WILEY
Publishers Since 1807

A method for retrospective estimation of natural head movement during structural MRI

Journal:	*Journal of Magnetic Resonance Imaging*
Manuscript ID	Draft
Wiley - Manuscript type:	Original Research
Classification:	Quality assurance and phantoms < Imaging technology and safety < Basic Science, Imaging techniques and processing < Imaging technology and safety < Basic Science, Neuro-imaging < Clinical Science
Manuscript Keywords:	Head motion, brain structural MRI, brain morphometry, Parkinson's disease, average edge strength, entropy

SCHOLARONE™
Manuscripts

FOR REVIEW ONLY

A method for retrospective estimation of natural head movement during structural MRI

Abstract

Background

Head motion during brain structural MRI scans biases brain morphometry measurements but quantitative retrospective methods estimating head motion from structural MRI have not been evaluated.

Purpose/Hypothesis

We hypothesized that two metrics retrospectively computed from MR images 1) average edge strength (AES, reduced with image blurring) and 2) entropy (ENT, increased with blurring and ringing artifacts) could be sensitive to in-scanner head motion during acquisition of T1-weighted MR images. Specific goals were to: i) Evaluate if these metrics differentiated two populations with expected natural head motion differences: Healthy Control (HC) and cognitively normal Parkinson Disease (PD) patients; ii) Investigate whether within the PD group, AES and ENT were associated with clinical tremor score (TS); iii) Test whether AES and ENT predict local or distributed brain morphometry parameters, including cortical thickness (CT), gray-white matter contrast (GWC) and gray matter density (GMx).

Study Type

Retrospective

Population/Subjects/Phantom/Specimen/Animal Model

83 HC and 120 PD patients

Field Strength/Sequence

FOR PEER REVIEW ONLY

3D MPRAGE images at 3T

Assessment

We 1) compared AES and ENT distribution between HC and PD, 2) evaluated the correlation between TS and AES (or ENT) in PD and 3) investigated cortical regions showing an association between AES (or ENT) and local and network-level covariance measures of CT, GWC and GMx.

Statistical Tests

1) Mann-Whitney (for AES) and T-student (for Ent) test. 2) Spearman's rank correlation. 3) General Linear Model and Partial Least Square analysis

Results

AES, but not ENT, differentiated HC and PD ($z=-1.72$, $p=0.04$). In PD AES correlated negatively with TS ($rho=-0.21$, $p=0.02$) and showed a significant relationship ($|Z|>3$, $p<0.001$) with structural covariance of CT and GWC in 54 out of 68 cortical regions.

Data Conclusion

In clinical populations prone to head motion AES can provide a reliable retrospective index of motion during structural scans, identifying brain areas whose morphometric measures co-vary with motion.

Key words: Head motion; brain structural MRI; brain morphometry; Parkinson's disease; average edge strength; entropy

FOR PEER REVIEW ONLY

INTRODUCTION

Head motion during the acquisition of structural brain images can produce artifacts such as image blurring or ringing (1). These artifacts may be particularly problematic when using structural T1-weighted MRI for morphometric analyses (2-4), and particularly when trying to understand morphometric changes in clinical populations. Recent studies, where participants were instructed to perform head movement during structural MRI acquisition, show that head motion significantly biases brain morphometry estimates and test-retest reliability (5-7). There is, therefore, an interest in identifying head motion and characterizing its potential effects such that morphometry biases can be minimized.

When considering prospective studies, several novel procedures can reduce head motion artifacts during the acquisition of structural MRI data. These include increasing participant cooperation through training, using advanced acquisition sequences that enable shorter MRI acquisition times (8), applying k-space trajectories optimized to mitigate motion artifacts (9), using online prospective motion correction tools such as MR navigator (10), or external motion tracking devices (11). These correction methods reduce motion-related biases and variance in voxel based morphometry (VBM) and cortical thickness (CT) studies (12).

However, in standard retrospective structural MRI studies performed without online head motion correction, image quality control (QC) tends to be qualitative and aimed at classifying whether the data is of sufficient quality (e.g. good, moderate, bad) to be further processed based on visual assessment of head motion related image artifacts (13,14). A more quantitative approach was suggested by recent studies is to relate head motion estimations derived from functional MRI scans with CT biases from structural MRI data acquired in the same session (13,15,16). This method may only be applied if both structural and functional data are available and assumes similar levels of motion across functional and structural scans.

FOR PEER REVIEW ONLY

A previous study introduced two metrics for estimating whole-brain motion from the full 3D structural T1 volume: 1) Average Edge Strength (AES) and 2) image Entropy (ENT) (17). AES quantifies the average intensity of the contrast at the intensity edges of an image, so it is sensitive to the amount of blurring. ENT quantifies the distribution of the image energy over all voxels in an image. ENT increases when artifacts produce blurring, ghosting and ringing. Both metrics have been shown to successfully quantify improvement in image quality after prospective motion correction (17). However, in those studies participants were instructed to perform specific head movements during the MRI acquisitions. It remains unclear whether AES and ENT are sufficiently sensitive for capturing natural head motion effects from structural MRI.

To the best of our knowledge, no studies have investigated whether these retrospective head motion metrics can be applied to evaluate if certain populations are associated with greater head motion and if they co-vary with brain morphometric measures. This issue is fundamental for structural analyses: on the technical level, identifying participants, sessions, or groups for which there was strong motion during structural acquisition is essential for improving the quality of structural data. On the theoretical level, being able to regress out effects of head motion from univariate and multivariate analyses of structural data will undoubtedly result in more valid conclusions about the meaningful (vs. nuisance) differences in participants' brain anatomy. This issue is becoming of central importance, for example, in the field of structural-networks where conclusions about group differences are made based on differences in brain anatomy covariance structures (18,19). If such morphometric differences are shown to be linked to magnitude of head motion (or alternatively, to be independent of it) then this directly impacts the validity of the conclusions.

In this study, we evaluated AES and ENT as retrospective head motion metrics in a multicenter study including two clinical populations with expected different natural head

FOR PEER REVIEW ONLY

motion characteristics (20): cognitively normal Parkinson's Disease (PD) patients and Healthy Controls (HC). Importantly, the PD group had clinical tremor scores (TS). The main goals of the study were the following: i) test if AES or ENT can be used to differentiate HC and PD groups; ii) test if the clinical TS in PD correlate with the AES or ENT metrics; and iii) investigate the sensitivity of commonly used morphometric measures to these motion artifacts (we examined Freesurfer derived CT and gray-white matter contrast as well as SPM derived gray matter density maps and their structural covariance networks).

MATERIALS AND METHODS

Participants

This study's datasets were obtained from the Parkinson's Progressions Marker Initiative website (http://www.ppmi-info.org/). Written informed consent was obtained from all subjects. The study was approved by Institutional Review Boards/Independent Ethics Committees.

In order to make our analysis sensitive to head motion effects while minimizing potentially unrelated confounds for brain morphometry, we analyzed MRI and clinical data acquired on 120 PD subjects complying with the following three criteria: 1) diagnosis of resting tremor and 2) cognitive state classified as normal at the baseline visit; 3) T1-weighted MR images acquired during baseline visit at identical magnetic field and MRI system (3T Siemens Trio) using the same image acquisition sequence (3D-MPRAGE). We included in this study also data from 83 HC, who were selected to match the PD group on age and gender (see independent sample T-test and a Chi-square test in Table 1) and fulfilled the MR scan requirements described above for PD. Supplemental Table 1 reports the list of subject IDs included in the study.

FOR PEER REVIEW ONLY

MRI acquisition protocol

All participants included in this study underwent a structural MRI scan on a 3T Siemens Trio scanner (Siemens Medical Solutions, Erlangen, Germany). 3D-MPRAGE T1-weighted images were acquired at 13 different sites using three protocols with identical acquisition plane (sagittal), TI=900 ms, FA=9° and minor differences in other acquisition parameters (Table 2).

Clinical tremor scores

Following Jankovic et al. (21) we calculated a clinical tremor score (TS) for each PD patient by averaging 11 scores from the part II and III of the Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) reported in Table 3. Each score assesses the tremor at rest or during movement of different parts of the body with a 0 (no tremor) to 4 (severe tremor) rating system evaluated by an expert observer (with the exception of the MDS-UPDRS II patient's "self-assessment of tremor" score). This aggregated clinical TS for each PD participant was used as an independent clinical marker of head motion magnitude for his/her structural MRI scan.

Retrospective head motion estimates

Motion during acquisition of brain structural MRI causes blurring and aliasing (i.e. ghosting), which, in turn, can spread the image energy from one pixel to multiple pixels. Therefore, MRI measures of image contrast across edges (AES) and entropy (ENT) have been proposed as retrospective estimates of motion (17). Here we used the same definitions after normalizing the image intensity between the 5th and 95th percentile in order to remove potential biases related to intensity differences from the multisite data.

FOR PEER REVIEW ONLY

Average Edge Strength calculation

Average Edge Strength (AES) is a metric defined on 2D images. Because artifacts are more visible along the phase encoding directions (sagittal and coronal in our case) we calculated AES along the axial slices (22). We first performed skull stripping using the optiBET in FSL 5.0 (23). Then for each brain slice we calculated a binary mask of edges using the Canny edge detector. This algorithm first applies a Gaussian smoothing (kernel size=$\sqrt{2}$) on the original 2D image to obtain results robust against noise. Subsequently it computes the 2D gradient image on which one low and one high threshold are automatically determined. All the pixels above the high threshold are included in the edge mask and the pixels with gradient values between the low and the high threshold are added to the mask if connected to pixels above the high threshold. Then for each slice we calculated AES as follows (17):

$$AES\,(k) = \frac{\sqrt{\sum_{i,j} E\left(I_{i,j}^{k}\right)\left[\left(G_x\left(I_{i,j}^{k}\right)\right)^2 + \left(G_y\left(I_{i,j}^{k}\right)\right)^2\right]}}{\sum_{i,j} E\left(I_{i,j}^{k}\right)}$$

where k: slice number , *i,j*: pixel index; *E(I$_{i,j}$)* is the binary mask of Edges; G_x and G_y are gradient kernels of 1 pixel size in each direction ([−1 −1 −1; 0 0 0; 1 1 1] and [−1 0 1; −1 0 1; −1 0 1] for G_x and G_y respectively). AES provides a quantification of the average blurring at the edges detected on a 2D image: when blurring increases, for example as a result of increased head motion, AES decreases.

Entropy calculation

Image entropy (ENT) was defined as follows (17):

FOR PEER REVIEW ONLY

$$ENT = -\sum_{\rho=1}^{n} \frac{I_\rho}{I_{tot}} ln\left(\frac{I_\rho}{I_{tot}}\right)$$

where ρ is the image voxel index, n is the number of voxels in the image, Iρ is image intensity in pixel ρ. I_{tot} is the total image energy:

$$I_{tot} = \sqrt{\sum_{\rho=1}^{n} I_\rho^2}$$

ENT is independent of the image size, and we therefore derived a single value for the entire 3D image. ENT is minimum (i.e. equal to 0) when all the image energy is concentrated in one voxel. When the image energy is uniformly distributed, i.e. all voxels have the same gray scale value, its entropy is maximal and equal to:

$$Emax = \frac{1}{2}\sqrt{n_\rho}\,\ln(n_\rho)$$

Increased head motion will increase image entropy as described above. Thus, lower entropy implies fewer motion artifacts. Because ENT is sensitive to head motion artifacts extending outside the brain, such as ghosting and ringing, we calculated it on the whole image before skull stripping.

Retrospective head motion metrics: group and clinical TS effects

We tested the hypothesis that T1-weighted images are more affected by head motion related image artifacts in PD than HC by comparing AES and ENT between the two groups. AES values were not normally distributed (across all the slices and all the participants; Lilliefors test, K=0.351, p<001 for HC and K=0.304, p<=0.001 for PD). We therefore compared the distribution of AES values between HC and PD using the non-parametric Mann-Whitney test.

FOR PEER REVIEW ONLY

For ENT (a single value per participant) after verifying that it was normally distributed for HC and PD we compared the two groups means using an independent samples T-test.

We then assessed whether AES and ENT can be used as proxy of head motion quantification by measuring their Spearman's correlation with clinical TS (in PD patients only). Since AES is calculated separately for each image slice, it was necessary to select one AES value to represent the whole brain volume of each subject and its corresponding clinical TS. The mean AES across slices was discarded because AES did not show a normal distribution. Instead, since AES was found to differ significantly between HC and PD when considering all slices (Results), we determined the AES percentile for each subject by computing the empirical cumulative distribution function (eCDF) of AES, separately for HC and PD. From these, the percentile AES value that better discriminated the eCDFs of the two groups was chosen (Supplemental Figure 1). We found that the AES eCDF diverged for the two groups (HC and PD) above the 90th percentile. This value was therefore chosen as a representative AES for each participant. A comparison it between HC and PD using a Mann-Whitney test confirmed, that using the 90^{th} percentile, AES gave a statistically significant lower AES for PD than HC ($z=-2.06$, $p=0.02$, $dof=201$, Supplemental Figure 2). We therefore used the 90^{th} percentile AES value of each subject to compute the correlation between AES and clinical TS in PD. All statistical tests were performed in Matlab 8.1.0.

Brain morphometry with Freesurfer and VBM

For brain morphometry estimates from the structural MRI of each subject we used two software tools: Freesurer (5.1.0) for CT and Gray to White Matter (GWC) measurements and SPM12 for Gray Matter density (GMx) maps (24-26).

For each participant we obtained the CT and GWC vertex-wise values on the FreeSurfer average template and their average in each ROIs of the Desikan cortical

FOR PEER REVIEW ONLY

parcellation atlas (27). GWC was calculated at each vertex along the cortical surface as [(white − gray)/(white + gray)] at 0.5 mm above versus below the gray/white interface with trilinear interpolation of the images (24). GWC values were between 0 and 1, with values closer to 0 indicating less contrast and thus more blurring of the gray/white boundary. For subsequent group analyses we registered both CT and GWC maps to an average template after smoothing (with 15 mm FWHM and 10 mm FWHM respectively for CT and GWC) across the surface.

GMx values were obtained voxel-wise on the Dartel template and their average for each cortical ROIs of the AAL atlas (28). For GMx maps computation we corrected the raw images for bias-field inhomogeneity and segmented into gray matter, white matter and cerebrospinal fluid using standard default settings (2). Then we aligned the GMx maps to common space using a high-dimensional DARTEL normalization modulating for nonlinear effects and smoothed using a 9-mm FWHM Gaussian smoothing kernel.

Brain morphometry: head motion effects

The aim of this analysis was to investigate the effect of head motion as retrospectively measured by AES on CT, GWC and GMx. The 90th percentile AES value was chosen as representative for each subject, as done in the correlation analysis with clinical TS. We will simply refer to this value as AES for the rest of morphometry analysis description.

We did not evaluate the effects of ENT on brain morphometry metrics because ENT did not correlate with clinical TS in PD patients and the ENT distributions for HC and PD were not significantly different (Results).

Univariate Whole-Brain Analysis

FOR PEER REVIEW ONLY

We identified cortical regions where AES correlated with CT (or GWC) and where the correlation was different for HC and PD. We used the following linear regression model, identical for CT and GWC (here shown for CT), on each vertex of the reconstructed cortical surface:

$$CT = \beta 0 + \beta 1 Group + \beta 2 Age + \beta 3 AES + \beta 4 (Group * AES) + \varepsilon$$

The model attempts to explain the group variance of CT (or GWC) by assigning weights to a linear combination of multiple explanatory terms: subject group (HC or PD), age, head motion (AES) and an interaction between head motion and group. The error ε denotes the unexplained variance. The regressors of interest were β3 (AES correlation slope) and β4 (group-by-AES interaction, or the slope difference between the groups). We obtained vertex-wise regressors with their p-value, applied a single vertex $p<0.01$ threshold and performed cluster-wise correction for multiple comparison on the surface at the $p<0.05$ level (29).

For GMx maps the same linear regression model as above was applied in SPM12 adding as a covariate the total GM volume to control for individual differences (30). Voxels with GM density < 0.2 were excluded from the regression analysis. We thresholded the regressors of interest (β3 and β4) at $p<0.05$ level with Family Wise Error correction for multiple comparisons.

Multivariate analysis

Univariate analyses may not detect subtle effects at a distributed topographic-level (31). For this reason, we investigated the hypothesis that a network-level covariance pattern of CT, GWC and GMx predicts AES using a Partial Least Squares (PLS) approach. Specifically, we used the CT (and GWC) average values in the 68 cortical regions of the

FOR PEER REVIEW ONLY

Desikan Atlas to generate an *N* (number of participants either in HC or PD group) x 68 (number of cortical ROIs) matrix representing the explanatory data. We implemented PLS using the package available in *R* software (32).

We first determined, using a leave-one-out validation procedure, that the first component of the explanatory data predicted most of the covariance between CT (GWC) and AES. To this aim, we constructed 500 permutations of the explanatory data keeping the original CT matrix while randomly assigning AES to the participants in each permutation and verified that the prediction error of AES by CT was significantly below what would be expected by chance ($p<0.001$). Subsequently, to identify which brain regions reliably predicted the CT (GWC) vs. AES relationship, we used a bootstrapping procedure that was run 100 times (33). In each instance, we bootstrapped, with replacement, rows from the original [$N \times 68$] CT (GWC)-value matrix, to populate a proxy [$N \times 68$] matrix. For each proxy matrix, PLS was run, and the loadings for the estimated first component was retrieved. This loading matrix was rotated to match the direction of the loadings in the original data via a Procrustes Rotation, and the Y-loadings (AES loadings) for the first component saved. Finally, we calculated the standard deviation of the 100 bootstrap loadings, per region, and then obtained a Z-score per region [region loading/sd(loading)]. Only regions that passed a Z-score of ± 3.0 were considered significantly ($p<0.001$) "salient."

We performed identical analysis to investigate the prediction of AES by GMx. We obtained the explanatory data matrix for GMx calculating average GMx values in 82 cortical regions of the AAL atlas. We run the PLS analysis described above for CT, GWC and GMx separately for HC and PD.

FOR PEER REVIEW ONLY

RESULTS

We used visual inspection of representative data to qualitatively evaluate the relation between AES and ENT scores in images with different degree of visually recognized motion artifacts. Figure 1 shows the sensitivity of AES and its inverse relation to image blurring (Figure 1a, left HC with lower blurring, right PD with higher blurring) and the sensitivity of ENT to motion artifacts like ringing (Figure 1b, left HC with no artifacts and right PD with ringing artifact).

Retrospective MRI-derived head motion metrics: sensitivity to clinical differences

The group analysis demonstrated a subtle yet statistically significant reduction of AES (Mann Whitney $z=-1.72$, $p=0.04$, dof=28541), measures (slice by slice) in PD relative to HC. This result indicates increased head motion in the PD group as compared to the HC group: the reduced edge intensity contrast (i.e. AES) is consistent with higher level of image blurring at the edges, a common motion related artifact.

To evaluate correlations between AES and TS scores of individuals in the PD group, it was necessary to define a single AES per participant. AES distributions were found to be non-normal and skewed in both groups (PD: Lilliefors Test $K=0.351$, $p<0.001$; HC: Lilliefors Test $K=0.304$, $p<0.001$). Therefore, a percentile-determined definition was preferred as a single-subject AES summary. We chose the 90th percentile because the main differences in AES eCDF emerged above this threshold (Supplemental Figure 1) and we verified that this value was statistically significantly different between HC and PD (Supplemental Figure 2). The Spearman correlation between the AES 90th percentile and clinical TS in PD was statistically significant ($rho=-0.21$, $p=0.02$, dof=118, Figure 2), consistent with the hypothesis that AES is sensitive to motion, and that PD patients with a higher clinical TS were more prone to head

FOR PEER REVIEW ONLY

motion in the scanner. We note that the main purpose of this analysis was to identify an AES percentile that tracks a tremor score in order to then evaluate the relationship between these AES values and morphometric data.

As opposed to AES, ENT was not statistically different between HC and PD (t=0.72, p=0.47, dof=201). In addition, there was no significant correlation between ENT and TS of PD patients (rho=-0.12, p=0.19, dof=118). For this reason only the AES 90th percentile (to which we will simply refer as AES) was considered in the morphometry analyses.

Retrospective MRI-derived head motion metrics: effects on brain morphometry

The results of the whole-brain univariate analysis are shown in Figure 3. For the relationship between AES and CT, we found a negative association in the right superior parietal gyrus, supramarginal gyrus, lateral occipital gyri and the left paracentral gyrus and a positive association the right fusiform gyrus. We did not find any region where the relationship between AES and CT was significantly different for HC and PD (Group-by-AES interaction). There were no regions where the relationship between AES and GWC or between AES and GMx was statistically significant. We did not find any region where the interaction between AES and group was statistically significant for either GWC or GMx.

The multivariate PLS analysis identified, fort the PD group, a set of 54 regions, out of 68, of the Desikan atlas (Figure 4), where the loadings for the first CT component predicted AES with strong salience ($|z| > 3$, $p < 0.001$). This set of ROIs included all the regions, except the right fusiform gyrus, where significant association between AES and CT was found by the univariate analysis and also extended to a vast portion of the cortical surface.

A comparable analysis revealed a set of regions where GWC covariance predicted AES in the PD population. These included the ones found for CT in addition to 9 others (Figure 5).

FOR PEER REVIEW ONLY

As opposed to the above-presented findings for PD, for HC the covariance of CT or GWC did not predict AES. Finally we did not find any network-level covariance pattern for GM that predicted AES either in PD or HC.

DISCUSSION

The main findings of this study are: i) AES (and not ENT) gave subtle but significant differences between the PD and HC groups, consistently with higher head motion in the PD group; ii) AES (and not ENT) correlated negatively with clinical tremor scores, consistently associating higher head motion with higher tremor scores; and iii) in PD, AES revealed patterns of CT and GWC covariance across a large portions of the cortical surface.

The lower AES values (more image blurring) found in PD patients are consistent with previous results showing that AES is sensitive to head motion during structural MRI (17,34). Indeed, the PD patients in the current study who were not taking medication suffered from tremor or dystonia, a condition that prevents them from lying absolutely still in the scanner (20). While it is generally accepted that PD patients do not have head tremor, they experience head movement as a result of trunk or limb tremor transmitted to the head (35). This explains the higher blurring (i.e. lower AES) measured on T1-weighted MRI in the PD patients relative to the HC group. Previous studies using prospective motion correction tools on structural MR images demonstrated the sensitivity of AES to intra-subject head motion variability because AES decreased after motion correction (17,34). Here we extend those findings showing that AES is also sensitive for detecting natural head motion group differences between PD and age and gender matched HCs. This result suggests that AES could be used to test the level of head-motion differences across populations in cross-sectional brain morphometry studies.

FOR PEER REVIEW ONLY

We found a significant negative correlation between AES and the clinical TS in PD, with an effect size rho=-0.21. This finding is consistent with a previous study that showed how individual differences in impulsivity score predict head motion retrospectively measured on resting state fMRI data (36). In our study we show, for the first time, an association between a clinical motion metric and an image based retrospective estimation of head motion for structural MR images. This finding thus suggests that AES is a reliable QC measure to detect inter-individual differences in head motion. However, the small effect size in correlation with TS reported in this study suggests that these findings need to be replicated using different datasets before AES can be widely adopted as a QC metric for head motion estimation on structural images.

Image entropy did not demonstrate as much head motion sensitivity as AES. The higher sensitivity of AES relative to entropy may be related to the fact that AES is calculated on 2D slices and we chose the slice orientation most sensitive to motion from our 3D images (axial slices, which had two phase encoding directions) (22). Entropy was calculated from the whole 3D volume potentially making the measure less sensible to subtle motion.

Investigating the biases of AES as proxy of head motion on morphometry measures we found that AES was a predictor of CT in five regions of the brain. This result marginally confirms previous studies that have demonstrated a strong association between CT and GMx measures and the amount of head motion in larger portions of the cerebral cortex (6,12). The weaker linear relationship between morphometry measures and head motion found in our dataset may be due to study design: the above mentioned studies were conducted with the participants guided to perform substantial head motion during image acquisition whereas the images analyzed in this study were acquired on subjects instructed to stay still as much as they could.

FOR PEER REVIEW ONLY

We found that despite the head motion differences between the PD and HC groups, no brain areas showed significant interactions between head motion and morphometry data using the univariate analyses. This suggests that the morphometry tools used are overall robust in comparing morphometry measures between populations with small but statistically significant different head motion characteristics ($z=-1.72$, $p=0.04$ in our case). However future studies performed on different datasets could investigate the sensitivity of AES to different types of motion (e.g. nodding and free) and quantify the relationship between measured head motion and AES. This may help determine the minimum AES necessary to induce morphometry group differences caused by image artefacts.

The results of the multivariate analysis indicate that in a clinical population with expected motion degraded MR image quality, a brain distributed covariance pattern for both CT and GWC can emerge that is related to head motion. This is in line with previous literature findings that have demonstrated that although multivariate connectivity methods can detect connections which cannot be observed in standard univariate analysis, they are more sensitive to voxel specific noise (37). Specifically, a study by Geerligs et al. showed association with increased head motion for a multivariate (Distance correlation) and not for a univariate (Pearson's correlation) measure of functional connectivity (38). In our study we extend those findings demonstrating also the strong sensitivity of multivariate structural connectivity analyses to head motion.

Our results thus underline the need for using QC tools to verify the reliability of structural covariance measurements and indicate that AES could be used retrospectively for this purpose. This is potentially relevant for the neuroimaging community because an increasing number of studies are looking at large-scale structural networks effects rather than local morphometric measures to investigate potential anatomical imaging biomarkers for PD (18) and other brain pathologies (19).

FOR PEER REVIEW ONLY

This work has a number of limitations: first of all we compared two retrospective head motion metrics, AES and ENT because they have been defined to be sensitive to two very commonly found head motion related image artifacts (i.e. blurring and ringing). However further image based metrics exist in the literature, such as the gray level co-occurrence matrix, that could be tested in future studies alone or in combination with AES to the same aim of discerning motion from disease due structural differences between different clinical populations (34). Second, we could have included in our analysis an additional morphometry measure, the Brain Boundary Shift Integral (BBSI), that computes volume changes due to the shifting of boundaries between brain tissue and cerebral spinal fluid over time (39). By its definition BBSI should be more sensitive than CT, GWC or GM to head motion because it directly maps voxel intensity changes over time, i.e., repeated measures. However, the first longitudinal structural images available in the PPMI database were performed 12 months after the baseline scan, at a time when disease evolution or Parkinson's drug effects may have caused cortical changes (40). For this reason we did not use the longitudinal PD data available in the database, which will make difficult to disentangle head motion effects related to disease progression and/or with treatment effects. Third, we correlated AES with the clinical TS that provides an overall assessment of tremor in different parts of the body. It would have been beneficial to compare AES with a more specific behavioral measure of head motion, such as the lip/jaw TS that is one of the MDS-UPDRS part III ratings used in the calculation of the TS. However, head tremor is in general absent in PD patients (35) and so it was in our study where only 7 patients in the database we analyzed had a lip/jaw TS different than 0. Future studies in patients with Parkinson's disease or different clinical populations correlating AES with a more specific or possibly direct measure of head motion should be performed to confirm its specificity for quantifying head motion. Nevertheless, the TS showed better

FOR PEER REVIEW ONLY

correlation with AES than the MDS-UPDRS Part III score (results not shown) because the ratings used to calculate the TS mainly pertain to postural and rest tremor. The MDS-UPDRS III score instead consists of a more global evaluation of the motor signs of PD including features like patient's flowing speech, rising from the chair or finger tapping. Therefore, the TS is more likely to be indicative of the tremor scenario during the MR scan.

In conclusion the results of this work suggest that: i) AES may provide a reliable metric to retrospectively and quantitatively estimate head motion occurred during the acquisition of T1-weighted structural images, thus offering an objective metric for normalizing head motion across populations; ii) AES can be used as a head-motion related QC metric for structural MRI data studies investigating large-scale structural covariance effects. This evaluation will be particularly important when head motion may not be easy to match across clinical groups.

FOR PEER REVIEW ONLY

References

1. Bellon EM, Haacke EM, Coleman PE, Sacco DC, Steiger DA, Gangarosa RE. MR artifacts: a review. AJR Am J Roentgenol 1986;147(6):1271-1281.
2. Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage 2007;38(1):95-113.
3. Douaud G, Smith S, Jenkinson M, et al. Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. Brain 2007;130(Pt 9):2375-2386.
4. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci U S A 2000;97(20):11050-11055.
5. Ducharme S, Albaugh MD, Nguyen TV, et al. Trajectories of cortical thickness maturation in normal brain development--The importance of quality control procedures. Neuroimage 2016;125:267-279.
6. Reuter M, Tisdall MD, Qureshi A, Buckner RL, van der Kouwe AJ, Fischl B. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. Neuroimage 2015;107:107-115.
7. Watanabe K, Kakeda S, Igata N, et al. Utility of real-time prospective motion correction (PROMO) on 3D T1-weighted imaging in automated brain structure measurements. Sci Rep 2016;6:38366.
8. Prakkamakul S, Witzel T, Huang S, et al. Ultrafast Brain MRI: Clinical Deployment and Comparison to Conventional Brain MRI at 3T. Journal of neuroimaging : official journal of the American Society of Neuroimaging 2016;26(5):503-510.
9. Pipe JG. An optimized center-out k-space trajectory for multishot MRI: comparison with spiral and projection reconstruction. Magn Reson Med 1999;42(4):714-720.
10. Tisdall MD, Hess AT, Reuter M, Meintjes EM, Fischl B, van der Kouwe AJ. Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. Magn Reson Med 2012;68(2):389-399.
11. Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: A complex problem with many partial solutions. Journal of magnetic resonance imaging : JMRI 2015;42(4):887-901.
12. Tisdall MD, Reuter M, Qureshi A, Buckner RL, Fischl B, van der Kouwe AJW. Prospective motion correction with volumetric navigators (vNavs) reduces the bias and variance in brain morphometry induced by subject motion. Neuroimage 2016;127:11-22.
13. Pardoe HR, Kucharsky Hiess R, Kuzniecky R. Motion and morphometry in clinical and nonclinical populations. Neuroimage 2016;135:177-185.
14. Shaw P, Eckstrand K, Sharp W, et al. Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. Proc Natl Acad Sci U S A 2007;104(49):19649-19654.
15. Alexander-Bloch A, Clasen L, Stockman M, et al. Subtle in-scanner motion biases automated measurement of brain anatomy from in vivo MRI. Hum Brain Mapp 2016;37(7):2385-2397.
16. Savalia NK, Agres PF, Chan MY, Feczko EJ, Kennedy KM, Wig GS. Motion-related artifacts in structural brain images revealed with independent estimates of in-scanner head motion. Hum Brain Mapp 2017;38(1):472-492.
17. Aksoy M, Forman C, Straka M, Cukur T, Hornegger J, Bammer R. Hybrid prospective and retrospective head motion correction to mitigate cross-calibration errors. Magn Reson Med 2012;67(5):1237-1251.
18. Chou KH, Lin WC, Lee PL, et al. Structural covariance networks of striatum subdivision in patients with Parkinson's disease. Hum Brain Mapp 2015;36(4):1567-1584.
19. Hasson U, Andric M, Atilgan H, Collignon O. Congenital blindness is associated with large-scale reorganization of anatomical networks. Neuroimage 2016;128:362-372.
20. Hacker CD, Perlmutter JS, Criswell SR, Ances BM, Snyder AZ. Resting state functional connectivity of the striatum in Parkinson's disease. Brain 2012;135(Pt 12):3699-3711.

FOR PEER REVIEW ONLY

21. Jankovic J, McDermott M, Carter J, et al. Variable expression of Parkinson's disease: a baseline analysis of the DATATOP cohort. The Parkinson Study Group. Neurology 1990;40(10):1529-1534.
22. Malamateniou C, Malik SJ, Counsell SJ, et al. Motion-compensation techniques in neonatal and fetal MR imaging. AJNR American journal of neuroradiology 2013;34(6):1124-1136.
23. Lutkenhoff ES, Rosenberg M, Chiang J, et al. Optimized brain extraction for pathological brains (optiBET). PLoS One 2014;9(12):e115551.
24. Blackmon K, Halgren E, Barr WB, et al. Individual differences in verbal abilities associated with regional blurring of the left gray and white matter boundary. J Neurosci 2011;31(43):15257-15263.
25. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage 1999;9(2):195-207.
26. Fischl B, Sereno MI, Tootell RB, Dale AM. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum Brain Mapp 1999;8(4):272-284.
27. Desikan RS, Segonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006;31(3):968-980.
28. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 2002;15(1):273-289.
29. Hagler DJ, Jr., Saygin AP, Sereno MI. Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. Neuroimage 2006;33(4):1093-1103.
30. Peelle JE, Cusack R, Henson RN. Adjusting for global effects in voxel-based morphometry: gray matter decline in normal aging. Neuroimage 2012;60(2):1503-1516.
31. Bergfield KL, Hanson KD, Chen K, et al. Age-related networks of regional covariance in MRI gray matter: reproducible multivariate patterns in healthy aging. Neuroimage 2010;49(2):1750-1759.
32. Mevik BH, Wehrens, R. The PLS package: principal component and partial least squares regression in R. J Stat Softw 2007;18(2):24.
33. McIntosh AR, Lobaugh NJ. Partial least squares analysis of neuroimaging data: applications and advances. Neuroimage 2004;23 Suppl 1:S250-263.
34. Pannetier NA, Stavrinos T, Ng P, et al. Quantitative framework for prospective motion correction evaluation. Magn Reson Med 2016;75(2):810-816.
35. Roze E, Coelho-Braga MC, Gayraud D, et al. Head tremor in Parkinson's disease. Mov Disord 2006;21(8):1245-1248.
36. Kong XZ, Zhen Z, Li X, et al. Individual differences in impulsivity predict head motion during magnetic resonance imaging. PLoS One 2014;9(8):e104989.
37. Anzellotti S, Caramazza, A., Saxe, R. Multivariate pattern connectivity. bioRxiv 046151 2016.
38. Geerligs L, Cam C, Henson RN. Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation. Neuroimage 2016;135:16-31.
39. Freeborough PA, Fox NC. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. IEEE Trans Med Imaging 1997;16(5):623-629.
40. Jia X, Liang P, Li Y, Shi L, Wang D, Li K. Longitudinal Study of Gray Matter Changes in Parkinson Disease. AJNR American journal of neuroradiology 2015;36(12):2219-2226.

FOR PEER REVIEW ONLY

Table 1: Summary demographic information of the HC and PD patients included in this study. The statistical scores in the bottom row indicate that the groups were matched for age and gender.

Groups	**Age** (years)	**Gender**
HC: n=83	59.6 ± 10.9	M 53/F 30
PD: n=120	60.4 ± 9.8	M 75/F 45
Group difference statistics	T=-0.44, p=0.66	X^2=0.10, p=0.75

M: male. F: female. HC: Healthy controls. PD: Parkinson's disease subjects.

FOR PEER REVIEW ONLY

Table 2: Main structural brain MRI acquisition parameters for the 3 PPMI protocols used to acquire the 3D T1-weighted images analyzed in this study.

Protocol name	TR (ms)	TE (ms)	FOV (mm)	Voxel size (mm^3)	Sites per protocol	Samples per protocol
MPRAGE GRAPPA	2300	2.98	240x256x176	1.0x1.0x1.0	12	173 (HC=66,PD=107)
MPRAGE T1 SAG	2300	2.52	256x256x176	0.98x0.98x1.0	1*	10 (HC=5,PD=5)
SAG T1 3D MPRAGE	1900	2.27	256x256x176	0.98x0.98x1.0	1	20 (HC=12,PD=8)

PPMI: Parkinson's Progressions Marker Initiative. HC: Healthy controls. PD: Parkinson's disease subjects.

*In site #32, according to the PPMI database, 15 datasets were acquired with the MPRAGE GRAPPA protocol and 10 with the MPRAGE T1 SAG protocol

FOR PEER REVIEW ONLY

Table 3: Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) evaluations (parts II and III) used to calculate the clinical tremor score for PD patients.

Index	MDS-UPDRS Part	Acronym	Name
1	II	NP2TRMR	Self-assessment of tremor
2	III	NP3PTRMR	Postural tremor of right hand
3	III	NP3PTRML	Postural tremor of left hand
4	III	NP3KTRMR	Kinetic tremor of right hand
5	III	NP3KTRML	Kinetic tremor of left hand
6	III	NP3RTARU	Rest tremor amplitude of right upper limbs
7	III	NP3RTALU	Rest tremor amplitude of left upper limbs
8	III	NP3RTARL	Rest tremor amplitude of right lower limbs
9	III	NP3RTALL	Rest tremor amplitude of left lower limbs
10	III	NP3RTALJ	Rest tremor amplitude of lip/jaw
11	III	NP3RTCON	Constancy of rest tremor

PD: Parkinson’s disease subjects.

FOR PEER REVIEW ONLY

Figure Legends

Figure 1: Head motion artifact examples in healthy controls (HC) and Parkinson's disease (PD) subjects with the corresponding quantitative characterization given by the average edge strength (AES, upper panel) and image entropy (ENT, lower panel) metrics derived from the corresponding structural T1-weighted images. Figure 1a (upper panel) shows an example where the head motion differences are quantitatively dominated by lower AES (32% lower due to higher image blurring) in the PD (right) relative to the HC (left), while both images have comparable ENT values across the two examples (5% difference). Figure 1b (lower panel) shows the converse, an example where head motion quantitative differences are dominated by higher ENT (11% higher due to more ringing artifacts) in the PD (right) relative to the HC (left), while AES was comparable across the two examples (3% difference).

Figure 2: Binned scatter plot of the average edge strength (AES, 90^{th} percentile), as quantitative head motion metric derived from structural T1-images, and clinical tremor scores (TS) in Parkinson's disease patients. To avoid overlap for participants with similar clinical TSs, participants are binned into equal sized 10 TS groups. Average AES and TS for each group are plotted. Error bars indicate standard error of the mean (s.e.m). Clinical TSs showed a significant negative correlation between AES and clinical TS, indicating increased image blurring (consistent with higher motion artifacts) with tremor severity.

Figure 3: Univariate analysis showing gray matter clusters where cortical thickness (CT) has a significant linear relationship with average edge strength (AES, $p<0.05$ corrected for multiple comparisons). Group results are shown on the inflated cortex and are not different between healthy controls and Parkinson's disease patients. Red/yellow indicates thickness

FOR PEER REVIEW ONLY

gain and blue indicates thickness loss with increased AES. R: right hemisphere. L: left hemisphere. Upper panel: lateral views. Lower panel: medial views.

Figure 4: Multivariate partial least square (PLS) covariance analysis of cortical thickness (CT) and average edge strength (AES) in Parkinson's disease (PD) The regions shown are where CT predicts AES (CT loadings showed low variance across AES bootstrap solutions ($|z|>3$, $p<0.001$). R: right hemisphere. L: left hemisphere. Upper panel: lateral views. Lower panel: medial views.

Figure 5: Multivariate partial least square (PLS) covariance analysis of gray-white contrast (GWC) and average edge strength (AES) in Parkinson's disease (PD). The regions shown are those where GWC predicts AES (GWC loadings showed low variance across AES bootstrap solutions ($|z|>3$, $p<0.001$). R: right hemisphere. L: left hemisphere. Upper panel: lateral views. Lower panel: medial views.

FOR PEER REVIEW ONLY

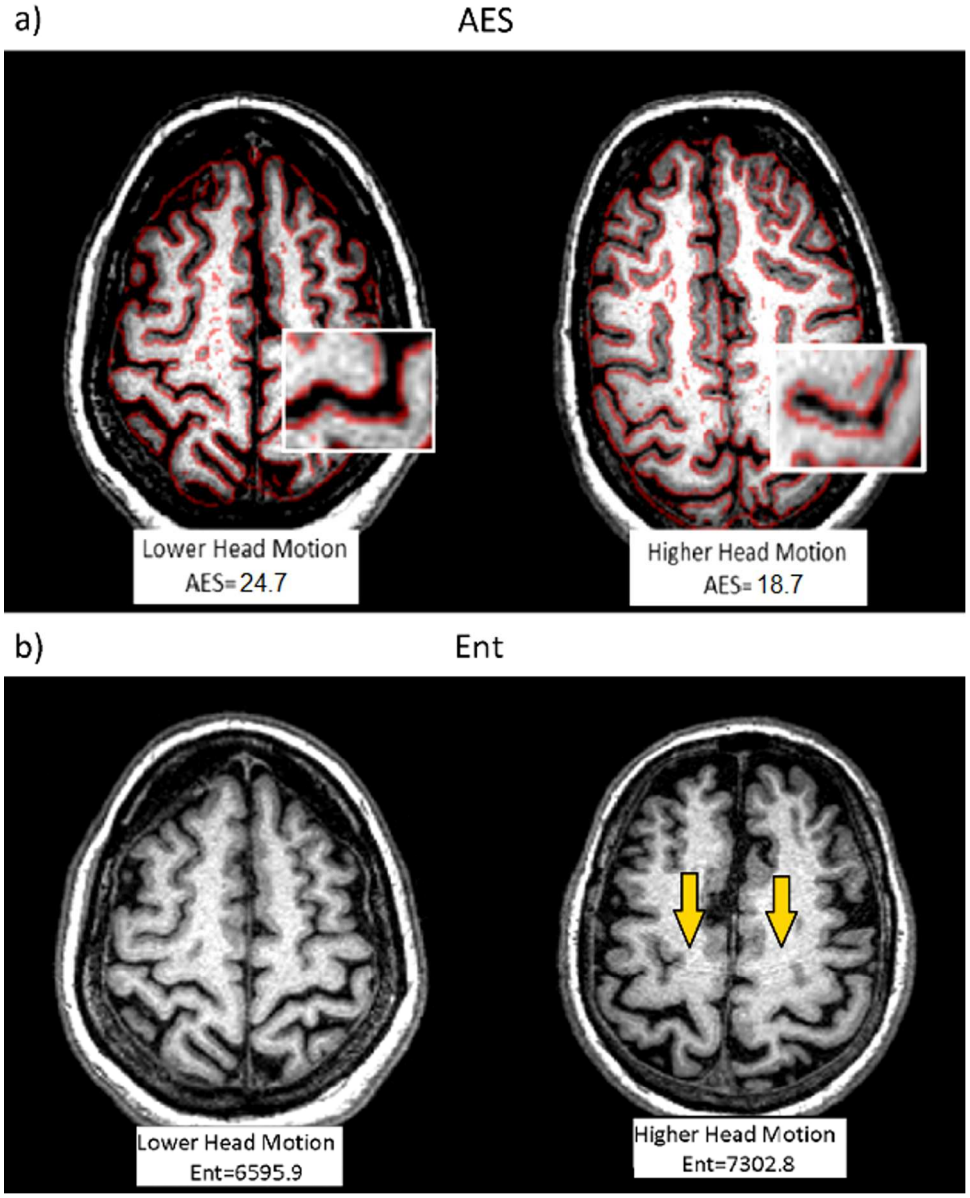


Figure 1: Head motion artifact examples in healthy controls (HC) and Parkinson's disease (PD) subjects with the corresponding quantitative characterization given by the average edge strength (AES, upper panel) and image entropy (ENT, lower panel) metrics derived from the corresponding structural T1-weighted images. Figure 1a (upper panel) shows an example where the head motion differences are quantitatively dominated by lower AES (32% lower due to higher image blurring) in the PD (right) relative to the HC (left), while both images have comparable ENT values across the two examples (5% difference). Figure 1b (lower panel) shows the converse, an example where head motion quantitative differences are dominated by higher ENT (11% higher due to more ringing artifacts) in the PD (right) relative to the HC (left), while AES was comparable across the two examples (3% difference).

190x237mm (300 x 300 DPI)

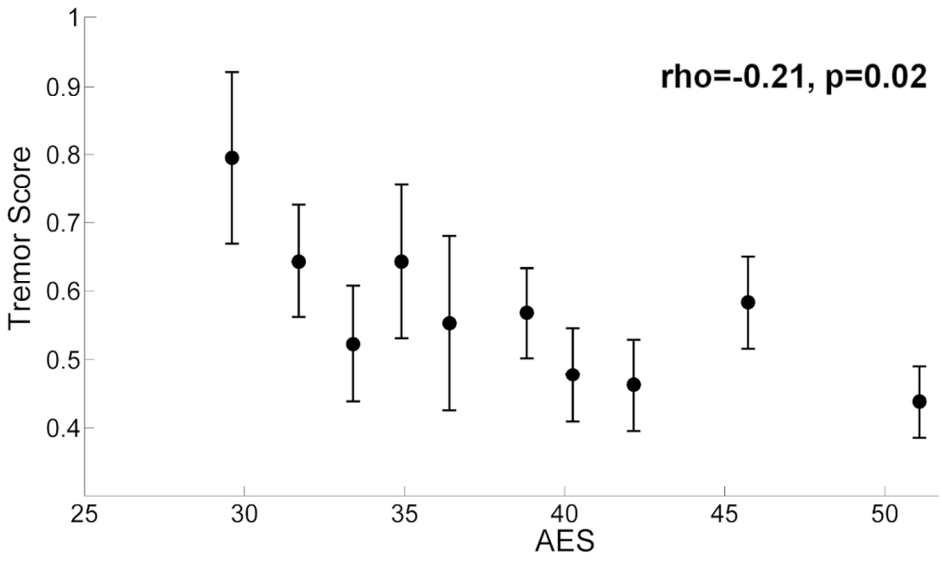


Figure 2: Binned scatter plot of the average edge strength (AES, 90th percentile), as quantitative head motion metric derived from structural T1-images, and clinical tremor scores (TS) in Parkinson's disease patients. To avoid overlap for participants with similar clinical TSs, participants are binned into equal sized 10 TS groups. Average AES and TS for each group are plotted. Error bars indicate standard error of the mean (s.e.m). Clinical TSs showed a significant negative correlation between AES and clinical TS, indicating increased image blurring (consistent with higher motion artifacts) with tremor severity.

99x55mm (300 x 300 DPI)

VIEW ONLY

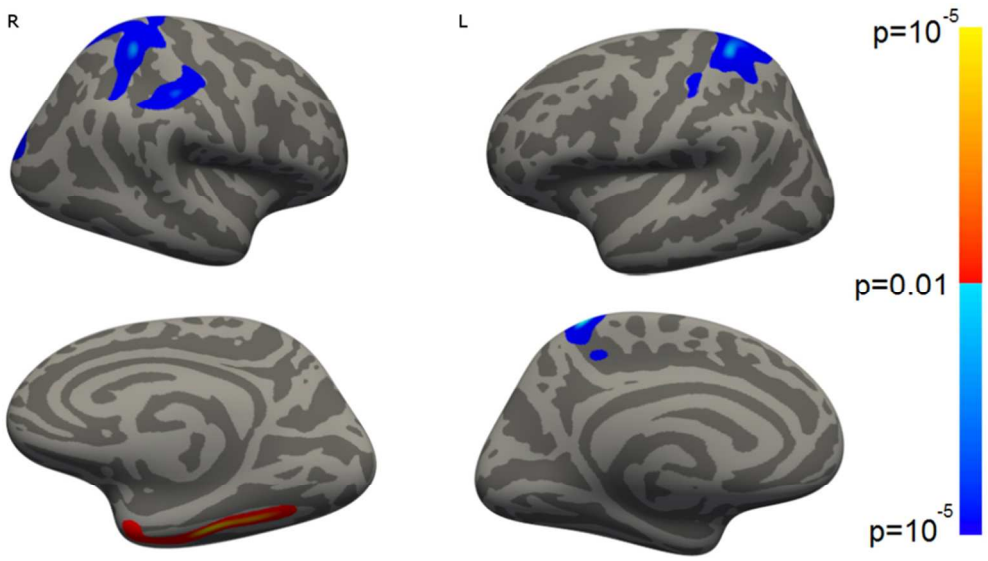


Figure 3: Univariate analysis showing gray matter clusters where cortical thickness (CT) has a significant linear relationship with average edge strength (AES, $p<0.05$ corrected for multiple comparisons). Group results are shown on the inflated cortex and are not different between healthy controls and Parkinson's disease patients. Red/yellow indicates thickness gain and blue indicates thickness loss with increased AES. R: right hemisphere. L: left hemisphere. Upper panel: lateral views. Lower panel: medial views.

87x50mm (300 x 300 DPI)

For Peer Review Only

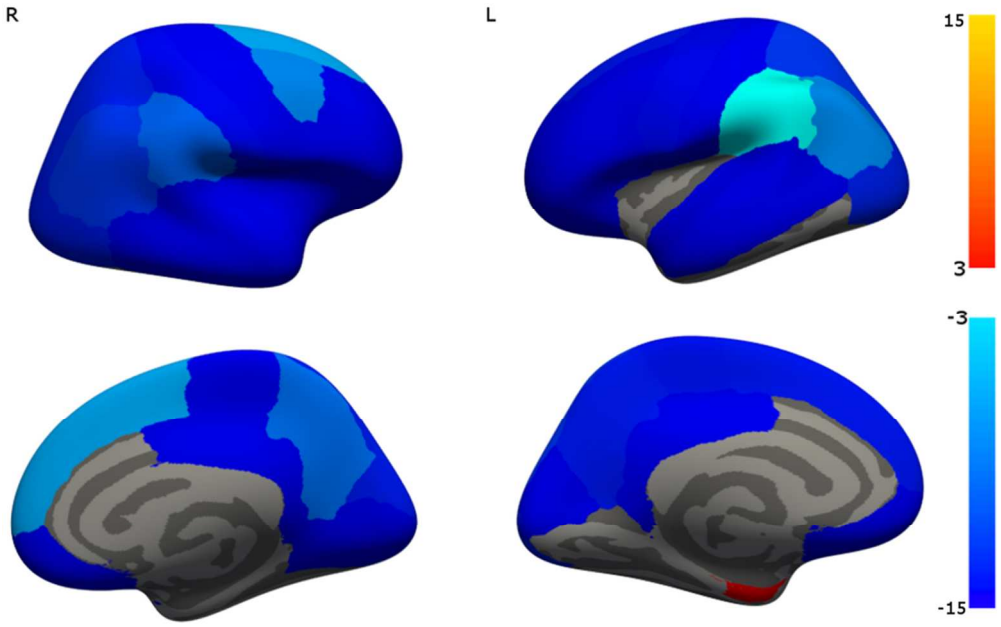


Figure 4: Multivariate partial least square (PLS) covariance analysis of cortical thickness (CT) and average edge strength (AES) in Parkinson's disease (PD) The regions shown are where CT predicts AES (CT loadings showed low variance across AES bootstrap solutions (|z|>3, p<0.001). R: right hemisphere. L: left hemisphere. Upper panel: lateral views. Lower panel: medial views.

96x60mm (300 x 300 DPI)

VIEW ONLY

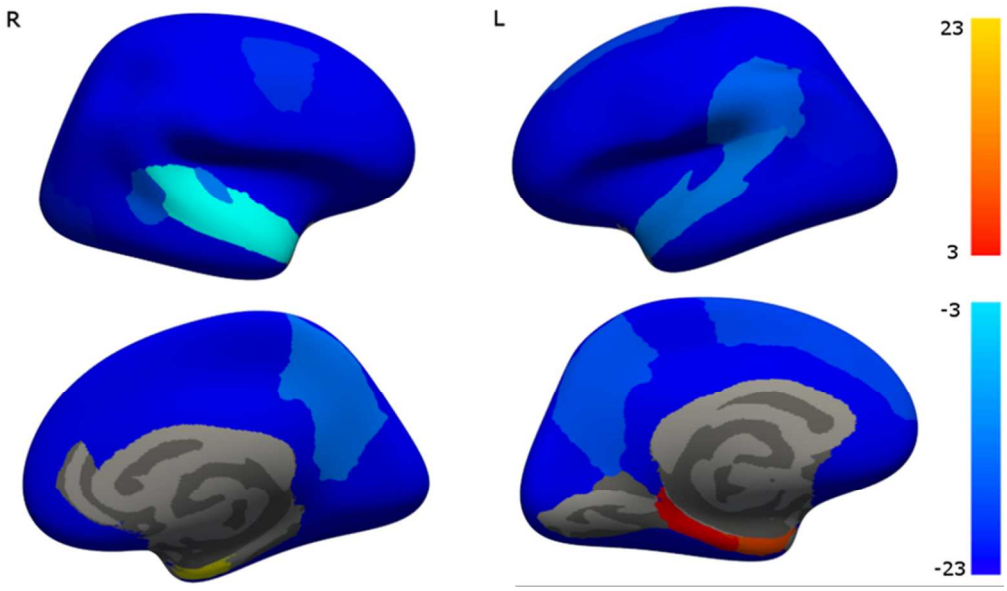


Figure 5: Multivariate partial least square (PLS) covariance analysis of gray-white contrast (GWC) and average edge strength (AES) in Parkinson's disease (PD). The regions shown are those where GWC predicts AES (GWC loadings showed low variance across AES bootstrap solutions (|z|>3, p<0.001). R: right hemisphere. L: left hemisphere. Upper panel: lateral views. Lower panel: medial views.

88x51mm (300 x 300 DPI)

For Peer Review ONLY

Supplemental Table 1: ID code and clinical group for the subjects included in the study obtained from the Parkinson's Progression Marker Initiative (PPMI) dataset: PD: Parkinson's disease. HC: healthy controls.

ID	Group
3101	PD
3104	HC
3105	PD
3106	HC
3108	PD
3111	PD
3112	HC
3113	PD
3114	HC
3115	HC
3118	PD
3119	PD
3122	PD
3123	PD
3124	PD
3125	PD
3127	PD
3128	PD
3129	PD
3130	PD
3131	PD
3132	PD
3150	PD
3151	HC
3154	PD
3156	HC
3160	HC
3161	HC
3165	HC
3168	PD
3169	HC
3170	PD
3171	HC
3172	HC
3173	PD
3175	PD
3176	PD
3181	PD
3182	PD
3188	HC
3189	PD
3190	PD

ID	Group
3191	HC
3300	HC
3301	HC
3304	PD
3305	PD
3307	PD
3309	PD
3310	HC
3311	PD
3316	HC
3320	HC
3322	PD
3324	PD
3325	PD
3326	PD
3328	PD
3332	PD
3350	HC
3353	HC
3354	PD
3355	HC
3357	HC
3358	HC
3360	PD
3361	HC
3362	HC
3364	PD
3365	PD
3367	PD
3368	HC
3369	HC
3370	HC
3372	PD
3373	PD
3375	PD
3377	PD
3380	PD
3383	PD
3386	PD
3389	HC
3390	HC
3392	PD

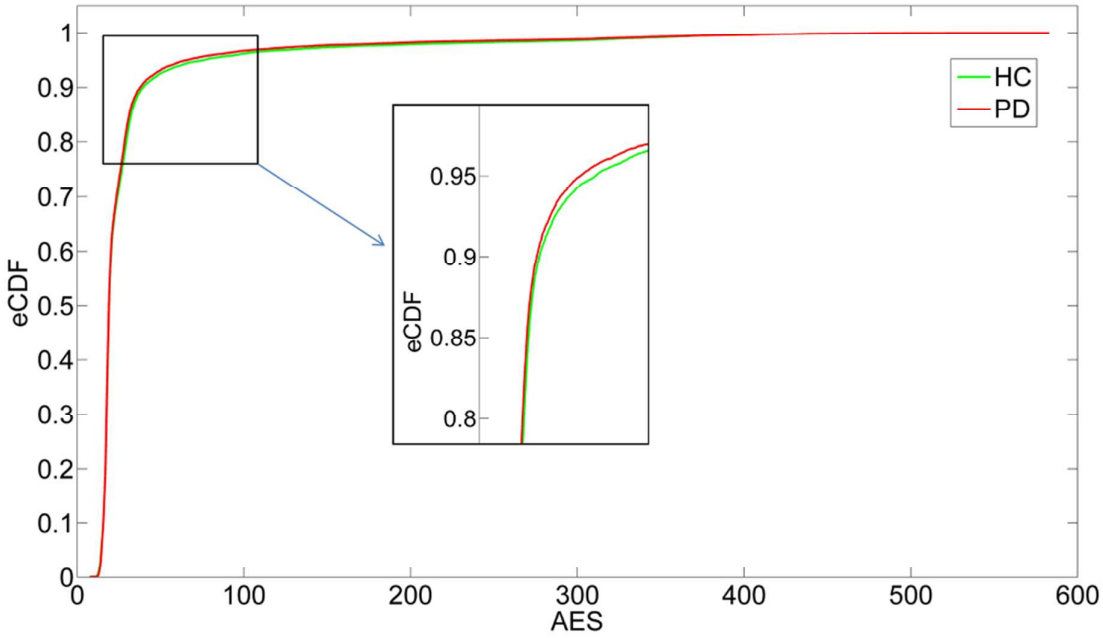
ID	Group
3550	PD
3551	HC
3554	HC
3555	HC
3556	PD
3565	HC
3567	PD
3569	HC
3570	HC
3572	HC
3573	PD
3574	PD
3575	PD
3577	PD
3580	PD
3581	PD
3582	PD
3584	PD
3585	PD
3588	PD
3591	PD
3593	PD
3600	HC
3603	PD
3604	PD
3607	PD
3608	PD
3610	HC
3611	HC
3612	PD
3613	HC
3614	HC
3615	HC
3617	PD
3620	HC
3622	PD
3624	HC
3627	HC
3634	PD
3635	HC
3636	HC
3637	HC

ID	Group
3756	HC
3759	HC
3765	HC
3767	HC
3768	HC
3769	HC
3770	PD
3775	PD
3778	PD
3779	HC
3780	PD
3781	PD
3784	PD
3788	PD
3789	PD
3790	PD
3803	HC
3804	HC
3805	HC
3806	HC
3807	HC
3808	PD
3809	HC
3811	HC
3812	HC
3813	HC
3815	PD
3816	HC
3817	HC
3820	PD
3822	PD
3823	PD
3824	PD
3825	PD
3826	PD
3828	PD
3829	PD
3830	PD
3831	PD
3836	PD
3850	HC
3851	HC

ID	Group
3852	HC
3853	HC
3854	HC
3855	HC
3857	HC
3860	PD
3869	PD
4001	PD
4004	HC
4005	PD
4010	HC
4012	PD
4018	HC
4020	PD
4021	PD
4022	PD
4023	PD
4024	PD
4025	PD
4026	PD
4029	PD
4030	PD
4031	PD
4032	HC
4035	PD
4064	PD
4066	PD
4067	HC
4081	PD
4082	PD
4083	PD
4084	PD
4085	HC
4136	PD
4139	HC

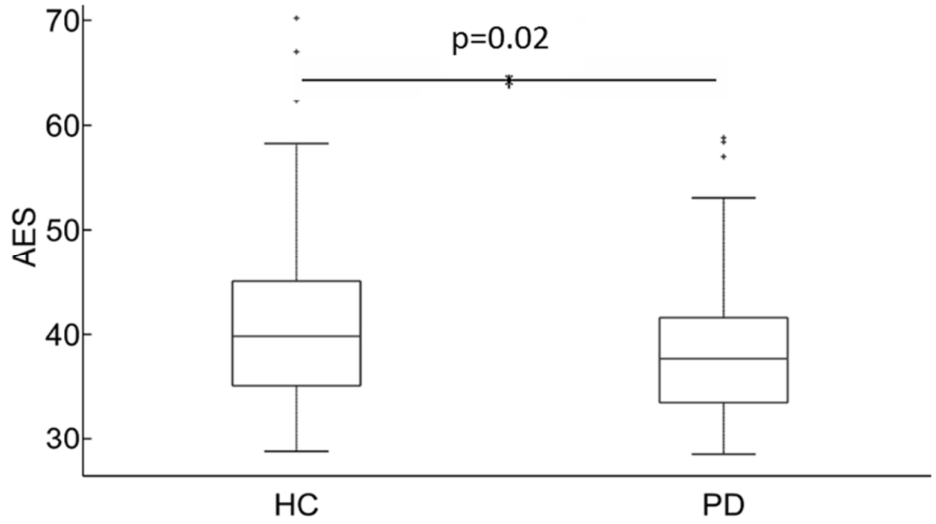
For Peer Review Only

Supplemental Figure 1: Empirical cumulative distributive function (eCDF) of the average edge strength (AES) computed from T1-weighted structural images from healthy controls (HC, green) and Parkinson's disease patients (PD, red). The differences across the two groups emerge when AES is at the 90th percentile (value below which 90% of the AES values may be found).



REVIEW ONLY

Supplemental Figure 2: Head motion group differences between Parkinson's disease (PD) patients and healthy controls (HC, age and gender matched) as measured by the average edge strength (AES, 90th percentile) computed from T1-weighted structural scans. The results are consistent with a relative increased image blurring (lower AES) in PD patients relative HC.



FOR REVIEW ONLY