Eur J Nucl Med Mol Imaging
DOI 10.1007/s00259-014-2882-8

ORIGINAL ARTICLE

Machine learning models for the differential diagnosis of vascular parkinsonism and Parkinson's disease using [^{123}I]FP-CIT SPECT

I. Huertas-Fernández · F. J. García-Gómez · D. García-Solís · S. Benítez-Rivero · V. A. Marín-Oyaga · S. Jesús · M. T. Cáceres-Redondo · J. A. Lojo · J. F. Martín-Rodríguez · F. Carrillo · P. Mir

Received: 30 April 2014 / Accepted: 28 July 2014

Abstract
Purpose The study's objective was to develop diagnostic predictive models using data from two commonly used [^{123}I]FP-CIT SPECT assessment methods: region-of-interest (ROI) analysis and whole-brain voxel-based analysis.
Methods We included retrospectively 80 patients with vascular parkinsonism (VP) and 164 patients with Parkinson's disease (PD) who underwent [^{123}I]FP-CIT SPECT. Nuclear-medicine specialists evaluated the scans and calculated bilateral caudate and putamen [^{123}I]FP-CIT uptake and asymmetry indices using BRASS software. Statistical parametric mapping (SPM) was used to compare the radioligand uptake between the two regions at the voxel level. Quantitative data from these two methods, together with potential confounding factors for dopamine transporter availability (sex, age, disease duration and severity), were used to build predictive models following a tenfold cross-validation scheme. The performance of logistic regression (LR), linear discriminant analysis and support vector machine (SVM) algorithms for ROI data, and their penalized versions for SPM data (penalized LR, penalized discriminant analysis and penalized SVM), were assessed.
Results Significant differences were found in the ROI analysis after covariate correction between VP and PD patients in [^{123}I]FP-CIT uptake in the more affected side of the putamen and the ipsilateral caudate. Age, disease duration and severity were also found to be informative in feeding the statistical model. SPM localized significant reductions in [^{123}I]FP-CIT uptake in PD with respect to VP in two specular clusters comprising areas corresponding to the left and right striatum. The diagnostic predictive accuracy of the LR model using ROI data was 90.3 % and of the SVM model using SPM data was 90.4 %.
Conclusion The predictive models built with ROI data and SPM data from [^{123}I]FP-CIT SPECT provide great discrimination accuracy between VP and PD. External validation of these methods is necessary to confirm their applicability across centres.

Keywords Vascular parkinsonism · Parkinson's disease · [^{123}I]FP-CIT SPECT · Statistical parametric mapping · Predictive models

I. Huertas-Fernández · S. Benítez-Rivero · S. Jesús · M. T. Cáceres-Redondo · J. F. Martín-Rodríguez · F. Carrillo · P. Mir (✉)
Unidad de Trastornos del Movimiento, Servicio de Neurología y Neurofisiología Clínica, Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Avda. Manuel Siurot s/n., 41013 Seville, Spain
e-mail: pmir@us.es

F. J. García-Gómez · D. García-Solís · V. A. Marín-Oyaga · J. A. Lojo
Servicio de Medicina Nuclear, UDIM, Hospital Universitario Virgen del Rocío, Seville, Spain

D. García-Solís · P. Mir
Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Seville, Spain

Introduction

Vascular parkinsonism (VP) is a parkinsonian syndrome resulting from cerebrovascular lesions and is characterized by the presence of gait difficulties, symmetrically lower body bradykinesia and postural instability, and the absence of resting tremor [1–3]. Although recent neuropathology and epidemiological studies have identified hallmarks distinguishing VP from idiopathic Parkinson's disease (PD), overlap in symptom presentation is not rare and their differentiation is still a clinical challenge, especially in early stages [4–8].

Visualization of the dopamine transporter (DAT) through the use of [^{123}I]FP-CIT SPECT is a commonly used tool that may help in differentiating VP and PD. However, the status of the striatal DAT in VP patients is controversial due to its heterogeneity, and accuracy in the differential diagnosis is still poor [6, 9–14]. This heterogeneity was reflected in a recent study involving a large cohort of patients with VP in which the [^{123}I]FP-CIT SPECT scans in about one-third of the patients were normal, while the scans in the other two-thirds were abnormal, and the imaging pattern in a small percentage of patients overlapped the typical pattern seen in PD [14]. Furthermore, it has been suggested that a normal scan in patients with VP may be associated with negative responsiveness to levodopa treatment [14], although this association was not seen in another study [9].

The majority of studies including patients with VP have evaluated [^{123}I]FP-CIT SPECT imaging through visual assessment according to standardized scales [15] or semiquantification of striatal ligand uptake involving region-of-interest (ROI) analysis. Such methods may be suboptimal mainly because of first a certain degree of subjectivity in visual interpretation and in manual ROI delineation and second the focus primarily on DAT uptake in the striatum, thus missing the extent of radioligand binding to the DAT, serotonin and noradrenergic transporters in other brain regions. In contrast, voxel-based analysis has proven to be a reliable and unbiased tool for the analysis of whole-brain imaging. Statistical parametric mapping (SPM) is one of the most popular tools for whole-brain voxel-based analysis and some studies have used it with success in the differentiation of PD from other neurodegenerative diseases [16–19]. However, voxel-based studies including VP patient series are still lacking.

We have recently reported the results of a detailed clinical study in a large cohort of VP and PD patients [20]. We presented a newly developed visual scoring system with an accuracy in the differentiation of VP and PD greater than 94 % and a clustering method using ROI data with an accuracy of 82 %. The first objective of the present study was to build a diagnostic predictive model using the ROI data from the same dataset with improved performance and applying a more suitable methodology for the problems of classification from the machine learning theory. The second objective was to conduct a whole-brain voxel-based comparison of imaging data between VP and PD patients using SPM and following the same strategy as for the ROI data to build a predictive model with the voxel data.

Materials and methods

Patients

We included a total of 80 patients with VP and a control group of 164 patients with PD seen at our centre from 2006 to 2011. This is the subset of patients with [^{123}I]FP-CIT SPECT scans available from our previous work, and detailed clinical information for the whole population of patients was given in the report of that study [20]. For this study, the features sex, age, disease duration and severity measured according to the Hoehn & Yahr (H&Y) scale were reviewed when carrying out SPECT (Table 1). The diagnosis of VP was made according to the diagnostic criteria proposed by Zijlmans et al. [5] and the diagnosis of PD was made according to the UK Parkinson's Disease Society Brain Bank clinical diagnostic criteria [4]. Patients gave written informed consent for the [^{123}I]FP-CIT SPECT scan after a full discussion of possible risks and benefits as is the general practice in our hospital. This study was approved by the local ethics committee and conducted in accordance with the Declaration of Helsinki.

SPECT imaging

Patients underwent a brain SPECT scan with a dual-head rotating gamma camera (Philips Axis) fitted with LEHR fan-beam collimators. In order to block the thyroid uptake of free radioactive iodide, patients were given potassium perchlorate 500 mg orally 30 min before intravenous injection of 185 MBq of [^{123}I]FP-CIT (Ioflupane, Datscan®; GE Healthcare). Image acquisition was started between 3 and 4 h after radioligand injection. A total of 120 projections of 30 s each over a 360° circular orbit were acquired on a 128× 128 matrix (zoom 1.5). Reconstruction was performed by filtered back-projection using a Butterworth filter without attenuation or scatter correction and further reorientation to obtain transaxial slices.

ROI analysis

An automated semiquantitative analysis was performed to evaluate specific-to-nondisplaceable [^{123}I]FP-CIT binding potential (BP_{ND}) using HERMES-BRASS software (version 3.5). ROIs were constructed around the left and right striatum, the striatal subregions caudate and putamen, and in the background brain (occipital cortex). The automated method in HERMES-BRASS first normalized the patient scans to a [^{123}I]FP-CIT mean template and then delineated the regions using the standardized 3-D volume-of-interest (VOI) maps. Further details of the procedure and creation of the [^{123}I]FP-CIT mean template and standardized 3-D VOI maps have been provided by Koch et al. [21]. [^{123}I]FP-CIT BP_{ND} for the left and right putamen and caudate were calculated by normalizing the subregional radioactivity counts by the background counts (for each striatal subregion: BP_{ND}=[(striatal subregion counts−occipital counts)/occipital counts]). We defined the more affected

JrnlID 259_ArtID 2882_Proof# 1 - 05/08/2014

t1.1 **Table 1** Demographic and main clinical features of VP and PD patients

t1.2		VP patients (n=80)	PD patients (n=164)	P value (intergroup comparison)			t1.3
				t test	Mann–Whitney test	Logistic regression	
t1.4	Sex (M/F), n	57/23	101/68				
t1.5	Age (years), mean±SD	75.11±6.70	60.26±10.84	<0.001		<0.001	
t1.6	Disease duration (years), median (interquartile range)	4 (2, 8)	2 (1, 4)	<0.001		<0.01	
t1.7	H&Y stage, median (interquartile range)	2.5 (2.5, 3)	2 (2, 2.5)		<0.001	<0.05	

side as the hemisphere with the lower putamen BP_{ND}, and the ROI variables were defined as those from the putamen and caudate ipsilateral to the more affected side (Put_I, Cau_I), and the putamen and caudate contralateral to the more affected side (Put_C, Cau_C). The asymmetry index (AI) was calculated using the following formula [9]:

AI=[(contralateral striatum binding−ipsilateral striatum binding)/(contralateral striatum binding+ipsilateral striatum binding)]×2×100.

SPM analysis

A semiquantitative whole-brain voxel-based analysis was performed using SPM8 (Wellcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) running under a Matlab environment (MathWorks, Sherborn, MA).

SPECT images were first manually reoriented, setting the anterior commissure as the origin of the coordinates. Each scan was then spatially normalized into the standard stereotactic MNI (Montreal Neurological Institute) space using a [^{123}I]FP-CIT template developed by our group (available at http://www.nitrc.org/projects/spmtemplate) [22]. Next, spatially normalized images were smoothed using an isotropic 8-mm full-width at half-maximum isotropic gaussian kernel (FWHM). For further details about processing underlying the normalization and smoothing steps, the reader is referred to the SPM manual (http://www.fil.ion.ucl.ac.uk/spm/doc/manual.pdf). For the analysis stage, to account for the interindividual variability of [^{123}I]FP-CIT uptake, the option "proportional scaling" was enabled to intensity-normalize each scan. Also, since DAT densities are known to be low in the occipital lobe and the cerebellum, a general brain mask for those areas was created using an automated anatomical labelling atlas and applied to all images for statistical comparison. A total of 152,673 voxels were analysed. Clusters of a minimum of 16 (twice FWHM of the gaussian filter) 3-D contiguous voxels with a threshold of P_{FWE}<0.05 corrected for multiple comparisons based on family-wise error (FWE) were considered to be statistically significant.

Data analysis

Statistical analyses were performed using IBM SPSS Statistics 20.0 software and the free software environment R (http://www.r-project.org/). Descriptive statistics are reported with percentages, means and standard deviations and medians and interquartile ranges when appropiate. Univariate analyses were first performed to compare the demographic and clinical features, and the ROI variables between VP and PD patients. Sex distribution was compared using the chi-squared test. Scale variables (i.e. age, disease duration, H&Y stage and ROI variables) were compared using the t-test (parametric) or the Mann–Whitney test (nonparametric). To decide whether to use parametric or nonparametric tests with the scale variables, we assessed the assumptions for a normal distribution using the Shapiro-Wilk test and the homogeneity of the variance (homoscedasticity) using the Levene test. To assess differences in the ROI variables between the VP and PD patients taking into account the effects of the demographic and clinical features, we further performed multivariate analysis using logistic regression (LR) introducing the ROI variables as factors and sex, age, disease duration and H&Y stage as covariates. Additionally, since these covariates are known to influence radioligand uptake [23, 24], interaction terms were included to check their role as effect modifiers. In the SPM analysis, the [^{123}I]FP-CIT uptake was compared between VP and PD patients with a two-sample t-test contrast (VP>PD).

Moreover, due to its clinical relevance and in order to clarify the inconsistencies apparent in the literature [9, 14], subanalyses of VP patients comparing levodopa responders and nonresponders were performed using LR for ROI data, and again using a two-sample t-test contrast in SPM (VP nonresponder>VP responder).

Predictive models

Predictive models using the quantitative data from the ROI and SPM analyses were built for diagnosis classification. The clinical diagnosis, as defined in the section Patients, was considered the gold standard in this study and was used as the dependent variable in building the models. For the models

JnlID 259_ArtID 2882_Proof# 1 - 05/03/2014

t2.1 **Table 2** Mean regional putamen and caudate [^{123}I]FP-CIT BP_{ND} in VP and PD patients

t2.2			VP patients (n=80)	PD patients (n=164)	P value (intergroup comparison)			t2.3
					t test	Mann–Whitney test	Logistic regression	
t2.4	[^{123}I]FP-CIT BP_{ND}, mean±SD	Cau_I	1.54±0.54	1.06±0.46	<0.001		<0.001	
t2.5		Cau_C	1.64±0.49	1.23±0.50	<0.001			
t2.6		Put_I	1.20±0.52	0.53±0.30	<0.001		<0.001	
t2.7		Put_C	1.43±0.52	0.78±0.37	<0.001			
t2.8	Asymmetry index, median (interquartile range)		7.04 (3.00, 18.94)	20.69 (8.82, 39.00)		<0.001		

using ROI data, the independent variables were the significant factors from the LR, while for the models using SPM data, the independent variables were the intensity values of the voxels (after normalization and smoothing) contained in the significant clusters from the SPM contrasts. Significant covariates were also included in the models.

Since the number of independent variables for the models using SPM data was large (a few hundred voxels), regularized algorithms were used. These algorithms weight the independent variables according to their information content, priorizing some and penalizing others through tunable shrinkage functions. We opted for comparing three algorithms recommended elsewhere [25]: penalized LR (PLR), penalized discriminant analysis and penalized support vector machine (SVM). The ROI data were analysed using equivalent methods: LR, linear discriminant analysis (LDA) and SVM. Tuning parameters for the algorithms were chosen based on the package default grid of iterations.

The models were assessed using a tenfold cross-validation scheme, which randomly split the dataset into ten parts (K=10), 90 % used for training and the remaining 10 % for testing, for every kth=1, 2, …, 10. This strategy prevented overfitting the model with our dataset, thus allowing model generalization of data from other centres. The final model and performance results were obtained from averaging the ten runs, which were given in terms of area under the receiver operating characteristic curve, accuracy, sensitivity and specificity. All calculations were done using the R package "caret" (http://caret.r-forge.r-project.org/).

Results

Demographic and clinical features

Age and disease duration fulfilled normality and homoscedasticity assumptions and were compared in the univariate analysis using a t-test. The H&Y stage did not fulfil the assumptions and was compared using the Mann–Whitney test. There were significant differences between VP and PD patients in terms of age, disease duration and H&Y stage (P<0.001; see Table 1). These associations were consistent in the regression analysis. As already described in our previous work [20], our VP patients were older, with longer disease duration and higher H&Y stage than our PD patients.

Discrimination between VP and PD patients using ROI analysis

Regional [^{123}I]FP-CIT uptake and AI values of the VP and PD patients and intergroup statistics are shown in Table 2. The variables Put_I, Cau_I, Put_C and Cau_C uptake values fulfilled normality and homoscedasticity assumptions and were compared in the univariate analysis using a t-test. AI did not fulfil the assumptions and was compared using the Mann–Whitney test. Univariate analyses showed significantly lower [^{123}I]FP-CIT BP_{ND} values for all four regions along with a higher AI in PD patients than in VP patients (P<0.001). Regression analysis indicated that these findings were consistent after covariate correction for the more affected

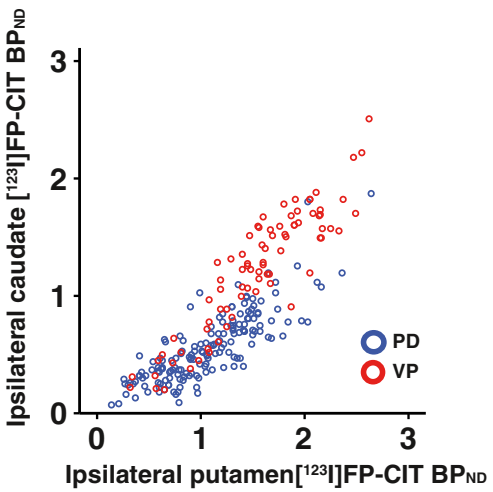

Fig. 1 Ipsilateral caudate [^{123}I]FP-CIT BP_{ND} as function of ipsilateral putamen [^{123}I]FP-CIT BP_{ND} (*red circles* VP patients, *blue circles* PD patients)

t3.1 **Table 3** Average tenfold cross-validation performance results (mean± standard deviation) for the diagnostic predictive models built with ROI data, given as area under the ROC curve (AUC), accuracy, sensitivity and specificity . The methods tested were logistic regression (LR), linear discriminant analysis (LDA) and support vector machine (SVM)

t3.2 Method	AUC	Accuracy	Sensitivity	Specificity	Parameters
t3.3 LR	0.951±0.046	0.903±0.058	0.944±0.062	0.794±0.142	–
t3.4 LDA	0.940±0.042	0.898±0.065	0.963±0.049	0.775±0.138	–
t3.5 SVM	0.950±0.045	0.899±0.049	0.947±0.061	0.784±0.145	C=1

hemisphere regions (Cau_I, $P<0.001$; Put_I, $P<0.001$). None of the interaction terms reached significance.

These significant variables, along with the covariates age, disease duration and H&Y stage, were further used to build the predictive models. Figure 1 displays the scatter plot of the two input factors, where Cau_I [^{123}I]FP-CIT BP_{ND} is plotted as a function of Put_I [^{123}I]FP-CIT BP_{ND}. Most of the VP patients grouped separately from most of the PD patients, and the decision boundaries between the two entities could be fitted with linear algorithms. Hence, LR, LDA and SVM fed by the first-order terms of the input factors were adequate approaches. The cross-validation results for the three methods are shown in Table 3. LR demonstrated slightly better discrimination accuracy than SVM and LDA (accuracy 0.903, 0.899 and 0.898, respectively), and its equation is given by the following formula:

$$\begin{aligned}\text{logit(diagnosis)} = &-14.55 - 3.92 \times \text{Cau\_I} + 7.29 \times \text{Put\_I} \\ &+ 0.18 \times \text{age} + 0.75 \times \text{H\&Y} - 0.28 \\ &\times \text{DisDur}\end{aligned}$$

where logit represents the logarithm of $p/(1-p)$ and p is the probability of being a VP patient, and DisDur is disease duration.

These results indicate that, despite being a very good model, a small percentage of scans were misdiagnosed. To improve the discrimination accuracy, we established a cut-off of 80 % for the class probability. In other words, we assigned a diagnosis only if the probability of belonging to that class applying the formula was above 80 %. We tested the LR model in the whole dataset and the accuracy was increased to 95 %, although the data from 17 % of the patients were under the threshold and their diagnosis remained tagged as "doubtful" (Fig. 2).

Discrimination between VP and PD patients using SPM analysis

Voxel-based analysis of [^{123}I]FP-CIT SPECT scans supported the results of the striatal ROI analysis. SPM contrasts revealed decreased intensity values in PD patients compared with VP patients in two specular clusters (1,113 and 1,320 voxels) that comprised areas corresponding to, respectively, the left and right striatum (Table 4).

The predictive models were built using all intensity values of voxels contained in the significant clusters as independent variables, and the same covariates as in the ROI analysis (age, disease duration and H&Y stage). The cross-validation results from the penalized methods are summarized in Table 5. SVM showed slightly better accuracy in discriminating between VP and PD than PLR and LDA (accuracy 0.904, 0.887 and 0.884, respectively).

Comparison between levodopa responders and nonresponders

Neither LR with ROI data nor SPM analysis revealed an association between [^{123}I]FP-CIT uptake and levodopa responsiveness in VP patients.

Discussion

In this study, we investigated the accuracy of methods for distinguishing between VP and PD using [^{123}I]FP-CIT SPECT. We developed predictive models using the

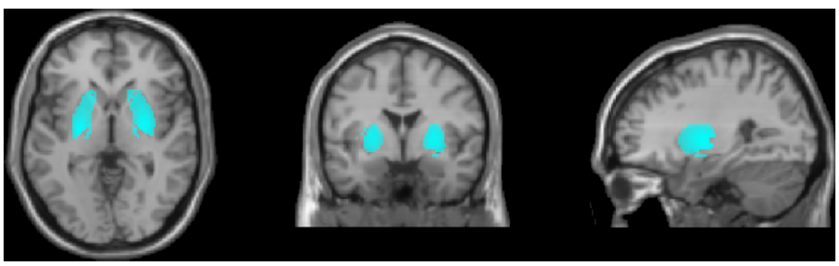

Fig. 2 Voxel clusters representing significant decreases in [^{123}I]FP-CIT uptake in PD patients with respect to VP patients. The areas include the putamen and caudate nucleus, and are represented in MNI-normalized MRI scans

Table 4 Significant findings of the SPM comparison of VP and PD patients (VP>PD)

Cluster localization	Cluster size	MNI coordinates			*T* value	*Z* value	p_{FWE} value
		x	*y*	*z*			
Left striatum	1,113	−26	−10	2	8.31	7.77	<0.001
		−20	14	2	5.59	5.42	<0.001
		−22	−18	4	5.46	5.29	<0.001
Right striatum	1,320	28	−6	2	8.08	7.59	<0.001
		22	−16	0	5.32	5.17	0.001
		20	10	12	5.27	5.12	0.001

semiquantitative data from the SPECT evaluations of a large cohort of patients using two widely differing methods: striatal ROI analysis and whole-brain voxel-based analysis. Our previous study [20], as well as a similar multicentre study performed in Italy [14], confirmed what previous studies have indicate: VP is a different and distinguishable entity from PD, but clinical manifestations and imaging patterns are heterogeneous. [^{123}I]FP-CIT SPECT is a widely available tool helping the physician in the diagnosis of VP, and numerous studies have investigated visual assessment and ROI quantification using [^{123}I]FP-CIT SPECT [6, 26]. Some authors have found significant differences in the AI in PD patients [9, 12], but these studies had small sample sizes and their sensitivity was as low as 50 % [12]. These results have led to questioning the accuracy of [^{123}I]FP-CIT SPECT in the diagnosis of VP, and indeed, a very recent study considered the inclusion of cardiac [^{123}I]MIBG SPECT and the use of the smell identification UPSIT test in the differential diagnosis [13].

In our previous study we used the [^{123}I]FP-CIT BP_{ND} values of the more affected side of the putamen and the ipsilateral caudate and the AI in a clustering method, and achieved an accuracy of 82 %. However, this approach did not exploit all the information available from the patient and contained in the image, nor did it provide a generalizable mathematical formula for use by other groups. In contrast, other studies have successfully applied elegant methods for distinguishing atypical parkinsonisms and other diseases from PD using DAT SPECT imaging [16–19]. Scherfler et al. used ROI analysis and SPM to extract mean voxel cluster values and introduced their parameters into a stepwise discriminant analysis [17]. Some years later, the same group elaborated a computer-assisted image algorithm (CAIA) using voxel data that outperformed a multinomial regression using ROI data [18]. In this study, we investigated images from VP patients using these types of approaches. In the ROI analysis, in agreement with the findings of previous studies [10, 12], we found that in PD patients, in comparison with VP patients, the striatal DAT availability is markedly reduced and the AI is significantly higher. LR revealed that the more affected side of the putamen and the ipsilateral caudate, along with the covariates age, disease duration and H&Y stage, were informative in feeding the predictive model. Cross-validation procedures demonstrated that the algorithms LR, LDA and SVM were excellent classifiers using these variables. In the case of LR, the model achieved a diagnostic accuracy of 90.3 %. Moreover, the results could be improved to 95 % accuracy by thresholding the class probability and creating a pool of patients with a doubtful diagnosis. For these patients, we assumed that the ROI analysis of the [^{123}I]FP-CIT SPECT scans was inconclusive and that it would be necessary to evaluate their clinical profile and structural neuroimaging to determine a more reliable diagnosis.

Despite the diagnostic accuracy for the newly developed visual scoring system in our previous work that reached above 94 %, we acknowledge that the application of this system requires highly trained nuclear medicine specialists, and that the intraobserver and interobserver rates are not perfect. Although we strongly encourage specialists to learn and apply the new visual scoring system, we believe that the application of the LR formula could be used more easily to achieve diagnostic accuracies above 90 %.

Table 5 Average tenfold cross-validation performance results (mean± standard deviation) for the diagnostic predictive models built with SPM data, given as area under the ROC curve (AUC), accuracy, sensitivity and specificity. The methods tested were penalized logistic regression (PLR), penalized discriminant analysis (PDA) and support vector machine (SVM)

Method	AUC	Accuracy	Sensitivity	Specificity	Parameters
PLR	0.960±0.039	0.887±0.049	0.981±0.034	0.704±0.144	α=0.1, λ=0.1
PDA	0.878±0.073	0.884±0.069	0.944±0.058	0.769±0.135	λ=3
SVM	0.954±0.057	0.904±0.059	0.954±0.056	0.801±0.172	C=1

Regarding SPM, the comparison gave significant differences in [^{123}I]FP-CIT uptake in two specular clusters of voxels including areas of the striatum. We took advantage of this high level of information by introducing all voxel values together with the covariates into a penalized classification algorithm, and found that SVM was able to achieve 90.4 % accuracy. Furthermore, as for the ROI data, it would also be possible to raise this accuracy by restricting the allocation to high class probabilities. This method demonstrated that the use of whole-brain voxel data is a powerful alternative with two great advantages with respect to the previous method, i.e. no a priori assumptions about the location of the ligand uptake and more importantly, the method is conducted in an unbiased and automated fashion.

It is also important to note that our models made use of basic clinical information, namely age, disease duration and H&Y stage. Differences in these factors are not uncommon between VP and PD cohorts. Antonini et al. obtained differences in H&Y stage in a large cohort of patients [14]. Other studies have also shown differences in age [9, 10, 12, 13] and disease duration [13]. These differences give these factors predictive ability to differentiate VP from PD, and they are also potential confounders for determining differences in [^{123}I]FP-CIT uptake. For these reasons, it is necessary to incorporate these variables in the discriminative models using [^{123}I]FP-CIT SPECT that seek to be applicable to the general populations of VP and PD patients. However, previous studies using [^{123}I]FP-CIT SPECT for differentiating VP from PD did not fully take into account this information. Age, disease duration and H&Y stage were not quantitatively included in the studies that used visual assessment [13, 14], and other studies that used ROI semiquantification have shown differences in the ROI variables between these groups of patients without accounting for them [9–12]. In this study we observed that these factors, apart from directly influencing radioligand uptake per se, were simple, accessible and very informative for differentiating VP from PD. Hence, we recommend their inclusion in the models.

We also sought to determine if there was higher striatal ligand uptake in VP patients with a negative response to levodopa treatment than in positive responders, as found by Antonini et al. [14]. Our results were all negative for this association indicating that the [^{123}I]FP-CIT uptake is not a good predictor of responsiveness to dopamine replacement therapy.

Finally, it is interesting to speculate as to why these models did not reach 100 % accuracy. In our opinion, a major limitation influencing the accuracy might have been that our gold standard was based on clinical criteria that did not take into account the SPECT findings, and perhaps a few patients were wrongly diagnosed. Some of the patients who were diagnosed as having VP, even though they fulfilled the criteria for VP when included in the study, had a PD-like scan pattern. It is possible that some of these patients truly had VP with a [^{123}I]FP-CIT SPECT scan pattern indistinguishable from that in PD, while others had in reality underlying PD accompanied by cerebrovascular damage. In this case, updating our models would have resulted in an increase in the accuracy and therefore a boost in the credibility of the SPECT-aided diagnosis. Nevertheless, to confirm this hypothesis it would be necessary to perform a long-term follow-up to verify how these patients evolve clinically, or preferably, an MRI scan or an anatomopathological examination in the most misleading cases.

In conclusion, this study provided accuracies above 90 % in discriminating between VP and PD using two common methods for SPECT scan evaluation: ROI analysis and SPM. We provide a mathematical formula for the ROI analysis model for evaluation by other groups. We also introduce a method for processing voxel-based data: the use of penalized algorithms implemented in R packages. This approach provides an automated and therefore objective, fast and efficient solution that would be very beneficial for decision-making in nuclear medicine. Future work will investigate the method including more types of parkinsonism and its implementation in a distributable application for external evaluation.

Acknowledgments This work was supported by grants from the Ministerio de Economía y Competitividad de España (SAF2007-60700), the Instituto de Salud Carlos III (PI10/01674, CP08/00174, PI13/01461), the Consejería de Economía, Innovación, Ciencia y Empresa de la Junta de Andalucía (CVI-02526, CTS-7685), the Consejería de Salud y Bienestar Social de la Junta de Andalucía (PI-0377/2007, PI-0741/2010, PI-0437-2012), the Sociedad Andaluza de Neurología, the Fundación Alicia Koplowitz, the Fundación Mutua Madrileña and the Jaques and Gloria Gossweiler Foundation.

Conflicts of interest None.

References

1. Rektor I, Rektorova I, Kubova D. Vascular parkinsonism - an update. J Neurol Sci. 2006;248:185–91.
2. Fitzgerald PM, Jankovic J. Lower body parkinsonism: evidence for vascular etiology. Mov Disord. 1989;4:249–60.
3. Winikates J, Jankovic J. Clinical correlates of vascular parkinsonism. Arch Neurol. 1999;56:98–102.
4. Hughes AJ, Daniel SE, Kilford L, Lees AJ. The accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinicopathologic study. J Neurol Neurosurg Psychiatry. 1992;55:181–4.
5. Zijlmans JC, Daniel SE, Hughes AJ, Revesz T, Lees AJ. Clinicopathological investigation of vascular parkinsonism, including clinical criteria for diagnosis. Mov Disord. 2004;19:630–40.
6. Kalra S, Grosset DG, Benamer HT. Differentiating vascular parkinsonism from idiopathic Parkinson's disease: a systematic review. Mov Disord. 2010;25:149–56.
7. Jellinger KA. Prevalence of cerebrovascular lesions in Parkinson's disease. A post-mortem study. Acta Neuropathol. 2003;105:415–9.

8. Zijlmans J, Katzenschlager R, Daniel S, Lees A. The L-dopa response in vascular parkinsonism. J Neurol Neurosurg Psychiatry. 2004;75:545–7.
9. Zijlmans J, Evans A, Fontes F, Katzenschlager R, Gacinovic S, Lees AJ, et al. [123I]FP-CIT SPECT study in vascular parkinsonism and Parkinson's disease. Mov Disord. 2007;22:1278–85.
10. Gerschlager W, Bencsits G, Pirker W, Bloem BR, Asenbaum S, Prayer D, et al. [123I]beta-CIT SPECT distinguishes vascular parkinsonism from Parkinson's disease. Mov Disord. 2002;17:518–23.
11. Lorberboym M, Djaldetti R, Melamed E, Sadeh M, Lampl Y. 123I-FP-CIT SPECT imaging of dopamine transporters in patients with cerebrovascular disease and clinical diagnosis of vascular parkinsonism. J Nucl Med. 2004;45:1688–93.
12. Contrafatto D, Mostile G, Nicoletti A, Dibilio V, Raciti L, Lanzafame S, et al. [123I]FP-CIT-SPECT asymmetry index to differentiate Parkinson's disease from vascular parkinsonism. Acta Neurol Scand. 2012;126:12–6.
13. Navarro-Otano J, Gaig C, Muxi A, Lomeña F, Compta Y, Buongiorno MT, et al. (123)I-MIBG cardiac uptake, smell identification and (123)I-FP-CIT SPECT in the differential diagnosis between vascular parkinsonism and Parkinson's disease. Parkinsonism Relat Disord. 2014;20:192–7.
14. Antonini A, Vitale C, Barone P, Cilia R, Righini A, Bonuccelli U, et al. The relationship between cerebral vascular disease and parkinsonism: the VADO study. Parkinsonism Relat Disord. 2012;18:775–80.
15. Benamer TS, Patterson J, Grosset DG, Booij J, de Bruin K, van Royen E, et al. Accurate differentiation of parkinsonism and essential tremor using visual assessment of [123I]-FP-CIT SPECT imaging: the [123I]-FP-CIT study group. Mov Disord. 2000;15:503–10.
16. Colloby SJ, O'Brien JT, Fenwick JD, Firbank MJ, Burn DJ, McKeith IG, et al. The application of statistical parametric mapping to 123-I-FPCIT SPECT in dementia with Lewy bodies, Alzheimer's disease and Parkinson's disease. Neuroimage. 2004;23:956–66.
17. Scherfler C, Seppi K, Donnemiller E, Goebel G, Brenneis C, Virgolini I, et al. Voxel-wise analysis of [123I]β-CIT SPECT differentiates the Parkinson variant of multiple system atrophy from idiopathic Parkinson's disease. Brain. 2005;128:1605–12.
18. Goebel G, Seppi K, Donnemiller E, Warwitz B, Wenning GK, Virgolini I, et al. A novel computer-assisted image analysis of [123I]β-CIT SPECT images improves the diagnostic accuracy of parkinsonian disorders. Eur J Nucl Med Mol Imaging. 2011;38(4):702–10.
19. Nocker M, Seppi K, Donnemiller E, Virgolini I, Wenning GK, Poewe W, et al. Progression of dopamine transporter decline in patients with the Parkinson variant of multiple system atrophy: a voxel-based analysis of [123I]β-CIT SPECT. Eur J Nucl Med Mol Imaging. 2012;39:1012–20.
20. Benitez-Rivero S, Marin-Oyaga VA, Garcia-Solis D, Huertas-Fernández I, García-Gómez FJ, Jesús S, et al. Clinical features and 123I-FP-CIT SPECT imaging in vascular parkinsonism and Parkinson's disease. J Neurol Neurosurg Psychiatry. 2013;84:122–9.
21. Koch W, Radau P, Hamann C, Tatsch K. Clinical testing of an optimized software solution for an automated, observer-independent evaluation of dopamine transporter SPECT studies. J Nucl Med. 2005;46:1109–18.
22. García-Gómez FJ, García-Solís D, Luis-Simón FJ, Marín-Oyaga VA, Carrillo F, Mir P, et al. Elaboration of the SPM template for the standardization of SPECT images with 123I-Ioflupane. Rev Esp Med Nucl Imagen Mol. 2013;32:350–6.
23. Varrone A, Dickson JC, Tossici-Bolt L, Sera T, Asenbaum S, Booij J, et al. European multicentre database of healthy controls for [123I]FP-CIT SPECT (ENC-DAT): age-related effects, gender differences and evaluation of different methods of analysis. Eur J Nucl Med Mol Imaging. 2013;40:213–27.
24. Erro R, Pappatà S, Amboni M, Vicidomini C, Longo K, Santangelo G, et al. Anxiety is associated with striatal dopamine transporter availability in newly diagnosed untreated Parkinson's disease patients. Parkinsonism Relat Disord. 2012;18:1034–8.
25. Friedman J, Hastie T, Tibshirani R. The Elements of Statistical Learning - Data Mining, Inference, and Prediction, 2nd Edition. Springer Series in Statistics. Berlin: Springer; 2009. p. 1–745.
26. Tatsch K, Poepperl G. Nigrostriatal dopamine terminal imaging with dopamine transporter SPECT: an update. J Nucl Med. 2013;54:1331–8.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES.

Q1. Please check the change made here if correct.
Q2. Please check this URL if correct.