Parkinson´s Disease

Vera S Kerstens[1], Patrik Fazio[1], Mathias Sundgren[2,3], Granville J Matheson[1], Erika Franzén[4,5], Christer Halldin[1], Simon Cervenka[1], Per Svenningsson[2,3], Andrea Varrone[1]

[1] Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, & Stockholm Health Care Services, Region Stockholm, Karolinska University Hospital, Sweden.

[2]Department of Clinical Neuroscience, Division of Neuro, Karolinska Institutet, Stockholm, Sweden.

[3]Neurology Department, Karolinska University Hospital, Stockholm, Sweden.

[4]Department of Neurobiology, Care Sciences and Society, Division of Physiotherapy, Karolinska Institutet, Stockholm, Sweden.

[5]Function Area Occupational Therapy & Physiotherapy, Allied Health Professionals Function, Karolinska University Hospital, Stockholm, Sweden.

Corresponding author

Vera Kerstens, MD, PhD-student

vera.kerstens@ki.se

Abstract

Background. Reliable quantification of dopamine transporter (DAT), a biomarker for Parkinson's disease (PD), is essential for diagnostic purposes as well as for evaluation of potential disease-modifying treatment. Due to degeneration of dopaminergic neurons and thus lower expected radioligand binding to DAT, higher measurement variability in PD patients might be expected than earlier reproducibility results in healthy controls. Therefore, we aimed to examine the test-retest properties of [^{18}F]FE-PE2I-PET in PD patients.

Methods. Nine patients with PD (Hoehn & Yahr stage < 3) were included (men/women: 6/3; mean age 65.2±6.8y). Each patient underwent two [^{18}F]FE-PE2I-PET measurements within 7–28 days. The outcome measure was non-displaceable binding potential generated using wavelet-aided parametric imaging with cerebellum as reference region. We assessed test-retest performance using estimates of reliability and repeatability. Regions for primary analysis were caudate, putamen, ventral striatum, and substantia nigra. Exploratory analysis was performed for functional subdivisions of the striatum. We also compared the more vs. less-affected side.

Results. [^{18}F]FE-PE2I showed absolute variability estimates of 5.3–7.6% in striatal regions and 11% in substantia nigra, and ICCs of 0.74—0.97 (median: 0.91). The absolute variability for functional striatal subdivisions was 6.0–9.6% and ICCs of 0.76–0.91 (median: 0.91). The less affected substantia nigra exhibited greater consistency than the more affected side. According to power calculations based on the current sample size, DAT changes of 5–11% in the striatum and 28% in the substantia nigra can be detected with a power of 0.8 ($p<0.0125$).

Conclusion. DAT-PET measurements with [^{18}F]FE-PE2I in PD patients showed good repeatability and reliability. The slightly lower reliability in the substantia nigra in patients may be explained by lower DAT density and smaller anatomical size. Power calculations suggest that [^{18}F]FE-PE2I PET of DAT is a suitable disease progression marker in PD.

Trial registration number: EudraCT 2017-003327-29

Keywords: reliability, test-retest, [^{18}F]FE-PE2I, dopamine transporter, Parkinson's disease

Background

Parkinson's disease (PD) is a movement disorder characterized by progressive degeneration of the dopaminergic system, affecting both the cell bodies in the substantia nigra and its projections to the striatum. Within the dopaminergic system, the dopamine transporter regulates the synaptic dopamine levels by dopamine reuptake, and its density reflects presynaptic functioning.

[^{18}F]FE-PE2I, developed in 2009 [1,2], has, through several validity studies, proven to be a valuable positron emission tomography (PET) radioligand for dopamine transporter (DAT) imaging [1–8]. A clinical PET study with [^{18}F]FE-PE2I in 20 patients with early-stage PD showed an *in vivo* striato-nigral gradient of DAT-loss [9], in agreement with post-mortem studies in patients, and animal model studies [10,11]. After initial clinical validation [6], additional studies showed that simplified quantification of [^{18}F]FE-PE2I -PET can be achieved with a shortened imaging protocol, making clinical implementation realistic [7,12]. The test-retest reliability of [^{18}F]FE-PE2I has been previously studied in twelve young, healthy men [13], showing low variability and good reliability. However, in order to evaluate the suitability of [^{18}F]FE-PE2I as a disease progression marker, it is critical to assess reliability also in patient samples. Due to the degeneration of striatal projections, patients with PD have lower DAT-availability, which could lead to lower measurement reliability compared with healthy controls.

In the majority of PET studies of the striatal dopaminergic system, the regions of interest are the striatum, divided into caudate, putamen, and ventral striatum (nucleus accumbens), and the substantia nigra (midbrain). This anatomical subdivision of the striatum, although useful, might not represent the functional organization of the striatum. Instead, subdivision based on the connectivity between the basal ganglia and the neocortex can be used, where striatum is divided into a limbic, associative and sensorimotor striatum (respectively ventral striatum; caudate and ventrolateral putamen; and dorsolateral putamen) [14]. The functional regions have shown to be useful for molecular imaging studies examining correlations with behavioral and clinical outcome measures [15–17].

The assessment of the test-retest reliability of DAT-PET in patients with PD is relevant for several reasons. First, the knowledge on the natural variability is essential for interpretation of longitudinal follow-up study results; second, the measured variability can be used to estimate the minimum effect size on DAT needed for disease-modifying treatment trials; and third, for the purpose of power calculations for future clinical studies investigating longitudinal treatment efficacy.

The primary objective of this study was, therefore, to assess the test-retest reliability of [^{18}F]FE-PE2I measurements in the main striatal areas and substantia nigra in patients with PD. The hypothesis was that the reliability in the striatum would be similar to that observed in healthy subjects, and the reliability in the substantia nigra lower than in the striatum considering the lower DAT density in the substantia nigra. The secondary objective was to evaluate the test-retest reliability of three connectivity-based functional subdivisions of the striatum in view of future PET-analyses.

Materials and methods

STUDY POPULATION

Eleven patients with PD, Hoehn and Yahr (H&Y) stage < 3, were recruited via advertisement on the Swedish Parkinson Foundation website and via two specialist outpatient clinics in Stockholm (Academic Specialist Centre, Karolinska University Hospital). None of the subjects had clinically relevant somatic comorbidities, cognitive decline, history of psychiatric disease, illicit drug abuse or alcoholism, as assessed by a structured interview. Physical examination, electrocardiography, and routine blood tests were normal. One patient had to be excluded from the PET analysis because the cerebellum was partly out of the PET axial field of view. Demographic details are shown in Table 1.

DATA COLLECTION

Activity monitor and disease severity assessment

An activity monitor (Actigraph GT3X+) was worn on the hip for 5—7 days before each PET measurement. Average amount of steps and magnitude of movement per day were measured as a supportive measure of clinical stability [18,19]. Only days with minimal 540 minutes wear time were included in the calculation [20]. As measure of disease severity, the Movement Disorder Society

Unified Parkinson's Disease Rating Scale part 3 motor function (MDS-UPDRS-III) was done, including H&Y staging. All MDS-UPDRS-III assessments were performed on the same time of day by the same physician (VSK) in practically defined "OFF" (see below). Symptom duration was defined as the time from reported onset of first motor symptoms.

TABLE 1 Demographic and clinical characteristics of the patients

Subject	Age (y) / sex	Symptom duration (y)	MDS-UPDRS-III	H&Y	LEDD (mg)	Days between PET 1 - 2	Avg. steps/day*	Avg. magnitude counts/minute*
1	67 / F	8	1: 16 2: 16	1	425	7	1: 5133 2: 6925	1: 378.5 2: 422.3
2	68 / M	7	1: 10 2: 9	1	500	7	1: 11904 2: 7254	1: 777.8 2: 567.2
3	56 / F	14	1: 16 2: 21	1	680	7	1: 6496 2: 4429	1: 378.9 2: 377.8
4	71 / M	5	1: 18 2: 28	1	560	7	1: 4361 2: 4770	1: 225.7 2: 248.8
5	69 / M	8	1: 30 2: 30	2	675	20	1: 5428 2: 6854	1: 426.5 2: 502.1
6	54 / M	3	1: 10 2: 7	2	650	28	1: 7232 2: 6930	1: 449.3 2: 455.5
7	74 / M	8	1: 25 2: 20	2	400	7	1: 3976 2: 6152	1: 426.2 2: 536.7
8	47 / M	2.5	1: 3 2: 6	1	320	17	1: 7084 2: 7166	1: 546.8 2: 532.7
9	67 / F	5	1: 40 2: 36	2.5	300	7	1: 3091 2: 5971	1: 278.0 2: 361.7
10	61 / M	2.5	1: 16 2: 16	2	150	14	1: 12020 2: 11801	1: 726.3 2: 895.8
Mean ± SD	63.4 ± 8.6	6.3 ± 3.5	1: 18.2 ± 10.8 2: 18.7 ± 10.2		465 ± 180	-	-	1: 461.4 ± 177.3 2: 490.0 ± 172.5
Median				1.5		7	1: 5962 2: 6889	

***Assessed with the activity monitor (Actigraph GT3X+)
MDS-UPDRS-III: Movement Disorder Society Unified Parkinson's Disease Rating Scale, part 3 motor function (range 0-72). Assessed before PET1 and PET2, respectively. **H&Y**: Hoehn & Yahr (range 1-5); **LEDD**: Levodopa equivalent daily dose.

MRI acquisition

Using a 3 Tesla MRI system (General Electric, Discovery MR750), T2-weighted images were acquired to exclude clinically significant pathology, and 3D T1-weighted images were acquired for co-registration with PET and delineation of the regions of interest (ROI). This last sequence has 176 slices of 1 mm thickness, field of view 256 x 256 mm, resolution 1x1x1 mm, inversion time 450 ms, echo time 3.18 ms, repetition time 8.16 ms.

PET acquisition

[^{18}F]FE-PE2I was prepared as previously described [21]. Two 93-minute [^{18}F]FE-PE2I PET measurements were acquired in each subject within an interval of 7–28 days (see PET injection

characteristics, Supplementary Table S1). PET measurements were done on the same time of day, around 1:30 pm. Patients were asked to be practically defined ”OFF”, meaning a withdrawal of levodopa-medication for at least 12 hours and other dopaminergic medication for at least 24 hours. Also, abstinence of caffeine 3 hours before PET, nicotine on day of PET, alcohol 48 hours before PET, and intense physical training 96 hours before PET were requested. An individually made plaster helmet was used for head fixation in the PET camera.

PET-measurements were acquired with a high-resolution research tomograph (HRRT, Siemens Molecular Imaging) after an intravenous bolus injection of [^{18}F]FE-PE2I. Details can be found in Supplementary Table S1. A 6-minute transmission scan with a Caesium-137 source was obtained for attenuation correction. Due to technical reasons, the transmission scan for one patient could not be acquired on the day of first PET measurement, so the transmission scan acquired before the second PET measurement was used for attenuation correction of the first PET measurement.

List mode PET data were reconstructed into 37 frames (8x10, 5x20, 4x30, 4x60, 4x180, 12x360 seconds) using 3D OP OSEM with 10 iterations and 16 subsets, including modeling of the PSF [22]. Frame-to-frame realignment was performed as previously described [23], with the only difference that the first two minutes instead of the first minute were used as reference frame for PET-realignment.

PET motion correction

Head motion was evaluated by patient observation during data acquisition as well as during image analysis by reviewing the realignment plots and brain time activity curves (TACs). Translation of more than 3 mm on the realignment plots led to additional motion correction using an in-house developed automatic procedure. Description of the method is given in Supplementary Text 1.

IMAGE ANALYSIS

Using SPM12, the T1-weighted 3D MRI sequence was first realigned to the AC-PC-plane (anterior commissure-posterior commissure), after which the PET was realigned and co-registered to the realigned MRI. The following regions of interest were then delineated automatically on the T1-weighted images with Freesurfer version 6.0.0 (http://surfer.nmr.mgh.harvard.edu/): whole striatum

(STR), caudate (CAU), putamen (PUT), ventral striatum (VS), and cerebellum. For substantia nigra (SN), the functional molecular template, as created in the research group [9], was used. As exploratory outcome, three functionally subdivided striatal areas [14] were added to the analysis (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases/striatumconn). For all regions, regional non-specific binding potentials (BP_{ND}) of [^{18}F]FE-PE2I were generated using wavelet-aided parametric imaging (WAPI) [24] with t^* = 27 min and cerebellum as reference region.

As described above, for one subject, only one transmission scan could be acquired, which had to be used for attenuation correction of both PETs. This technical issue introduced a bias in BP_{ND} due to misaligned attenuation correction in the reference region. See Supplementary Text 2 for analysis and explanation. It was, therefore, decided to exclude the subject for the main analysis and report the results including the outlier as supplementary material. We believe that the test-retest metrics calculated without the outlier are more representative of the study sample.

STATISTICAL ANALYSIS

For statistical analysis, R version 3.4.3, was used with the package *relfeas* (https://github.com/mathesong/relfeas). [^{18}F]FE-PE2I measurement reproducibility was determined with calculation of repeatability (absolute intrasubject variability, AbsVar; and the minimum detectable difference, MDD) and reliability (intraclass correlation coefficient, ICC), as per recommendation of Weir, Baumgartner, and Matheson [25–27]. Absolute variability was calculated as: (test-retest)/(mean test and retest)*100. For ICC, the two-way random effects, absolute agreement, single rater/measurement was used, corresponding to:

$$\frac{MS_S - MS_E}{MS_S + (k-1)MS_E + k(MS_T - MS_E)/n}$$

with MS_S, subjects mean square; MS_E, error mean square; k, number of trials; MS_T, trials mean square; and n, number of subjects.

The ICC represents the proportion of the variability not attributable to measurement error. As such, an ICC of 1 indicates perfect measurement reliability with all observed variability being due to true

(biological) differences and none to measurement variability (error), while an ICC of 0.5 indicates that the variability is comprised of true differences and measurement error in equal measure. Different interpretations of the ICC exist: as proposed by Portney & Watkins [28] and suggested by Matheson [27], we regard an ICC < 0.5 as low, 0.5—0.75 moderate, 0.75—0.9 good, and > 0.9 excellent. Measurement reliability with an ICC > 0.9 is recommended as a lowest acceptable standard for measurements from which diagnostic decisions are made, ICC > 0.7 for research purposes, with 0.95 and 0.8 considered as adequate, respectively [29].

The agreements between measurements in each region were plotted with the Bland Altman plots. Power plots were generated with the jamovi software (https://www.jamovi.org/). The results of study variables are expressed as mean ± standard deviation unless otherwise stated.

Results

All subjects completed the study according to the protocol, with exception of the subject with only one transmission acquisition. The MDS-UPDRS-III and Actigraph outcomes support the subject's clinical stability during the study period (Table 1), with the exception of two patients who had a week of influenza and a week of holiday respectively explaining the lower Actigraph outcome. One subject showed a large difference in test/retest MDS-UPDRS-III (18 vs. 28), based on 1 point increases spread over the different domains, with the left side being mildly symptomatic in the second assessment versus not symptomatic in the first assessment. The difference could not be explained by patient self-report and is probably due to either natural symptom fluctuations and/or intra-observer variability.

Representative test-retest BP_{ND} images of [^{18}F]FE-PE2I are shown in Figure 1. Table 2 shows the individual test-retest BP_{ND} values for the four main regions of interest, with the highest BP_{ND} in the CAU and VS, and lowest in PUT and SN (Figure 2). The Bland-Altman plots (Figure 3) showed good agreement between the test-retest measurements. Supplementary figure S1 shows the Bland Altman plots including the outlier described earlier. Test-retest calculations are reported in Table 3. The striatal regions displayed low variability (AbsVar 5.3%—7.5%) and high ICC (0.89—0.97). SN showed relatively higher absolute variability (11%), with a moderate ICC (0.74).

TABLE 2 Individual binding potential (BP_{ND}) values of [^{18}F]FE-PE2I in striatal regions and substantia nigra

	Striatum		Caudate		Putamen		Ventral striatum		Substantia nigra	
Subject	**PET 1**	**PET 2**	**PET 1**	**PET 2**	**PET 1**	**PET 2**	**PET 1**	**PET 2**	**PET 1**	**PET 2**
1	1.763	1.827	2.428	2.562	1.264	1.275	1.736	1.722	0.73	0.849
2	1.051	1.065	0.914	0.91	1.095	1.121	1.623	1.687	0.635	0.602
3	1.892	1.91	2.339	2.388	1.516	1.513	2.317	2.319	0.813	0.798
4	1.835	1.789	1.697	1.872	1.824	1.607	3.131	2.717	1.016	0.854
5	1.306	1.372	1.43	1.496	1.15	1.21	1.848	2.021	0.668	0.666
6	1.273	1.497	1.876	2.125	0.672	0.898	1.758	1.838	0.397	0.602
7	1.339	1.217	1.356	1.245	1.203	1.104	2.441	2.011	0.752	0.754
8	2.801	2.256	3.268	2.588	2.332	1.888	3.054	2.714	0.894	0.716
9	1.537	1.426	1.928	1.744	1.166	1.112	2.211	2.191	0.653	0.603
10	1.978	1.948	2.564	2.582	1.423	1.37	2.855	2.686	0.796	0.75

PET 1: Test**; PET 2:** Retest.

TABLE 3 Test-retest metrics of [^{18}F]FE-PE2I PET measurements (n = 9)

Region	PET 1 (BP_{ND}) COV (%)	PET 2 (BP_{ND}) COV (%)	AbsVar (%)	ICC	MDD	Powered detectable % change
Striatum	1.55 ± 0.33 21.1	1.56 ± 0.32 20.5	5.3	0.95 (0.82-0.989)	0.195	-8.3
Caudate	1.84 ± 0.55 29.9	1.88 ± 0.59 31.4	6.0	0.97 (0.89-0.99)	0.264	-8.9
Putamen	1.26 ± 0.32 25.3	1.25 ± 0.22 17.8	7.5	0.91 (0.68-0.98)	0.225	-11.9
Ventral Striatum	2.21 ± 0.53 23.9	2.13 ± 0.38 17.9	6.5	0.89 (0.61-0.97)	0.426	-12.1
Substantia nigra	0.72 ± 0.17 23.2	0.72 ± 0.1 14.5	10.6	0.74 (0.24-0.93)	0.194	-17.9

COV: Coefficient of variability (SD/mean *100); **AbsVar:** Absolute variability; **ICC:** Intraclass correlation coefficient; **MDD**: Minimum detectable difference covering 95% of the distribution of test-retest differences; **Powered detectable % change:** based on measured effect size and power 0.8.

Bland Altman plots of the exploratory regions of interest are presented in Supplementary Figure S2, and the repeatability and reliability results in Table S3. Low absolute variability and good ICC (> 0.75) were observed in the functional striatal subdivisions.

For transparency, test-retest metrics calculated including the outlier are reported in Supplementary Table S2.

Additionally, the reliability was assessed for the outcomes of the less vs. more affected hemisphere (Supplementary Table S3). This analysis showed similar test-retest consistency for both the more and

less affected hemispheres. The less affected SN exhibited numerically better reliability and repeatability, although not significant.

Discussion

This study was designed to assess the test-retest reproducibility of [^{18}F]FE-PE2I PET-measurements of DAT in Parkinson patients (H&Y stage < 3). The results showed good repeatability and reliability of the measurements, providing support that [^{18}F]FE-PE2I can be used as DAT biomarker for disease progression in PD.

The [^{18}F]FE-PE2I test-retest study in twelve young, healthy controls [13], observed comparable repeatability (CAU 4.8%, PUT 5.6%, SN 9.7%) and reliability (CAU 0.83, PUT 0.88, SN 0.71). This shows that the lower DAT-availability as consequence of PD, at least in H&Y stages <3, does not substantially influence the consistency of its measurements.

Test-retest reproducibility of other DAT-PET and –SPECT radioligands

Hirvonen et al. [30] showed in their same-day test-retest study in five healthy subjects an absolute variability and reliability of ^{11}C-PE2I measurements in manually delineated ROIs for ventral striatum and midbrain of 7.2 ± 4.4%, 6.5 ± 5.2%, and ICC of 0.81, 0.83, respectively. The higher ICC in midbrain compared to Suzuki et al., is probably due to higher absolute BP_{ND} values in the midbrain observed for ^{11}C-PE2I and measured with an HRRT. Nurmi et al. [31] found the intraclass correlations of ^{18}F-CFT uptake values in eight healthy controls to be 0.91, 0.94, and 0.86 for caudate, anterior putamen, and posterior putamen, respectively (manual ROIs). In the seven *de novo* PD patients, the intraclass correlation was 0.97, 0.95 and 0.96, respectively, which is comparable to our results. Scans were performed 2.5—3 months apart, with the second scan in PD patients being after a three-month levodopa treatment, showing that even after initiated levodopa treatment, the DAT-measurements have high reproducibility in PD-patients in this time-range.

Test-retest results of clinical DAT-SPECT radioligands ^{123}I-beta-CIT and ^{123}I-FP-CIT showed measurement variability in PD patients of 7.4—16.8 % in STR and 12.2% in striatal subdivisions [32–

34], with corresponding ICCs of 0.59—1.00 [33,34]. Results vary because of different ROI definition methods and outcome measures. The striatal test-retest variability of ^{123}I-PE2I in seven healthy subjects [35] was 5.2 ± 4.5 (STR), 9.4 ± 7.0 (CAU), and 10.3 ± 5.1% (PUT), with corresponding ICC of 0.92, 0.83 and 0.84. Our test-retest results are thus comparable to the clinically implemented DAT-SPECT radioligands. Given the advantages of PET compared to SPECT, the results confirm that ^{18}F-FE-PE2 is a strong candidate for clinical applications as well.

DAT-measurement in smaller regions

The higher resolution of PET compared with SPECT permits a better assessment of low-binding regions, such as the SN. Test-retest repeatability and reliability in this region were inferior to test-retest metrics in the striatum. The smaller size of the SN and the smaller numerical value of BP_{ND} in this region are likely reasons for its greater variability. Despite this relative limitation, a previous study on DAT availability in the nigrostriatal system of early PD-patients as quantified with [^{18}F]FE-PE2I [9], showed changes of DAT availability in the SN compared to healthy controls that were still larger than the measured test-retest repeatability, confirming that the assessment of the nigrostriatal system with [^{18}F]FE-PE2I PET provides a comprehensive assessment of the PD pathophysiology *in vivo*.

The low variability and high reliability observed in the connectivity-based functional subdivisions of the striatum furthermore indicates that [^{18}F]FE-PE2I PET is a reliable research tool. This is relevant for studies using the functional rather than anatomical striatal subdivisions in assessing correlations of DAT availability to specific clinical variables of cognitive function or behavior.

Power calculation

Using the variability in test-retest differences within each region, we estimated the size of within-individual changes which could be detected with a power of 80%, using a significance threshold of 0.0125 (making use of Bonferroni correction for multiple comparisons for the four primary regions of interest, 0.05/4). These estimates suggest that a study with a sample size of 9 patients is sufficiently sensitive to detect within-individual differences of greater than between 5% and 11% for different regions within the striatum, and greater than 28% for the SN (Supplementary Table S2). The relationship between sample size and effect size was examined as a function of statistical power and is

presented in Figure 4. Yearly DAT-decline in PD-patients has been estimated to be between 5 and 13% in striatal regions [31,36,37], meaning that [^{18}F]FE-PE2I PET is well suited for measuring biological DAT-differences in striatal regions in a longitudinal follow-up study with a typical sample size for PET studies.

Conclusion

[^{18}F]FE-PE2I -measurements of DAT have good reliability in Parkinson patients (H&Y: 1-2.5) even in the small anatomical areas with lower DAT density, such as the substantia nigra. The test-retest metrics were equal-to-superior to other DAT-radioligands. Thus, this study further supports the suitability of [^{18}F]FE-PE2I as imaging marker for longitudinal follow-up studies and the evaluation of future disease-modifying treatment in PD.

List of abbreviations

[18F]FE-PE2I: 18F-(E)-N-(3-Iodoprop-2-Enyl)-2b- Carbofluoroethoxy-3b-(49-Methyl-Phenyl) Nortropane 3D OP OSEM: three dimensional Ordered Poisson Ordered Subset Expectation-Maximization; AbsVar: Absolute variability; AC-PC: anterior commissure-posterior commissure; BPnd: non-displaceable binding potential; CAU: caudate; COV: coefficient of variability; DAT: dopamine transporter; H&Y: Hoehn and Yahr disease severity score; ICC: intraclass correlation coefficient; LEDD: levodopa equivalent daily dose; MDD: minimum detectable difference; MDS-UPDRS-III: Movement Disorder Society Unified Parkinson's Disease Rating Scale, part 3, motor function; MRI: magnetic resonance imaging; PD: Parkinson's disease; PET: positron emission tomography; PSF: point-spread function; PUT: putamen; ROI: region of interest; SN: substantia nigra; SPECT: Single-photon emission computed tomography; SPM12: Statistical Parametric Mapping software, current version; STR: striatum; TAC: time activity curve; VS: ventral striatum; WAPI: wavelet-aided parametric imaging;

Declarations

Ethics approval and consent to participate

The study was approved by the Swedish Ethical Review Authority, the Radiation Safety Committee of the Karolinska University Hospital, and the Swedish Medicinal Product Agency (EudraCT 2017-003327-29). Patients provided written informed consent for study participation.

Consent for publication

Patients provided written informed consent for publication.

Availability of data and material

The dataset of the current study is available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests

Funding

This study was supported by funds from the Swedish Science Council and Åhlén Foundation. VSK's salary was supported by KID funds for doctoral students at Karolinska Institutet. S.C. is supported by the Swedish Research Council.

Authors' contributions

AV, SC and PS contributed to the design of the study. VSK, PF, MS and PS recruited subjects. VSK, PF, CH and AV performed the data collection. VSK performed the imaging analysis. VSK, GJM and AV performed the statistical analysis. EF analyzed the activity monitor data. VSK and AV drafted the manuscript. All authors reviewed and approved the manuscript.

Acknowledgements

The authors would like to thank Malena Kjellén for help with data collection, Göran Rosenqvist for help with PET reconstruction, and Zsolt Cselényi for the development of an automated PET-analysis pipeline. The Actigraphs (GT3X+) were lent to us by uMove Core facility/Strategic Research Area in Health Care Science at Karolinska Institutet.

References

1. Schou M, Steiger C, Varrone A, Guilloteau D, Halldin C. Synthesis, radiolabeling and preliminary in vivo evaluation of [18F]FE-PE2I, a new probe for the dopamine transporter. Bioorganic & Medicinal Chemistry Letters. 2009;19:4843-5.
2. Varrone A, Steiger C, Schou M, et al. In vitro autoradiography and in vivo evaluation in cynomolgus monkey of [18F]FE-PE2I, a new dopamine transporter PET radioligand. Synapse. 2009;63:871–880.
3. Varrone A, Toth M, Steiger C, et al. Kinetic analysis and quantification of the dopamine transporter in the nonhuman primate brain with 11C-PE2I and 18F-FE- PE2I. J Nucl Med. 2011;52:132–139
4. Varrone A, Gulyás B, Takano A, Stabin MG, Jonsson C, Halldin C. Simplified quantification and whole-body distribution of 18F-FE-PE2I in nonhuman primates: prediction for human studies. Nucl Med Biol. 2012;39:295–303.
5. Sasaki T, Ito H, Kimura Y, et al. Quantification of dopamine transporter in human brain using PET with 18F-FE-PE2I. J Nucl Med. 2012;53:1065–1073.
6. Fazio P, Svenningsson P, Forsberg A, et al. Quantitative analysis of 18F-(E)-N-(3-iodoprop-2-enyl)-2 -carbofluoroethoxy- 3-(4′-methyl-phenyl) nortropane binding to the dopamine transporter in Parkinson disease. J Nucl Med. 2015;56:714–20.
7. Sonni I, Fazio P, Schain M, et al. Optimal acquisition time window and simplified quantification of dopamine transporter availability using 18F-FE-PE2I in healthy controls and Parkinson's disease patients. J Nucl Med. 2016;57:1529–34.
8. Lizana H, Johansson L, Axelsson J, et al. Whole-body biodistribution and dosimetry of the dopamine transporter radioligand 18fF-Fe-PE2I in human subjects. J Nucl Med. 2018;59;1275-1280.
9. Fazio P, Svenningsson P, Cselényi Z, Halldin C, Farde L, Varrone A. Nigrostriatal dopamine transporter availability in early Parkinson's disease. Mov Disord. 2018;33:592–9.

10. Kordower, J. H., Olanow, C. W., Dodiya, H. B., et al. Disease duration and the integrity of the nigrostriatal system in Parkinson's disease. Brain, 2013;136;2419–2431.
11. Karimi, M., Tian, L., Brown, C.A., et al. Validation of nigrostriatal positron emission tomography measures: Critical limits. Ann Neurol. 2013;73;390-396.
12. Jakobsson Mo S, Axelsson J, Jonasson L, et al. Dopamine transporter imaging with [18F]FE-PE2I PET and [123I]FP-CIT SPECT-a clinical comparison. EJNMMI Res. 2018;8:100.
13. Suzuki M, Ito H, Kodaka F, et al. Reproducibility of PET measurement for presynaptic dopaminergic functions using L-[b-11C]DOPA and [18F]FE-PE2I in humans. Nucl Med Commun. 2014;35:231–237.
14. Tziortzi AC, Haber SN, Searle GE, et al. Connectivity-Based Functional Analysis of Dopamine Release in the Striatum Using Diffusion-Weighted MRI and Positron Emission Tomography. Cerebral Cortex. 2014; 24;1165–1177.
15. Cervenka S, Bäckman L, Cselényi Z, Halldin C, Farde L. Associations between dopamine D2-receptor binding and cognitive performance indicate functional compartmentalization of the human striatum. Neuroimage. 2008;40:1287-1295.
16. Kübler D, Schroll H, Buchert R, Kühn AA. Cognitive performance correlates with the degree of dopaminergic degeneration in the associative part of the striatum in non-demented Parkinson's patients. J Neural Transm. 2017;124:1073-1081.
17. Martín-Bastida A, Lao-Kaim NP, Roussakis AA, et al. Relationship between neuromelanin and dopamine terminals within the Parkinson's nigrostriatal system. Brain. 2019;142:2023-2036.
18. Nero H, Benka Wallén M, Franzén E, Conradsson D, Ståhle A, Hagströmer M. Objectively Assessed Physical Activity and its Association with Balance, Physical Function and Dyskinesia in Parkinson's Disease. J Parkinsons Dis. 2016;6:833-840.
19. Benka Wallén M, Franzén E, Nero H, Hagströmer M. Levels and Patterns of Physical Activity and Sedentary Behavior in Elderly People With Mild to Moderate Parkinson Disease. Phys Ther. 2015;95:1135-1141

20. Choi L, Liu Z, Matthews CE, et al. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sport Exer. 2011; 43: 357–364
21. Stepanov V, Krasikova R, Raus L, Loog O, Hiltunen J Halldin C. An efficient one-step radiosynthesis of [18F]Fe-PE2I, a PET radioligand for imaging of dopamine transporters. J. Label Compd. Radiopharm. 2012;55:206-210.
22. Varrone A, Sjöholm N, Eriksson L, Gulyás B, Halldin C, Farde L. Advancement in PET quantification using 3D-OP-OSEM point spread function reconstruction with the HRRT. Eur J Nucl Med Mol Imaging. 2009;36:1639-1650.
23. Schain M1, Tóth M, Cselényi Z, et al. Quantification of serotonin transporter availability with [11C]MADAM--a comparison between the ECAT HRRT and HR systems. Neuroimage. 2012; 60:800-807.
24. Cselényi Z, Olsson H, Halldin C, Gulyas B, Farde L. A comparison of recent parametric neuroreceptor mapping approaches based on measurements with the high affinity PET radioligands [11C]FLB 457 and [11C]WAY 100635. Neuroimage. 2006;32:1690–1708.
25. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005;19:231–40.
26. Baumgartner R, Joshi A, Feng D, Zanderigo F, Ogden T. Statistical evaluation of test-retest studies in PET brain imaging. EJNMMI Research. 2018;8:13.
27. Matheson GJ. We need to talk about reliability: making better use of test-retest studies for study design and interpretation. PeerJ 2019;7:e6918
28. Portney, LG, and Watkins MP. Foundations of clinical research: applications to practice. FA Davis. 2015.
29. Nunnally JC. Psychometric theory. New York: McGraw-Hill; 1978: 245-246.
30. Hirvonen J, Johansson J, Teräs M et al. Measurement of striatal and extrastriatal dopamine transporter binding with high-resolution PET and 11C-PE2I: quantitative modeling and test-retest reproducibility. Journal of Cerebral Blood Flow & Metabolism. 2008;28:1059-1069.

31. Nurmi E, Bergman J, Eskola O, et al. Reproducibility and effect of levodopa on dopamine transporter function measurements: a [18F]CFT PET study. J Cereb Blood Flow Metab. 2000; 20:1604-9.
32. Seibyl JP, Marek K, Sheff K, et al. Test/retest reproducibility of iodine-123-betaCIT SPECT brain measurement of dopamine transporters in Parkinson's patients. J Nucl Med. 1997;38:1453-9.
33. Booij J, Habraken JB, Bergmans P, et al. Imaging of dopamine transporters with iodine-123-FP-CIT SPECT in healthy controls and patients with Parkinson's disease. J Nucl Med. 1998; 39:1879-84.
34. Tsuchida, T., Ballinger, J.R., Vines, D., et al. Reproducibility of dopamine transporter density measured with123I-FPCIT SPECT in normal control and Parkinson's disease patients. Ann Nucl Med 2004;18:609-616.
35. Ziebell M, Thomsen G, Knudsen GM, et al. Reproducibility of 123I-PE2I binding to dopamine transporters with SPECT. EJNMMI. 2007;34:101-109.
36. Winogrodzka A, Bergmans P, Booij J, Van Royen EA, Stoof JC, Wolters EC. [123I]β-CIT SPECT is a useful method for monitoring dopaminergic degeneration in early stage Parkinson's disease. J Neurol Neurosurg Psychiatry. 2003;74:294-298.
37. Sung C, Lee JH, Oh JS, et al. Longitudinal Decline of Striatal Subregional [18F]FP-CIT Uptake in Parkinson's Disease. Nucl Med Mol Imaging. 2017;51:304-31

Figure titles and legends

Figure 1 Representative test and retest parametric BP_{ND} images of [^{18}F]FE-PE2I.

Figure 2 Individual BP_{ND} values between scan 1 and scan 2. Striatal regions and substantia nigra, with and without the exclusion of the outlier are displayed.

Figure 3 Bland Altman plots of main regions of interest. The yellow lines correspond the upper and lower 2SD line; red line: bias.

Figure 4 Relationship between sample size and effect size for caudate, putamen, ventral striatum, and substantia nigra. The arrow indicates the sample size needed to detect a statistically significant difference with 0.8 power, based on the effect size corresponding to a 10% change.