Abnormal metabolic network activity in Parkinson's disease: test–retest reproducibility

Yilong Ma[1,2], **Chengke Tang**[1,2], **Phoebe G Spetsieris**[1,2], **Vijay Dhawan**[1,2], and **David Eidelberg**[1,2]

[1]Center for Neurosciences, Feinstein Institute for Medical Research, North Shore-Long Island Jewish Health System, Manhasset, New York, USA

[2]Department of Neurology and Medicine, New York University School of Medicine, Manhasset, New York, USA

Abstract

Parkinson's disease (PD) is associated with an abnormal pattern of regional brain function. The expression of this PD-related covariance pattern (PDRP) has been used to assess disease progression and the response to treatment. In this study, we validated the PDRP network as a measure of parkinsonism by prospectively computing its expression (PDRP scores) in ^{15}O-water (H$_2$ ^{15}O) and ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET) scans from PD patients and healthy volunteers. The reliability of this measure was also assessed within subjects using a test–retest design in mildly affected and advanced PD patients scanned at baseline and during treatment with levodopa or deep brain stimulation (DBS). We found that PDRP expression was significantly elevated in PD patients ($P < 0.001$) relative to controls in a prospective analysis of brain scans obtained with either H$_2$ ^{15}O or FDG PET. A significant correlation ($R^2 = 0.61$; $P < 0.001$) was evident between PDRP scores computed from H$_2$ ^{15}O and FDG images in PD subjects scanned with both tracers. Test–retest reproducibility was very high (intraclass correlation coefficient (ICC) > 0.92) for PDRP scores measured both within PET session and between sessions separated by up to 2 months. This high reproducibility was observed in both early stage and advanced PD patients scanned at baseline and during treatment. The within-subject variability of this measure was less than 10% for both unmedicated and treated conditions. These findings suggest that the PDRP network is a reproducible and stable descriptor of regional functional abnormalities in parkinsonism. The quantification of PDRP expression in PD patients can serve as a potential biomarker in PET intervention studies for this disorder.

Keywords

cerebral blood flow; glucose metabolism; PCA; PD; test–retest reliability

Correspondence: Dr Y Ma, Center for Neurosciences, Feinstein, Institute for Medical Research, North Shore-Long Island Jewish, Health System, 350 Community Drive, Manhasset, New York, 11030, USA. yma@nshs.edu.

Introduction

Brain imaging techniques have been increasingly used to link the clinical outcome of therapy to alterations in cerebral function. However, because of substantial variability in measurements of cerebral blood flow and metabolism in single brain regions, these assessments have not entered into routine use as descriptors of the treatment response. An alternative approach (e.g., Moeller and Strother, 1991; Alexander and Moeller, 1994) uses multivariate analysis to identify disease-related spatial covariance patterns, that is, brain networks with abnormal expression in patients relative to healthy subjects. This method has proved useful in imaging studies of neurodegenerative disorders in which a focal pathology affects the activity of broadly distributed neural systems (Eidelberg *et al*, 2000; Eckert and Eidelberg, 2005). Moreover, this form of network analysis can be applied prospectively to quantify pattern expression in individual subjects on a single scan basis (Trošt *et al*, 2002, 2006; Spetsieris *et al*, 2006).

Spatial covariance analysis has been used extensively in the study of Parkinson's disease (PD) (Carbon *et al*, 2003; Eckert and Eidelberg, 2005). We have found that this disease is associated with a specific spatial covariance pattern involving metabolic abnormalities in basal ganglia thalamocortical functional/anatomic pathways (Eidelberg *et al*, 1994, 1997). Indeed, this PD-related covariance pattern (PDRP) has been detected in multiple independent patient populations scanned in the resting condition (Moeller *et al*, 1999; Feigin *et al*, 2002; Lozza *et al*, 2004; Asanuma *et al*, 2005).

Significant reductions in PDRP expression have been noted with successful pharmacologic and stereotaxic surgical interventions for PD (Eckert and Eidelberg, 2005; Trošt *et al*, 2006). We have also noted that network expression increases linearly with disease progression (Huang *et al*, 2005, 2006). However, such findings can be difficult to interpret in the absence of information concerning the reliability of these measurements within and between imaging sessions. In this study, we validate PDRP expression as a disease-related biomarker in prospective patient cohorts scanned at rest with cerebral blood flow and glucose metabolism positron emission tomography (PET) techniques. We also provide data on the stability of these measures obtained in test–retest studies of patients scanned at different disease stages in the untreated condition and during pharmacological or neurosurgical interventions. This information will be useful in the design of interventional PET studies of novel therapies for PD and related neurodegenerative disorders.

Materials and methods

Subjects

Repeat PET imaging was conducted in several cohorts of subjects (Table 1). For within-session reliability, repeat scanning was performed in both healthy volunteers and PD patients using ^{15}O-water (H$_2$ ^{15}O) and PET to create blood flow images. For between-session reliability, we compared images of cerebral glucose metabolism in PD patients obtained with ^{18}F-fluorodeoxyglucose (FDG) and PET. In both comparisons, we assessed the Unified Parkinson's Disease Rating Scale (UPDRS) in PD patients who were scanned at least 12 h after the cessation of oral antiparkinsonian medications (i.e., a practically defined

'off' condition). In the within-session comparison, we also evaluated test–retest reliability in PD patients scanned in the baseline and acutely treated conditions, receiving either a stable intravenous levodopa infusion or during deep brain stimulation (DBS) of the internal globus pallidus (GPi) or the subthalamic nucleus (STN). In the between-session comparison, we also studied a group of PD patients who were scanned on stable doses of dopaminergic medications in both PET sessions.

In all subjects, a diagnosis of PD was made if the patients had 'pure' parkinsonism without a history of known causative factors such as encephalitis or neuroleptic treatment, and did not have dementia, supranuclear gaze abnormalities, or ataxia. Ethical permission for the studies was obtained from the Institutional Review Board of North Shore University Hospital. Written consent was obtained from each subject after detailed explanation of the procedures.

Positron Emission Tomography

Positron emission tomography imaging was conducted in a fasting state. H_2O and FDG studies were performed in three-dimensional (3D) mode on a GE Advance PET tomograph (GE Medical Systems, Milwaukee, WI, USA) producing 35 image slices over the whole brain with an axial coverage of 15cm and an intrinsic 3D resolution of 4.2mm. Subjects were positioned in the gantry of the scanner using a stereographic head-holder in conjunction with 3D laser alignment. To reduce repositioning errors, we used the same stereoadapter settings in both PET sessions. Photon attenuation was corrected using a 10 min 2D transmission scan with three rotating ^{68}Ge rod sources. Image acquisition was performed in a resting state in a quiet and dimly lit room. Blood flow scans were obtained over 90 secs after a bolus injection of 10 mCi H_2 ^{15}O. Metabolism scans were obtained over 10 min beginning at 35 min after intravenous injection of 5 mCi of FDG.

Within-session study—Two resting H_2 ^{15}O PET scans were acquired within 1 h in 14 normal subjects (age 52 ± 14 years; mean $\pm$ standard deviation (s.d.)) and 36 PD patients (age 59 ± 11 years; UPDRS 24 ± 16). The PD cohort consisted of 19 mild patients (Hoehn and Yahr (H&Y) Stages 1 and 2) and 17 more advanced patients (H&Y Stages 3 and 4). We also acquired repeat H_2 ^{15}O PET scans in 24 out of 36 patients who underwent PET imaging before and during acute antiparkinsonian intervention with either levodopa infusion or DBS (see Table 1). In these individuals, test–retest H_2 ^{15}O PET imaging was conducted on each of two consecutive days in which 1 day was assigned arbitrarily to the OFF condition and the other day to the ON condition. Nine of these subjects received intravenous levodopa infusion, and 15 received DBS (GPi $n = 8$; STN $n = 7$). In all subjects, treatment resulted in at least 20% improvement in UPDRS motor ratings. The details of these PET intervention studies have been provided elsewhere (Feigin *et al*, 2001; Fukuda *et al*, 2001*a*, *b*; Asanuma *et al*, 2006).

Between-session study—Two resting FDG scans were acquired one day apart in five mildly affected PD patients (age 63 ± 7 years; UPDRS 15 ± 5). These scans were conducted in a practically defined 'off' condition to assess test–retest reliability in the 2-day experimental design that we used to evaluate the metabolic effects of short-term antiparkinsonian interventions (Feigin *et al*, 2001; Fukuda *et al*, 2001*b*). Twenty other PD

patients (age 67 ± 10 years; UPDRS 30 ± 11) were scanned at rest with FDG PET while on stable oral dopaminergic therapy. This cohort included 8 mild and 12 more advanced PD patients. To evaluate the test–retest reliability of network expression during long-term treatment, we repeated FDG PET imaging 8 weeks later in each of these patients. Both scans were performed in a stable 'on' condition without alteration in daily medications.

Image Processing

Preprocessing of imaging data was performed by SPM99 software (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK; website: http://www.fil.ion.ucl.ac.uk/spm). Each type of PET images from the same subject was first realigned to correct possible motion and create a mean image between two time points. Individual images were then spatially normalized into Talairach brain space using the mean image to improve the accuracy of this procedure. Blood flow and glucose metabolism scans were transformed separately with the parameters determined after spatial realignment. The normalized images were then smoothened by a Gaussian filter (15mm for H_2 ^{15}O and 10mm for FDG) over a 3D space to increase signal to noise ratio before statistical analysis.

Network Computations

In this study, we used an automated software package available on our website (http://www.neuroscience-nslij.org) to rapidly perform principal component analysis (PCA) on groups of brain images transformed into a common anatomic space. The computational procedures of our implementation were recently described in detail (Spetsieris *et al*, 2006). Briefly, the algorithm applies PCA to the subject residual profile (Moeller and Strother, 1991):

$$\mathbf{SRP}_j = \mathbf{P}_j - GMR_j\mathbf{I} - GMP = \Sigma_k \sqrt{(\lambda_k)} SSF_{kj}\mathbf{GIS_k} \quad (1)$$

where $\mathbf{P}_j$ refers to the log-transformed brain image matrix of subject j and GMR_j is the global mean rate of $\mathbf{P}_j$. $\mathbf{I}$ is the identity matrix. GMP is the group mean profile computed by averaging the difference between $\mathbf{P}_j$ and GMR_j over subjects. This is a characteristic of the group measuring the mean residual of all log-transformed brain images. $\mathbf{SRP}_j$ of each subject j is then transformed into a weighted sum of a set of principal component images $\mathbf{GIS_k}$ multiplied by the corresponding subject scalar factor $\mathbf{SSF_{kj}}$ and the square root of the eigenvalue λ_k. Each network $\mathbf{GIS_k}$ represents a group invariant subprofile of spatial covariance where the percent variance accounted for by each component is given by $vaf_k = 100\lambda_k / \lambda_k$. A disease-related pattern was defined when subject scores for the resulting principal components (singly or in linear combination) discriminated patients from controls at $P < 0.001$ (Moeller *et al*, 1999).

The expression of a given network pattern in an individual subject can be quantified prospectively using a voxel-based topographic profile rating (TPR) algorithm. Subject score for a network pattern k is calculated by

$$SCORE_{kj} = \mathbf{SRP}_j^T \mathbf{GIS_k} \quad (2)$$

where $\mathbf{SRP}_j^{\mathrm{T}}$ is the transpose of SRP_j for subject j. This matrix product is summed and divided by the total number of voxels within the brain to yield a subject-specific scalar factor. SRP_j is generally computed according to the left-hand side of equation (1) based on GMP from a new group of subjects. We have introduced a simple method to obtain $\mathbf{SRP}_j$ for a single subject by using the GMP image associated with the derivation of the original pattern. This allows us to establish an objective and reproducible scale for network expression in brain images of any individual subjects. All computations described above are usually performed within a user-defined brain mask to remove areas of least interest and noise.

In this study, we assessed the test–retest reproducibility of a PD-related spatial covariance pattern analogous to that reported by us previously (Asanuma *et al*, 2005; Trošt *et al*, 2006). We identified a PDRP on a voxel basis by applying PCA to resting FDG PET images from 33 PD patients (age 57 ± 8 years; UPDRS 32 ± 16) and 33 age-matched normal controls (age 55 ± 13 years). The scans from these subjects were used solely for pattern identification. These subjects were entirely different from those mentioned above in whom the PDRP was prospectively quantified in test–retest scan pairs.

The pattern (Figure 1A) was characterized by increases in pallidothalamic, pontine, and cerebellar metabolic activity associated with relative reductions in the lateral premotor cortex, supplementary motor area, and posterior association cortices. This pattern was topographically equivalent to that described by us previously in the analysis of 20 PD patients and 20 controls (Asanuma *et al*, 2005). A highly significant correlation was observed between the region weights for the two PDRP networks ($R^2 = 0.92$, $P < 0.0001$) as well as between their associated subject scores ($R^2 = 0.98$, $P < 0.0001$)). In this analysis, the PDRP was identified as the first principal component, accounting for 21% of subjects × voxel variance. Subject scores for this pattern discriminated PD patients from control subjects ($P < 0.0001$; see Figure 1B). We validated this pattern by computing its expression in the FDG PET scans from 32 subsequent PD patients (age 59 ± 8 years; UPDRS 36 ± 18). These network calculations were performed on a prospective case basis using the computational algorithm described above. The PDRP scores of these subjects were compared with those of the original subjects whose scans were used to identify the disease-related pattern.

We also determined whether PDRP quantification can be achieved with H_2 ^{15}O PET. We first assessed the comparability of network quantification of metabolism-based patterns in blood flow data. To this end, we studied nine healthy volunteers (age 56 ± 14 years) and 36 PD patients (58 ± 10 years; UPDRS 21 ± 17). These subjects were scanned at rest in a practically defined 'off' condition with the two tracers administered within 24 h of each other. We computed PDRP scores on a prospective basis in both scans and correlated the values obtained with blood flow and metabolism scans for the same subjects. We then determined whether PDRP scores computed in the H_2 ^{15}O PET scans of PD patients were abnormally elevated relative to values computed in healthy controls. This was achieved by comparing the PDRP scores obtained in the H_2 ^{15}O PET scans of the 36 PD patients with those obtained in the corresponding scans of a subset of the 14 normal subjects used in the

within-session test–retest study (see above). This control group was comprised of 11 healthy volunteers (57 ± 11 years) matched in age to the PD group (59 ± 11 years; $P = 0.6$).

Having established that PDRP quantification can be performed using either FDG or $H_2 {}^{15}O$ PET imaging, we assessed the reproducibility of network expression in the test–retest scan pairs of the subjects who underwent repeat testing. As noted above, the within-session reproducibility of pattern expression was based on repeat $H_2 {}^{15}O$ PET imaging; the between-session validation involved repeat FDG PET scanning.

Statistical Procedures

Parkinson's disease-related covariance pattern scores computed in the blood flow and metabolism scan data from the same subjects were compared using regression analysis. Differences in PDRP scores between groups were assessed by one-way analysis of variance (ANOVA) followed by *post hoc* tests with Bonferroni correction. All comparisons were considered significant with $P < 0.05$.

We assessed the within-subject reproducibility of PDRP scores in the test–retest data for each group by computing the within-subject standard deviation (WSD), the within-subject coefficient of variation (WCOV), and the intraclass correlation coefficient (ICC) (Shrout and Fleiss, 1979; Bland and Altman, 1996; Floyd *et al*, 2003; Coles *et al*, 2006). To use these measures of reliability, we tested the following two assumptions in each group. The Shapiro–Wilk *W*-test was performed to determine whether the normal distribution applied to the differences between the test and retest PDRP scores. The nonparametric rank correlation test (Kendall's τ_b) was also performed to ensure that no correlation existed between the differences of the test and retest PDRP scores and the magnitude of these scores.

The following calculations were then conducted for each group:

1. The subject variance of the two test–retest PDRP scores was the square of the difference divided by two, and the subject s.d. was the square root of the variance.

2. The WSD was the square root of the mean subject variance.

3. The WCOV was $100 \times$ (the mean of the (subject s.d./subject mean PDRP score)).

The ICC and its 95% confidence interval (CI) for each group were computed as described previously in test–retest studies of striatal dopaminergic imaging measures (Vingerhoets *et al*, 1996; Nurmi *et al*, 2000; Tsuchida *et al*, 2004). All statistical analyses were performed using SPSS program for Windows (SPSS Inc., Chicago, IL, USA).

Results

Validation in Prospective Cohorts

Glucose metabolism—To validate PDRP expression as a disease marker, we computed subject scores for this pattern in the FDG PET scans of 32 subsequent patients. One-way ANOVA revealed a significant difference in PDRP expression across the three groups (Figure 1B, $F(2,97) = 44.7$; $P < 0.001$). *Post hoc* testing showed that PDRP scores were

abnormally elevated ($P < 0.001$) in both the original and the prospective PD groups. Mean network expression did not differ ($P = 0.95$) between the two PD groups.

Cerebral blood flow—Parkinson's disease-related covariance pattern scores computed in $H_2 {}^{15}O$ and FDG PET scans were highly intercorrelated in the PD group ($R^2 = 0.61$; $P < 0.001$; Figure 2) and in the combined group of PD patients and healthy volunteers ($R^2 = 0.62$; $P < 0.001$). Network scores obtained with both tracers were also significantly intercorrelated in the normal group ($R^2 = 0.42$; $P = 0.05$). We also found that PDRP expression was elevated in the $H_2 {}^{15}O$ PET scans of PD patients relative to age-matched controls ($t = 3.5$, $P < 0.001$; unpaired Student's t-test).

Test–Retest Reliability

The differences between the test–retest PDRP scores were normally distributed in all subject groups ($P > 0.09$). The correlation between the difference and the magnitude of the test–retest PDRP scores was not significant in any of the groups ($P > 0.08$). Therefore, our data satisfied the assumptions for the measures of test–retest reliability used in this study as described above. These reproducibility measures are present in Table 2 and Figure 3.

Within-session reliability—The computed PDRP scores exhibited excellent within-session reliability in all groups scanned with repeat $H_2 {}^{15}O$ PET imaging (Table 2). Specifically, low within-subject variability and high test–retest reproducibility were observed in the healthy volunteer group ($n = 14$: WSD = 1.2; WCOV = 9.9%; ICC = 0.94 with 95% CI: 0.84 to 0.98) and in the unmedicated PD group ($n = 36$: WSD = 1.6; WCOV = 6.4%; ICC = 0.96 with 95% CI: 0.92 to 0.98). Test–retest reliability was similar for the mild and advanced PD subgroups.

Test–retest reliability was also excellent for PD patients scanned before and during acute antiparkinsonian interventions. PD-related covariance pattern scores for scans acquired off treatment showed low within-subject variability (WSD < 2; WCOV < 7%) and high reproducibility (ICC = 0.95 with 95% CI: 0.89 to 0.98). PD-related covariance pattern scores for PD patients treated with levodopa infusion or with either GPi or STN stimulation were also associated with low within-subject variability (WSD < 1.5; WCOV < 5%) and high reproducibility (ICC = 0.98 with 95% CI: 0.96 to 0.99). The test–retest reliability of PDRP scores measured before and during acute therapy was comparable for the levodopa group and the two DBS treatment groups.

Between-session reliability—The computed PDRP scores exhibited excellent between-session reliability. Network expression was found to be highly reproducible in the unmedicated PD patients who underwent repeat FDG PET sessions a day apart (WSD = 0.9; WCOV = 4.9%; ICC = 0.99 with 95% CI: 0.94 to 0.99). PD-related covariance pattern scores were also highly reproducible in the medicated PD patients undergoing repeat FDG PET imaging 2 months apart (WSD = 2.3; WCOV = 8.7%; ICC = 0.96 with 95% CI: 0.89 to 0.98). Mild and advanced patients exhibited a similar degree of high reliability. However, within-subject variability was relatively greater for the group with the longer between-scan interval.

Discussion

In this study, we evaluated the reproducibility of a specific disease-related spatial covariance pattern as an imaging biomarker of Parkinson's disease. In the first part of the study, we showed that elevated PDRP expression was not limited to the original group of PD patients in whom the pattern was originally identified. Indeed, we found that abnormal network activity was present in prospective patient cohorts scanned with either FDG or H_2 ^{15}O PET techniques. These findings validate the PDRP as a generalizable marker of disease across populations.

We note that PDRP activity can be assessed in functional images of cerebral blood flow as well as glucose metabolism. Although absolute quantification of cerebral blood flow and glucose utilization was not performed in this study, a high correlation was observed between PDRP scores obtained from resting H_2 ^{15}O and FDG PET scans acquired in the same patients. The corresponding correlation in healthy volunteer subjects was also significant but of lower magnitude than for the patients. We attribute this to the relatively narrow range of PDRP values in healthy subjects. The presence of abnormal increases in PDRP expression in the blood flow images of PD patients is consistent with our previous study showing comparable elevations in patients scanned with ^{99m}Tc ECD SPECT perfusion techniques (Feigin *et al*, 2002). Indeed, preliminary data from our laboratory suggest that similar group differences in PDRP expression are present in patients and healthy volunteers scanned with perfusion-weighted MRI techniques.

Despite the similarity of PDRP expression in PET images of blood flow and glucose metabolism, the correlation of these values is not perfect. Indeed, our data suggest that approximately 40% of the intersubject variability is not accounted for between these measures (Figure 2). While this discrepancy may be attributed to differences in the signal-to-noise characteristics of the two PET tracer methods, we note that the within-subject COVs are similar in PDRP scores obtained in the blood flow and metabolism scans. Alternatively, it is possible that these differences are biologic in origin. Specifically, an uncoupling of blood flow and energy demand may be present in one or more regions within the PDRP network in patients scanned in the resting condition. Indeed, we have recently observed a dissociation of cerebral blow flow and glucose utilization in the globus pallidus and premotor cortex of unmedicated PD patients. The clinical relevance of these findings is currently not understood.

In the second part of this study, we assessed the within-subject reproducibility of the PDRP network measures. We found that prospectively computed PDRP scores showed excellent test–retest reliability in H_2 ^{15}O PET scans conducted within 1 h of each other. Indeed, network scores were highly reproducible in healthy volunteer subjects as well as in PD patients at mild and more advanced clinical stages scanned in the untreated condition. Moreover, these values were also similarly reliable during treatment, whether during a stable intravenous levodopa infusion or with either GPi or STN stimulation. The high within-session reproducibility shows that PDRP scores obtained from cerebral blood flow (CBF) data can potentially be used to assess the acute effects of novel antiparkinsonian interventions with functional brain imaging.

We also found that the test–retest reproducibility of this measure was excellent across FDG PET sessions separated by as long as 2 months. In PD patients scanned off medication, the reproducibility of PDRP scores was comparably high over a between-session interval of 1 day. Despite the small sample size, this finding is also in close agreement with the within-session $H_2{}^{15}O$ PET data for scanning intervals of approximately 1 h. Indeed, these results lend credence to the finding of PDRP suppression during acute dopaminergic therapy or DBS, in which the treatment effects on network activity were assessed over 1 to 2 days (Feigin *et al*, 2001; Fukuda *et al*, 2001*a*, *b*; Asanuma *et al*, 2006). In contrast to FDG PET, PDRP quantification in $H_2{}^{15}O$ PET scans may not be as stable over comparatively longer time intervals. To illustrate this point, we measured the test–retest reliability of PDRP scores obtained in the resting $H_2{}^{15}O$ PET scans of the 20 patients used to assess between-session reproducibility in the chronically medicated state. In these subjects, PDRP expression measured in CBF images was less reliable (ICC = 0.81 with 95% CI: 0.58 to 0.92) than network values computed in FDG PET scans separated by the same 2-month interval. The better between-session reproducibility observed with FDG PET supports the use of this approach in evaluating the long-term effects of treatment on network activity.

The stability of PDRP expression contrasts with the general degree of within-subject variability that has been observed with routine measurements of regional and global CBF and glucose utilization (Bartlett *et al*, 1988; Maquet *et al*, 1990; Matthew *et al*, 1993; Coles *et al*, 2006). Indeed, we assessed the test–retest reliability of individual regions of interest (ROIs) as well as global measures in the same scan pairs that were used for the PDRP assessments. We found that PDRP scores had better test–retest reproducibility than ROI or global measures obtained in the $H_2{}^{15}O$ PET data (ICC = 0.84 to 0.88 for PD and 0.71 to 0.87 for controls as compared with PDRP scores that had ICC = 0.96 for PD and 0.94 for controls) as well as in the FDG PET data (ICC = 0.53 to 0.68 versus PDRP scores with ICC = 0.96 for PD). Likewise, the reproducibility of network expression was superior to that of globally normalized ROI values, whether evaluated within or between PET sessions. These findings suggest that the effects of pathology (i.e., PDRP expression) are greater than other factors that influence the variation in regional brain function that is observed within individual subjects.

The within-subject variability of the computed PDRP scores in our population was under 10%, analogous to that reported with dopaminergic ligands to quantify nigrostriatal function with PET or SPECT imaging (Seibyl *et al*, 1997; Booij *et al*, 1998; Nurmi *et al*, 2000; Hwang *et al*, 2004; Tsuchida *et al*, 2004). In addition, we note that the variability of the network measure in the 20 PD patients scanned over a 2-month period was somewhat greater than for the groups with shorter between-scan intervals. Nonetheless, even with this comparatively large between-scan separation, the PDRP scores were quite stable, with ICC values between 0.92 and 0.99. Indeed, these values were comparable to those observed with dopaminergic ligands for the assessment of parkinsonism (Nurmi *et al*, 2000; Hwang *et al*, 2004; Tsuchida *et al*, 2004).

Despite similarly high test–retest reliability, imaging measures of presynaptic dopaminergic dysfunction and disease-related network descriptors are not interchangeable biomarkers of PD. Dopaminergic radiotracer imaging can provide information on the effects of potential

disease-modifying agents relating specifically to the nigrostriatal system (Ravina *et al*, 2005). By contrast, network quantification can be useful in objectively assessing interventions in which treatment modulates an entire neural system (Eckert and Eidelberg, 2005; Trošt *et al*, 2006). However, it is important to note that although treatment-mediated changes in PDRP expression correlate significantly with improvement in UPDRS ratings, the metabolic and clinical disease descriptors are not interchangeable. Indeed, intersubject differences in these measures generally have less than 50% of their variability in common, as observed in correlations between PDRP scores and motor UPDRS ratings obtained at baseline or during therapy (e.g., Feigin *et al*, 2002; Lozza *et al*, 2004; Trošt *et al*, 2006). Thus, it is likely that the objective imaging-based network measure may provide unique information about disease severity or the treatment response that is not captured by the simpler and more accessible clinical rating scales.

The use of network quantification in gauging the effects of potential disease-modifying interventions is more challenging. This approach would have limited value in neuroprotection trials if the effects of symptomatic treatment on network expression proved to be irreversible. However, we have recently found such effects to be short-lived, with PDRP scores returning to baseline off-state values within 18 h after levodopa administration. Sufficient data are not currently available to exclude a long duration effect of dopaminergic therapy on PDRP suppression. However, it has been possible to detect highly significant longitudinal changes in network expression in medicated patients scanned with FDG PET after only 12 to 18 h of washout (Huang *et al*, 2005, 2006). In keeping with our findings in the baseline condition and with treatment (see above), the rate of increase in PDRP expression over a 4-year period correlated significantly ($R^2 \sim 0.36$, $P < 0.01$) with concurrent longitudinal changes in putamen DAT binding and motor UPDRS ratings. These findings support the use of network quantification methods in studies of the natural history of PD, and potentially in the evaluation of neuroprotection strategies to modify the course of disease. The assessment of the relative sensitivities of the different clinical and imaging descriptors to disease progression remains a topic of investigation.

In summary, this study validates the use of the PDRP metabolic network as a potential functional imaging-based biomarker for PD. We found that the expression of this brain network is highly reproducible both within subjects and across patient populations. Moreover, PDRP scores proved to be reliable irrespective of disease stage and treatment status. The excellent reproducibility of this disease-related network measure within- and between-PET sessions supports its use in the assessment of novel symptomatic treatments for PD as well as its potential application in studies of the natural history of this disorder.

Acknowledgements

References

Alexander G, Moeller J. Application of the scaled subprofile model to functional imaging in neuropsychiatric disorders: a prinicipal component approach to modeling brain function in disease. Hum Brain Mapping. 1994; 2:1–16.

Asanuma K, Ma Y, Huang C, Carbon M, Edwards C, Raymond D, Bressman S, Moeller JR, Eidelberg D. The metabolic pathologyof dopa-responsive dystonia. Ann Neurol. 2005; 57:596–600. [PubMed: 15786454]

Asanuma K, Tang C, Ma Y, Dhawan V, Mattis P, Edwards C, Kaplitt MG, Feigin A, Eidelberg D. Network Modulation in the Treatment of Parkinson's Disease. Brain. 2006 in press.

Bartlett EJ, Brodie JD, Wolf AP, Christman DR, Laska E, Meissner M. Reproducibility of cerebral glucose metabolic measurements in resting human subjects. J Cereb Blood Flow Metab. 1988; 8:502–512. [PubMed: 3260593]

Bland JM, Altman DG. Measurement error. BMJ. 1996; 313:744. [PubMed: 8819450]

Booij J, Habraken JB, Bergmans P, Tissingh G, Winogrodzka A, Wolters EC, Janssen AG, Stoof JC, van Royen EA. Imaging of dopamine transporters with iodine-123-FP-CIT SPECT in healthy controls and patients with Parkinson's disease. J Nucl Med. 1998; 39:1879–1884. [PubMed: 9829575]

Carbon M, Edwards C, Eidelberg D. Functional brain imaging in Parkinson's disease. Adv Neurol. 2003; 91:175–181. [PubMed: 12442676]

Coles JP, Fryer TD, Bradley PG, Nortje J, Smielewski P, Rice K, Clark JC, Pickard JD, Menon DK. Intersubject variability and reproducibility of (15)O PET studies. J Cereb Blood Flow Metab. 2006; 26:48–57. [PubMed: 15988475]

Eckert T, Eidelberg D. Neuroimaging and therapeutics in movement disorders. Neurorx. 2005; 2:361–371. [PubMed: 15897956]

Efron, B.; Tibshirani, R. An introduction to the bootstrap. New York: CRC Press, LLC; 1994.

Eidelberg, D.; Edwards, C.; Mentis, M.; Dhawan, V.; Moeller, J. Movement disorders: Parkinson's disease. In: Mazziotta, JC.; Toga, AW.; Frackowiak, R., editors. Brain mapping: the disorders. San Diego: Academic Press; 2000. p. 241-261.

Eidelberg D, Moeller JR, Dhawan V, Spetsieris P, Takikawa S, Ishikawa T, Chaly T, Robeson W, Margouleff D, Przedborski S, et al. The metabolic topography of parkinsonism. J Cereb Blood Flow Metab. 1994; 14:783–801. [PubMed: 8063874]

Eidelberg D, Moeller JR, Kazumata K, Antonini A, Sterio D, Dhawan V, Spetsieris P, Alterman R, Kelly PJ, Dogali M, Fazzini E, Beric A. Metabolic correlates of pallidal neuronal activity in Parkinson's disease. Brain. 1997; 120:1315–1324. [PubMed: 9278625]

Feigin A, Antonini A, Fukuda M, De Notaris R, Benti R, Pezzoli G, Mentis MJ, Moeller JR, Eidelberg D. Tc-99m ethylene cysteinate dimer SPECT in the differential diagnosis of parkinsonism. Mov Disord. 2002; 17:1265–1270. [PubMed: 12465066]

Feigin A, Fukuda M, Dhawan V, Przedborski S, Jackson-Lewis V, Mentis MJ, Moeller JR, Eidelberg D. Metabolic correlates of levodopa response in Parkinson's disease. Neurology. 2001; 57:2083–2088. [PubMed: 11739830]

Floyd TF, Ratcliffe SJ, Wang J, Resch B, Detre JA. Precision of the CASL-perfusion MRI technique for the measurement of cerebral blood flow in whole brain and vascular territories. J Magn Reson Imaging. 2003; 18:649–655. [PubMed: 14635149]

Fukuda M, Mentis M, Ghilardi MF, Dhawan V, Antonini A, Hammerstad J, Lozano AM, Lang A, Lyons K, Koller W, Ghez C, Eidelberg D. Functional correlates of pallidal stimulation for Parkinson's disease. Ann Neurol. 2001a; 49:155–164. [PubMed: 11220735]

Fukuda M, Mentis MJ, Ma Y, Dhawan V, Antonini A, Lang AE, Lozano AM, Hammerstad J, Lyons K, Koller WC, Moeller JR, Eidelberg D. Networks mediating the clinical effects of pallidal brain stimulation for Parkinson's disease: a PET study of resting-state glucose metabolism. Brain. 2001b; 124:1601–1609. [PubMed: 11459751]

Huang C, Carbon M, Mattis P, Eidelberg D. Metabolic patterns associated with cognitive function in Parkinson's disease. Mov Disord. 2006; 21:S104.

Huang C, Feigin A, Ma Y, Eidelberg D. Imaging measures of longitudinal change in Parkinson's disease. Neurology. 2005; 64:A235.

Hwang WJ, Yao WJ, Wey SP, Ting G. Reproducibility of ^{99m}Tc-TRODAT-1 SPECT measurement of dopamine transporters in Parkinson's disease. J Nucl Med. 2004; 45:207–213. [PubMed: 14960637]

Lozza C, Baron JC, Eidelberg D, Mentis MJ, Carbon M, Marie RM. Executive processes in Parkinson's disease: FDG-PET and network analysis. Hum Brain Mapp. 2004; 22:236–245. [PubMed: 15195290]

Maquet P, Dive D, Salmon E, von Frenckel R, Franck G. Reproducibility of cerebral glucose utilization measured by PET and the [^{18}F]-2-fluoro-2-deoxy-d-glucose method in resting, healthy human subjects. Eur J Nucl Med. 1990; 16:267–273. [PubMed: 2351175]

Matthew E, Andreason P, Carson RE, Herscovitch P, Pettigrew K, Cohen R, King C, Johanson CE, Paul SM. Reproducibility of resting cerebral blood flow measurements with H$_2$(15)O positron emission tomography in humans. J Cereb Blood Flow Metab. 1993; 13:748–754. [PubMed: 8360281]

Moeller JR, Nakamura T, Mentis MJ, Dhawan V, Spetsieres P, Antonini A, Missimer J, Leenders KL, Eidelberg D. Reproducibility of regional metabolic covariance patterns: comparison of four populations. J Nucl Med. 1999; 40:1264–1269. [PubMed: 10450676]

Moeller JR, Strother SC. A regional covariance approach to the analysis of functional patterns in positron emission tomographic data. J Cereb Blood Flow Metabol. 1991; 11:A121–A135.

Nurmi E, Bergman J, Eskola O, Solin O, Hinkka SM, Sonninen P, Rinne JO. Reproducibility and effect of levodopa on dopamine transporter function measurements: a [^{18}F]CFT PET study. J Cereb Blood Flow Metab. 2000; 20:1604–1609. [PubMed: 11083235]

Ravina B, Eidelberg D, Ahlskog JE, Albin RL, Brooks DJ, Carbon M, Dhawan V, Feigin A, Fahn S, Guttman M, Gwinn-Hardy K, McFarland H, Innis R, Katz RG, Kieburtz K, Kish SJ, Lange N, Langston JW, Marek K, Morin L, Moy C, Murphy D, Oertel WH, Oliver G, Palesch Y, Powers W, Seibyl J, Sethi KD, Shults CW, Sheehy P, Stoessl AJ, Holloway R. The role of radiotracer imaging in Parkinson disease. Neurology. 2005; 64:208–215. [PubMed: 15668415]

Seibyl JP, Marek K, Sheff K, Baldwin RM, Zoghbi S, Zea-Ponce Y, Charney DS, van Dyck CH, Hoffer PB, Innis RB. Test/retest reproducibility of iodine-123-betaCIT SPECT brain measurement of dopamine transporters in Parkinson's patients. J Nucl Med. 1997; 38:1453–1459. [PubMed: 9293807]

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psych Bull. 1979; 86:420–428.

Spetsieris P, Ma Y, Dhawan V, Moeller JR, Eidelberg D. Highly automated computer-aided diagnosis of neurological disorders using functional brain imaging. Proc SPIE: Med Imag. 2006; 6144:5M1–5M12.

Trošt M, Carbon M, Edwards C, Raymond D, Mentis M, Moeller JR, Bressman SB, Eidelberg D. Primary dystonia: is abnormal functional brain architecture linked to genotype? Ann Neurol. 2002; 52:853–856. [PubMed: 12447944]

Trošt M, Su S, Su PC, Yen R-F, Tseng H-M, Barnes A, Ma Y, Eidelberg D. Network modulation by the sub-thalamic nucleus in the treatment of Parkinson's disease. NeuroImage. 2006; 31:301–307. [PubMed: 16466936]

Tsuchida T, Ballinger JR, Vines D, Kim YJ, Utsunomiya K, Lang AE, Ichise M. Reproducibility of dopamine transporter density measured with ^{123}I-FPCIT SPECT in normalcontrol and Parkinson's disease patients. Ann Nucl Med. 2004; 18:609–616. [PubMed: 15586635]

Vingerhoets FJ, Schulzer M, Ruth TJ, Holden JE, Snow BJ. Reproducibility and discriminating ability of fluorine-18-6-fluoro-l-Dopa PET in Parkinson's disease. J Nucl Med. 1996; 37:421–426. [PubMed: 8772636]
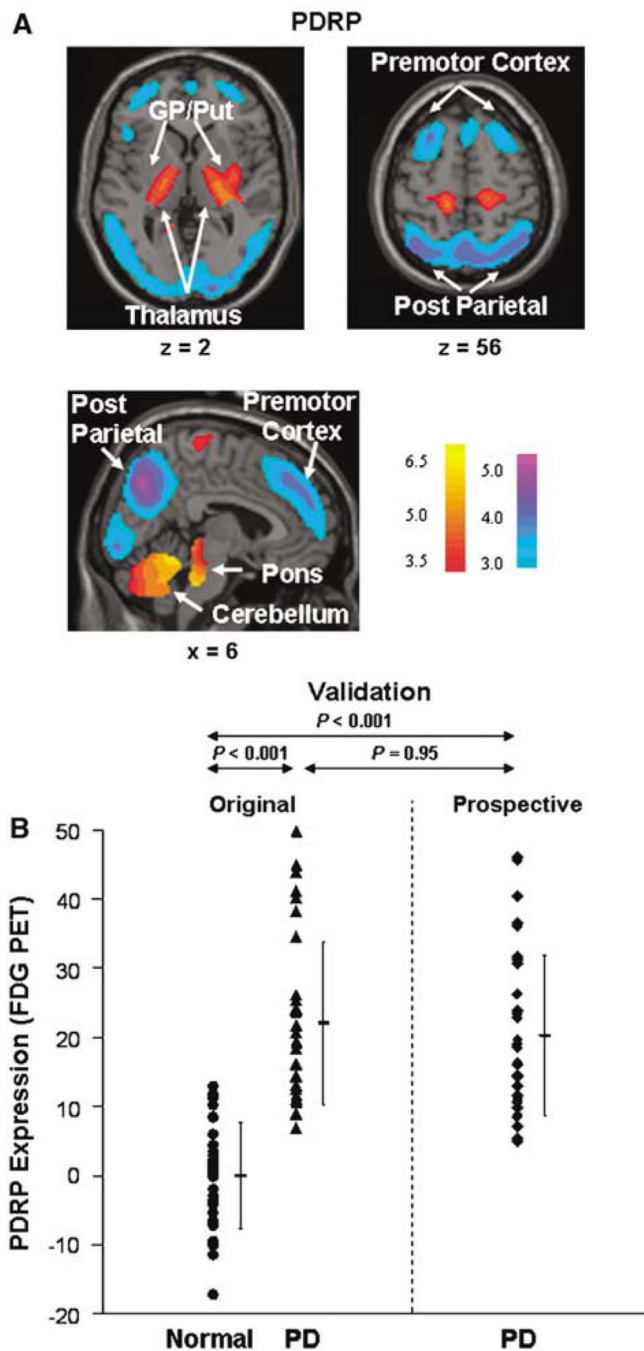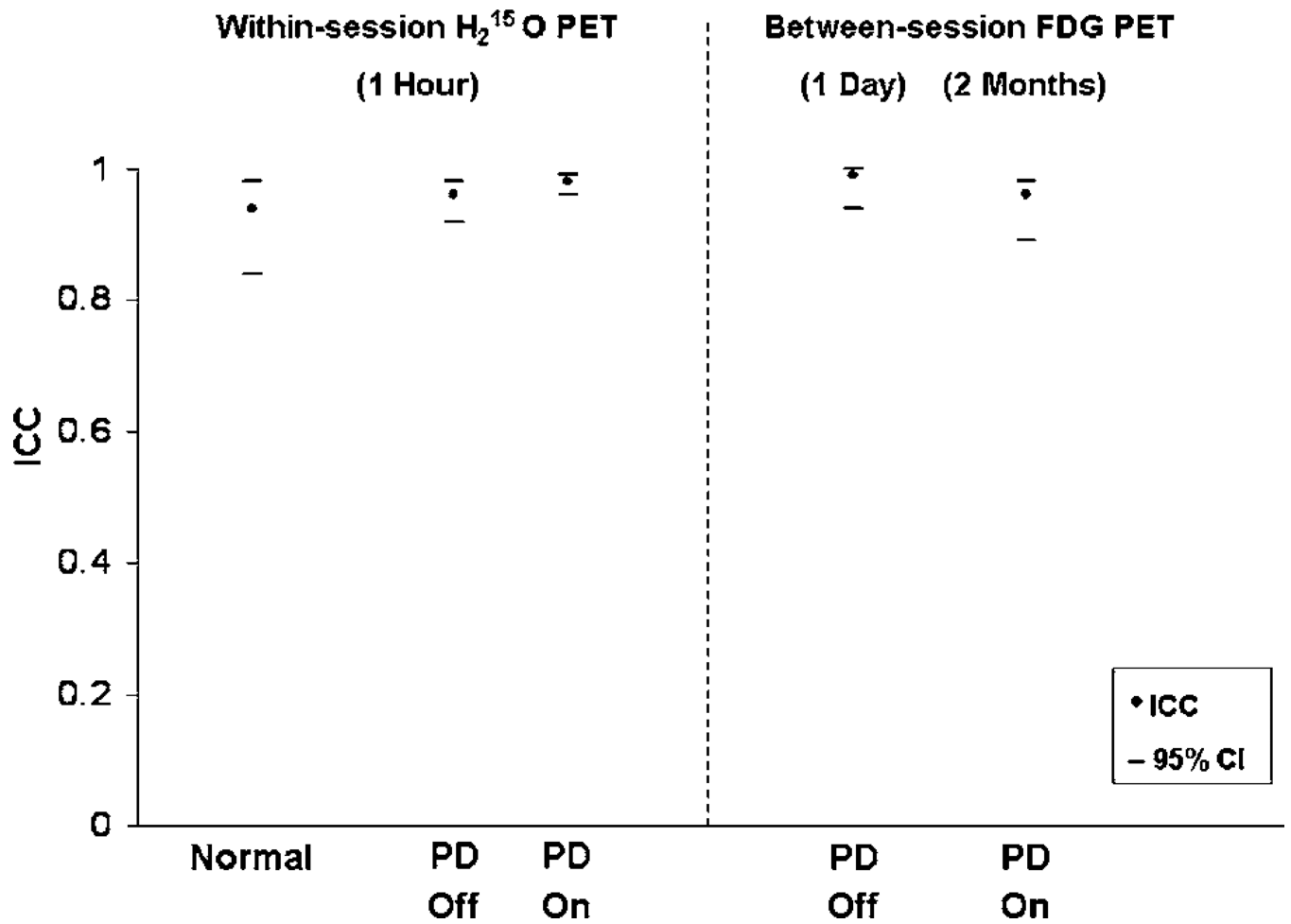
Figure 1.

(**A**) Parkinson's disease-related pattern (PDRP) identified by network analysis of FDG PET scans from 33 PD patients and 33 age-matched normal volunteers (see text). This spatial covariance pattern was characterized by relative increases in pallidothalamic, pontine, and cerebellar metabolism, associated with decreases in the premotor and posterior parietal areas. (The display represents voxels that contribute significantly to the network at $P = 0.001$ and that were showed to be reliable ($P < 0.001$) by bootstrap estimation (Efron and Tibshirani, 1994). Voxels with positive region weights (metabolic increases) are color coded

from red to yellow; those with negative region weights (metabolic decreases) are color coded from blue to purple.) (**B**) Parkinson's disease-related pattern expression (subject scores) was increased in the PD patients (left, $P < 0.001$) relative to the normal subjects. It remained high in a new group of PD patients (right, $P < 0.001$) not included in the identification of the PDRP shown in Figure 1A. Error bars represent s.d.'s.
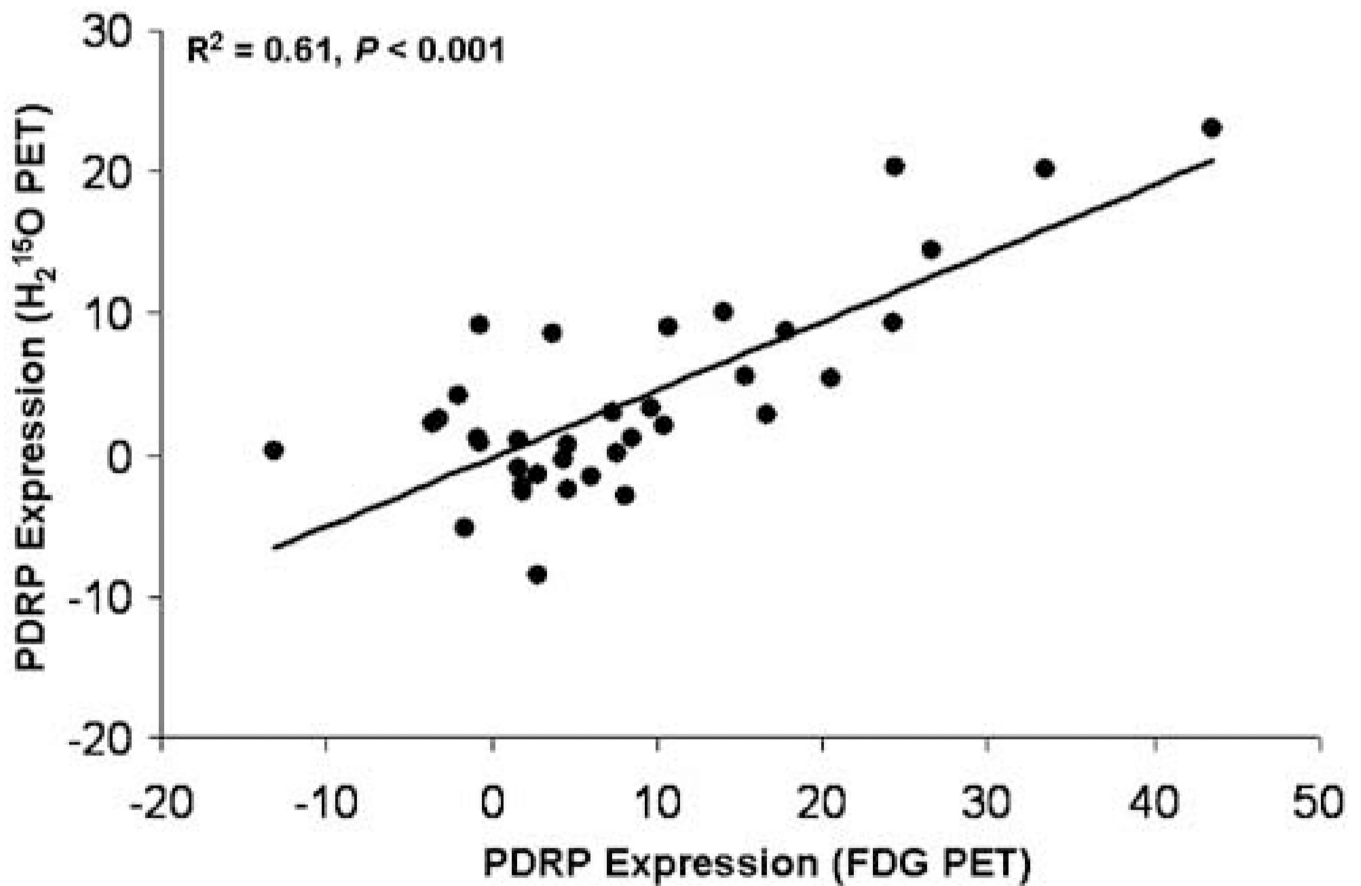
Figure 2.
Results of regression analysis showing a significant linear relationship between PDRP subject scores computed from $H_2\,^{15}O$ and FDG brain scans (see text). The plot showed data from 36 PD patients scanned off medication ($F(1,34) = 52.4$; linear equation $Y = -0.22 + 0.48X$). A less-significant correlation was seen in nine normal controls.

Figure 3.
Summary plot of high test–retest reliability of PDRP subject scores showing intraclass correlation coefficients (filled symbols) and their 95% confidence intervals (horizontal bars) from normal controls and four subgroups of PD patients (see Table 2). The within-session data were from H_2 ^{15}O PET scans repeated twice within 1 h for each treatment condition. The between-session data were obtained from FDG PET scans repeated 1 day (off medications) and 2 months (on medications) apart.

Table 1

Demographic characteristics of patients and controls

	N	Age (years)	Gender (M/F)	UPDRS (motor)
Within-session (H$_2$^{15}O PET)				
Off				
Normal	14	51.8 ± 13.8	6/8	
PD (mild)	19	55.9 ± 10.6	14/5	14.8 ± 12.4
PD (advanced)	17	62.1 ± 10.1	12/5	34.5 ± 14.1
PD (mild+advanced)	36	58.8 ± 10.7	26/10	24.1 ± 16.4
PD (LD)	9	61.7 ± 8.0	7/2	25.8 ± 8.9
PD (GPi)	8	51.4 ± 10.9	5/3	42.9 ± 16.7
PD (STN)	7	67.0 ± 8.4	5/2	28.0 ± 10.1
PD (LD+GPi+STN)	24	59.8 ± 10.9	17/7	32.1 ± 14.2
On				
PD (LD)	8	59.6 ± 5.5	7/1	16.8 ± 6.2
PD (GPi)	9	52.1 ± 10.4	6/3	30.4 ± 11.7
PD (STN)	6	65.2 ± 7.5	5/1	19.3 ± 7.5
PD (LD+GPi+STN)	23	58.1 ± 9.5	18/5	22.8 ± 10.7
Between-session (FDG PET)				
Off				
PD (mild)	5	62.5 ± 7.1	3/2	15.1 ± 5.4
On				
PD (mild)	8	59.9 ± 9.5	5/3	19.7 ± 3.9
PD (advanced)	12	70.8 ± 7.8	6/6	35.8 ± 9.2
PD (mild+advanced)	20	66.5 ± 9.9	11/9	29.8 ± 10.9

Age and UPDRS motor ratings are given as mean ± s.d. Mild disease refers to Stages 1 and 2 according to the Hoehn and Yahr Rating Scale for Parkinson's disease (PD). Advanced disease refers to Stages 3 and 4 according to this scale. The within-session study included PD patients scanned twice before and during acute therapy with levodopa (LD) or with deep brain stimulation of the internal globus pallidus (GPi) or the subthalamic nucleus (STN).

Table 2

Test–retest reproducibility of PDRP network activity

	WSD	WCOV (%)	ICC	95% CI
Within-session ($H_2^{15}O$ PET)				
Off				
Normal	1.2	9.9	0.94	(0.84, 0.98)
PD (mild)	1.2	6.2	0.97	(0.92, 0.99)
PD (advanced)	1.9	6.6	0.92	(0.81, 0.97)
PD (mild+advanced)	1.6	6.4	0.96	(0.92, 0.98)
PD (LD)	1.6	6.8	0.96	(0.83, 0.99)
PD (GPi)	1.9	7.0	0.95	(0.80, 0.99)
PD (STN)	1.8	6.1	0.94	(0.75, 0.99)
PD (LD+GPi+STN)	1.8	6.7	0.95	(0.89, 0.98)
On				
PD (LD)	1.3	4.6	0.98	(0.91, 0.99)
PD (GPi)	0.9	3.8	0.99	(0.94, 0.99)
PD (STN)	0.9	3.4	0.99	(0.94, 0.99)
PD (LD+GPi+STN)	1.1	4.0	0.98	(0.96, 0.99)
Between-session (FDG PET)				
Off				
PD (mild)	0.9	4.9	0.99	(0.94, 0.99)
On				
PD (mild)	1.9	9.7	0.98	(0.89, 0.99)
PD (advanced)	2.5	8.1	0.94	(0.82, 0.98)
PD (mild+advanced)	2.3	8.7	0.96	(0.89, 0.98)

The columns represent within-subject standard deviation (WSD), within-subject coefficient of variation (WCOV), intraclass correlation coefficient (ICC), and its 95% confidence interval (CI). The within-session study included PD patients scanned twice before and during acute therapy with levodopa (LD) or with deep brain stimulation of the internal globus pallidus (GPi) or the subthalamic nucleus (STN). Parkinson's disease-related covariance pattern scores were obtained on a prospective individual case basis using an automated routine, masked to diagnostic category and treatment condition (see text).