



# Unveiling relevant non-motor Parkinson's disease severity symptoms using a machine learning approach



Rubén Armañanzas<sup>a,\*</sup>, Concha Bielza<sup>a</sup>, Kallol Ray Chaudhuri<sup>b</sup>,  
Pablo Martinez-Martin<sup>c</sup>, Pedro Larrañaga<sup>a</sup>

<sup>a</sup> Computational Intelligence Group, Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, Campus de Montegancedo, 28660 Boadilla del Monte, Spain

<sup>b</sup> National Parkinson Foundation Centre of Excellence, King's College Hospital, Denmark Hill, London SE59RS, UK

<sup>c</sup> Unidad de Investigación Proyecto Alzheimer y CIBERNED, Instituto de Salud Carlos III, Centro Alzheimer Fundación Reina Sofía, C/ Valderrebollo 5, 28031 Madrid, Spain

## ARTICLE INFO

### Article history:

Received 30 May 2012

Received in revised form 12 March 2013

Accepted 7 April 2013

### Keywords:

Estimation of distribution algorithms

Feature subset selection

Severity indexes

Parkinson's disease

## ABSTRACT

**Objective:** Is it possible to predict the severity staging of a Parkinson's disease (PD) patient using scores of non-motor symptoms? This is the kickoff question for a machine learning approach to classify two widely known PD severity indexes using individual tests from a broad set of non-motor PD clinical scales only.

**Methods:** The Hoehn & Yahr index and clinical impression of severity index are global measures of PD severity. They constitute the labels to be assigned in two supervised classification problems using only non-motor symptom tests as predictor variables. Such predictors come from a wide range of PD symptoms, such as cognitive impairment, psychiatric complications, autonomic dysfunction or sleep disturbance. The classification was coupled with a feature subset selection task using an advanced evolutionary algorithm, namely an estimation of distribution algorithm.

**Results:** Results show how five different classification paradigms using a wrapper feature selection scheme are capable of predicting each of the class variables with estimated accuracy in the range of 72–92%. In addition, classification into the main three severity categories (mild, moderate and severe) was split into dichotomic problems where binary classifiers perform better and select different subsets of non-motor symptoms. The number of jointly selected symptoms throughout the whole process was low, suggesting a link between the selected non-motor symptoms and the general severity of the disease.

**Conclusion:** Quantitative results are discussed from a medical point of view, reflecting a clear translation to the clinical manifestations of PD. Moreover, results include a brief panel of non-motor symptoms that could help clinical practitioners to identify patients who are at different stages of the disease from a limited set of symptoms, such as hallucinations, fainting, inability to control body sphincters or believing in unlikely facts.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Parkinson's disease (PD) is characterized by the loss of dopaminergic neurons mainly in the pars compacta of the substantia nigra [1]. The exact cause of this neuronal death is still unknown. The direct consequence is that the levels of dopamine in the striatal region of the brain drop sharply. This neurotransmitter shortage is the main cause of the disease's classical motor symptoms, such as tremor or hypokinesia [2,3]. In addition to these classical motor

symptoms, a number of non-motor symptoms also occur in PD patients, e.g. cognitive impairment, mood disorders, sleep disturbances, gastrointestinal and urinary dysfunction. They are probably related to serotonergic and noradrenergic denervation, as well as dopamine.

The objective of this study is to quantitatively analyze the inner relationships between both motor and non-motor symptoms. To do this, we propose a supervised classification task in which two clinical indexes used for assessing global PD severity are predicted from a combination of non-motor clinical symptoms only. By checking the symptoms most often selected for classification, we will be able to relate different non-motor symptoms to disease progression. The two severity indexes are well established in PD clinical practice. Hoehn & Yahr (HY) staging [4] is a severity index based purely on motor aspects [5], and clinical impression of severity index for PD

\* Corresponding author. Tel.: +34 913363753; fax: +34 913524819.

E-mail addresses: [r.armananzas@upm.es](mailto:r.armananzas@upm.es) (R. Armañanzas), [mcbielza@fi.upm.es](mailto:mcbielza@fi.upm.es) (C. Bielza), [ray.chaudhuri@nhs.net](mailto:ray.chaudhuri@nhs.net) (K.R. Chaudhuri), [pmartinez@fundacioncien.es](mailto:pmartinez@fundacioncien.es) (P. Martinez-Martin), [pedro.larranaga@fi.upm.es](mailto:pedro.larranaga@fi.upm.es) (P. Larrañaga).

(CISI-PD) staging [6] is based on the assessment of four different domains.

A dataset including information on 410 PD patients has been used. Information on HY and CISI-PD staging and from a battery of tests on non-motor symptoms was measured *a priori* and available for each patient. Clinical severity indexes employ a test for each individual symptom. After all tests have been completed, the whole stage is computed by applying a simple arithmetic operation, usually an addition. To explore beyond this linear constraint, all the individual symptom tests within each non-motor criterion were put together. A selection process was then performed to find which subset best classifies either HY or CISI-PD stages.

We discuss not only the numeric performance but also the relationship of the non-motor symptoms to the neurodegenerative progression of the disease. The discovery of key relationships between changes in non-motor symptoms and the advance of PD could potentially help to identify disease subtypes and make the clinical evaluation faster and more accurate.

The content of the paper is divided as follows. Section 2 extends the rationale behind this study and surveys the state of the art of this new approach to non-motor symptoms in PD. Section 3 details the dataset, PD severity indexes and methodological approaches used throughout the study. Section 4 reports the quantitative and qualitative results. The results are discussed in Section 5, and conclusions and future lines are outlined in Section 6.

## 2. Rationale and background

The most common PD manifestations are motor symptoms, such as bradykinesia, rigidity, rest tremor, and disorders of gait or posture [2,3]. It is now universally accepted, however, that a wide range of non-motor symptoms are also clear manifestations of PD. Some of these non-motor symptoms may even precede the classical motor manifestations [7] and are related to the degeneration of olfactory and lower medulla structures [8,9].

Although motor symptoms are more psychologically debilitating for PD patients due to public embarrassment, current drugs significantly improve and alleviate these manifestations throughout the disease's course. By contrast, non-motor symptoms are very prevalent [10] and tend to accumulate and increase in severity with disease progression [11]. In the long run, non-motor symptoms become the most important problem for the quality of life of long-term survivors [12,13]. Nevertheless, these non-motor aspects have been kept separate from the motor disorder and have only recently received the attention of clinicians and researchers. Furthermore, even patients have overlooked [14] and failed to declare symptoms to doctors and health professionals [15], thereby compromising treatment.

As discussed in Section 3.2, motor impairment and disability in PD are divided into stages by the HY scale, and this scale is used universally to describe patients and select participants for PD studies, including clinical trials. There is no similar alternative for describing and classifying PD patients taking into account non-motor symptoms. This would be very helpful considering the importance of this aspect of the disease. Following this rationale, the objective of this study is to explore whether such a system would be possible using scores and cutoffs from already valid and established non-motor symptom scales (see Section 3.3 for a detailed description).

State-of-the-art literature in regard to this issue is limited. One of the first papers using computerized models to relate external symptoms and neurogenesis and internal central nervous system flows in PD was [16]. Moving onto classification, Refs. [17–19] predicted general PD prognosis using speech analysis features and a battery of classifiers. All these studies are limited by the number of patients, usually just a few dozen. Regarding other features, Ref.

[20] used neuropsychological features in order to find personality markers of early PD diagnosis, i.e. comparing patients and control individuals. This is the classical approach when trying to predict disease from healthy samples. However, as far as we know, our research is the first attempt to analyze and relate motor and non-motor manifestations in PD using machine learning as the research tool.

## 3. Materials and methods

### 3.1. Patients

The sample consists of consecutive patients diagnosed as having PD by neurologists with competence in movement disorders, who applied international criteria [21]. Patients that were unable to understand or answer questionnaires or had any comorbidity or disorder interfering with or impeding assessment of PD manifestations were excluded. The database for this study, sourced from an international collaboration [11], was prepared by one of the authors (PMM). The study included 410 patients (males, 61.3%). Age (mean  $\pm$  sd) was  $64.48 \pm 9.91$  years and duration of disease,  $8.07 \pm 5.75$  years. Treatment was 81.02% levodopa; 61.36% dopamine agonists (49.5% in combination); 6.44% selegiline; 5.42% rasagiline; and 38.64% other antiparkinsonian drugs. Detailed information is presented in Table 1 listed by patient severity stage.

### 3.2. Severity indexes

Here we briefly introduce each of the indexes used as class variables in the supervised classification problems. The results reported in Sections 4.3 and 4.4 include outputs considering all three possible stages, or analyzing only the closest two.

- The HY scale is a classical instrument used to categorize patients according to PD stages [4]. HY is based on motor impairment only and recognizes five stages. Despite the original formulation, the HY index is usually reconfigured in three stages, namely *mild*, *moderate* and *severe*. This is a classical adaptation in the PD state of the art [5,22,23], and the translation is straightforward.<sup>1</sup>
- CISI-PD [6]. Known as clinical impression of severity index for PD, CISI-PD extends the evaluated motor symptoms criteria to more complex aspects like the patients' cognitive state. It records the clinician's global impression of severity, and it is composed of four different items that cover the motor signs, disability issue, possible motor complications and decline in cognitive state. These four criteria are scored on a scale from 0, none, to 6, in the worst cases. The range of values for a patient is thus from 0 to 24. This continuous formulation was categorized into HY-like *mild*, *moderate* and *severe* stages to assure equivalent severity across indexes.

### 3.3. Non-motor symptom scales

A total of 87 individual tests collected from the following five non-motor scales form the predictor variables of the classification problem. All patients completed all tests, although not all values for all tests could be collected (see Section 3.4). Further details on each individual test are available as *supplementary content*. An introduction to the five non-motor scales follows.

- Scales for outcomes in Parkinson's disease-cognition or SCOPA-COG [22]. This rating scale is the result of the sum of 10 cognitive tests, covering symptoms from various domains, such as

<sup>1</sup> *mild*  $\Rightarrow$  HY=1 or 2; *moderate*  $\Rightarrow$  HY=3; *severe*  $\Rightarrow$  HY=4 or 5.

**Table 1**

Extended information on the dataset patients and prescribed drugs. The Age and PD duration – Parkinson's disease duration – rows include average values (in years) with their associated standard deviation.

	HY stage				
	1	2	3	4	5
Age	61.00 ± 9.33	62.39 ± 9.78	66.17 ± 9.27	70.42 ± 8.59	77.28 ± 7.09
PD duration	4.19 ± 2.35	6.57 ± 4.19	11.77 ± 26.21	13.35 ± 6.99	19.29 ± 10.18
Treatment					
None	6	6			
Levodopa	6	37	21	4	
Selegiline	1				
Rasagiline	2	2			
DA agonist	15	18	2	1	
Amantadine	1	1			
Other	4	1	1		
Combination of drugs	31	98	108	37	7
Totals	62 (15.12%)	166 (40.49%)	132 (32.20%)	43 (10.49%)	7 (1.71%)

memory, attention, executive functions and visuo-spatial abilities. The scale ranges from 0 to 43, where the patient's cognitive condition is worse, the lower the total score.

- Non-motor symptoms scale (NMSS) [23]. This scale evaluates 30 non-motor symptoms frequently experienced by PD patients. These items cover nine different domains: cardiovascular, sleep/fatigue, mood/apathy, perceptual problems/hallucinations, attention/memory, gastrointestinal, urinary, sexual, and miscellaneous. The score for each symptom is computed as the product of its frequency (1, rarely or less than once a week, to 4, very frequently/daily) and severity degree (from 0, none, to 3, very severe). Each domain score is the sum of its items, and the NMSS total score is calculated by adding all domain scores. This way, the full NMSS ranges from 0 to 360. The highest scores should map to the highest non-motor symptomatic burden, whereas the lowest scores match the mildest non-motor manifestations.
- Parkinson's disease sleep scale or PDSS [24] is used to assess sleep disorders in PD. All 15 PDSS symptoms are focused on nocturnal sleep, except one which addresses daytime sleepiness. Patients mark their response to each symptom on a visual analog scale running from worst (0) to best (10). The PDSS total score (0–150) is the sum of individual ratings.
- Another two severity indexes based on non-motor features were used: SCOPA-PC [25] and SCOPA-AUT [26]. These scales were developed as specific instruments to assess psychiatric complications (7 symptoms), and autonomic dysfunction (25 symptoms), respectively, in PD patients. As in the case of NMSS, the final ratings are computed by adding the individual item scores. The higher the score, the worse the state of the patient is. SCOPA-PC ranges from 0 to 21, whereas SCOPA-AUT has a wider range, from 0 to 69.

### 3.4. Data analysis

In order to test how classical PD severity indexes can be predicted using knowledge from non-motor symptoms, hereinafter features, we configured the following data matrix: 10 features from SCOPA-COG, 30 features from NMSS, 15 features from PDSS, 7 features from SCOPA-PC and 25 features from SCOPA-AUT. The last variable of the data matrix is the so-called class variable of a supervised classification problem, which is ultimately the target of the classification process. This pattern recognition process was enacted twice, using HY and CISI-PD, respectively, as the class variable. The number of instances available for the data mining analysis differed slightly from the original number of patients. Patients who were missing values for some symptoms were removed from the dataset. Finally, we had 371 and 370 instances for HY and CISI-PD prediction, respectively.

The values for CISI-PD had to be categorized into three bins following the same policy as for HY. This process is not so straightforward because we need to map a quantitative value to a categorical division. Binning was done by searching for the optimal cutoff points that maximize the association between HY and CISI-PD. An exhaustive search was performed, and the three optimal intervals were identified (see Section 4.1).

After categorizing both HY and CISI-PD into three stages, the classification problem was tackled as a feature subset selection problem for a supervised classification task. Five different supervised classification paradigms were used: naïve Bayes [27] (NB), k-nearest neighbors [28] (k-NN), linear discriminant analysis [29] (LDA), C4.5 decision trees [30] (C4.5) and artificial neural networks [31] (ANN).

The main assumption of the NB paradigm is the conditional independence of predictors given the value of the class. The model is relatively immune to inclusion of irrelevant variables, whereas redundant variables reduce its classification performance. NB is well suited to physiological datasets since humans tend to choose independent or orthogonal clinical tests in order to solve medical problems. k-NN performs well when patient groups are large and homogeneous. However, it has no explicit model so all the calculations have to be repeated in order to classify a new case. The root hypothesis of LDA is that the conditional probability density function of the predictors follows a normal distribution given the class value. LDA is able to capture statistical dependencies among the predictor variables. However, its performance decreases significantly when these dependencies are not linear. If the class is binary, both LDA's and NB's decision region is a hyperplane. Classification trees such as C4.5 emulate human reasoning: they are fully interpretable both as a set of rules or as a tree-like flow of information. A tree works well in medical analyses when the problem is internally divided into hierarchical levels or categories. If enough data is available to properly train ANNs, classification performance is usually good. As this is not the case in many physiological problems, they tend to overfit to the available data, losing generalizability. In addition, ANNs are black-box models which are not suited for knowledge discovery by exploring variable dependencies. As explained earlier, this list of classifiers was chosen because they cover a broad mathematical groundwork. This ensures that the results are reliable provided they agree.

Lastly, we decided to use estimation of distribution algorithms (EDA) as the feature subset selection method [32]. EDAs are well suited to tackling feature selection processes within large datasets in an affordable time, and they also guarantee a (near)optimal performance. The selection process is guided using a wrapper evaluation score with an internal 10-fold cross-validation scheme.

**Table 2**

Contingency matrix for the optimal correspondence of clinical impression of severity index for Parkinson's disease (CISI-PD) numeric values and Hoehn & Yahr (HY) stages.

	CISI-PD		
	[0,9]	[10,16]	[17,24]
HY			
mild	210	18	0
moderate	44	82	5
severe	4	26	20

## 4. Results

### 4.1. CISI-PD categorization

Instead of using its numeric value, the CISI-PD index was binned into three categories. Since the aim is to get HY-like intervals, both indexes were confronted by looking for the optimal cutoff points to bin CISI-PD into the three HY categories [33]. We found these cutoffs by testing all combinations of binning points for adjustment by means of an exhaustive search across all possibilities. The optimal cutoff was found to be the index marks of 9 and 16, configuring the intervals [0,9], [10,16] and [17,24]. The accuracy between these intervals and HY values is 76.30%. The respective contingency matrix is shown in Table 2.

### 4.2. Setting up EDA

The particular parameters of the EDA algorithm for all classification experiments were the univariate marginal distribution, 100 individuals for each population, a truncation factor of 0.5, and 100 populations or perfect classification as the stopping criteria. The initial subset of selected features is computed at random, where each item has a probability of inclusion of 0.1 and the objective function is the accuracy of the model [34].

### 4.3. Predicting HY severity index

Table 3 (top) shows both the quantitative and qualitative results when all the patients are taken into account. The estimated accuracy for this three-class task varied from 66.85% to 72.51%. Two

features were selected by all the classifiers (scaulcho, scpc1) and another two were selected by four out of the five (nms19, scpc2).

A patient at a mild stage is easy to differentiate from another one at a severe stage in clinical practice. Therefore, classification into neighboring stages gains importance. Accordingly, two more classification problems were tackled: mild vs moderate, and moderate vs severe.

In the first problem, the number of patients was as follows: 213 mild and 112 moderate. Performance varied from 74.46% to 78.77%. Table 4 (top) shows that two features were selected by all the classifiers (scau7sto, scpc1); scaulcho was selected four times, and another four items (sccog8, scau5con, scau15st, scau19li) were selected by three out of the five models.

For the second problem, the cardinalities were 112 moderate and 46 severe cases. Performance accuracy peaks at 86.08% for k-NN and C4.5. Of the selected features (see Table 5), again only two features are jointly selected by four classifiers, specifically from the non-motor symptoms scale (nms19, nms25).

### 4.4. Predicting CISI-PD severity index

As in Section 4.3 above, the first-listed results correspond to the three-class classification problem. Performance accuracy varies from 74.05% to 80.00%, which is slightly higher than the values for HY classification. The subset of items and respective accuracies are listed in Table 3 (bottom). Two items were selected by all the classification algorithms (nms14, scpc1), whereas another seven were selected by four out of the five (scaulcho, scau4ful, scau7sto, scau17pe, sccog6, sccog8, scpc2).

Following on from the above discussion, we reanalyzed the classification into mild and moderate cases. Cardinalities for this problem are 229 mild and 117 moderate. Numerical performance peaks at an estimated accuracy of roughly 83.24%. This performance and the respective subsets of items are shown in Table 4 (bottom). We found that three items (scpc1, scau7sto, scaulcho) were selected by four classifiers and another two (sccog8, nms2) by at least three.

Again using pairwise comparisons, we tackled the problem of classifying moderate vs severe cases from the CISI-PD. Because of the binning points reported in Section 4.1, the number of severe cases is 24, whereas the number of moderate cases is 117, as for mild vs moderate above. Subsets and performances are shown in

**Table 3**  
Percentage of estimated accuracies (Acc.), subsets of selected non-motor items and subset size found by the wrapper approach when classifying HY<sup>a</sup> (top) and CISI-PD<sup>b</sup> (bottom) in three stages (mild, moderate and severe). Item numbers refer to the order of the item within each of the non-motor indexes in use.

Model	Acc.	Selected items for HY – 3 classes	No.
NB	69.81	nms2, nms6, nms14, nms19, nms21, scaulcho, scau3stu, scau7sto, scau11ur, scau16fa, scau17pe, sccog6, sccog8, pdss3, pdss12, scpc1, scpc2, scpc5	18
k-NN	72.51	nms19, nms20, scaulcho, scau2dri, scau3stu, scau5con, scau6sto, scau7sto, scau16fa, scau17pe, scau19li, sccog2, sccog5, sccog7, scpc1, scpc2, scpc4	17
LDA	69.00	nms2, nms14, nms18, nms19, scaulcho, scau4ful, scau5con, scau19li, sccog5, sccog7, pdss12, scpc1, scpc2, scpc7	14
C4.5	69.00	scaulcho, scau23ej, scpc1	3
ANN	66.85	nms19, scaulcho, scau22er, sccog8, pdss7, pdss12, scpc1, scpc2	8
Model	Acc.	Selected items for CISI-PD – 3 classes	No.
NB	79.19	nms2, nms14, nms29, scaulcho, scau3stu, scau5con, scau7sto, scau10ur, scau13ur, scau17pe, sccog6, sccog8, pdss10, pdss14, scpc1, scpc2	16
k-NN	80.00	nms4, nms6, nms12, nms14, scau3stu, scau4ful, scau7sto, scau14st, scau17pe, scau22er, sccog3, sccog6, sccog7, sccog9, scpc1, scpc2, scpc5	17
LDA	77.03	nms2, nms14, nms16, nms24, scaulcho, scau4ful, scau7sto, scau17pe, scau18pe, sccog6, sccog8, pdss2, scpc1, scpc2	14
C4.5	78.11	nms12, nms14, scaulcho, scau4ful, scau22er, sccog5, sccog8, scpc1	8
ANN	74.05	nms1, nms11, nms14, nms24, scaulcho, scau4ful, scau7sto, scau17pe, sccog6, sccog8, scpc1, scpc2, scpc5	13

<sup>a</sup> Hoehn & Yahr index.

<sup>b</sup> Clinical impression of severity index for Parkinson's disease.



**Table 4**

Percentage of estimated accuracies (Acc.), subsets of selected non-motor items and subset size found by the wrapper approach when classifying by HY<sup>a</sup> (top) and CISI-PD<sup>b</sup> (bottom) stages of mild and moderate.

Model	Acc.	Selected items for HY – mild vs moderate	No.
NB	78.77	scau3stu, scau5con, scau7sto, scau15st, scau19li, sccog8, pdss12, scpc1, scpc2, scpc7	10
k-NN	78.77	nms13, nms17, nms26, scaulcho, scau4ful, scau5con, scau7sto, scau9uri, scau19li, sccog2, pdss7, pdss12, scpc1, scpc4, scpc6	15
LDA	77.54	nms16, nms18, nms30, scaulcho, scau4ful, scau5con, scau7sto, scau15st, sccog8, scpc1	10
C4.5	76.00	nms11, scaulcho, scau7sto, scau19li, sccog5, scpc1	6
ANN	74.46	nms9, nms18, scaulcho, scau7sto, scau15st, sccog5, sccog8, pdss2, scpc1, scpc2	10
Model	Acc.	Selected items for CISI-PD – mild vs moderate	No.
NB	82.08	nms24, scaulcho, scau3stu, scau17pe, sccog8, pdss2, pdss13, scpc1, scpc2, scpc5	10
k-NN	80.06	nms2, nms5, nms30, scaulcho, scau3stu, scau4ful, scau5con, scau7sto, scau13ur, sccog4, sccog5, sccog7, sccog9, sccog10, pdss7, pdss14, scpc1	17
LDA	83.24	nms2, nms24, nms29, scaulcho, scau3stu, scau7sto, scau9uri, scau11ur, scau17pe, sccog8, pdss2, scpc1, scpc2	13
C4.5	81.50	nms2, nms12, nms13, nms14, nms15, scaulcho, scau7sto, scau10ur, scpc1	9
ANN	64.74	scau7sto, scau13ur, scau15st, sccog7, sccog8, pdss6	6

<sup>a</sup> Hoehn & Yahr index.

<sup>b</sup> Clinical impression of severity index for Parkinson's disease.

**Table 5** (bottom). According to the results performance was very promising: three out of the five models had accuracies ranging from 92.20% to 92.91%. Regarding the selected features, nms4 is the most often selected item, whereas scpc2 and nms2 were selected by three out of the five models.

## 5. Discussion

In clinical practice, the differentiation between a mild and a severe patient is apparently obvious when only motor manifestations are considered (mild motor disease may be accompanied by no obvious severe non-motor symptoms). Therefore, it would be worthwhile distinguishing close severity states considering motor and non-motor disorders. For this last problem, the non-motor items achieved competitive accuracies. Using only 10 items, estimated classifier accuracy reached a maximum of almost 79% when classifying mild and moderate samples using HY as the reference scale. In the case of CISI-PD, performance accuracy is higher at 83.24% using 13 items (see Table 4).

An in-depth analysis of the commonly selected non-motor items gives clues about the underlying progress of neurological damage. Table 6 lists the most commonly selected items when classifying mild and moderate cases. These items were selected by at least three out of the five models in each feature selection process. Table 6 also includes the average value of the above symptoms for

the respective subset of patients, i.e. values for mild and moderate patients using HY or CISI-PD as classes, respectively.

Interestingly, five out of eight of these very relevant items belong to the autonomic index SCOPA-AUT, whereas the others belong to other scales. This suggests that there could be a border in terms of autonomic symptoms that divides mild-stage from moderately severe patients. Being a psychiatric complication, however, scpc1 was constantly selected as very relevant. It maps the appearance of hallucination episodes, and its values show an almost five-fold difference on average. This suggests that the advent of hallucination may be considered a clear sign of worsening disease.

Another big difference in values is found for fainting episodes, nms2, and involuntary loss of stools, scau7sto. Involuntary loss of stools is related to dysregulation of the intestinal activity and impairment of the control of body sphincters in advanced conditions. Fainting episodes in PD respond to hypotensive faints due to dysfunction of the autonomic nervous system. The complex pathophysiology of autonomic dysfunction is related to structural and functional damage of brain stem nuclei (dorsal vagal, ambiguous and other medullary nuclei). Cholinergic, monoaminergic and serotonergic nuclei degeneration causes abnormal functioning of the central autonomic system. Very related to these fainting episodes is the symptom of dizziness or light-headedness after prolonged standing, scau15st. Differences in average scau15st values also show a significant increase of such episodes. There is less variability

**Table 5**

Percentage of estimated accuracies (Acc.), subsets of selected non-motor items and subset size found by the wrapper approach when classifying by HY<sup>a</sup> (top) and CISI-PD<sup>b</sup> (bottom) stages of moderate and severe.

Model	Acc.	Selected items for HY – moderate vs severe	No.
NB	81.01	nms19, nms25, nms29, scau16fa, scau22er, scau23ej, sccog6, scpc2, scpc7	9
k-NN	86.08	nms9, nms12, nms14, nms16, nms19, nms23, nms24, nms25, nms27, scau13ur, scau15st, sccog9, pdss1, pdss4, pdss8, pdss11	16
LDA	79.75	nms9, nms13, nms14, nms15, nms19, nms24, nms25, sccog7, scpc1	9
C4.5	86.08	nms19, nms23, nms25, scau10ur, scau11ur, scpc2	6
ANN	71.34	nms8, nms9, nms16, nms23, nms24, nms27, scau8uri, scau12ur, scau17pe, sccog6, pdss3, pdss5, pdss11, scpc2, scpc3, scpc5, scpc7	17
Model	Acc.	Selected items for CISI-PD – moderate vs severe	No.
NB	92.20	nms2, nms11, nms14, nms27, scau3stu, scau5con, scau23ej, sccog1, sccog7, sccog10, scpc2, scpc6	12
k-NN	92.91	nms14, scau10ur, scau12ur, scau15st, scau23ej, pdss8, pdss9, pdss11, pdss15, scpc1, scpc2, scpc5	12
LDA	90.07	nms2, nms14, scau7sto, sccog1, sccog9, pdss2, scpc2	7
C4.5	92.20	nms2, nms14, sccog3	3
ANN	57.45	nms1, nms9, scau5con, scau11ur, sccog5, pdss1, pdss6	7

<sup>a</sup> Hoehn & Yahr index.

<sup>b</sup> Clinical impression of severity index for Parkinson's disease.

**Table 6**  
Individual non-motor symptoms most commonly selected by the feature selection process in the classification of *mild* and *moderate* instances. The *selection* column lists the number of times each item was selected for HY<sup>a</sup> and CISI-PD<sup>b</sup> classes, respectively. The *mild* and *moderate* columns report the average value of each item for the respective cases and problems. Statistically significant differences between the values of the two groups for each classification problem using a signed rank sum test with  $\alpha = 0.01$  are marked with the † symbol.

Item	Description	Selection		Mild		Moderate	
		HY	CISI-PD	HY	CISI-PD	HY	CISI-PD
scpc1	Hallucinations	5	4	0.0845	0.0830	0.4196	0.4274†
scau7sto	Involuntary loss of stools	5	4	0.0657	0.0742	0.2321	0.3162†
scau1cho	Difficulty swallowing or choked	4	4	0.3756	0.3930	0.8214	0.8974†
sccog8	Dice	3	3	2.4366	2.4541	1.9286	1.8803†
scau5con	Constipation	3	1	0.9484	0.9869	1.4018	1.3333†
scau15st	Light-headed	3	1	0.3756	0.4061	0.7589	0.7009†
nms2	Fainting	0	3	0.0657	0.0306	0.3482	0.3761†
scau19li	Over-sensitive to bright light	3	0	0.5728	0.5502	0.5179	0.5983

<sup>a</sup> Hoehn & Yahr index.

<sup>b</sup> Clinical impression of severity index for Parkinson's disease.

in swallowing and constipation episodes, *scau1cho* and *scau5con*, although both differences were also statistically significant.

In terms of pure cognitive decline, *sccog8* measures the ability to order numbers coming from a dice roll. Values for this test suggest a significant shrinkage of patients' numerical ability as the damage spreads. Lastly, hypersensitivity to bright light, *scau19li*, was selected by the machine learning processes, although its average values do not appear to be so different. A possible explanation is that the selection technique is multivariate and may need an auxiliary item to assure the goodness of fit of the selection for the classification task.

Classification accuracy for the *moderate* vs *severe* case rose to more than 90% for CISI-PD with a variety of non-motor items (see Table 5). This performance increase was expected because of the clinical difference between a patient at the *moderate* stage of the disease and the motor and non-motor symptoms of another patient at the *severe* stage. In addition, there appears to be a more marked tendency to select non-motor symptoms from a variety of domains, including physical and mental aspects (fatigue, dribbling of saliva, fainting-induced fall, abnormal interest in sex, illusions and misidentifications), than autonomic signs. The preponderant accumulation of non-motor symptoms in advanced phases of the disease has been previously emphasized [12].

The scenario of relevant symptoms for classifying *moderate* vs *severe* patients changes drastically. Autonomic dysfunctions from SCOPA-AUT were key points in the above discussion, whereas patients are now confined to care, and other non-motor manifestations from the NMSS scale (e.g. belief in unlikely facts, nervousness, drooling, etc.) take over.

Two of the most often selected symptoms reported in Table 7 stand out above all others: false beliefs (delusions), *nms14*, and

misidentification of already known people, *scpc2*. The differences between *moderate* and *severe* values for these scores are large and suggest that the neuronal damage has spread from the temporal mesocortex to association neocortical areas, the prefrontal cortex and other brain fields [9,35]. In addition, factors like dopaminergic treatment, disease duration, older age and cognitive impairment can influence the presence of psychotic manifestations at advanced stages of Parkinson's disease [36,37].

Usually benign, visual hallucinations are relatively common (up to 40%) and may be present in moderate disease, but other psychotic symptoms, such as delusions and paranoid ideation, become more frequent in advanced phases and have adverse prognostic connotations [8]. Although there is a clear increase in average values for *nms9*, corresponding to anxiety manifestations, they are also associated with high variance because of the wide range of cases and, hence, are not statistically significant.

The joint selection of *nms14* and *scpc2* points to the co-existence of modalities of psychotic symptoms that are more characteristic of the advanced phases of PD and usually, albeit not always [38], associated with cognitive deterioration [39]. As a whole, these psychotic manifestations are a common feature of synucleinopathies (PD, Lewy body dementia) and other neurodegenerative diseases [40].

In Parkinson's disease, sexual functioning may be reduced as a consequence of dysautonomia, testosterone deficiency, disability, depression and other factors. On the contrary, aberrant sexual behavior and hypersexuality may be a side effect of dopaminergic treatment and occur in susceptible patients as a part of impulse control disorders. Also sexual fantasy may increase in patients with longer disease duration [8,41]. Therefore, values reported in Table 7 for the *nms25* item make complete sense.

**Table 7**  
Individual non-motor symptoms most commonly selected by the feature selection process in the classification of *moderate* and *severe* instances. The *selection* column lists the number of times each item was selected for HY<sup>a</sup> and CISI-PD<sup>b</sup> classes, respectively. The *moderate* and *severe* columns report the average value of each item for the respective cases and problems. Statistically significant differences between the values of the two groups for each classification problem using a signed rank sum test with  $\alpha = 0.01$  and  $\alpha = 0.05$  are marked with the † and ‡ symbols, respectively.

Item	Description	Selection		Moderate		Severe	
		HY	CISI-PD	HY	CISI-PD	HY	CISI-PD
scpc2	Illusions and misidentification of persons	3	3	0.2679	0.2991	0.6957	1.2083†
nms14	Does the patient believe in unlikely facts	2	4	0.5179	0.6667	1.9565	3.9583†
nms9	Nervousness or frightened for no reason	3	1	2.5268	3.0085	3.9565	5.0417
nms19	Drooling during the day	4	0	2.0000	2.6410	4.8478	6.0000†
nms25	Altered interest in sex	4	0	2.5893	2.7009	1.5435	1.7917
nms2	Fainting	0	3	0.3482	0.3761	1.5000	2.9583†
nms23	Void within 2 h of last voiding	3	0	2.8482	3.2650	5.0000	5.2500‡
nms24	Pass urine regularly at night	3	0	3.6161	4.6068	6.5870	6.5833‡

<sup>a</sup> Hoehn & Yahr index.

<sup>b</sup> Clinical impression of severity index for Parkinson's disease.

Increased salivation is significantly more frequent in patients with Parkinson's disease than in controls [42]. Drooling is likely to result from saliva pooling in the mouth due to decreased swallowing frequency and postural changes (antecollis). Dribbling of saliva has serious emotional consequences and a negative social impact in advanced phases of the disease. Orthostatic hypotension usually develops late in the disease, but cardiac sympathetic denervation may be an early phenomenon in Parkinson's disease [43,44]. Symptoms of autonomic dysfunction (nms2, nms19, nms23 and nms24) correlate with disease duration and severity, increasing age, and antiparkinsonian medication (levodopa, dopamine agonists, L-MAO B) [45].

Lastly, note that no items related to memory and attention impairment were selected as relevant. This is something that we expected. First, the dataset is comprised of patients without overt dementia [11]. It is often impossible to put together a battery of non-motor items in patients suffering from mental impairment because of physical impediment. Even if this is possible, results are far from reliable precisely because of that impairment. Even considering possible mental problems, studies on newly diagnosed PD patients found cognitive impairment in only over one third of the sample [46,47]: around 20% of patients with early, untreated PD have mild cognitive impairment [48]. Therefore, the presence of cognitive deficits as of the earlier stages of the disease may disable this kind of dysfunction as a marker of progression from one severity level to the next. Also, memory may be defective as a consequence of depression, a very prevalent disorder across all PD stages, and this circumstance can also disable memory and attention as a marker of global disease severity.

## 6. Conclusions

A breadth of motor and non-motor symptoms exists across all stages of PD, although diagnosis is typically made when the classical motor features of akinesia and tremor become evident. Current research is, however, signaling the importance of non-motor signs possibly occurring before motor signs for both patients' and carers' quality of life. As neurodegeneration spreads, motor and non-motor symptoms can develop differently due to differences in individual brain dendritic innervations. Classical severity indexes usually focus on only a fraction of all symptoms, as they are elementary metrics based mostly on simple additive measures. This linearity constraint can potentially mask relevant individual symptoms and their relation to disease progression. Non-motor symptoms should therefore be considered in the diagnostic framework of PD.

In this paper, we have explored the synergies between motor and non-motor symptoms for the first time. To do this, we have used supervised classification and feature subset selection to predict two severity indexes, HY and CISI-PD, in terms of just non-motor symptoms. Making use of advanced evolutionary computation techniques, namely EDAs, we have identified which non-motor symptoms are more relevant for which PD clinical stage.

Specifically, classification performance suggests a partial correlation between motor and non-motor symptoms, with accuracy estimators ranging from 72% to 92%. As we used a set of classification algorithms covering a broad mathematical groundwork, results were highly consistent. Hence, reported items should be considered reliable and widespread. The panel of non-motor symptoms presented in Tables 6 and 7 includes findings that can be directly mapped to the clinical evaluation of a PD patient. The appearance of hallucinations, fainting, inability to control body sphincters or belief in unlikely facts suggests major neurological degeneration. The differences in these symptoms between close severity stages turned out to be statistically significant in our database.

From the results in Tables 6 and 7, clinical practitioners find that autonomic symptoms are much more prevalent in moderate patients with respect to mild or untreated PD patients. All the signs identified throughout our experiments can be directly mapped to routine clinical practice. Treatments, advice to family or carers and patient-directed visitation can be quickly adjusted if the clinician detects any of the patterns reported here.

Future work should address the combination of relevant non-motor items with other subsets of relevant motor symptoms. This could help to build an accurate severity index for the staging of PD. The new index will possibly be based on fewer tests than the current indexes, and it will reflect a broader vision of the patient's clinical state.

## Acknowledgements

This work has been partially supported by Spanish Ministry of Economy and Competitiveness (MINECO) projects TIN2010-20900-C04-04 and the Cajal Blue Brain Project, Spanish partner of the Blue Brain Project initiative from EPFL. R.A. is supported by a Juan de la Cierva postdoctoral fellowship (MINECO).

## Appendix A. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.artmed.2013.04.002>.

## References

- [1] Trétiakoff MC. Contribution à l'étude de l'anatomie pathologique du locus niger de Soemmering avec quelques déductions relatives à la pathogénie des troubles du tonus musculaire et de la maladie de Parkinson. PhD thesis, University of Paris; 1919.
- [2] Berardelli A, Rothwell JC, Thompson PD, Hallett M. Pathophysiology of bradykinesia in Parkinson's disease. *Brain* 2001;124:2131–46.
- [3] Turner RS, Grafton ST, McIntosh AR, DeLong MR, Hoffman JM. The functional anatomy of Parkinsonian bradykinesia. *Neuroimage* 2003;19:163–79.
- [4] Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurology* 1967;17:427–42.
- [5] Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, Counsell C, et al. Movement disorder society task force report on the Hoehn and Yahr staging scale: status and recommendations. *Movement Disorders* 2004;19:1020–8.
- [6] Martínez-Martin P, Forjaz MJ, Cubo E, Frades B, Cuesta JdP, the ELeP Project Members. Global versus factor-related impression of severity in Parkinson's disease: a new clinimetric index (CISI-PD). *Movement Disorders* 2006;21:208–14.
- [7] Stern MB, Lang A, Poewe W. Toward a redefinition of Parkinson's disease. *Movement Disorders* 2012;27:54–60.
- [8] Chaudhuri KR, Healy DG, Schapira AHV. Non-motor symptoms of Parkinson's disease: diagnosis and management. *Lancet Neurology* 2006;5:235–45.
- [9] Braak H, Tredici KD, Rüb U, de Vos RAI, Steur ENHJ, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging* 2003;24:197–211.
- [10] Martínez-Martin P, Schapira AH, Stocchi F, Sethi K, Odin P, MacPhee G, et al. Prevalence of nonmotor symptoms in Parkinson's disease in an international setting: study using nonmotor symptoms questionnaire in 545 patients. *Movement Disorders* 2007;22:1623–9.
- [11] Martínez-Martin P, Rodríguez-Blázquez C, Abe K, Bhattacharyya KB, Bloem BR, Carod-Artal FJ, et al. International study on the psychometric attributes of the non-motor symptoms scale in Parkinson disease. *Neurology* 2009;73:1584–91.
- [12] Hely MA, Reid WG, Adena MA, Halliday GM, Morris JC. The Sydney multicenter study of Parkinson's disease: the inevitability of dementia at 20 years. *Movement Disorders* 2008;23:837–44.
- [13] Martínez-Martin P, Rodríguez-Blázquez C, Kurtis MM, Chaudhuri KR, NMSS Validation Group, the impact of non-motor symptoms on health-related quality of life of patients with Parkinson's disease. *Movement Disorders* 2011;26:399–406.
- [14] Shulman LM, Taback RL, Rabinstein AA, Weiner WJ. Non-recognition of depression and other non-motor symptoms in Parkinson's disease. *Parkinsonism & Related Disorders* 2002;8:193–7.
- [15] Chaudhuri KR, Prieto-Jurcynska C, Naidu Y, Mitra T, Frades-Payo B, Tluk S, et al. The non declaration of non motor symptoms of Parkinson's disease to health care professionals: an international study using the non motor symptoms questionnaire. *Movement Disorders* 2010;25:704–9.

- [16] Toffano-Nioche C, Beroule D, Tassin J. A functional model of some Parkinson's disease symptoms using a guided propagation network. *Artificial Intelligence in Medicine* 1998;14:237–58.
- [17] Ozcift A. SVM feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of Parkinson disease. *Journal of Medical Systems* 2012;36:2141–7.
- [18] Tsanas A, Little MA, McSharry PE, Ramig LO. Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *Journal of the Royal Society of Interface* 2012;8:842–55.
- [19] Halawani SM, Ahmad A. Ensemble methods for prediction of Parkinson disease. In: Yin H, Costa JA, Barreto G, editors. 13th International Conference on Intelligent Data Engineering and Automated Learning. Natal., 2002, p. 516–21.
- [20] Navío M, Aguilera J, del Jesús MJ, González R, Herrera F, Iríbar C. Feature selection algorithms applied to Parkinson's disease. In: Crespo J, Maojo V, Martín F, editors. Second International Symposium on Medical Data Analysis. Madrid., 2001, p. 195–200.
- [21] Lees AJ, Hardy J, Revesz T. Parkinson's disease. *Lancet* 2009;373:2055–66.
- [22] Marinus J, Visser M, Verwey NA, Verhey FR, Middelkoop HA, Stiggelbout AM, et al. Assessment of cognition in Parkinson's disease. *Neurology* 2003;61:1222–8.
- [23] Chaudhuri KR, Martinez-Martin P, Brown RG, Sethi K, Stocchi F, Odin P, et al. The metric properties of a novel non-motor symptoms scale for Parkinson's disease: results from an international pilot study. *Movement Disorders* 2007;22:1901–11.
- [24] Chaudhuri KR, Pal S, DiMarco A, Whately-Smith C, Bridgman K, Mathew R, et al. The Parkinson's disease sleep scale: a new instrument for assessing sleep and nocturnal disability in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* 2002;73:629–35.
- [25] Visser M, Verbaan D, van Rooden SM, Stiggelbout AM, Marinus J, van Hilten JJ. Assessment of psychiatric complications in Parkinson's disease: the SCOPA-PC. *Movement Disorders* 2007;22:2221–8.
- [26] Visser M, Marinus J, Stiggelbout AM, van Hilten JJ. Assessment of autonomic dysfunction in Parkinson's disease: the SCOPA-AUT. *Movement Disorders* 2004;19:1306–12.
- [27] Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. *Machine Learning* 1997;29:131–64.
- [28] Aha DW, Kibler D, Albert MK. Instance-based learning algorithms. *Machine Learning* 1991:37–66.
- [29] McLachlan GJ. Discriminant analysis and statistical pattern recognition. New Jersey: Wiley-Interscience; 2004.
- [30] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. Belmont: Chapman and Hall/CRC; 1984.
- [31] Cochoki A, Unbehauen R. Neural networks for optimization and signal processing. New York: John Wiley & Sons; 1993.
- [32] Armañanzas R, Inza I, Santana R, Saeys Y, Flores JL, Lozano JA, Van de Peer Y, et al. A review of estimation of distribution algorithms in bioinformatics. *BioData Mining* 2008;1.
- [33] Martinez-Martin P, Arroyo S, Rojo-Abuin JM, Rodríguez-Blázquez C, Frades B, Cuesta JP, et al. Burden, perceived health status, and mood among caregivers of Parkinson's disease patients. *Movement Disorders* 2008;23:1673–80.
- [34] Santana R, Bielza C, Larrañaga P, Lozano JA, Echegoyen C, Mendiburu A, et al. Mateda: a Matlab package for the implementation and analysis of estimation of distribution algorithms. *Journal of Statistical Software* 2010;35:1–30.
- [35] Jellinger KA. Neuropathology of sporadic Parkinson's disease: evaluation and changes of concepts. *Movement Disorders* 2012;27:8–30.
- [36] Weintraub D, Burn DJ. Parkinson's disease: the quintessential neuropsychiatric disorder. *Movement Disorders* 2011;26:1022–31.
- [37] Martinez-Martin P, Frades-Payo B, Agüera-Ortiz L, Ayuga-Martinez A. A short scale for evaluation of neuropsychiatric disorders in Parkinson's disease: first psychometric approach. *Journal of Neurology* 2012;259:2299–308.
- [38] Lee AH, Weintraub D. Psychosis in Parkinson's disease without dementia: common and comorbid with other non-motor symptoms. *Movement Disorders* 2012;27:858–63.
- [39] Pagonabarraga J, Llebaria G, García-Sánchez C, Pascual-Sedano B, Gironell A, Kulisevsky J. A prospective study of delusional misidentification syndromes in Parkinson's disease with dementia. *Movement Disorders* 2008;23:443–8.
- [40] Jellinger KA. Cerebral correlates of psychotic syndromes in neurodegenerative diseases. *Journal of Cellular and Molecular Medicine* 2012;16.
- [41] Yu M, Roane DM, Miner CR, Fleming F, Rogers JD. Dimensions of sexual dysfunction in Parkinson disease. *American Journal of Geriatric Psychiatry* 2004;12:221–6.
- [42] Siddiqui MF, Rast S, Lynn MJ, Auchus AP, Pfeiffer RF. Autonomic dysfunction in Parkinson's disease: a comprehensive symptom survey. *Parkinsonism & Related Disorders* 2002;8:277–84.
- [43] Senard JM, Raï S, Lapeyre-Mestre M, Brefel C, Rascol O, Rascol A, et al. Prevalence of orthostatic hypotension in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* 1997;63:584–9.
- [44] Goldstein DS, Holmes C, Li ST, Bruce S, Metman LV, ROC III. Cardiac sympathetic denervation in Parkinson disease. *Annals of Internal Medicine* 2000;133:338–47.
- [45] Allcock LM, Kenny RA, Burn DJ. Clinical phenotype of subjects with Parkinson's disease and orthostatic hypotension: autonomic symptom and demographic comparison. *Movement Disorders* 2008;21:1851–5.
- [46] Foltynie T, Brayne CEG, Robbins TW, Barker RA. The cognitive ability of an incident cohort of Parkinson's patients in the UK. the CamPaGN study. *Brain* 2004;127.
- [47] Kandiah N, Narasimhalu K, Lau PN, Seah SH, Au WL, Tan LCS. Cognitive decline in early Parkinson's disease. *Movement Disorders* 2009;24.
- [48] Aarsland D, Brønnick K, Larsen JP, Tysnes OB, Alves G. the Norwegian ParkWest Study Group. Cognitive impairment in incident, untreated Parkinson disease: the Norwegian ParkWest study. *Neurology* 2009;72:1121–6.