) increased (Fig. 3-B). In addition, within-group variance in caudate volume loss was comparable between techniques (Table 2).

Group separation on the basis of caudate atrophy rate

Raw caudate atrophy rates by group are depicted in Fig. 4. Mean (SD) rates measured manually were 0.2 (1.0) %/year in controls, 1.7 (1.2) %/year in premanifest and 3.4 (1.7) %/year in early HD, and measured with the CBSI were 0.1 (1.0) %/year in controls, 1.2 (0.9) %/year in premanifest and 2.9 (1.6) %/year in early HD. After adjusting for age and gender differences between groups, atrophy rates were significantly higher in the premanifest group relative to the control group; this finding was consistent between measurement techniques (both $p<0.01$, see Table 3). Caudate atrophy rates were significantly higher in the early HD group relative to both controls and premanifest subjects (all $p<0.001$), again this was consistent between techniques (Table 3). Table 3 also presents effect sizes (defined as the ratio of the between-group difference to the residual standard deviation) for the pair-wise comparisons between groups. For each comparison these were of similar magnitude for both techniques. Differences between effect sizes for the two techniques were not statistically significant (all $p>0.05$).

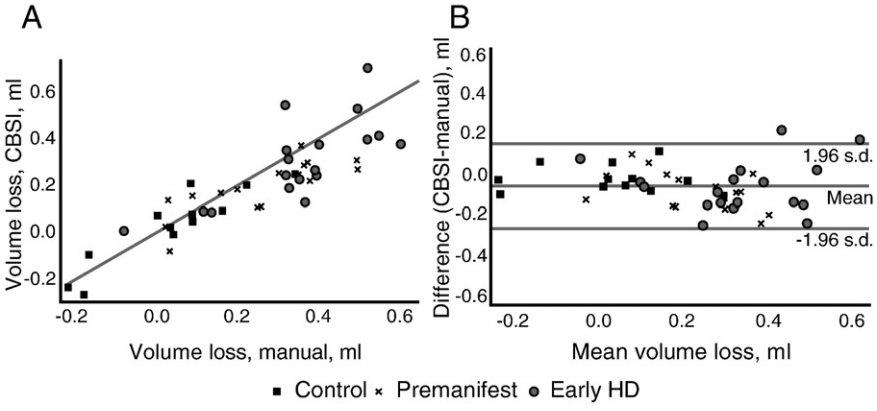

Fig. 3. (A) Scatter plot illustrating the association between manual- and CBSI-derived measures of caudate volume loss measured over 2 years. Line of identity is illustrated. (B) Bland–Altman plot of mean volume loss against difference in volume loss measurements between techniques, with limits of agreement indicated.

Table 2
Caudate volume loss measured from two MRIs 2 years apart using a detailed manual measure and the automated CBSI technique.

	Manual caudate volume loss	CBSI caudate volume loss	Difference in mean caudate volume loss (Manual − CBSI)	Ratio of variance of caudate volume loss (Manual/CBSI)
	Mean (SD), ml	Mean (SD), ml	Mean (95% CI), ml	Mean (95% CI)
Control ($N=12$)	0.04 (0.17)	0.03 (0.17)	0.01 (−0.03 to 0.06), $p=0.6$	1.02 (0.77 to 1.35), $p=0.9$
Premanifest HD ($N=16$)	0.26 (0.16)	0.19 (0.12)	0.07 (0.02 to 0.12), $p=0.01$	1.33 (0.98 to 1.91), $p=0.06$
Early HD ($N=18$)	0.36 (0.17)	0.31 (0.18)	0.05 (−0.01 to 0.12), $p=0.1$	0.94 (0.66 to 1.34), $p=0.7$

Paired *t*-tests were used to compare mean volume loss between techniques for each group. Pitman's variance ratio test was used to test for differences in variance between techniques for each group.

Associations with clinical variables

In the premanifest group higher caudate atrophy rates were associated with increased proximity to predicted disease onset and longer CAG repeat length using both techniques (both age- and gender- adjusted) (Table 4). There was no evidence of an association between atrophy rate and CAG repeat length in the early HD group (age- and gender- adjusted). We were unable to show evidence of a relationship between caudate atrophy rate and change in UHDRS motor, independence or total functional capacity score. These findings were consistent between techniques.

CBSI additivity

As expected, CBSI values measured from a single pair of scans (A→C), were highly correlated with those measured from two pairs of scans spanning the same interval (A→B→C), see Fig. 5-A. There was no significant difference in mean caudate volume loss between estimations ($p=0.5$), and there was no evidence of bias, see Bland–Altman plot, Fig. 5-B.

Processing and rater requirements

For the manual measure the average delineation time was 30 min per caudate, 180 min per subject for this study (left plus right caudate at three time-points). The CBSI required the manual segmentation of the baseline caudates only, requiring 60 min of rater time per subject. The local registration step and generation of the CBSI were both automated steps requiring approximately 5 min of processor time per scan pair using a dual-processor Intel® Pentium 4 Xeon™ CPU 3.06 GHz computer. Visual assessment of the local registration and the BSI region took 3 min per subject.

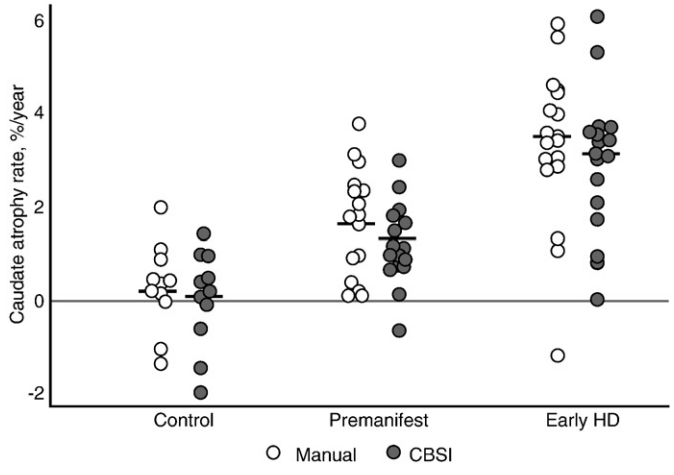

Fig. 4. Raw caudate atrophy rates (unadjusted for age or gender differences between groups in the figure). Caudate volume loss was measured from two MRI scans 2 years apart using a detailed manual measure and the automated CBSI measure, converted to a percentage of the baseline caudate volume and annualized. Solid lines indicate the mean rate per group.

Discussion

We developed and evaluated an automated technique for measuring the volume of caudate atrophy from locally-registered serial MRI in HD. The technique (CBSI) uses a boundary shift integral approach at both the caudate–CSF and caudate–white matter boundaries to give a direct estimate of caudate volume change (atrophy) over time.

We found a close linear relationship in atrophy measurements between the CBSI and a detailed manual measure over a two-year interval. Although the CBSI had an undesirable tendency to under-report change relative to the manual measure, this underestimation was independent of the amount of atrophy and therefore did not compromise the ability of the CBSI to track disease progression. We found good agreement between CBSI estimates of caudate volume loss measured from a single pair of scans, CBSI A→C, and the sum of estimates obtained from two pairs of scans, CBSI A→B→C, both spanning a total of 2 years, with no evidence of bias. Such agreement is important in the context of multiple time-point studies particularly where subjects may not have equal scanning intervals or where some scans are of inadequate quality for analysis.

An important test for the CBSI technique is its ability to distinguish between subject groups on the basis of caudate atrophy rate. We found caudate atrophy rates (CBSI) to be approximately 0.1%/year in controls, 1.2%/year in premanifest subjects and 2.9%/year in early HD subjects, slightly lower but comparable to the rates found manually. Both measurement techniques distinguished controls, premanifest and early HD groups on the basis of caudate atrophy rate with a stepwise increase in rates across groups.

Within the premanifest HD group subjects closer to predicted disease onset had significantly higher caudate atrophy rates than those further from estimated onset; an increase in 5-year onset probability 0.1 was associated with an increase in caudate atrophy rate of approximately 0.6%/year (CBSI). Within this group higher atrophy rates were also associated with a longer CAG repeat length, with a one triplet increase predicting approximately 0.5%/year increase in atrophy rate (CBSI). These data suggest that it may be advantageous

Table 3
Group differences in caudate atrophy rate (%/year) measured over 2 years using manual delineation and the CBSI, with adjustment for age and gender differences between groups.

	Group differences in caudate atrophy rate, %/year		
	PM − control	Early HD − control	Early HD − PM
Manual measurement			
Mean (95% CI)	1.5 (0.7 to 2.3)	3.3 (2.5 to 4.3)	1.9 (1.0 to 2.7)
p value	$p<0.001$	$p<0.001$	$p<0.001$
Effect size (95% CI)	1.1 (0.3 to 1.7)	2.5 (1.2 to 3.4)	1.4 (0.5 to 2.1)
CBSI measurement			
Mean (95% CI)	1.1 (0.4 to 1.7)	3.0 (2.1 to 3.9)	1.9 (1.1 to 2.7)
p value	$p=0.002$	$p<0.001$	$p<0.001$
Effect size (95% CI)	0.9 (0.3 to 1.4)	2.6 (1.7 to 3.2)	1.6 (0.7 to 2.2)

PM = premanifest HD. Effect sizes defined to be adjusted between-group differences divided by the pooled residual standard deviation.

Table 4
Associations between caudate atrophy rate (%/year) and clinical measures with age- and gender- adjustment.

Variable	Group	Manual measurement	CBSI measurement
		Slope (95% CI) p value	Slope (95% CI) p value
5-year onset probability[a]	Premanifest HD	1.0 (0.3 to 1.7) $p=0.01$	0.6 (0.1 to 1.2) $p=0.04$
CAG repeat length	Premanifest HD	0.80 (0.39 to 1.20) $p=0.001$	0.48 (0.15 to 0.82) $p=0.01$
	Early HD	0.32 (−0.27 to 0.90) $p=0.3$	0.12 (−0.41 to 0.65) $p=0.6$

Slope indicates the predicted effect a 1 unit increase in the variable of interest will have on caudate atrophy rate (%/year), except for onset probability where it is the predicted effect of a 0.1 increase in 5-year onset probability.
[a] Calculated from age at baseline assessment using the equation from Langbehn et al. (2004).

for future studies to stratify premanifest subjects based on clinical characteristics such as CAG repeat length in order to improve group homogeneity. We were not able to show an association between caudate atrophy rate and change in the functional/motor scores over a two-year interval. This may be due to high variability in either the caudate or the clinical measure, or a lack of sensitivity to change. In fact we found the clinical change scores tended to show both higher variability and lower sensitivity to change than the caudate measure, especially in the premanifest HD group. This is one of the reasons that structural markers of progression are of such interest in HD. Further work over longer scanning intervals and with larger datasets is required to assess relationships between functional/motor abilities and structural brain changes.

An advantage of automated or semi-automated techniques such as the CBSI is the reduction in labour requirements relative to a manual measure. This is particularly beneficial with multiple time-point studies. For a five time-point study the CBSI would take 1/5 of the operator time, saving 4 h per subject. Reducing rater requirements in this way is both time- and cost-effective, and potentially reduces the number of raters required. Furthermore, it may be possible to eliminate rater input completely by incorporating an automated caudate segmentation algorithm into the CBSI process stream.

In contrast to our clinical findings, a previous MRI study found no differences in the amount of caudate volume change over time between premanifest, mild and moderate HD subjects suggesting that caudate atrophy progresses uniformly with disease (Aylward et al., 2000). This discrepancy is most likely to be due to differences in the disease severity of the subjects studied. Aylward et al. (2000) included premanifest subjects much closer to predicted motor onset and therefore likely to be more similar to affected HD subjects than controls. We found caudate atrophy rates to be approximately 3%/year in early HD, slightly lower but comparable to a previous study that found a rate of 4.9%/year (Aylward et al., 2003). Again discrepancies may be accounted for by differences in the subjects studied (disease severity, CAG repeat length, and age), variation in measurement techniques, and differences in image quality and acquisition.

We demonstrated increased atrophy rates of 1.2%/year in a premanifest cohort which was on average 17 years prior to predicted motor onset (range 9 to 30 years). Aylward et al. (2004) found a higher rate of 4.3%/year in a cohort the subjects of which were all within 11 years to predicted onset. The studies differed in the subjects studied and the estimates of onset applied; Aylward et al. (2004) used a formula which took parental age at onset into account, whereas we applied the probabilistic survival model of Langbehn et al. (2004). However, results are consistent with a hypothesis of caudate atrophy initialising many years prior to onset (Paulsen et al., 2008), and atrophy rates increasing with closer proximity to onset.

The strength of this study is in the level of detail in the manual measure. Caudates were delineated in 'standard space' enabling the development and use of a robust segmentation protocol with detailed arbitrary cut-offs. The effects of 'segmentor drift' (an increase in intra-rater variation with time) were controlled for by segmenting scans for each subject simultaneously, blinded to time-point, diagnosis and left–right orientation. We also controlled for changes in voxel size over time caused by alterations in scanner calibration by segmenting caudate nuclei on registered (i.e. scaled) scan pairs instead of in native space (Whitwell et al., 2001). The retrospective study design we employed enabled us to produce our most precise manual change results, which was important since these results were used to validate a novel measurement technique. However, it is notable that for large drug trials it would be very difficult to delay all segmentations until the end of the trial due to high manual requirements. In this situation segmentations would probably be performed continuously throughout the trial by several raters, and hence inter- and intra-rater reliability may be lower. In this study we reviewed the baseline and registered-repeat images for each subject simultaneously to assess whether the scan pair was of adequate quality for analysis. This is especially important in studies of HD to ensure that group differences are not influenced by a higher prevalence of motion artefact in the scans of HD subjects rather than in controls or premanifest HD subjects; such careful quality control may not be feasible in much larger studies and hence results may be more 'noisy'.

Further work will be needed to determine whether conclusions from this relatively small, single-site study will be applicable to larger, multi-centre projects. For example, as a single-site study the scans used here were of consistent quality but this may be harder to guarantee in a large multi-centre study. Additionally, the scans in this study were acquired on a 1.5-Tesla scanner. Future studies are likely to

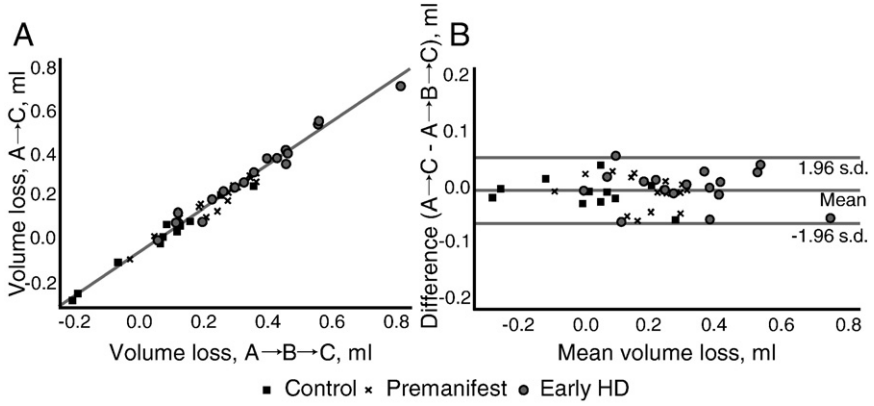

Fig. 5. (A) Association between estimates of caudate volume loss measured from a single pair of scans, A→C, and the sum of estimates obtained from two pairs of scans, A→B→C, both spanning 2 years. 'A' represents a baseline scan, 'B' represents a 1-year scan scan and 'C' represents a 2-year scan. (B) Bland–Altman plot showing the agreement between the two CBSI estimates, with limits of agreement indicated.

use higher field strengths and achieve improved MRI acquisition and this may improve the performance of contrast-based measures such as the CBSI. A limitation of the CBSI is its dependency on consistent image quality between time-points. Intensity non-linearities localised to the caudate boundary are likely to have more impact on the CBSI than the manual measure since a human rater can be guided by landmarks rather than signal intensity where necessary. However, longitudinal differences in global contrast and geometric distortion are likely to affect both the manual and CBSI measures.

We compared our CBSI measures with our manual measures of change. Although we aimed to produce the most precise measures of manual change possible, the adoption of thresholding as an initialisation step could potentially lead to bias in segmentation. Once the threshold-dependent outline was produced, the rater may be less-inclined to change this contour compared with an entirely manual technique. However, we used these thresholds consistently throughout the study on every scan and every contour was manually checked and changed if it was deemed not to follow anatomical boundaries.

We applied consistent CBSI parameters (window width and window centre) across the dataset. One limitation of this approach is that the parameter choice for a given subject is dependent on the characteristics of the rest of the study population. However, parameter selection was entirely automated and cohort specific. This automated procedure could be applied to any cohort and would enable this technique to be applicable to new scan acquisitions and different field strengths.

Computer-assisted segmentation techniques have shown promise cross-sectionally (for example Iosifescu et al., 1997; Khan et al., 2008), and produce caudate segmentations which have approximately 80% spatial overlap with manually-outlined regions, and similarity coefficients of approximately 90% (Iosifescu et al., 1997). Such techniques may also be of use longitudinally although it is yet to be shown if independent automated segmentations at multiple time-points can provide sensitive measures of longitudinal volume change.

Conclusion

While manual measures remain the ‘gold standard’ for caudate volumetry, the development of automated techniques such as the CBSI will become increasingly important as compounds showing promise in model systems of HD require clinical testing in large multi-centre trials. This study has shown the CBSI to be a promising biomarker candidate. However, further validation using larger datasets and multi-site scans is essential.

Acknowledgments

The authors would like to thank the patients and controls who took part in this study. This study was funded by the Cure Huntington's Disease Initiative (CHDI), Inc. (a non-profit organisation). NCF is funded by the Medical Research Council UK. JB is funded by an Alzheimer's Research Trust UK research fellowship with the kind support of the Kirby Liang Foundation. KL is supported by Technology Strategy Board Grant UK (M1638A). This work was undertaken at UCLH/UCL which received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme. This work was supported by a NIHR BRC award to Addenbrooke's Hospital. The Dementia Research Centre is an Alzheimer's Research Trust Co-ordinating Centre. We gratefully acknowledge help and suggestions from Dr Sebastien Ourselin.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2009.06.003.

References

Anderson, V.M., Bartlett, J.W., Fox, N.C., Fisniku, L., Miller, D.H., 2007. Detecting treatment effects on brain atrophy in relapsing remitting multiple sclerosis: sample size estimates. J. Neurol. 254, 1588–1594.
Aylward, E.H., 2007. Change in MRI striatal volumes as a biomarker in preclinical Huntington's disease. Brain Res. Bull. 72, 152–158.
Aylward, E.H., Li, Q., Stine, O.C., Ranen, N., Sherr, M., Barta, P.E., Bylsma, F.W., Pearlson, G.D., Ross, C.A., 1997. Longitudinal change in basal ganglia volume in patients with Huntington's disease. Neurology 48, 394–399.
Aylward, E.H., Codori, A.M., Rosenblatt, A., Sherr, M., Brandt, J., Stine, O.C., Barta, P.E., Pearlson, G.D., Ross, C.A., 2000. Rate of caudate atrophy in presymptomatic and symptomatic stages of Huntington's disease. Mov. Disord. 15, 552–560.
Aylward, E.H., Rosenblatt, A., Field, K., Yallapragada, V., Kieburtz, K., McDermotte, M., Raymondf, L.A., Almqvist, E.W., Haydenh, M., Ross, C.A., 2003. Caudate volume as an outcome measure in clinical trials for Huntington's disease: a pilot study. Brain Res. Bull. 62, 137–141.
Aylward, E.H., Sparks, B.F., Field, K.M., Yallapragada, V., Shpritz, B.D., Rosenblatt, A., Brandt, J., Gourley, L.M., Liang, K., Zhou, H., Margolis, R.L., Ross, C.A., 2004. Onset and rate of striatal atrophy in preclinical Huntington disease. Neurology 63, 66–72.
Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1, 307–310.
Efron, B., Tibshirani, R.J., 1993. An introduction to the Bootstrap. Chapman and Hall.
Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der, K.A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355.
Freeborough, P.A., Fox, N.C., 1997a. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. IEEE Trans. Med. Imaging 16, 623–629.
Freeborough, P.A., Fox, N.C., Kitney, R.I., 1997b. Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans. Comput. Methods Programs Biomed. 53, 15–25.
Gutekunst, C., Norflus, F., Hersch, S.M., 2002. The neuropathology of Huntington's disease. In: Bates, G., Harper, P., Jones, L. (Eds.), Huntington's Disease. Oxford University Press, Oxford.
Halliday, G.M., McRitchie, D.A., Macdonald, V., Double, K.L., Trent, R.J., McCusker, E., 1998. Regional specificity of brain atrophy in Huntington's disease. Exp. Neurol. 154, 663–672.
Henley, S.M., Frost, C., MacManus, D.G., Warner, T.T., Fox, N.C., Tabrizi, S.J., 2006. Increased rate of whole-brain atrophy over 6 months in early Huntington disease. Neurology 67, 694–696.
Huntington Study Group, 1996. Unified Huntington's Disease Rating Scale: reliability and consistency. Huntington Study Group. Mov. Disord. 11, 136–142.
Iosifescu, D.V., Shenton, M.E., Warfield, S.K., Kikinis, R., Dengler, J., Jolesz, F.A., McCarley, R.W., 1997. An automated registration algorithm for measuring MRI subcortical brain structures. NeuroImage 6, 13–25.
Khan, A.R., Wang, L., Beg, M.F., 2008. FreeSurfer-initiated fully-automated subcortical brain segmentation in MRI using Large Deformation Diffeomorphic Metric Mapping. NeuroImage 41, 735–746.
Langbehn, D.R., Brinkman, R.R., Falush, D., Paulsen, J.S., Hayden, M.R., 2004. A new model for prediction of the age of onset and penetrance for Huntington's disease based on CAG length. Clin. Genet. 65, 267–277.
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. The International Consortium for Brain Mapping (ICBM). NeuroImage 2, 89–101.
Paulsen, J.S., Langbehn, D.R., Stout, J.C., Aylward, E., Ross, C.A., Nance, M., Guttman, M., Johnson, S., MacDonald, M., Beglinger, L.J., Duff, K., Kayson, E., Biglan, K., Shoulson, I., Oakes, D., Hayden, M., 2008. Detection of Huntington's disease decades before diagnosis: the Predict-HD study. J. Neurol. Neurosurg Psychiatry 79, 874–880.
Paviour, D.C., Price, S.L., Jahanshahi, M., Lees, A.J., Fox, N.C., 2006. Longitudinal MRI in progressive supranuclear palsy and multiple system atrophy: rates and regions of atrophy. Brain 129, 1040–1049.
Schott, J.M., Price, S.L., Frost, C., Whitwell, J.L., Rossor, M.N., Fox, N.C., 2005. Measuring atrophy in Alzheimer disease: a serial MRI study over 6 and 12 months. Neurology 65, 119–124.
Shoulson, I., Fahn, S., 1979. Huntington disease: clinical care and evaluation. Neurology 29, 1–3.
Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans. Med. Imaging 17, 87–97.
Wallius, E., Tohka, J., Hirvonen, J., Hietala, J., Ruotsalainen, U., 2008. Evaluation of the automatic three-dimensional delineation of caudate and putamen for PET receptor occupancy studies. Nucl. Med. Commun. 29, 53–65.
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. AJNR Am. J. Neuroradiol. 22, 1483–1489.
Woods, R.P., Grafton, S.T., Holmes, C.J., Cherry, S.R., Mazziotta, J.C., 1998. Automated image registration: I. General methods and intrasubject, intramodality validation. J Comput. Assist. Tomogr. 22, 139–152.
Xia, Y., Bettinger, K., Shen, L., Reiss, A.L., 2007. Automatic segmentation of the caudate nucleus from human brain MR images. IEEE Trans. Med. Imaging 26, 509–517.