Understanding cognitive heterogeneity in Parkinson's disease:

Gerrits, N.J.H.M.

2015

document version
Publisher's PDF, also known as Version of record

Chapter 8

Cortical thickness, surface area and subcortical volume differentially contribute to cognitive heterogeneity in Parkinson's disease

Authors
Niels J.H.M. Gerrits*
Anita C. van Loenhoud*
Stan van den Berg
Henk W. Berendse
Elisabeth M.J. Foncke
Martin Klein
Diederick Stoffers
Ysbrand D. van der Werf#
Odile A. van den Heuvel#

*, # Both authors contributed equally

Under revision

ABSTRACT

Parkinson's disease (PD) is often associated with cognitive deficits, although their severity varies considerably between patients. Recently, we used voxel-based morphometry (VBM) to show that individual differences in gray matter (GM) volume relate to cognitive heterogeneity in PD. VBM does, however, not differentiate between cortical thickness (CTh) and surface area (SA), which might be independently affected in PD. We therefore re-analyzed our cohort using the surface-based method FreeSurfer, and investigated (i) CTh, SA, and (sub)cortical GM volume differences between 93 PD patients and 45 matched controls, and (ii) the relation between these structural measures and cognitive performance on six neuropsychological tasks within the PD group. We found cortical thinning in PD patients in the occipital, parietal and frontal cortex. Within the PD group, we found negative correlations between (i) CTh of occipital areas and performance on a verbal memory task, (ii) SA of the orbitofrontal cortex and interference susceptibility, (iii) SA and volume of the frontal cortex and visuospatial memory performance, and, (iv) volume of the right thalamus and scores on two verbal fluency tasks. Our primary findings illustrate that i) CTh and SA are differentially affected in PD, and ii) VBM and FreeSurfer yield non-overlapping results in an identical dataset. We argue that this discrepancy is due to technical differences and the subtlety of the PD-related structural changes. We recommend to use both techniques to obtain information on cortical PD-related pathology from FreeSurfer, and more subtle volume effects in (sub)cortical structures from VBM.

INTRODUCTION

In addition to typical motor symptoms such as tremor, bradykinesia, rigidity, and postural instability, patients with Parkinson's disease (PD) often experience non-motor symptoms. Among these non-motor symptoms are cognitive deficits, which predominantly exist in the domain of executive functions, memory and visuospatial performance [12, 14]. Cognitive deficits are common, even in early stage PD [14], and up to 80% of all patients suffer from dementia at the end-stage of the disease [23] . The onset and rate of cognitive decline, however, differs considerably between patients.

Recently, we showed that differences in brain structure may contribute to cognitive heterogeneity in PD [91]. In this VBM study patients had relatively small areas of decreased gray matter (GM) volume in cortical areas such as the parietal, temporal, and frontal cortex, and in the cerebellum. Within the PD group, we found positive correlations between GM volume and cognitive performance for (i) parahippocampal gyrus and occipital lobe and verbal memory, (ii) medial temporal lobe and putamen and visuospatial memory, (iii) middle temporal gyrus and frontal lobe and verbal fluency, and (iv) inferior parietal lobe and cognitive flexibility. These VBM results suggest that in addition to the diffuse structural changes that affect the PD population in general, between-patient differences in regional GM volume may play a role in cognitive heterogeneity.

Despite the advantages of this voxel-based technique, VBM suffers from a major drawback: it does not distinguish between different cortical morphological properties [283]. GM volume is the product of cortical thickness (CTh) and surface area (SA) [284]. There is evidence to suggest that CTh and SA are differentially affected in normal aging [285] and Alzheimer's disease [286]. Similarly, recent studies suggest that a separate consideration of these two components of GM volume may also be more informative in the context of PD [287-289]. We therefore employed FreeSurfer, a surface-based technique, to measure CTh, SA, and (sub)cortical GM volume in the PD and HC groups originally analyzed with VBM [91]. This approach provided the opportunity to i) investigate specific structural changes related to PD, ii) study the contribution of different aspects of brain structure to cognitive heterogeneity in PD, and iii) compare the use of two common neuroimaging techniques for structural analyses in an identical dataset. We hypothesized to find a decrease in structural measures in PD patients, which could be (partly) explained by differences in CTh and SA. Similarly, we expected to find correlations between task performance and structural measures in brain areas that would (partly) overlap with those found in our previous VBM study within the PD sample. Although VBM and FreeSurfer are complementary (i.e. they do not measure the same

(sub)cortical characteristics) we expected to replicate the most robust cortical and subcortical effects we found in our previous study.

METHODS

Participants

A detailed description of the selection procedure of our participants is provided in Gerrits et al (2013). Briefly, we selected 93 idiopathic PD patients from a large, well-documented cohort of the outpatient clinic for movement disorders at the VU University medical center (VUmc), as well as 46 demographically age- and gender-matched HC. Magnetic resonance imaging (MRI) scans and demographic information, such as age and gender, were collected for the entire sample. Due to incorrect cortical reconstruction, we excluded one control participant, resulting in a sample that is almost, but not entirely, identical to the sample used in the VBM study [91]. Within the PD group, we evaluated education level using a scaled Dutch classification system ranging from 1 (did not finish primary school) to 7 (university degree) [290]. We assessed severity of motor symptoms and stage of illness with the motor subscore of the Unified Parkinson's Disease Rating Scale (UPDRS-III) and Hoehn & Yahr scales [102], respectively. Disease duration was defined as the subjective time interval between the first reported classical motor symptoms and the moment of clinical assessment. We evaluated mood and anxiety symptoms with the Beck Depression Inventory (BDI) [105] and the Beck Anxiety Inventory (BAI) [106], respectively. The cognitive status of our PD cohort was assessed by trained neuropsychologists as part of the standard diagnostic procedure. Of the 93 PD patients, 75 patients did not show evident cognitive impairments, eight patients fulfilled the criteria for mild cognitive impairment and four patients were diagnosed with PD dementia. Six patients could not be classified. All participants gave informed consent to the protocol, which was approved by the local ethics committee of the VUmc.

Neuropsychological assessment

Neuropsychological data were available only for the PD group, and not all patients participated in each cognitive task (see table 8.1). To evaluate global cognitive status, we used the mini-mental status examination (MMSE) [241, 291]. We assessed verbal memory with the Dutch version of the Rey auditory verbal learning task (RAVLT) and measured both the total number of immediately recalled items after five presentations and the number of items retrieved after a delay [292]. The delayed recall condition of the Rey-Osterrieth complex figure test (ROCFT) was used to evaluate visuospatial memory [293]. We administered the Category fluency task

(naming as many animals as possible in 60 seconds) to examine semantic fluency and the Letter fluency task (naming as many words possible starting with D, A and T in 3 trials of 60 seconds each) to assess phonemic verbal fluency. We examined executive functioning with the Stroop color word test [294] and the Trail making test [295]. Interference susceptibility was measured as the time needed for card III of the Stroop Color-Word Test minus the average completion time of Card I (speed of word reading) and II (speed of color naming). We subtracted the completion time on TMT-A from the completion time of TMT-B (TMTB-A) to obtain a measure of cognitive flexibility. The procedures for neuropsychological assessment followed those described by Lezak and colleagues [296].

Table 8.1 Demographic and clinical features of the PD and HC group, and PD subgroups for each neuropsychological test. Data represent mean ± SD or median (range).

	PD total	RAVLT	ROCFT	Stroop	TMTB-A	Category Fluency	Letter Fluency	HC	p-value (PD total vs HC)
Number of participants	93	88	83	86	79	85	80	45	-
tGM	648 ± 66	648 ± 67	651 ± 67	648 ± 66	649 ± 67	650 ± 68	650 ± 65	652 ± 59	-
Sex (male) (%)	61 (65.6)	56 (63.6)	53 (63.9)	54 (62.8)	48 (60.8)	56 (65.9)	52 (65.0)	27 (60.0)	0.52[b]
Age (years) (range)	62.8 ± 10.3 (27-88)	62.5 ± 10.2 (27-88)	62.2 ± 10.4 (27-88)	62.5 ± 10.3 (27-88)	62.2 ± 9.9 (27-88)	62.5 ± 10.5 (27-88)	62.0 ± 10.1 (27-84)	60.6 ± 7.8 (47-77)	0.19[a]
Education (Verhage)	5 (1-7)	5 (1-7)	5 (1-7)	5 (1-7)	5 (1-7)	5 (1-7)	5 (1-7)	-	-
UPDRS-III score	25.5 ± 10.3	24.2 ± 9.6	23.4 ± 9.2	24.0 ± 9.6	24.1 ± 9.7	24.7 ± 9.9	23.9 ± 9.6	-	-
Hoehn and Yahrstage (range)	2 (1-4)	2 (1-3)	2 (1-3)	2 (1-3)	2 (1-3)	2 (1-4)	2 (1-3)	-	-
Disease duration (years)	3.0 ± 3.2	2.8 ± 2.7	2.8 ± 2.7	2.8 ± 2.7	2.7 ± 2.7	3.0 ± 3.2	2.7 ± 2.6	-	-
DRT (n) (%)	32 (34.4)	30 (34.1)	27 (32.5)	29 (33.7)	28 (35.4)	27 (31.8)	26 (32.5)	-	-
LEDD (mg/day)	509 (100-1590)	492 (100-1590)	522 (150-1590)	495 (100-1590)	509 (150-1590)	525 (100-1590)	490 (100-1590)	-	-
MMSE	27.7 ± 3.0	27.8 ± 2.7	27.7 ± 3.0	28.0 ± 2.1	28.0 ± 2.1	28.0 ± 2.2	28.1 ± 2.1	-	-
BDI	8 (0-32)	8 (0-28)	8 (0-28)	8 (0-28)	8 (0-28)	8 (0-26)	8 (0-28)	-	-
BAI	10 (0-45)	10 (0-45)	10 (0-45)	10 (0-45)	11 (1-45)	10 (0-45)	10 (0-45)	-	-

Abbreviations: *PD* Parkinson's disease patients; *HC* healthy controls; *RAVLT* Rey Auditory Verbal Learning Test immediate and delayed recall; *ROCFT* Rey Osterrieth Complex Figure Test delayed recall; *Stroop* Stroop word color task; *TMTB-A* Trail Making Test B-A; *tGM* total grey matter volume; *UPDRS-III* Unified Parkinson's Disease Rating Scale-III; *DRT* dopamine replacement therapy; *LEDD* Levodopa equivalent daily dose, computed in the group of medicated patients only; *MMSE* Mini-Mental State Examination; *BDI* Beck Depression Inventory; *BAI* Beck Anxiety Inventory

[a] Student t test [b] Chi squared test

MRI acquisition and preprocessing

High-resolution structural MRI scans were obtained at the VUmc, using a GE Signa HDxt 3.0-Tesla MRI-scanner (General Electric, Milwaukee, Wisconsin, USA) with an 8-channel head coil. We acquired structural MRI data using a sagittal 3-dimensional gradient-echo T1-weighted sequence (256 x 256 matrix; field of view = 25cm; slice thickness = 1mm; voxel size = 1 x 0.98 x 0.98 mm; TR = 7.8 ms; TE = 3.0 ms; view angle = 12º). Image analysis was carried out with the stable version (v.5.3.0) of the FreeSurfer software (http://surfer.nmr.mgh.harvard.edu) [297-299]. In short, the procedure included: motion correction, intensity normalization, Talairach registration, skull stripping, segmentation of subcortical white matter, tessellation of the GM/white matter (WM) boundary, automated topology correction, and surface deformation. We used a 10 mm (full-width at half-maximum) Gaussian kernel to smooth maps. Finally, FreeSurfer created a surface 3D model of the cortex using intensity and continuity information.

Cortical analysis

We visually checked the cortical reconstruction of each subject for inaccuracies and manually corrected major topological inaccuracies with vertex edits or control points and subsequently repeated the processing. CTh was calculated as the shortest distance between the GM/WM boundary and pial surface at each vertex across the cortical mantle, measured in millimeters (mm). In addition to vertex-based reconstruction, FreeSurfer automatically parcellated the cortex into 34 gyral-based regions-of-interest (ROIs) per hemisphere, according to the Desikan-Killiany atlas. For each parcellation, the average CTh (in mm) and total WM SA (in mm^2) were calculated. We used these measures to manually calculate the cortical GM volume (in mm^3) of each of the 68 cortical parcellations (the product of CTh and SA), in order to compare volumetric cortical measures between FreeSurfer and VBM. We, furthermore, performed post-hoc tests when areas differed in cortical thickness to gain further insight into the morphological properties (i.e. surface area, cortical volume) of that particular area.

Subcortical analysis

Subcortical volumes were calculated with FreeSurfer's automated procedure for volumetric measures. Each voxel in the normalized brain volume was assigned to one of 40 labels, using a probabilistic atlas obtained from a manually labeled training set [300]. The labels we used for further analysis were the putamen, caudate nucleus, globus pallidus, nucleus accumbens, brainstem, thalamus, amygdala, hippocampus, ventral diencephalon and the ventricular system. In contrast to our

VBM study, the cerebellum was excluded and volumetric measures of the ventricles were included. Last, a measure of total GM (tGM) (in mm^3) was also computed, consisting of both surface-based cortical GM volume calculations and subcortical voxel counts.

Statistical analyses

To assess differences in demographic variables between the PD and HC group and PD subgroups for each task, we performed t-tests (for continuous data) and chi-square tests (for categorical data). We checked assumptions of normality, homogeneity of variance and covariance with the Shapiro-Wilk test, Levene's test and Box's M test, respectively. To correct for non-normal distribution, all values of ventricle volume, the TMTB-A score, and Stroop color word test scores were log-transformed, and the RAVLT scores were square-root transformed. We used parametric tests since statistical assumptions were met for 74% of the data.

Group differences

A number of statistical tests was performed to assess between-group differences in structural measures. First, we performed a vertex-wise analysis of differences in CTh in FreeSurfer's statistical program QDEC 1.5, using Monte Carlo-simulations with 10.000 iterations to correct for multiple comparisons and a cluster-wise *p*-value of .05 to display results. Second, surface (i.e. SA per parcellation) and volumetric analyses (i.e. sub-cortical volume estimates calculated by FreeSurfer, and the manually calculated volume estimate per cortical parcellation) were performed in SPSS 20.0 (SPSS, Chicago, IL, USA). For SA, we performed multivariate analyses of variance (MANOVA) using the 68 parcellations (34 per hemisphere) as dependent variables and group as between-subject factor. Subcortical and cortical GM volume differences were assessed using multivariate analyses of co-variance (MANCOVAs), with as dependent variables the volume of the 23 automatically segmented subcortical regions, and the volume of the 68 cortical parcellations, respectively, group as between-subject factor, and tGM as a covariate. We applied a Bonferroni correction by dividing our *p*-value by the number of cortical areas per hemisphere ($p < (.05/34) = \sim.001$) and by the number of sub-cortical structures per hemisphere ($p < .05/13) = \sim.004$) for the post-hoc tests in order to correct for multiple comparisons.

Correlations with cognitive performance

Since neuropsychological data were only available for the PD patients, correlations between cognitive performance and structural measurements were restricted to this

group. We used a GLM model in QDEC 1.5 to correlate CTh at each vertex with scores on the six neuropsychological tasks, while including age, gender, and education as covariates, and applying Monte Carlo-simulations to correct for multiple comparisons. We used a 'different-onset-same-slope' model, which assumes that no gender*age interaction exists. For SA and (sub)cortical GM volume, we computed partial correlations in SPSS 20.0 using each segmentation/parcellation as criterion, task-performance score as predictor, and age, gender, and education level as covariates. For the volumetric measures, tGM was added as an additional covariate. Again, Monte Carlo-simulations (in QDEC) and Bonferroni corrections (in SPSS) were applied to correct for the multiple comparisons.

RESULTS

The PD and HC group were matched for age (p = .19) and gender (p = .52). In addition, the PD subgroups for each task were similar regarding education, disease-related variables (i.e. UPDRS III score, Hoehn and Yahr stage, disease duration, dopamine replacement therapy), global cognitive functioning and measures of mood (i.e. depression and anxiety level) (see table I). On average, patients had a UPDRS III score of 24, a Hoehn and Yahr stage of 2 and a median disease duration of 3 years. The majority of the PD group was still unmedicated at the time of scanning (i.e. only 34% received dopamine replacement therapy).

Group differences

The vertex-wise CTh analysis showed cortical thinning in PD patients compared with HC in the left pericalcarine gyrus, extending to the cuneus, precuneus and lingual areas, in the left inferior parietal cortex, bilateral rostral middle frontal cortex, and right cuneus (see table 8.2 and figure 1a-d). In addition, PD patients showed enlargement of the third, and bilateral lateral ventricles and left inferior lateral ventricle when compared with HC (see table 8.3). No group differences in SA or cortical GM volume were found. A post-hoc inspection of the results indicated that the significant difference in mean CTh (HC>PD) in the left pericalcarine gyrus and right cuneus was accompanied by a sub-threshold increase in SA (PD>HC), resulting in cortical GM volumes that did not differ between groups (see table 8.4).

Table 8.2 Vertex-wise cortical thickness group analysis. Data represent mean thickness in mm ± SD. Only effects with significant clusterwise-values after Monte Carlo simulations are presented.

Region	Cluster size (mm^2)	PD	HC	Peak MNI305 coordinates			CWP
				X	Y	Z	
L pericalcarine gyrus	1802	1.87 ± 0.13	2.00 ± 0.13	-5	-75	12	<.001
R rostral middle frontal	633	2.24 ± 0.12	2.36 ± 0.14	40	49	4	.001
R cuneus	597	1.92 ± 0.13	2.03 ± 0.12	7	-86	27	.002
L rostral middle frontal	588	2.16 ± 0.13	2.28 ± 0.15	-38	43	3	.002
L inferior parietal	419	2.37 ± 0.16	2.51 ± 0.15	-38	-62	27	.02

Abbreviations: *PD* Parkinson's disease patients; *HC* healthy controls; *CWP* clusterwise corecte *p*-value

Table 8.3 Subcortical volume group analysis. Data represent mean ± SD. Only effects with significant *p*-values after Bonferroni correction are presented. Abbreviations: *PD* Parkinson's disease patients; *HC* healthy controls

Measurement	Region	PD	HC	*t*	*p*
Subcortical volume (mm^3) [a]	3rd ventricle	3.20 ± 0.17	3.09 ± 0.17	14.64	<.001
	L lateral ventricle	4.15 ± 0.23	4.00 ± 0.22	14.35	<.001
	R lateral ventricle	4.11 ± 0.23	3.97 ± 0.23	11.12	.001
	L inferior lateral ventricle	2.71 ± 0.31	2.54 ± 0.27	9.43	.003

[a] All measurements of ventricle volume are log transformed.

Table 8.4 Cortical volume, mean cortical thickness and surface area of the left pericalcarine gyrus and right cuneus in the PD and HC group. Abbreviations: PD *Parkinson's disease patients*; *HC* healthy controls; *ns* non-significant

Measurement	Region	PD	HC	*p*-value
Cortical thickness (mm)	L pericalcarine gyrus	1.68	1.77	.001
	R cuneus	1.88	1.95	<.001
Surface area (mm^2)	L pericalcarine gyrus	1358	1317	.28 (ns)
	R cuneus	1497	1470	.41 (ns)
Cortical volume (mm^3)	L pericalcarine gyrus	2279	2333	.52 (ns)
	R cuneus	2812	2873	.40 (ns)

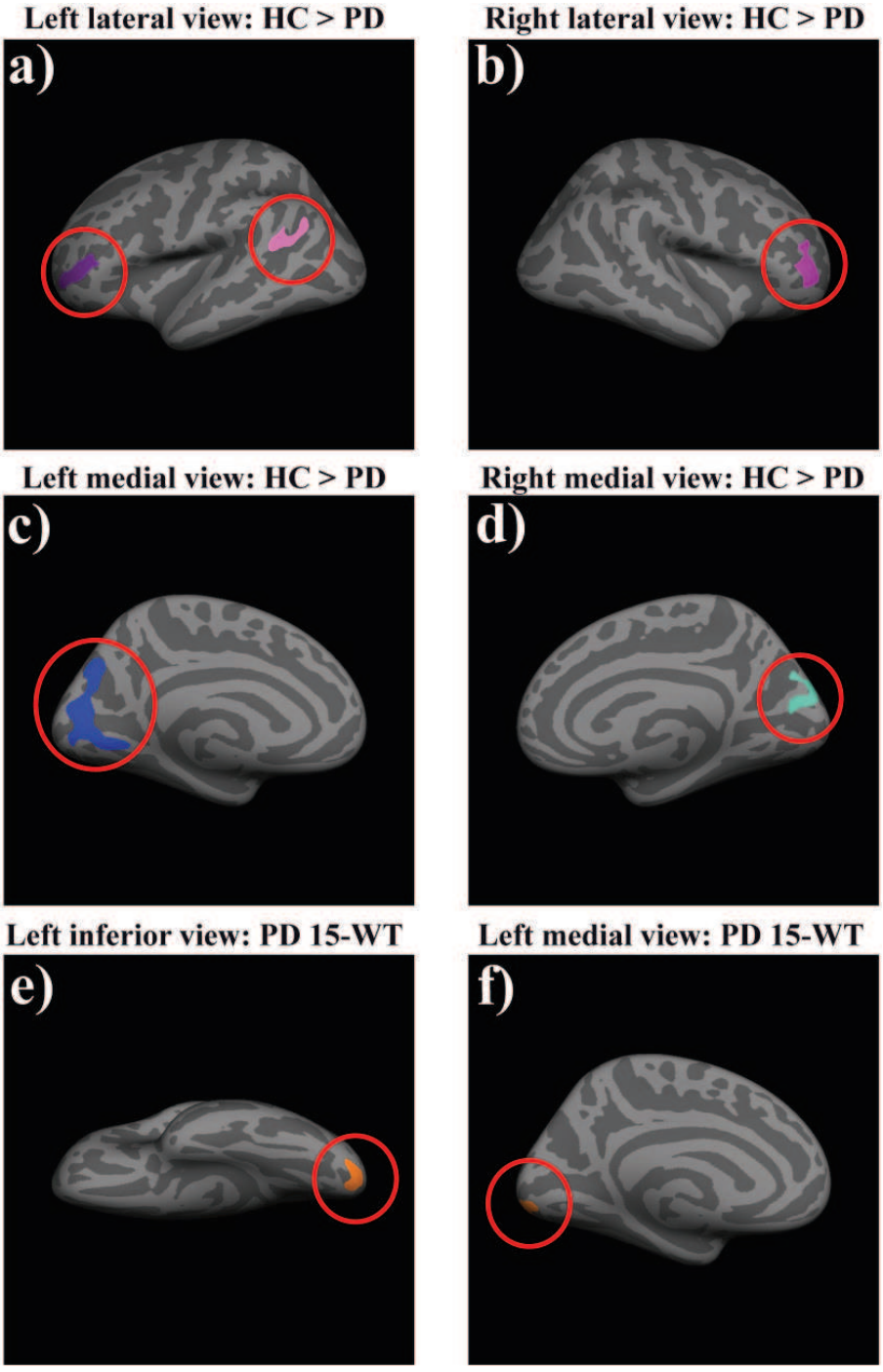

Figure 8.1 Between-group differences in cortical thickness and thickness and correlation with task performance

HC had increased cortical thickness in the left pericalcarine gyrus, extending to cuneus, precuneus and lingual areas left inferior parietal cortex, bilateral rostral middle frontal cortex, and right cuneus, when compared with PD patients (a-d). Within the PD sample, we found a negative correlation between the left lateral

occipital and lingual gyrus and performance on the RAVLT (e-f). Clusters were significant after multiple comparison correction with Monte Carlo simulations.

Correlations with cognitive performance

Vertex-wise analysis revealed a negative correlation between CTh in left lateral occipital and lingual areas and performance on the RAVLT immediate recall and delayed recall condition (see figure 1e,f). The volume of the right thalamus showed a negative correlation with performance on the Letter and Category fluency task. The volume and SA of the left pars opercularis correlated negatively with performance on the ROCFT. Furthermore, SA of the left medial orbitofrontal cortex correlated negatively with Stroop task performance (see table 8.5).

Table 8.5 Partial correlations of (sub)cortical volume and surface area with neuropsychological task performance, corrected for age, gender, and education. Abbreviations: *RAVLT* Rey Auditory Verbal Learning Task; *ROCFT* Rey Osterrieth Complex Figure Test; *Stroop* Stroop word color test; *r* Pearson's correlation coefficient; *CWP* clusterwise corrected *p*-vale; [a] Based on vertex-wise analysis [b] Based on parcellation-wise analysis

Measurement	Region	Task	*r*	Peak MNI305 coordinates			CWP
				X	Y	Z	
Cortical thickness (mm) [a]	L lateral occipital and lingual gyrus	RAVLT immediate recall	-.420	-19	-96	-15	.006
	L lateral occipital and lingual gyrus	RAVLT delayed recall	-.406	-8	-93	-10	.030
Subcortical volume (mm^3)[b]	R thalamus	Letter Fluency	-.338	-	-	-	.003
	R thalamus	Category Fluency	-.322	-	-	-	.003
Cortical volume (mm^3)[b]	L pars opercularis	ROCFT	-.374	-	-	-	.001
Surface area (mm^2) [b]	L pars opercularis	ROCFT	-.415	-	-	-	<.001
	L medial orbitofrontal cortex	Stroop	-.376	-	-	-	<.001

DISCUSSION

In this study, we used a surface-based analysis method to investigate structural brain changes in PD and the role of distinct morphological properties on cognitive heterogeneity among patients. Compared with controls, PD patients showed cortical thinning in the right cuneus, left lateral occipital areas, left inferior parietal cortex and the bilateral rostral middle frontal cortex, and ventricular enlargement. Within-group variance in volume of the thalamus, CTh of the left lateral occipital

and lingual areas, volume and SA of the pars opercularis, and SA of the left medial orbitofrontal cortex related to heterogeneity in verbal fluency performance, verbal memory, visuospatial memory and interference susceptibility, respectively. As in our VBM study, brain areas showing group differences in morphological properties did not overlap with brain areas in which structural changes were related to cognitive performance. Thus, while PD patients as a group showed atrophy in various regions compared with the HC, cognitive heterogeneity among patients was associated with between-patient structural differences in other regions. These differences may reflect subtle PD-related structural changes that affect only a subgroup of patients. Alternatively, they represent premorbid differences that may have caused some patients to be less vulnerable than others to cognitive impairment as a consequence of the PD-related structural changes observed at a group level.

Although we used the same dataset in the current study as in our previous VBM analysis [91], there was surprisingly little overlap between the studies in the areas in which we found significant effects. We believe that five issues may have contributed to this incongruity.

The first relates to differences in employed statistics. In our VBM study, we applied an uncorrected *p*-value of .001 with an extent threshold of 50 voxels, whereas the present study employs Monte Carlo simulations and Bonferroni corrections, which are statistically more stringent [301]. Indeed, when applying a similar significance and cluster threshold as used in the VBM study in our vertex-wise analysis, we find a group effect for CTh in the right superior frontal gyrus that has a near-identical coordinate correspondence with the GM volume effect found with VBM. The majority of the areas is, however, still non-overlapping.

A second factor concerns technical differences between the voxel-based approach in VBM and the atlas-based approach in FreeSurfer. Whereas FreeSurfer calculates the total volume of a cortical parcellation or subcortical segmentation, VBM assesses GM volume on a voxel-by-voxel basis. VBM might, therefore, be more sensitive to detect small local effects that may be 'averaged out' when measured over a larger area. Since our VBM results showed relatively small clusters of GM volume effects, these effects may have been too small to be detected by FreeSurfer. In our current study we found a negative correlation between verbal fluency and volume of the right thalamus, a relation which was not found in our previous study. Since both VBM [302] and FreeSurfer [303] have problems segmenting the thalamus from the surrounding WM, we advise caution when interpreting this finding, since these conflicting results could indicate a spurious finding. Replication in future research is therefore warranted.

A third issue relates to the fact that cortical GM volume as a measure of brain structure is different from CTh and SA. Our FreeSurfer results show, in accord-

ance with earlier studies [287-289], that CTh and SA are differentially affected in PD. Since the product of their combined influence is not uniform across the cortex, cortical GM volume may not show overlap with either measure. In theory, it is possible that two groups do not significantly differ on the CTh or SA in a certain area, but that this difference does reach significance when combined in GM volume. In contrast, significant CTh and SA effects may remain undetected when combined in GM volume when their effects are in opposite directions. As an example, we showed that in the right cuneus and left pericalcarine gyrus, significant decreases in CTh in the PD group were masked by non-significant increases in SA, leading to normal regional GM volume as measured with VBM. Increases in SA have been reported previously in the context of PD (Jubault et al 2011) and may relate to damage in the underlying white matter fiber tracts.

Fourth, we found a relation between cortical volume (as calculated from the product of SA and CTh) of the left pars opercularis and performance on the ROCFT, which we did not find in our VBM study. FreeSurfer calculates SA by measuring the surface area of the WM/GM boundary, thereby neglecting the surface area of the pial surface. It is therefore possible that VBM accounted for atrophy in the pial surface (consequently leading to a reduced GM volume) whereas FreeSurfer did not. Since our FreeSurfer results show that the left pars opercularis volume effect was mainly driven by a SA effect, we speculate that VBM did not detect this task-related association because it was driven by a (potentially biased) SA effect.

The fifth and last issue relates to the relatively subtle structural alterations observed at this early disease stage. Several other studies have investigated structural changes in early stage PD, and found little or no atrophy in cognitively preserved cohorts, comparable to ours [47, 48]. The areas in which atrophy was described varied considerably between studies, thus suggesting that the atrophy is subtle and topographically non-specific, in contrast with, for example, hippocampal atrophy in Alzheimer's disease. We argue that if the structural differences had been more pronounced, both techniques would have detected them. Our results confirm previous studies by showing that there is indeed atrophy in relatively early stage PD, but that it is, if anything, subtle and spread over various brain areas. Also the enlargement of the third and lateral ventricles indicates a diffuse and non-specific degenerative process.

Several results are consistent with previous data obtained using FreeSurfer in PD, mainly concerning CTh reductions in the bilateral rostral middle frontal cortex, bilateral cuneus and left inferior parietal areas [304-306], as well as the enlargement of the third and lateral ventricles [307, 308]. Also the positive correlation between SA of the left medial orbitofrontal cortex and Stroop task performance is in ac-

cordance with the involvement of this area in response inhibition [309]. In contrast, the negative correlation between verbal memory performance and CTh of the lateral occipital and lingual cortex is not in line with earlier findings. Pellicano et al. (2012) reported a positive correlation between verbal memory performance and thickness in occipital areas (i.e. the fusiform area) in PD [310]. Also the negative correlations between the left pars opercularis and the right thalamus with visuospatial memory and verbal fluency, respectively, are difficult to interpret, although numerous cognitive processes have been associated with these areas [311] [312-314]. Future studies should replicate these findings before any definite statements can be made.

To our knowledge, this is the first study that compared VBM and FreeSurfer data in the same cohort of PD patients to study the relation between brain structure and cognitive performance. Strengths of our study include our relatively large and well-powered [315] sample and the fact that we controlled for various confounding factors such as age, gender and education. An important limitation, however, is the absence of neuropsychological test scores from HC. Conclusions based on the correlations between brain structure and cognitive performance should therefore be interpreted with caution, as they may not be specific to PD. Future research should include a longitudinal approach to gain more insight into how structural changes relate to cognitive status over time. It would also be insightful to include patients with a more diverse cognitive profile to make the sample more heterogeneous, or subdivide the sample into subgroups based on cognitive status (e.g. cognitively not impaired; cognitively impaired; demented).

CONCLUSION

The results of the current study suggest that PD is associated with cortical thinning and ventricular enlargement, and that cognitive heterogeneity within the PD population is associated with subtle differences in CTh, SA, and (sub)cortical GM volume. Our results obtained with FreeSurfer support the hypothesis that CTh and SA are differentially affected by the disease, and have diverse associations with cognition. This underlines the necessity to take distinct morphological properties of brain areas into account in the context of PD. By directly comparing GM volume effects obtained with FreeSurfer and VBM, we have provided evidence that their methodological and technical differences can yield non-overlapping results in the same cohort of participants. We argue that FreeSurfer may be more sensitive to PD pathology in cortical areas showing isolated CTh or SA effects, or areas in which their effects are in opposite directions. VBM, however, may be able to detect more subtle effects in small cortical and subcortical GM areas, due to its

voxel-by-voxel approach. We thus recommend to use both techniques in complement to each other to obtain a comprehensive picture of the complex pathological process in PD.