

Multivariate analysis of 18F-DMFP PET data to assist the diagnosis of parkinsonism

Fermín Segovia^{1*}, Juan M. Gorriz¹, Javier Ramírez¹, Francisco Jesús Martínez-Murcia¹, Johannes Levin², Madeleine Schuberth², Matthias Brendel², Axel Rominger², Kai Boetzel², Gaëtan Garraux³, Christophe Phillips³

¹Department of Signal Theory, Networking and Communications, University of Granada, Spain, ²Department of Nuclear Medicine, University of Munich, Germany, ³University of Liège, Belgium

Submitted to Journal:
Frontiers in Neuroinformatics

ISSN:
1662-5196

Article type:
Original Research Article

Received on:
10 Dec 2016

Accepted on:
14 Mar 2017

Provisional PDF published on:
14 Mar 2017

Frontiers website link:
www.frontiersin.org

Citation:
Segovia F, Gorriz JM, Ramírez J, Martínez-murcia F, Levin J, Schuberth M, Brendel M, Rominger A, Boetzel K, Garraux G and Phillips C(2017) Multivariate analysis of 18F-DMFP PET data to assist the diagnosis of parkinsonism. *Front. Neuroinform.* 11:23. doi:10.3389/fninf.2017.00023

Copyright statement:
© 2017 Segovia, Gorriz, Ramírez, Martínez-murcia, Levin, Schuberth, Brendel, Rominger, Boetzel, Garraux and Phillips. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Provisional

Multivariate analysis of ^{18}F -DMFP PET data to assist the diagnosis of parkinsonism

Fermín Segovia^{1,2,*}, Juan M. Górriz¹, Javier Ramírez¹, Francisco J. Martínez-Murcia¹, Johannes Levin³, Madeleine Schuberth³, Matthias Brendel³, Axel Rominger³, Kai Bötzel³, Gaëtan Garraux² and Christophe Phillips²

¹*Department of Signal Theory, Networking and Communications, University of Granada, Granada, Spain*

²*Cyclotron Research Centre, University of Liège, Liège, Belgium*

³*Department of Nuclear Medicine, University of Munich, Munich, Germany*

Correspondence*:
Corresponding Author
fsegovia@ugr.es

2 ABSTRACT

An early and differential diagnosis of parkinsonian syndromes still remains a challenge mainly due to the similarity of their symptoms during the onset of the disease. Recently, ^{18}F -Desmethoxyfallypride (DMFP) has been suggested to increase the diagnostic precision as it is an effective radioligand that allows us to analyze postsynaptic dopamine $D_{2/3}$ receptors. Nevertheless, the analysis of these data is still poorly covered and its use limited. In order to address this challenge, this paper shows a novel model to automatically distinguish idiopathic parkinsonism from non-idiopathic variants using DMFP data. The proposed method is based on a multiple kernel support vector machine and uses the linear version of this classifier to identify some regions of interest: the olfactory bulb, thalamus and supplementary motor area. We evaluated the proposed model for both, the binary separation of idiopathic and non-idiopathic parkinsonism and the multigroup separation of parkinsonian variants. These systems achieved accuracy rates higher than 70%, outperforming DaTSCAN neuroimages for this purpose. In addition, a system that combined DaTSCAN and DMFP data was assessed.

Keywords: Multivariate analysis, ^{18}F -DMFP PET, Parkinson's disease, multiple kernel learning, support vector machine, multiple system atrophy, progressive supranuclear palsy

1 INTRODUCTION

Parkinson's disease (PD) has a lifetime risk of about 2%, making it the second most common neurodegenerative disease after Alzheimer's disease. A main risk factor for PD is aging and current demographic trends predict a doubling in the number of cases by 2050 (Bach et al., 2011). Clinically, the syndrome presents with the association of motor slowness (hypokinesia), with muscle rigidity and/or tremor and/or a postural instability (Greenberg et al., 2012). One of the neuropathological hallmarks of PD is a death of dopaminergic neurons in the substantia nigra (SN) *pars compacta*. Neuronal loss begins in the lateral ventral tier of the SN and throughout the illness this remains the most severely affected region. Post-mortem studies have shown that loss of cells from the SN in PD results mainly in profound dopamine (DA) depletion in the motor region of the striatum, nigral projections to the dorsal and caudal putamen being most affected.

Because a biochemical hallmark of PD is a deficiency of striatal DA, most imaging studies have focused on studying the problem directly, using a variety of methods (Antonini et al., 1997; Constantinescu et al., 2011; Garraux et al., 2013). Measurement of striatal ^{18}F -DOPA uptake with PET is still regarded by many as the "gold standard" for the diagnosis of PD. Several studies have reported that patients with sporadic PD have lost up to 50% of normal ^{18}F -DOPA uptake from the caudal putamen contralateral to the side

with the most severe symptoms compared with 20-30% on the ipsilateral side. The presynaptic striatal dopamine deficiency state can be assessed in vivo using other nuclear medicine techniques targeted to dopamine transporters (DaT) such as ^{123}I -ioflupane (also known by its tradename DaTSCAN) SPECT. These techniques are able to image a presynaptic striatal dopamine deficiency state shared by PD and atypical parkinsonian syndromes (APS) such as multiple system atrophy (MSA) (Wenning et al., 2013) and progressive supranuclear palsy (PSP) (Williams and Lees, 2009). They can be used to assist the clinicians to distinguish between these conditions and normal controls, essential and dystonic tremors, drug-induced and psychogenic parkinsonism. However, it is generally considered that imaging presynaptic striatal dopaminergic deficit does not reliably discriminate between PD and APS.

In comparison with PD, the latter are also characterized by a postsynaptic striatal impairment. This can be demonstrated using two main approaches. The first involves ^{18}F -FDG PET, which allows assessing the pattern of resting glucose metabolism throughout the brain. Several studies from independent groups consistently showed that FDG uptake pattern in the striatum and other regions can accurately discriminate APS from PD (Garraux et al., 2013; Ghaemi et al., 2002). These studies also showed that ^{18}F -FDG PET may also provide some clinically-relevant information for the distinction between APS subgroups. A postsynaptic striatal deficiency state can be specifically demonstrated using radioligands of $D_{2/3}$ striatal dopamine receptors. In PD, ^{11}C -Raclopride PET studies showed preserved (or even mildly increased in *de novo* cases) striatal binding. In contrast, striatal ^{11}C -Raclopride D_2 binding is reduced in APS although significant decreases might only present in around 50% of individuals. Aside from ^{11}C -Raclopride, ^{123}I -Iodobenzamide (^{123}I -IBZM) SPECT has also been used to image postsynaptic striatal dopamine $D_{2/3}$ receptors. However, a meta-analysis of IBZM-SPECT studies showed that its negative predictive value was low (Vlaar et al., 2007), while another study suggest that its diagnostic accuracy was considerably lower than that of ^{18}F -FDG PET (Hellwig et al., 2012). Finally, studies showed that ^{11}C -Raclopride PET and ^{18}F -FDG-PET have comparable discrimination accuracy performance for the distinction between PD and MSA (Antonini et al., 1997). One of the drawbacks of ^{11}C -Raclopride is its limited clinical availability due to the short physical half-life of ^{11}C (20 min), what requires to have a local cyclotron or radiochemistry unit in the neuroimaging center. This impediment to the wider use of PET for clinical dopamine receptor studies could be overcome through the use of suitable ^{18}F -labeled radioligands (Fougère et al., 2010). Thus, ^{18}F -desmethoxyfallypride (^{18}F -DMFP) has recently emerged as a possible alternative to ^{11}C -Raclopride. Its relatively long physical half-life (109 min) enables clinical applications for differentiating between PD and APS. A study found that differences in striatal binding in the posterior putamen between PD and APS provided the larger diagnostic performance using this technique (Fougère et al., 2010). Furthermore, as compared with IBZM and ^{11}C -Raclopride, ^{18}F -DMFP enables $D_{2/3}$ imaging in extrastriatal regions but this property has not been exploited for the distinction between PD and APS.

Neuroimaging data used to assist the diagnosis of parkinsonism is often analyzed by means of proprietary software to delimit regions of interest (ROIs) and quantify the radiopharmaceutical uptake (Tossici-Bolt et al., 2006; Morton et al., 2005). More sophisticated systems based on machine learning were proposed for FDG PET (Garraux et al., 2013) and DaTSCAN (Prashanth et al., 2014). For the latter modality, several machine learning approaches were presented. For example, in (Illán et al., 2012) the voxel intensities of DaTSCAN neuroimages were used as features along with several classifiers to separate PD patients and controls subjects. In (Towey et al., 2011) and (Segovia et al., 2012) two methods based respectively on Principal Component Analysis and Partial Least Squares were proposed to extract relevant features from DaTSCAN data. Structural data were also used to assist the diagnosis of PD (Rana et al., 2015), including the separation of PD and APS (Salvatore et al., 2014). However the validity of DMFP data to feed statistical classification procedures is still poorly covered. The intensity profile of this imaging modality is meaningfully different of those for DaTSCAN or FDG. In the latter case, the intensity is distributed over a large number of regions, conversely for DaTSCAN most of the signal is gathered in the striatum. DMFP is half-way between DaTSCAN and FDG in terms of its intensity profile: a large proportion of the activity is located in the striatum but other regions also have intensity values large enough to contain patterns allowing the inter-subject discrimination.

In this work, we present a model to automatically distinguish between idiopathic parkinsonism and the non-idiopathic variants using DMFP data. It is based on the application of a machine learning algorithm

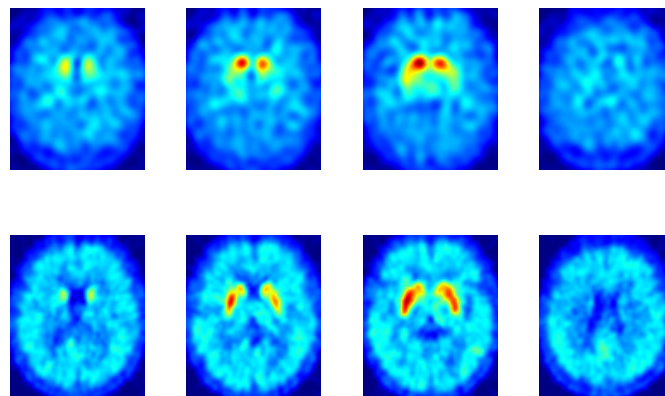


Figure 1. Comparison between a DaTSCAN (top row) and a DMFP (bottom row) neuroimage from a patient diagnosed with Parkinson's disease.

that consider several regions of interest by means of a multiple kernel approach. We evaluated the proposed model for both, the binary separation of idiopathic and non-idiopathic parkinsonism and the multigroup separation of PD, MSA and PSP patients. In addition, we studied the usefulness of combining DMFP and DaTSCAN data in a single computer system. This work is intended to be the baseline of multivariate analyses of DMFP data.

2 MATERIALS AND METHODS

2.1 Ethics Statement

Each patient (or a close relative) gave written informed consent to participate in the study and the protocol was accepted by the Ethics Committee of the University of Munich. All the data were anonymized by the clinicians who acquired them before being considered in this work.

2.2 Subjects

Data from 87 subjects showing parkinsonian movement disorders were used for testing purposes (demographic details and groups distribution are gathered in table 1). Subjects on medication with drugs that have (or are suspected to have) effects on dopaminergic transporters were excluded. A ^{18}F -DMFP PET image and a DaTSCAN SPECT image were collected for each subject during the first visit. The former images were acquired 60 minutes after the ^{18}F -DMFP injection. Three (3) frames of 10 minutes each were recorded using a Siemens/CTI camera. Images were reconstructed as 128×128 matrices of 2×2 mm voxels by filtered backprojection using a Hann filter with a cutoff frequency of 0.5 Nyquist. The scattered and random events as well as dead time issues were also corrected (Turkington, 2011). After verification of the absence of important head motion between frames (patients had their head immobilized during the emission recording), the 3 frames were summed and preprocessed (Fougère et al., 2010). DaTSCAN data were acquired using a GE Healthcare camera and according to widely accepted criteria (Koch et al., 2005).

All patients were followed clinically for approximately 2 years after SPECT and PET examinations, at which time the clinical differential diagnoses were assessed by clinicians on the basis of last observations and according to the United Kingdom Parkinson Disease Society Brain Bank Diagnostic Criteria for Parkinson Disease (Hughes et al., 2002) and the second consensus statement on the diagnosis of multiple-system atrophy (Gilman et al., 2008) as well as the established criteria for the diagnosis of progressive supranuclear palsy (Litvan et al., 1996). According to these criteria, 39 patients were labeled as idiopathic parkinsonism and the remaining 48 subjects either MSA or PSP. It is worth noting that all the images were acquired during the first examinations and, therefore, they correspond to early stages of the disorders.

Table 1. Demographic details of the patients considered in this work (μ and σ stand for the average and the standard deviation respectively).

	#	Sex		Age		
		M	F	μ	σ	range
PD	39	22	17	61.38	11.14	35-81
MSA	24	20	4	68.42	10.73	43-85
PSP	24	12	12	69.29	7.33	55-84

2.3 Data preprocessing

After the image reconstruction the images were spatially normalized using the template matching approach implemented in the Statistical Parametric Mapping (SPM) version 8 (Friston and *et al.*, 2007). This procedure ensures that any given voxel in different images refers to the same anatomical position across the brains. It was based on the affine part of the SPM normalization procedure, a method that assumes a general affine model with 12 parameters and a Bayesian framework that maximizes the product of the prior function (which is based on the probability of obtaining a particular set of zooms and shears) and the likelihood function (derived from the residual squared difference between the template and the source image) (Ashburner et al., 1997). In order to build the templates for DMFP and DaTSCAN data, only the neuroimages from the idiopathic group was used because of this group is more homogeneous, whereas the non-idiopathic one is the union of the MSA and PSP patients. Thus, idiopathic PD data were first registered to a randomly chosen one. The resulting images and their hemisphere midplane reflections (ensuring a symmetric template) were then averaged and smoothed (8 mm FWHM Gaussian kernel) before being used to spatially normalize the whole set of images. As a result, we got brain volumes with $79 \times 95 \times 68$ voxels of $2 \times 2 \times 2$ mm.

In addition, after the spatial normalization the intensity of the images were also normalized to a value I_{max} , obtained by averaging the 0.1% of the highest intensities per image, as described in (Saxena et al., 1998).

2.4 Multivariate analysis based on machine learning

Unlike univariate analyses where each voxel is independently analyzed, multivariate approaches analyze a neuroimage as a whole and explicitly consider the relationships between voxels (Schrouff et al., 2013). Effects comprised in the data, including activations and confounding and errors effects are assessed statistically both at each voxel and as interactions between voxels (Friston and Büchel, 2007).

A large proportion of multivariate analyses for neuroimaging data are based on statistical classification, such as Support Vector Machine (SVM) (Vapnik, 1998). During the training step, a binary statistical classifier builds a function $f: \mathbb{R}^N \rightarrow \pm 1$ using a set of labeled samples so that f is able to predict the label of new unseen samples. SVM calculates this function by estimating the hyperplane that has the largest distance to the closest sample with any label. Then, the label of a new sample is estimated according to the side of the hyperplane in which the new sample is. Mathematically, the decision hyperplane (also named maximal margin hyperplane) is defined as:

$$g(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0 = 0, \quad (1)$$

where \mathbf{w} is the weight vector, orthogonal to the decision hyperplane, and w_0 is the bias term. Calculating the hyperplane involves solving the following problem:

$$\begin{aligned} & \text{minimize} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \\ & \text{subject to} \quad y_i(\mathbf{w}^T \mathbf{x} + w_0) \geq 1 - \xi_i, \quad i = 1, 2, \dots, N \end{aligned} \quad (2)$$

where C , N and ξ_i stand for a predefined trade-off parameter between model simplicity and classification error, the number of training instances and the slack variables, respectively. This problem can be simplified by applying Lagrangian functions and results in:

$$\begin{aligned} & \text{maximize} \quad \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j) \\ & \text{subject to} \quad \sum_{i=1}^N \alpha_i y_i = 0 \\ & \quad \quad \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, N \end{aligned} \quad (3)$$

where α is the vector of dual variables corresponding to each separation constraint, and $k(\mathbf{x}_i, \mathbf{x}_j)$ is a function $\mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}$ known as “kernel” (Müller et al., 2001). For linear SVM, $k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i \mathbf{x}_j$.

2.5 Analysis based on multiple kernel learning

After the preprocessing steps, the data were automatically parceled using the well-known Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002). This procedure allows isolating specific regions such as the striatum or the olfactory bulb, and that way, they can be independently analyzed.

Neuroimaging from both DaTSCAN and DMFP radioligands gathers most of the total activation in the striatum. However, DMFP-based data also contain information in other regions (see Fig. 1). There is therefore a reasonable prospect that DMFP data contain useful information for a diagnosis in areas outside the striatum. This would be overlooked during the visual examination of the images, while automatic procedure (based on machine learning) may efficiently exploit this information. In order to better account for the different scales and relevances of the various regions, we used a multiple kernel learning (MKL) procedure along with a SVM classifier.

In a MKL procedure, two or more kernels are built from feature subsets and then combined by means of a predefined function (Gönen and Alpaydm, 2011). In this work, we used a kernel per each region of interest and a linear combination function that provides a good performance with relatively low computational burden:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sum_{m=1}^{N_k} q_m k_m(\mathbf{x}_i^m, \mathbf{x}_j^m) \quad (4)$$

where N_k is the number of kernels; q_m stands for the weight of kernel k_m (estimated through cross-validation); $\mathbf{x}_i, \mathbf{x}_j$ are two feature vectors and $\mathbf{x}_i^m, \mathbf{x}_j^m$ are subset of $\mathbf{x}_i, \mathbf{x}_j$ with only the features used for kernel k_m .

2.6 Identifying regions of interest

A linear SVM trained classifier allows us to analyze the weight assigned by the classifier to each feature included in the training set (related with importance of that feature in the separation problem). This information can be directly extracted from the weight vector, \mathbf{w} , defined in equation 1. Thus, using all the brain voxels as feature we estimated the importance of each voxel in the separation problem. Then, the weight/importance of each region was computed as

$$w_r = \frac{\sum_{i=1}^{N_r} w_{x_i}}{N_r} \quad \forall x_i \in r \quad (5)$$

where w_{x_i} is weight corresponding to the voxel x_i and $x_i, i = 1, \dots, N_r$ are the voxels in the region r .

A map containing the voxel weights (rearranged into brain form) is shown in figure 2. It was computed by training a binary SVM classifier (linear kernel) with DMFP data from PD patients in one group and

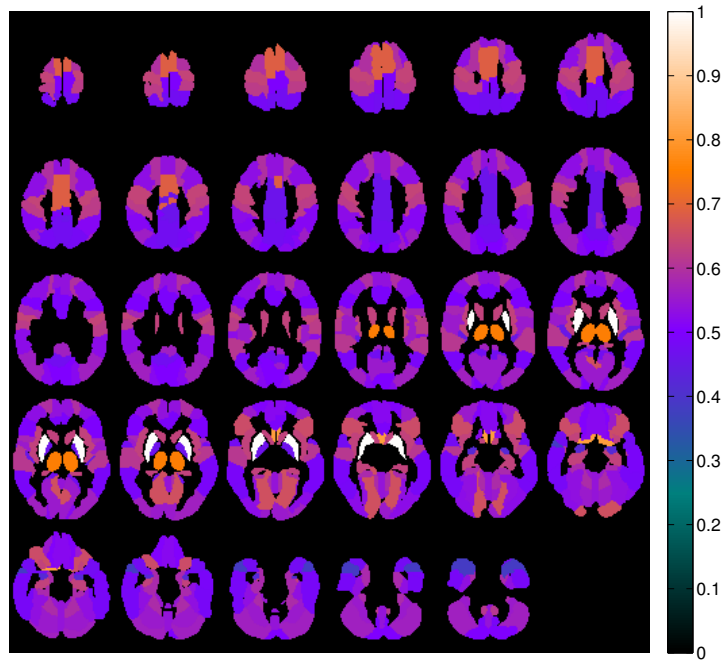


Figure 2. Importance of each brain region for the classification problem. The values were estimated from the voxels weight computed by a linear SVM classifier, and normalized to the range $[0, 1]$.

177 data from MSA and PSP patients in other group. As expected, the most important region is the striatum,
 178 specifically the putamen area. The high weight assigned to other regions such as the olfactory bulb or the
 179 thalamus and the supplementary motor area is also interesting. Figure 3 shows the mean intensity of these
 180 five areas for all the DMFP neuroimages in our dataset. Note that the mean intensity of the putamen is
 181 largely higher than those of the thalamus, olfactory bulb and supplementary motor area. These differences
 182 were successfully addressed by the MKL-based proposed approach. Figure 3 also suggests that there is no
 183 important differences between groups when only the mean of the preselected ROIs is analyzed.

3 EXPERIMENT AND RESULTS

184 Several experiments were conducted in order to evaluate the application of the multivariate approach
 185 described above to our DMFP dataset. Initially, we address the binary diagnosis problem, i.e., the
 186 separation of two groups: idiopathic and non-idiopathic (includes MSA and PSP disorders) parkinsonism.
 187 Table 2 shows the classification measures obtained in this case and compares the results obtained by the
 188 proposed MKL approach with the ones achieved by classical approaches such as using only the voxels
 189 at the striatum or using all the voxels inside the brain. The accuracy, sensitivity and specificity of the
 190 systems were estimated using a k -fold ($k = 10$) cross-validation scheme. All the systems were based on
 191 SVM classification and the parameters (trade-off parameter C and kernel weights q_m for MKL approaches)
 192 were estimated using a grid search (Varma and Simon, 2006). Values $C = 2^e, e = \{-3, -2, \dots, 5\}$ and
 193 $q_m = \{0.1, 0.2, \dots, 0.9\}$ satisfying $\sum_{m=1}^{N_k} q_m = 1$ were assessed.

194 In order to avoid biased results, the selection of the regions used in the MKL approach was performed
 195 using only the training dataset, i.e. in each iteration of the cross-validation loop, we selected the 5 regions
 196 with the most weight (note that the striatum is divided in two regions in the AAL atlas: putamen and
 197 caudate nucleus). The putamen, caudate nucleus, olfactory bulb and thalamus were among the 5 selected
 198 regions in all the iterations. The supplementary motor area was the fifth region in 92% of all cases. The
 199 pseudo-code corresponding to this experiment is shown in Algorithm 1.

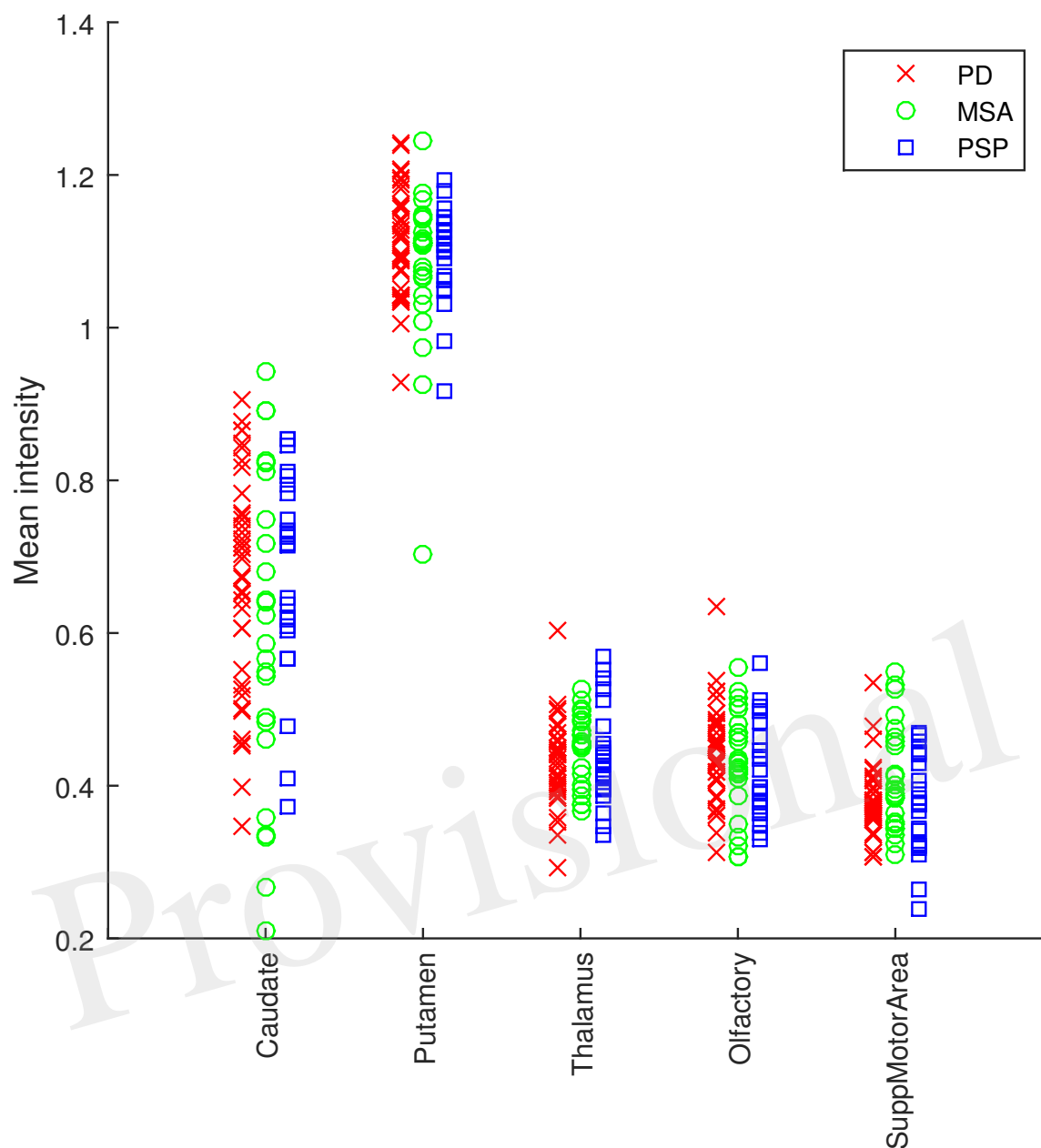


Figure 3. Analysis of the average intensity of the 5 regions of interest found in the analysis of DMFP data. Each patient is represented by 5 values: the average intensity of his DMFP neuroimage on the caudate, putamen, thalamus, olfactory and supplementary motor area. The values of each region are grouped by the patient group: PD (red crosses), MSA (green circles) and PSP (blue squares).

As shown in table 2, the proposed method based on multiple kernel learning achieved an accuracy rate of 73.56%. The statistical significance of this measure was assessed by means of a permutation test. In this procedure, the classification algorithm was run 1000 times using different label sets (generated as random permutations of the original label set). A p -value was then calculated as the fraction of the executions in which the accuracy was greater than or equal to the accuracy observed when using the correct labels (Pereira et al., 2009). As a result, a p -value of 0.001 was obtained. Figure 4 shows the histogram corresponding to the accuracy rates obtained in the procedure.

Algorithm 1: Evaluation procedure**Input:** The neuroimage dataset, $X = I_1, \dots, I_n$, and corresponding labels, $Y = Y_1, \dots, Y_n$ **Output:** Accuracy measures of the method

```

foreach fold  $f$  do                                     // Main cross-validation loop
     $[X_{trn}, X_{tst}, Y_{trn}, Y_{tst}] = \text{extract training and test sets from } X \text{ and } Y \text{ for fold } f$ 
     $g = \text{train a linear SVM with parameter } C = 1 \text{ using } X_{trn} \text{ and } Y_{trn}$ 
     $w = \text{extract weight map from } g$                                      // equation 1
    foreach region  $r \in \text{AAL atlas}$  do
         $w_r = \text{compute average weight for region } r \text{ in } w$                                      // equation 5
    end
     $\tilde{X}_{trn} = \text{select from } X_{trn} \text{ the voxels corresponding to the regions with highest } w_r$ 
    foreach parameter set  $P_i \in P$  do
        foreach fold  $f^*$  do                                     // Inner cross-validation loop for parameter selection
             $[X_{trn}^*, X_{tst}^*, Y_{trn}^*, Y_{tst}^*] = \text{extract training and test sets from } \tilde{X}_{trn} \text{ and } Y_{trn} \text{ for fold } f^*$ 
             $g = \text{train a multikernel SVM with parameters } P_i \text{ using } X_{trn}^* \text{ and } Y_{trn}^*$ 
             $\hat{Y}_{tst}^* = \text{evaluate } X_{tst}^* \text{ using } g$ 
        end
         $\hat{Y}^* = \text{combination of all } \hat{Y}_{tst}^* \text{ variables}$ 
         $A_{P_i} = \text{Estimate the accuracy of the model by comparing } Y_{trn} \text{ and } \hat{Y}^*$ 
    end
     $P_B = \text{select the parameter set } P_i \text{ with highest } A_{P_i}$ 
     $g = \text{train a multikernel SVM with parameters } P_B \text{ using } \tilde{X}_{trn} \text{ and } Y_{trn}$ 
     $\hat{Y}_{tst} = \text{evaluate } X_{tst} \text{ using } g$ 
end
 $\hat{Y} = \text{combination of all } \hat{Y}_{tst} \text{ variables}$ 
Estimate the accuracy of the model by comparing  $Y$  and  $\hat{Y}$ 

```

$P_i = \{C, q_1, q_2, \dots, q_{N_k}\}$ denotes the parameter set required by the multikernel SVM classifier. C stands for the trade-off parameter of the SVM algorithm (equation 2) whereas $q_m, m = 1, 2, \dots, N_k$ is the weight of the kernel k_m (equation 4). P gathers all P_i sets, i.e. all possible combinations of kernel weights and values for parameter C .

207 Subsequently, the tree-groups separation problem was addressed. In this case, the system was trained to
 208 differentiate between the disorders considered in our study: PD, MSA and PSP. To this end, an ensemble of
 209 SVM classifiers and the one-against-one strategy was used. The results are gathered in table 3.

210 In both classification approaches, binary and multiclass, the results obtained by using DMFP data were
 211 compared with the ones obtained by DaTSCAN neuroimages. For the latter neuroimage modality only
 212 the voxels at the striatum were considered as it is common in the field. Additionally, an experiment that
 213 combined both neuroimages modalities (DaTSCAN and DMFP) was performed. It allows us to evaluate if
 214 the information contained in DaTSCAN neuroimages can complement that contained in DMFP data. To
 215 this end, we extend the MKL approach to include 6 regions: the 5 regions selected for DMFP data and
 216 striatum from the DaTSCAN neuroimages. The results are included in tables 2 and 3.

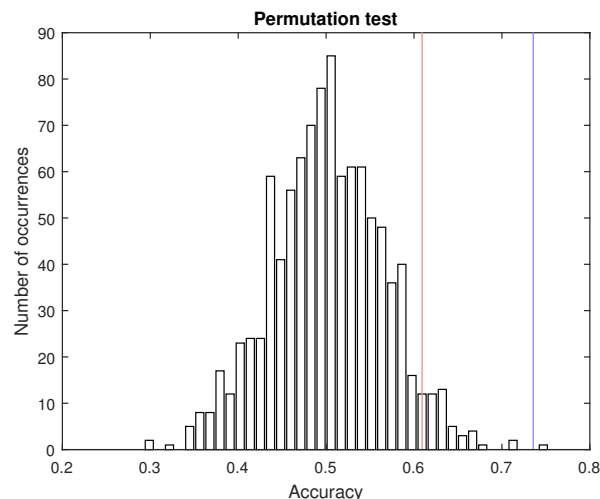


Figure 4. Pemutation test. Histogram of the accuracy rates achieved by using randomly generated label sets (1000 repetitions) and the proposed multikernel-based method. Red and blue lines are respectively the accuracy associated with a p-value of 0.05 and the accuracy obtained when using the true labels (73.56%).

Table 2. Accuracy, sensitivity and specificity obtained by the proposed approach and other classical approaches when separating neuroimaging data from idiopathic and non-idiopathic parkinsonism.

	Accuracy	Sensitivity	Specificity
Using DMFP data:			
Voxels in the striatum	68.96%	79.17%	56.41%
All the voxels inside the brain	67.82%	75.00%	58.97%
MKL approach (5 regions)	73.56%	77.08%	69.23%
Using DaTSCAN data:			
Striatum voxels	59.77%	62.50%	56.41%
Using DMFP and DaTSCAN data:			
Voxels in the striatum	70.11%	77.08%	61.54%
All the voxels inside the brain	63.22%	70.83%	53.85%
MKL approach (6 regions)	72.41%	75.00%	69.23%

3.1 Univariate analysis

For the sake of completeness, we compared the regions identified in the previous section with the ones obtained by means of a *t*-test. A smoothed version of the DMFP PET neuroimages (Gaussian filter of 8 mm FWHM) was analyzed using a 2-sample *t*-test in SPM. The results indicate that patients with idiopathic parkinsonism have lower dopamine levels than non-idiopathic patients in the striatum, thalamus and supplementary motor area. Specifically, clusters showing significant difference appeared in the thalamus, anterior cingulate cortex and putamen with *z*-values up to 4.74 (see Fig. 5). Some differences ($p < 0.001$) also appeared in the supplementary motor area (454 mm³) and caudate nucleus (22 mm³).

3.2 ROC analysis

Finally, receiver operating characteristic (ROC) curves were computed to assess the trade off between sensitivity and specificity provided by the developed systems. These curves provide an estimation of the performance of a classification procedure, not only in terms of sensitivity and specificity but also in terms

Table 3. Accuracy achieved by a multiclass classification system when separating PD, MSA and PSP neuroimages.

	Global acc.	PD acc.	MSA acc.	PSP acc.
Using DMFP data:				
Voxels in the striatum	56.32%	56.41%	66.67%	45.83%
All the voxels inside the brain	49.43%	64.10%	37.50%	37.50%
MKL approach (5 regions)	66.67%	82.05%	54.17%	54.17%
Using DaTSCAN data:				
Striatum voxels	44.83%	61.54%	29.17%	33.33%
Using DMFP and DaTSCAN data:				
MKL approach (6 regions)	62.07%	76.92%	54.17%	45.83%

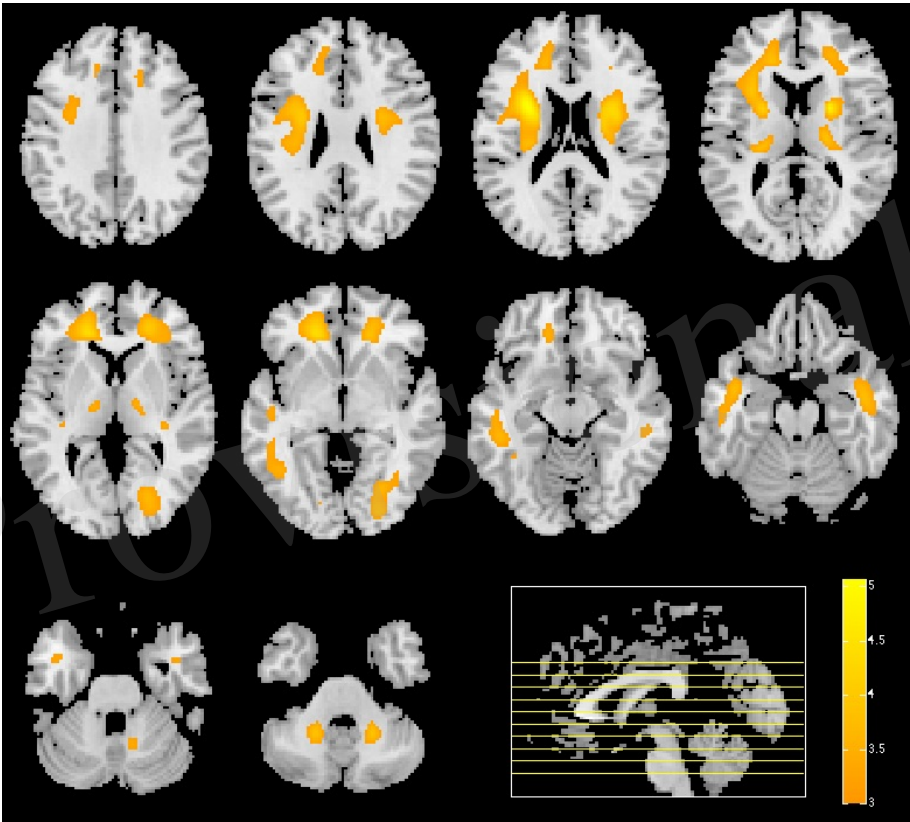


Figure 5. Result of the univariate analysis. *t*-test comparing patients with idiopathic and non-idiopathic parkinsonism. Regions in orange/yellow are significantly lower ($p < 0.001$, uncorrected) in idiopathic compared with non-idiopathic patients.

of the global performance by measuring the area under the curve (AUC). Curves and AUCs are shown in Fig. 6.

4 DISCUSSION AND CONCLUSIONS

Separating idiopathic and non-idiopathic parkinsonian patients in an early stage is a challenge because both groups can show similar clinical signs and symptoms (Litvan, 1999). In addition, neuroimaging techniques

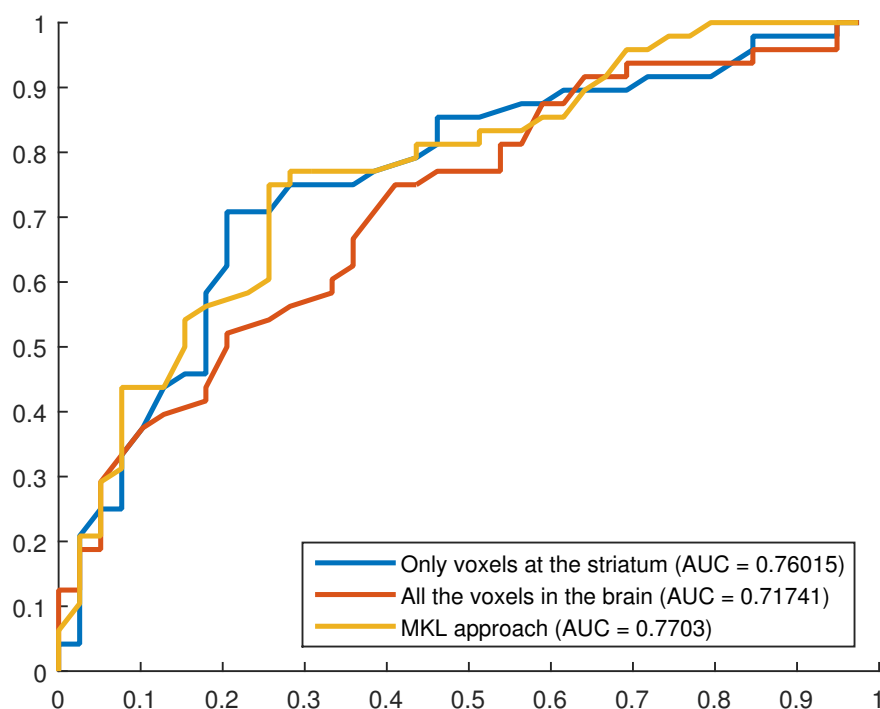


Figure 6. ROC curves for three binary classification systems using DMFP data: i) using only the voxels at the striatum, ii) using all the voxels in the brain, iii) using the proposed MKL-based approach. The AUC for each curve is shown in the legend.

such as DaTSCAN, widely used to assist the diagnosis of PD, may have difficulty to differentiate between PD and MSA or PSP, as shown in our experiments. In this work, we evaluated the use of ^{18}F -DMFP PET in computer systems to distinguish between three parkinsonian syndromes (PD, MSA and PSP), yielding accuracy rates about 70%. Using a multivariate analysis, we found that the striatum, olfactory bulb, thalamus and supplementary motor area are the most important regions to separate the groups. These findings were partly corroborated through univariate analysis and are consistent with previous results reported in recent studies. For example, in (Chen et al., 2014) the authors found that the volumes of the olfactory bulb and tract were significantly reduced in idiopathic PD patients compared to normal subjects and MSA patients. This corroborate the idea that the mechanisms that produce idiopathic PD also affect this region.

This result obtained by hypothesis free analysis of DMFP data is in line with clinical and histopathological observations. Clinically, olfactory testing can be used to identify patients in the premotor phase of Parkinson's disease (Berardelli et al., 2013). The olfactory deficits is caused by PD pathology in the olfactory bulb and can already be observed stage I of the Braak's staging of PD pathology (Braak et al., 2002). The key role of striatal pathology in MSA and PSP is widely established and also reflected in our hypothesis free approach (Litvan et al., 1997). On the other hand, evidences of the involvement of the motor cortex in PD were reported in (Lindenbach and Bishop, 2013; Vacherot et al., 2010), although results have been controversial (Viaro et al., 2011). Finally, the relationship between PD and thalamus is not yet proven on pathological bases and therefore might reflect functional disturbances on basal ganglia circuits that involve the thalamus. In short, our experiments corroborate that there exist differences in those regions (olfactory bulb, supplementary motor area and thalamus) between PD, MSA and PSP patients.

In this manuscript, we propose a methodology to analyze DMFP data for diagnostic purposes. The primary idea is to use not only the striatum region but additionally to take into account the other regions of the brain. In order to address the small sample size problem and overcome the differences in terms

of size between different regions, we propose to use a MKL approach, which is considered to be a good solution to combine heterogeneous data sources into one classification procedure (Gönen and Alpaydın, 2011; Segovia et al., 2014). In some sense, a MKL classification procedure as the one used in this work can be seen as a kind of regularization process that assigns the same weight to all the voxels in a specific region instead of a different weight per each voxel as a standard SVM classifier does.

Two classification approaches were addressed. The first one consisted on separating idiopathic and non-idiopathic parkinsonism (binary classification) whereas the second one was able to differentiate between PD, MSA and PSP data (multiclass classification). The proposed methodology based on MKL classification yields higher accuracy rates than classical approaches in both cases, however these results are limited by the relative small number of patients included in this study. Unfortunately the availability of these kind of data, especially the ones acquired using new or uncommon radiopharmaceuticals is limited. The significance of the accuracy rate achieved by the proposed method was estimated by means of permutation tests obtaining a p -value = 0.001. Recently, Noirhomme et al. demonstrated that permutation tests are more adequate than classical approaches as binomial assessment to calculate the significance of classification measures estimated using cross-validation schemes (Noirhomme et al., 2014). It is worth noting the relatively low accuracy rates obtained for MSA and PSP classes in the multiclass approach. These results suggest that the differences between these disorders do not appear in DMFP data. This is not unique to this data modality. Previous studies based on FDG data (Garraux et al., 2013) also faced the same issue. Finally, we included an experiment combining DMFP and DaTSCAN data in the same system. The results showed that including DaTSCAN data helps in the separation problem when only the striatum voxels were used however, when the additional DMFP regions proposed in this work (the putamen, caudate nucleus, olfactory bulb, thalamus and supplementary motor area) were included (in the MKL-based approach) DaTSCAN data did not increase the separation ability of the classifier. All in all, DaTSCAN data is probably not necessary to differentiate between PD, MSA and PSP patients if we use new systems based on DMFP data that take advantage of all the information contained in these neuroimages (not only the striatum).

ACKNOWLEDGMENTS

This work was supported by and the MINECO under the TEC2012-34306 and TEC2015-64718-R projects and the Ministry of Economy, Innovation, Science and Employment of the Junta de Andalucía under the Excellence Projects P09-TIC-4530 and P11-TIC-7103 and a Talent Hub project approved by the Andalucía Talent Hub Program launched by the Andalusian Knowledge Agency, co-funded by the European Union's Seventh Framework Program, Marie Skłodowska-Curie actions (COFUND – Grant Agreement no 291780) and the Ministry of Economy, Innovation, Science and Employment of the Junta de Andalucía. The work was also supported by the Universities of Liège (Belgium), Munich (Germany) and Granada (Spain). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

- Antonini, A., Leenders, K. L., Vontobel, P., Maguire, R. P., Missimer, J., Psylla, M., et al. (1997). Complementary PET studies of striatal neuronal function in the differential diagnosis between multiple system atrophy and parkinson's disease. *Brain* 120, 2187–2195. doi:10.1093/brain/120.12.2187
- Ashburner, J., Neelin, P., Collins, D. L., Evans, A., and Friston, K. (1997). Incorporating prior knowledge into image registration. *NeuroImage* 6, 344–352. doi:10.1006/nimg.1997.0299
- Bach, J., Ziegler, U., Deuschl, G., Dodel, R., and Doblhammer-Reiter, G. (2011). Projected numbers of people with movement disorders in the years 2030 and 2050. *Movement Disorders* 26, 2286–2290. doi:10.1002/mds.23878
- Berardelli, A., Wenning, G. K., Antonini, A., Berg, D., Bloem, B. R., Bonifati, V., et al. (2013). EFNS/MDS-ES recommendations for the diagnosis of parkinson's disease. *European Journal of Neurology* 20, 16–34. doi:10.1111/ene.12022
- Braak, H., Del Tredici, K., Bratzke, H., Hamm-Clement, J., Sandmann-Keil, D., and Rüb, U. (2002). Staging of the intracerebral inclusion body pathology associated with idiopathic parkinson's disease (preclinical and clinical stages). *Journal of Neurology* 249 Suppl 3, III/1–5
- Chen, S., Tan, H.-y., Wu, Z.-h., Sun, C.-p., He, J.-x., Li, X.-c., et al. (2014). Imaging of olfactory bulb and gray matter volumes in brain areas associated with olfactory function in patients with parkinson's

- disease and multiple system atrophy. *European Journal of Radiology* 83, 564–570. doi:10.1016/j.ejrad.2013.11.024
- Constantinescu, C. C., Coleman, R. A., Pan, M., and Mukherjee, J. (2011). Striatal and extra-striatal microPET imaging of D2/D3 dopamine receptors in rat brain with [18F]Fallypride and [18F]Desmethoxyfallypride. *Synapse (New York, N.Y.)* 65, 778–787. doi:10.1002/syn.20904
- Fougère, C. I., Pöpperl, G., Levin, J., Wängler, B., Böning, G., Uebleis, C., et al. (2010). The value of the dopamine d2/3 receptor ligand 18F-Desmethoxyfallypride for the differentiation of idiopathic and nonidiopathic parkinsonian syndromes. *Journal of Nuclear Medicine* 51, 581–587. doi:10.2967/jnumed.109.071811
- Friston, K. and Büchel, C. (2007). Chapter 37 - functional connectivity: Eigenimages and multivariate analyses. In *Statistical Parametric Mapping*, eds. K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny (London: Academic Press). 492–507
- Friston, K. and et al. (eds.) (2007). *Statistical Parametric Mapping: The Analysis of Functional Brain Images* (Academic Press)
- Garraux, G., Phillips, C., Schrouff, J., Kreisler, A., Lemaire, C., Degueldre, C., et al. (2013). Multiclass classification of FDG PET scans for the distinction between parkinson's disease and atypical parkinsonian syndromes. *NeuroImage: Clinical* 2, 883–893. doi:10.1016/j.nicl.2013.06.004
- Ghaemi, M., Hilker, R., Rudolf, J., Sobesky, J., and Heiss, W. (2002). Differentiating multiple system atrophy from parkinson's disease: contribution of striatal and midbrain MRI volumetry and multi-tracer PET imaging. *Journal of Neurology, Neurosurgery & Psychiatry* 73, 517–523. doi:10.1136/jnnp.73.5.517
- Gilman, S., Wenning, G. K., Low, P. A., Brooks, D. J., Mathias, C. J., Trojanowski, J. Q., et al. (2008). Second consensus statement on the diagnosis of multiple system atrophy. *Neurology* 71, 670–676. doi:10.1212/01.wnl.0000324625.00404.15
- Gönen, M. and Alpaydın, E. (2011). Multiple kernel learning algorithms. *J. Mach. Learn. Res.* 12, 2211–2268
- Greenberg, D., Aminoff, M., and Simon, R. (2012). *Clinical Neurology* (New York: McGraw-Hill Professional), 8 edition edn.
- Hellwig, S., Amtage, F., Kreft, A., Buchert, R., Winz, O. H., Vach, W., et al. (2012). 18F-FDG-PET is superior to 123I-IBZM-SPECT for the differential diagnosis of parkinsonism. *Neurology* 79, 1314–1322. doi:10.1212/WNL.0b013e31826c1b0a
- Hughes, A. J., Daniel, S. E., Ben-Shlomo, Y., and Lees, A. J. (2002). The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain* 125, 861–870. doi:10.1093/brain/awf080
- Illán, I. A., Górriz, J. M., Ramírez, J., Segovia, F., Jiménez-Hoyuela, J. M., and Lozano, S. J. O. (2012). Automatic assistance to parkinson's disease diagnosis in DaTSCAN SPECT imaging. *Medical Physics* 39, 5971–5980. doi:10.1118/1.4742055
- Koch, W., Radau, P. E., Hamann, C., and Tatsch, K. (2005). Clinical testing of an optimized software solution for an automated, observer-independent evaluation of dopamine transporter SPECT studies. *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine* 46, 1109–1118
- Lindenbach, D. and Bishop, C. (2013). Critical involvement of the motor cortex in the pathophysiology and treatment of parkinson's disease. *Neuroscience & Biobehavioral Reviews* 37, 2737–2750. doi:10.1016/j.neubiorev.2013.09.008
- Litvan, I. (1999). Recent advances in atypical parkinsonian disorders. *Current Opinion in Neurology August* 1999 12, 441–446
- Litvan, I., Agid, Y., Calne, D., Campbell, G., Dubois, B., Duvoisin, R. C., et al. (1996). Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. *Neurology* 47, 1–9
- Litvan, I., Campbell, G., Mangone, C. A., Verny, M., McKee, A., Chaudhuri, K. R., et al. (1997). Which clinical features differentiate progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome) from related disorders? a clinicopathological study. *Brain* 120, 65–74. doi:10.1093/brain/120.1.65
- Morton, R. J., Guy, M. J., Clauss, R., Hinton, P. J., Marshall, C. A., and Clarke, E. A. (2005). Comparison of different methods of DaTSCAN quantification. *Nuclear Medicine Communications* 26, 1139–1146
- Müller, K., Mika, S., Rätsch, G., Tsuda, K., and Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks* 12, 181–201. doi:10.1109/72.914517

- Noirhomme, Q., Lesenfants, D., Gomez, F., Soddu, A., Schrouff, J., Garraux, G., et al. (2014). Biased binomial assessment of cross-validated estimation of classification accuracies illustrated in diagnosis predictions. *NeuroImage: Clinical* 4, 687–694. doi:10.1016/j.nicl.2014.04.004
- Pereira, F., Mitchell, T., and Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *NeuroImage* 45, S199–S209. doi:10.1016/j.neuroimage.2008.11.007
- Prashanth, R., Dutta Roy, S., Mandal, P. K., and Ghosh, S. (2014). Automatic classification and prediction models for early Parkinson's disease diagnosis from SPECT imaging. *Expert Systems with Applications* 41, 3333–3342. doi:10.1016/j.eswa.2013.11.031
- Rana, B., Juneja, A., Saxena, M., Gudwani, S., Senthil Kumaran, S., Agrawal, R. K., et al. (2015). Regions-of-interest based automated diagnosis of Parkinson's disease using T1-weighted MRI. *Expert Systems with Applications* 42, 4506–4516. doi:10.1016/j.eswa.2015.01.062
- Salvatore, C., Cerasa, A., Castiglioni, I., Gallivanone, F., Augimeri, A., Lopez, M., et al. (2014). Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and Progressive Supranuclear Palsy. *Journal of Neuroscience Methods* 222, 230–237. doi:10.1016/j.jneumeth.2013.11.016
- Saxena, P., Pavel, D. G., Quintana, J. C., and Horwitz, B. (1998). An automatic thresholdbased scaling method for enhancing the usefulness of Tc-HMPAO SPECT in the diagnosis of Alzheimers disease. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI, Lecture Notes in Computer Science*. vol. 1496, 623–630
- Schrouff, J., Rosa, M. J., Rondina, J. M., Marquand, A. F., Chu, C., Ashburner, J., et al. (2013). PRoNTTo: pattern recognition for neuroimaging toolbox. *Neuroinformatics* 11, 319–337. doi:10.1007/s12021-013-9178-1
- Segovia, F., Bastin, C., Salmon, E., Górriz, J. M., Ramírez, J., and Phillips, C. (2014). Combining PET images and neuropsychological test data for automatic diagnosis of alzheimer's disease. *PLoS ONE* 9, e88687. doi:10.1371/journal.pone.0088687
- Segovia, F., Górriz, J. M., Ramírez, J., Álvarez, I., Jiménez-Hoyuela, J. M., and Ortega, S. J. (2012). Improved parkinsonism diagnosis using a partial least squares based approach. *Medical Physics* 39, 4395–4403. doi:10.1118/1.4730289
- Tossici-Bolt, L., Hoffmann, S., Kemp, P., Mehta, R., and Fleming, J. (2006). Quantification of [123I]FP-CIT SPECT brain images: an accurate technique for measurement of the specific binding ratio. *European Journal of Nuclear Medicine and Molecular Imaging* 33, 1491–1499. doi:10.1007/s00259-006-0155-x
- Towey, D. J., Bain, P. G., and Nijran, K. S. (2011). Automatic classification of 123i-FP-CIT (DaTSCAN) SPECT images. *Nuclear Medicine Communications August 2011* 32, 699–707. doi:10.1097/MNM.0b013e328347cd09
- Turkington, T. G. (2011). PET Imaging Basics. In *Clinical PET-CT in Radiology*, eds. P. Shreve and D. W. Townsend (Springer New York). 21–28. doi:10.1007/978-0-387-48902-5_2
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI Single-Subject brain. *NeuroImage* 15, 273–289. doi:10.1006/nimg.2001.0978
- Vacherot, F., Attarian, S., Eusebio, A., and Azulay, J. P. (2010). Excitability of the lower-limb area of the motor cortex in parkinson's disease. *Neurophysiologie Clinique/Clinical Neurophysiology* 40, 201–208. doi:10.1016/j.neucli.2010.04.002
- Vapnik, V. N. (1998). *Statistical Learning Theory* (New York: Wiley-Interscience), 1 edition edn.
- Varma, S. and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 7, 91. doi:10.1186/1471-2105-7-91
- Viaro, R., Morari, M., and Franchi, G. (2011). Progressive motor cortex functional reorganization following 6-hydroxydopamine lesioning in rats. *The Journal of Neuroscience* 31, 4544–4554. doi:10.1523/JNEUROSCI.5394-10.2011
- Vlaar, A. M., van Kroonenburgh, M. J., Kessels, A. G., and Weber, W. E. (2007). Meta-analysis of the literature on diagnostic accuracy of SPECT in parkinsonian syndromes. *BMC Neurology* 7, 27. doi:10.1186/1471-2377-7-27
- Wenning, G. K., Geser, F., Krismer, F., Seppi, K., Duerr, S., Boesch, S., et al. (2013). The natural history of multiple system atrophy: a prospective european cohort study. *The Lancet Neurology* 12, 264–274. doi:10.1016/S1474-4422(12)70327-7
- Williams, D. R. and Lees, A. J. (2009). Progressive supranuclear palsy: clinicopathological concepts and diagnostic challenges. *The Lancet Neurology* 8, 270–279. doi:10.1016/S1474-4422(09)70042-0