Regional gray matter changes and age predict individual treatment response in Parkinson's disease

Tommaso Ballarini[a], Karsten Mueller[a], Franziska Albrecht[a], Filip Růžička[b,d], Ondrej Bezdicek[b], Evžen Růžička[b], Jan Roth[b], Josef Vymazal[d], Robert Jech[b,d,1], Matthias L. Schroeter[a,c,e,*,1]

[a] Max-Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[b] Department of Neurology, Charles University, First Faculty of Medicine, Prague, Czech Republic
[c] Clinic for Cognitive Neurology, University Clinic, Leipzig, Germany
[d] Department of Radiology, Na Homolce Hospital, Prague, Czech Republic
[e] FTLD Consortium, Germany

ARTICLE INFO

ABSTRACT

We aimed at testing the potential of biomarkers in predicting individual patient response to dopaminergic therapy for Parkinson's disease. Treatment efficacy was assessed in 30 Parkinson's disease patients as motor symptoms improvement from unmedicated to medicated state as assessed by the Unified Parkinson's Disease Rating Scale score III. Patients were stratified into weak and strong responders according to the individual treatment response. A multiple regression was implemented to test the prediction accuracy of age, disease duration and treatment dose and length. Univariate voxel-based morphometry was applied to investigate differences between the two groups on age-corrected T1-weighted magnetic resonance images. Multivariate support vector machine classification was used to predict individual treatment response based on neuroimaging data. Among clinical data, increasing age predicted a weaker treatment response. Additionally, weak responders presented greater brain atrophy in the left temporoparietal operculum. Support vector machine classification revealed that gray matter density in this brain region, including additionally the supplementary and primary motor areas and the cerebellum, was able to differentiate weak and strong responders with 74% accuracy. Remarkably, age and regional gray matter density of the left temporoparietal operculum predicted both and independently treatment response as shown in a combined regression analysis. In conclusion, both increasing age and reduced gray matter density are valid and independent predictors of dopaminergic therapy response in Parkinson's disease.

1. Introduction

Dopaminergic therapy (DT) has been the foundation of Parkinson's disease therapy for decades. DT is based on two main pharmacological compounds: levodopa (L-3,4-dihydroxyphenylalanine) and dopamine receptor agonists. Despite its extensive application in clinical practice, DT efficacy is influenced by several factors. For example, DT response deteriorates after years of treatment, becoming more intermittent and leading to severe motor and non-motor complications (Connolly and Lang, 2014). Notably, a high individual variability in treatment response is a common finding in clinical practice and it has been often reported in the literature (Constantinescu et al., 2007). In particular, two studies investigated levodopa responsiveness in large series of autopsy-proven Parkinson's disease cases, showing a suboptimal treatment response in several patients (Hughes et al., 1993; Wenning et al., 2000). However, the sources of heterogeneous DT responses across subjects are still unclear. The aim of the study is to shed light on the potential of clinical and neuroimaging data in predicting clinical response to DT in Parkinson's disease.

2. Materials and methods

2.1. Participants

Thirty patients with Parkinson's disease participated in the study (Hughes et al., 1992) (age = 65.37 $\pm$ 7.05 years mean $\pm$ SD; Hoehn&

Yahr stages I-III; 13 females, disease onset after 45 years of age). Stable DT comprised levodopa either in monotherapy ($N = 5$) or combined with dopamine agonists (pramipexole and ropinirole, respectively $N = 8$ and $N = 17$). Differences in treatment response across DT types were tested with a one-way ANOVA. The assessment of motor symptoms, implemented with the Unified Parkinson's Disease Rating Scale (UPDRS) score III, was performed during medicated (DT-on) and unmedicated states (DT-off) in a randomized order (DT-on first in 16 and DT-off first in 14 patients). For DT-off state, dopamine agonists and levodopa were withdrawn, respectively, at least 72 and 12 h before examination. We considered the difference between the UPDRS-III scores in DT-on and DT-off conditions (i.e. *UPDRS-III-change*) as a proxy for treatment efficacy. The study was authorized by the Ethics Committee of the General University Hospital in Prague, Czech Republic, in line with the Declaration of Helsinki.

2.2. Clinical variables

Correlations between *UPDRS-III-change* and, in turn, age, disease duration, levodopa treatment duration and levodopa equivalent dose were assessed. Note that, for each pair, Pearson's partial correlations ($\alpha = 0.05$) were computed controlling for the remaining variables. Additionally, multivariate regression (enter mode) was conducted including these variables and setting *UPDRS-III-change* as dependent factor.

For MRI analyses, patients were stratified into weak and strong responders, using as cut-off the median *UPDRS-III-change* in our sample (15.50 points). Table 1 displays demographic and clinical characteristics of the groups, and their statistical comparison. Notably, age was significantly higher in the weak responders. Hence, age differences were controlled for in these and the following dichotomous analyses. In addition, to define the clinical significance of the treatment-induced UPDRS-III changes, we compared our results with the cut-off values provided by Shulman et al. (2010) defining minimal, moderate and large clinically important differences (i.e. differences that can be appreciated by the patients themselves). Accordingly, the majority of the patients ($n = 24$) reported large differences, five intermediate and only one minimal. At clinical evaluation, most patients presented with both akinesia/rigidity involvement and tremor. PD phenotypes were further defined according to Jankovic's classification to allow a better comparison with other studies (Jankovic et al., 1990). Three outcomes are possible here: postural instability and gait difficulty (PIGD), tremor dominant (TD) and intermediate. Namely, a ratio score, computed based on the UPDRS-III items in the DT-off state, was used for the

classification: ratio = mean(TD items)/mean(PIGD items). For a list of the items used for computing the mean values for TD and PIGD refer to Table 1 in Stebbins et al. (2013). In our PD cohort, 22 patients had predominant PIGD, 5 TD and 3 intermediate phenotypes. Differences in the distribution of the clinical phenotypes between strong and weak responders were assessed by means of chi-squared test.

2.3. Neuroimaging analysis

T1-weighted MR images (3D-MPRAGE, Magnetization Prepared Rapid Gradient Echo, TR 2.2 s, TE:2.43 ms, inversion time 900 ms, flip angle 8°, matrix 224 × 224) were acquired on a 3 T MAGNETOM Skyra (Siemens, Erlangen, Germany) at the Radiology Department, Na Homolce Hospital in Prague.

Images were preprocessed with the Computational Anatomy Toolbox (CAT-12). After spatial normalization, gray matter compartment was isolated and 8 mm smoothing was applied. Then, to take into account the effect of age, an age-correction procedure was applied based on data from 30 elderly healthy controls (age = 63.57 ± 8.09 years; 15 females) as previously described (Dukart et al., 2011). Thus, imaging data were normalized for age before all subsequent structural analyses. The efficacy of the procedure was further tested performing a whole-brain correlation analysis between age and gray matter density (GMD) before and after age correction and comparing the z-maps of the two analyses. Individual gray and white matter volumes were entered in the models as covariates.

2.3.1. Univariate analysis

Voxel-based morphometry (Ashburner and Friston, 2000) was run in CAT-12 and SPM (rev.12.6685). Age-corrected GMD images entered a univariate statistical model to compare GMD between weak and strong responders using gray and white matter volumes as covariates. Only clusters surviving $p < .05$ family-wise error (FWE) correction were deemed significant (uncorrected threshold, $p < .005$).

2.3.2. Multivariate analysis

Support vector machine (SVM) classification was applied on age-corrected GMD images using libsvm (release 3.18) in Matlab. This machine learning approach based on multivariate pattern analysis allows testing the prediction accuracy of GMD images for the classification of weak and strong DT responders. The classifier was based on a polynomial kernel and was validated using a leave-one-out cross-validation procedure. Different gray matter probability masks were tested to assure the stability of results. The relative weights, indicating the

Table 1
Demographic and clinical characteristics of Parkinson's disease patients stratified according to dopaminergic treatment response.

	Total	Weak responders	Strong responders	F (df)	p
N	30	15	15	–	–
Age (years)	65.37 ± 7.05	68.27 ± 7.56	62.47 ± 5.28	5.93 (28)	0.021*
Gender (male/female)	17/13	8/7	9/6	–	$\chi^2 = 0.14$, $p = .71$
Education (years)	14.1 ± 2.95	13.4 ± 2.82	14.8 ± 3	1.10 (27)	0.34
Disease duration (years)	11.4 ± 3.43	10.87 ± 3.81	11.93 ± 3.03	0.78 (27)	0.386
Levodopa duration (years)	8.529 ± 4.055	8.64 ± 4.2	8.42 ± 4.05	0.86 (27)	0.363
Levodopa equivalent dose (mg)	1385.98 ± 664.75	1155.65 ± 680.28	1600.97 ± 582.85	1.84 (27)	0.186
UPDRS-III DT-off	31.03 ± 10.31	28.27 ± 10.05	33.80 ± 10.13	1.48 (27)	0.235
UPDRS-III DT-on	14.77 ± 7.58	17.20 ± 8.12	12.33 ± 6.37	1.73 (27)	0.199
UPDRS-III-change	16.27 ± 7.06	11.07 ± 3.73	21.47 ± 5.59	23.93 (27)	< 0.0001**
UPDRS-IV	4.60 ± 3.65	3.33 ± 3.0	5.87 ± 3.871	2.61 (27)	0.118
NMSS total score	15.10 ± 19.18	13.53 ± 20.79	16.67 ± 18.015	0.50 (27)	0.485
MoCA	25.77 ± 2.34	25.53 ± 2.47	26.00 ± 2.27	0.073 (27)	0.789

Note: Mean ± SD are shown.
Abbreviations: DT-on medicated state; DT-off unmedicated state; NMSS non-motor symptoms scale; MoCA Montreal Cognitive Assessment; UPDRS Unified Parkinson's Disease Rating Scale (-III motor symptoms; -IV motor complications).
 * Significant p-value from ANOVA analysis.
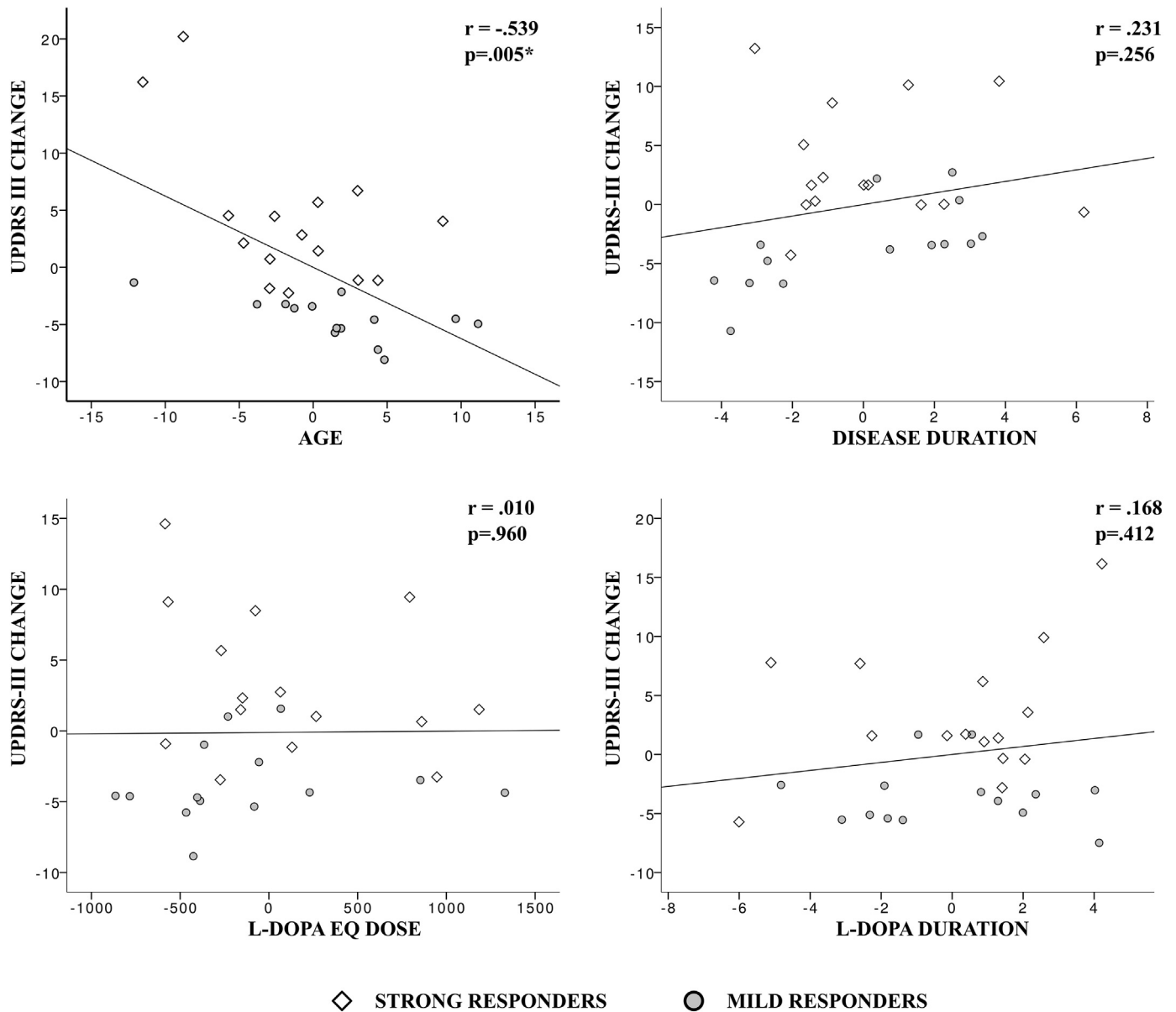 ** Significant p-value from ANCOVA analysis controlling for age differences.

Fig. 1. Partial correlations between clinical variables and *UPDRS-III-change* as proxy for treatment efficacy. Age is the only variable showing a statistically significant negative correlation (i.e. increasing age was associated with weaker treatment response). Pearson correlation coefficients (r) and *p*-values are displayed. Note that, for each correlation, x and y axes represent the two sets of unstandardized residuals from regressing the variables of interest on the confounding variables. Strong and weak responders, as selected for the neuroimaging analysis, are presented with different symbols.

contribution to the classification of each voxel, were then plotted in the original 3D brain space.

2.4. Disentangling impact of age and imaging parameters on treatment

To disentangle the influence of age and gray matter changes, the SVM model was run on the patients' images also before the age correction procedure. The rationale is that differences in SVM accuracy between pre and post age correction procedure would be attributable to the detrimental aging effect on GMD.

Moreover, to further address the influence of clinical and imaging data on the treatment response, a multivariate linear regression model was run including as predictors age, levodopa dose/duration, disease duration and the average GMD values from the consistent (i.e. over-laying) cluster from the univariate and multivariate analyses, centered on the x y z MNI coordinates − 33 -32 18 (represented in orange in Fig. 3).

3. Results

3.1. Clinical variables

The distribution of clinical phenotypes (PIGD, TD and intermediate) was comparable between strong and weak responders ($\chi 2 = 0.715$, $p = .699$). No difference in treatment response was found between different DT types (one-way ANOVA: F(29) = 0.33, $p = .723$).

A significant correlation was found between *UPDRS-III-change* and age ($r = -0.539, p < .005$). All the other tested correlations did not provide significant results: disease duration ($r = 0.231, p < .256$), levodopa treatment duration ($r = 0.168, p < .412$) and levodopa equivalent dose ($r = 0.01, p < .96$) (Fig. 1).

A multivariate regression model was applied to predict treatment efficacy from age, disease duration, levodopa equivalent dose and levodopa duration. The overall model was significant, F(4,24) = 4.408, $p < .008$, with an r^2 of 0.424. However, age was the only significant
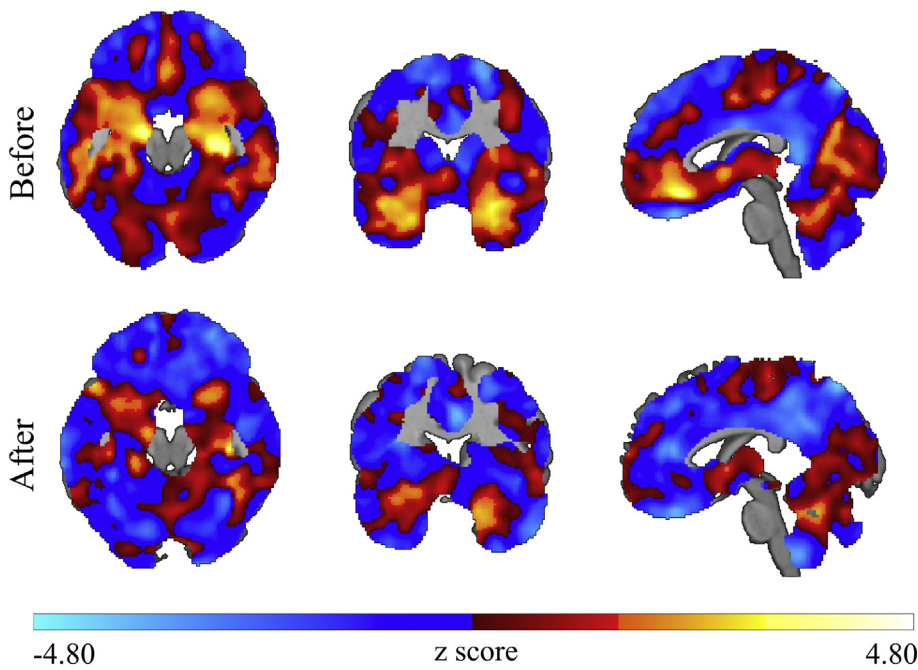
Fig. 2. Efficacy of the age correction procedure following Dukart et al. (2011). *Z*-maps showing the results of correlation analyses in the patient cohort between age and gray matter density maps before (upper row) and after (bottom row) the applied age-correction procedure. The pattern of negative correlations between age and gray matter density, mainly seen in medial temporal and medial frontal structures, is suppressed after age correction. Images shown in neurological orientation: left of the brain corresponds to the left of the image.
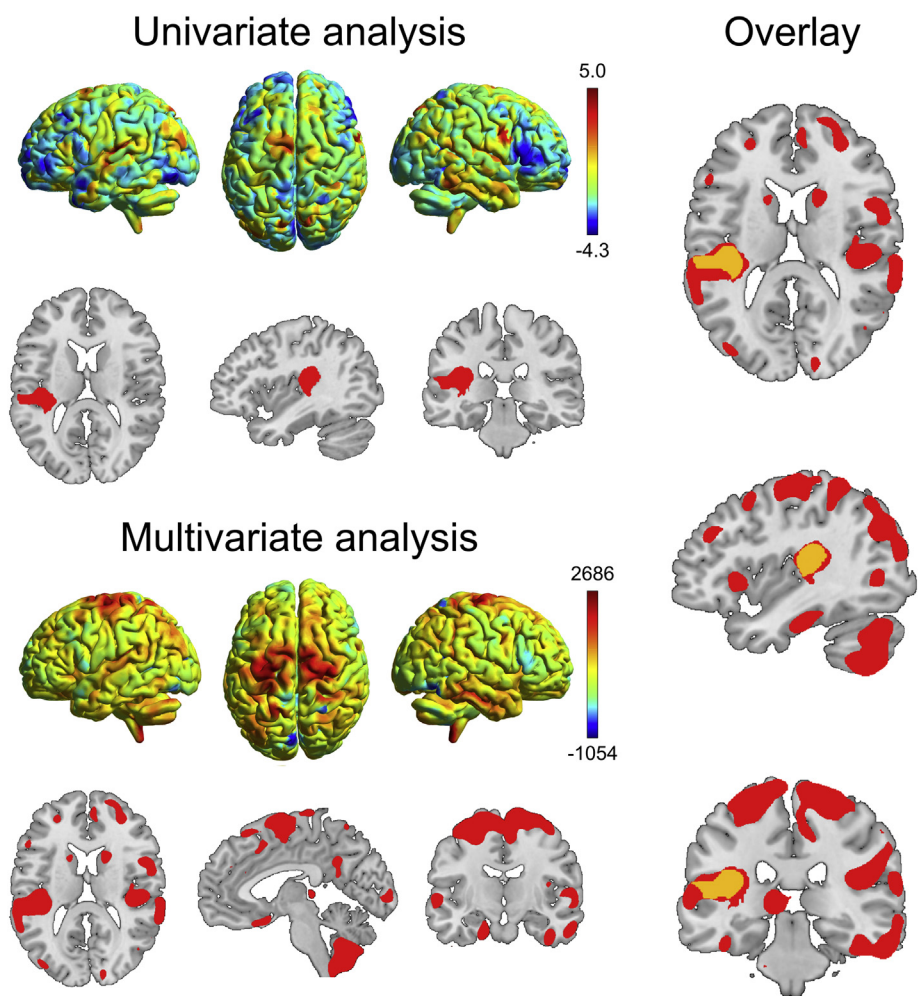

Fig. 3. Results of the univariate and multivariate analyses contrasting weak and strong responders. The 3D rendering shows, respectively, the unthresholded SPM-t map for the weak responders < strong responders univariate SPM comparison (warmer colors indicate lower gray matter density) and the voxel-wise weights obtained from the multivariate support vector machine classification (warmer colors regions contributing to weak responders' classification, arbitrary units). The slice views for the univariate analysis report in red the results surviving the cluster-level correction for multiple comparisons (*p* < .05 FWE), while for the multivariate analyses, the regions showing higher weights for the classification of the weak vs. strong responders are shown. On the right, the overlay between the two thresholded maps is shown in orange. Images shown in neurological orientation: left of the brain corresponds to the left of the image. 3D rendering plotted with BrainNet Viewer 1.61 (https://www.nitrc.org/projects/bnv/).

Table 2
Support vector machine (SVM) classification results (weak vs. strong responders) using different gray matter probability masks. Lower gray matter thresholds correspond to more inclusive masks.

Gray matter threshold	Accuracy in weak responders (%)	Accuracy in strong responders (%)	Total accuracy (%)	PPV	NPV
0.01	77.33	70.67	74.00	0.725	0.757
0.1	78.67	70.67	74.67	0.728	0.768
0.2	78.22	70.67	74.44	0.727	0.764
0.3	76.00	72.44	74.22	0.734	0.751
0.4	64.44	73.33	68.89	0.707	0.673
0.5	62.22	73.33	67.78	0.700	0.660

Note: NPV negative predictive value; PPV positive predictive value.

predictor in the model ($p = .005$). Namely, *UPDRS-III-change* decreased of -0.604 points for each additional year of age.

3.2. Neuroimaging analysis

The correlation analysis between age and GMD before the age-correction procedure shows a pattern of negative correlations mainly in medial temporal and medial frontal structures, as expected from a consistent body of literature (Giorgio et al., 2010; Tisserand et al., 2004). After the age correction, this prototypical correlation pattern was mostly suppressed as shown in Fig. 2.

3.2.1. Univariate analysis

Significant lower GMD was found in weak compared to strong responders in the left temporoparietal operculum after age correction (Fig. 3). Observing the unthresholded t-map for the weak < strong responders comparison, a trend towards lower GMD was detected in the right temporoparietal operculum, in supplementary motor area, superior frontal gyrus and cerebellum. The opposite contrast did not reveal significant findings.

3.2.2. Multivariate analysis

The SVM approach, based on whole brain data after age correction, was able to correctly classify 77.33% weak responders and 70.67% strong responders, thus with a total accuracy of 74.0%. Positive and negative predictive values were, respectively, 0.725 and 0.757 (Table 2). The regions that mostly contributed to the classification of weak responders included the supplementary motor and primary motor areas, superior frontal gyrus, the bilateral parietal opercula, and extensively the cerebellum (Fig. 3).

3.3. Disentangling impact of clinical and imaging parameters on treatment

The SVM model applied to the patients' images before the age correction procedure showed a total accuracy of 75.77%, correctly classifying 78.22% weak and 73.33% strong responders. An additional multivariate regression model was calculated to predict treatment efficacy from age and GMD in the left temporoparietal operculum to disentangle the contributions of both relevant factors. This analysis revealed significant effects, $F(5,23) = 6.56$, $p < .001$, with an increased r^2 of 0.588. Significant predictors in the model were both age (standardized $\beta = -0.496$; $p = .009$) and GMD in the left temporoparietal operculum (standardized $\beta = 0.456$; $p = .006$).

4. Discussion

The present findings represent an initial insight into clinical and neuroimaging features that can predict individual DT response in Parkinson's disease. We showed that both increasing age and lower GMD predict a weaker DT response. Of note, weak and strong DT responders were similarly impaired in the DT-off state, suggesting that treatment response does not depend on motor symptom severity. The relationship between aging and Parkinson's disease has been subject of previous investigations (Levy, 2007). Levy (2007) reported that age, more than disease duration, is associated with a more severe clinical progression. In addition, a previous study showed a negative association between age and the magnitude of DT response in Parkinson's disease (Durso et al., 1993). Taken together, these pieces of evidence are in line with our finding of a poorer treatment response in older Parkinson's disease patients.

Concerning neuroimaging results, structural between-group differences were identified with univariate and multivariate analyses after applying a stringent age-correction procedure. A good agreement was found between the two analytical approaches as exemplified in Fig. 3. Indeed, both revealed pronounced regional gray matter alterations in weak DT responders in the left temporoparietal operculum, compared to strong responders. The parietal operculum is a complex brain structure where different sensorimotor processes take place and, interestingly, it has been previously associated with sensorimotor recovery after moderate to severe strokes within the territory of the internal carotid artery (Hannanu et al., 2017). In addition, the SVM analysis showed also the involvement of the supplementary motor area, primary motor cortex and cerebellum in defining the weak responder group, all crucial regions in the motor pathway. In particular, an impaired activity of the supplementary motor area in Parkinson's disease has been previously associated with motor impairment and, accordingly, attempts to restore its function via transcranial magnetic stimulation showed modest but significant symptoms improvement (Hamada et al., 2008). It is thus reasonable that patients with more severe structural impairment in this area, as suggested by the univariate unthresholded results, present a weaker treatment response.

As for the relationship between age and GMD alterations, the linear regression model including both as predictors of *UPDRS-III change* explained 59% of the variance, as compared to the 42% of the model without imaging metrics. The comparison of standardized beta coefficients showed a similar effect of aging and GMD alterations in predicting the treatment response. This suggests that age and GMD are both significantly and independently contributing to the prediction of treatment response in PD. Additionally, in order to test the modulatory impact of age on GMD-based predictions, the SVM classification was run on the imaging data before and after applying the age correction procedure. This analysis revealed that the performance of the classifier was only minimally improved when images were uncorrected for age differences (overall accuracy pre and post age correction were respectively 75.77% and 74.0%). This again suggests a modest modulatory effect of age on the GMD-based classification, thus indicating that age and GMD are mostly independent predictors of DT treatment response in PD. Consequently, further research would be needed for the identification of mediators of the aging effect on treatment response. For example, it has been proposed that aging might influence neurodegeneration (and possibly treatment response) through the complex interaction between iron and neuromelanin accumulation and loss of dopamine neurons (Zucca et al., 2017).

In conclusion, we provide a new perspective into the issue of heterogeneous individual treatment response in Parkinson's disease. The implementation of SVM classification provides a new approach to the topic and might pave the way to further investigations including larger Parkinson's disease cohorts, atypical parkinsonian syndromes and different neuroimaging data.

References

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry—the methods. NeuroImage 11, 805–821.

Connolly, B.S., Lang, A.E., 2014. Pharmacological treatment of Parkinson disease: a review. JAMA 311, 1670–1683.

Constantinescu, R., Richard, I., Kurlan, R., 2007. Levodopa responsiveness in disorders with parkinsonism: a review of the literature. Mov. Disord. 22, 2141–2148 (quiz 2295).

Dukart, J., Schroeter, M.L., Mueller, K., Alzheimer's Disease Neuroimaging, I, 2011. Age correction in dementia—matching to a healthy brain. PLoS One 6, e22193.

Durso, R., Isaac, K., Perry, L., Saint-Hilaire, M., Feldman, R.G., 1993. Age influences magnitude but not duration of response to levodopa. J. Neurol. Neurosurg. Psychiatry 56, 65–68.

Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., Johansen-Berg, H., 2010. Age-related changes in grey and white matter structure throughout adulthood. NeuroImage 51, 943–951.

Hamada, M., Ugawa, Y., Tsuji, S., Effectiveness of rTms on Parkinson's Disease Study Group, J, 2008. High-frequency rTMS over the supplementary motor area for treatment of Parkinson's disease. Mov. Disord. 23, 1524–1531.

Hannanu, F.F., Zeffiro, T.A., Lamalle, L., Heck, O., Renard, F., Thuriot, A., Krainik, A., Hommel, M., Detante, O., Jaillard, A., Group, I.-H.S, 2017. Parietal operculum and motor cortex activities predict motor recovery in moderate to severe stroke.

Neuroimage Clin. 14, 518–529.

Hughes, A.J., Daniel, S.E., Kilford, L., Lees, A.J., 1992. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J. Neurol. Neurosurg. Psychiatry 55, 181.

Hughes, A.J., Daniel, S.E., Blankson, S., Lees, A.J., 1993. A clinicopathologic study of 100 cases of Parkinson's disease. Arch. Neurol. 50, 140–148.

Jankovic, J., McDermott, M., Carter, J., Gauthier, S., Goetz, C., Golbe, L., Huber, S., Koller, W., Olanow, C., Shoulson, I., 1990. Variable expression of Parkinson's disease a base-line analysis of the DAT ATOP cohort. Neurology 40, 1529.

Levy, G., 2007. The relationship of Parkinson disease with aging. Arch. Neurol. 64, 1242–1246.

Shulman, L.M., Gruber-Baldini, A.L., Anderson, K.E., Fishman, P.S., Reich, S.G., Weiner, W.J., 2010. The clinically important difference on the unified Parkinson's disease rating scale. Arch. Neurol. 67, 64–70.

Stebbins, G.T., Goetz, C.G., Burn, D.J., Jankovic, J., Khoo, T.K., Tilley, B.C., 2013. How to identify tremor dominant and postural instability/gait difficulty groups with the movement disorder society unified Parkinson's disease rating scale: comparison with the unified Parkinson's disease rating scale. Mov. Disord. 28, 668–670.

Tisserand, D.J., Van Boxtel, M.P., Pruessner, J.C., Hofman, P., Evans, A.C., Jolles, J., 2004. A voxel-based morphometry study to determine individual differences in gray matter density associated with age and cognitive change over time. Cereb. Cortex 14, 966–973.

Wenning, G., Ben-Shlomo, Y., Hughes, A., Daniel, S., Lees, A., Quinn, N., 2000. What clinical features are most useful to distinguish definite multiple system atrophy from Parkinson's disease? J. Neurol. Neurosurg. Psychiatry 68, 434–440.

Zucca, F.A., Segura-Aguilar, J., Ferrari, E., Muñoz, P., Paris, I., Sulzer, D., Sarna, T., Casella, L., Zecca, L., 2017. Interactions of iron, dopamine and neuromelanin pathways in brain aging and Parkinson's disease. Prog. Neurobiol. 155, 96–119.