



## Machine Learning for Parkinson's Disease and Related Disorders

Johann Fauzi, Olivier Colliot, and Jean-Christophe Corvol

### Abstract

Parkinson's disease is a complex heterogeneous neurodegenerative disorder characterized by the loss of dopamine neurons in the basal ganglia, resulting in many motor and non-motor symptoms. Although there is no cure to date, the dopamine replacement therapy can improve motor symptoms and the quality of life of the patients. The cardinal symptoms of this disorder are tremor, bradykinesia, and rigidity, referred to as parkinsonism. Other related disorders, such as dementia with Lewy bodies, multiple system atrophy, and progressive supranuclear palsy, share similar motor symptoms although they have different pathophysiology and are less responsive to the dopamine replacement therapy. Machine learning can be of great utility to better understand Parkinson's disease and related disorders and to improve patient care. Many challenges are still open, including early accurate diagnosis, differential diagnosis, better understanding of the pathologies, symptom detection and quantification, individual disease progression prediction, and personalized therapies. In this chapter, we review research works on Parkinson's disease and related disorders using machine learning.

**Key words** Clinical decision support, Deep learning, Disease understanding, Machine learning, Multiple system atrophy, Parkinson's disease, Parkinsonian syndromes, Parkinsonism, Precision medicine, Progressive supranuclear palsy

---

### 1 Introduction

Parkinson's disease (PD) is the second most frequent neurodegenerative after Alzheimer's disease, affecting more than six million individuals worldwide, a prevalence which is expected to double with the next 10 years [1]. It is characterized by the progressive degeneration of dopaminergic neurons in the *substantia nigra* associated with intracellular inclusions called Lewy bodies. These Lewy bodies are composed of protein aggregates enriched in  $\alpha$ -synuclein. Age is the greatest risk factor, but both environmental and genetic risk factors have been associated with PD. For instance, exposure to pesticides is a well-recognized risk factor for PD, whereas caffeine intake and smoking have been demonstrated to be protective [2]. Although commonly sporadic, rare genetic forms

of the disease have been described. More than 20 loci and associated genes have been identified to be responsible for autosomal dominant or recessive forms of the disease, and more than 90 genetic risk factors have been associated with sporadic PD [3]. Although rare, genetic forms of the disease have brought important insights on the causes and pathological mechanisms of PD [4]. Among them, aggregation and spreading of misfolded  $\alpha$ -synuclein, the protein enriched in Lewy bodies, is supposed to play a key role in the pathophysiology of the disease.

The loss of dopamine innervation of the basal ganglia network in the brain leads to the cardinal motor symptoms of the disease (parkinsonism): rest tremor, akinesia, and rigidity [2]. However, the spreading of the synucleinopathy (aggregation of  $\alpha$ -synuclein protein) and neuronal loss outside the dopaminergic pathway is associated with other non-motor symptoms like anosmia, sleep disorders, dysautonomia, and progressive cognitive decline. Some of these symptoms, particularly anosmia, constipation, and sleep disorders, can precede the motor phase during a long prodromal phase [5].

There is no cure for PD. The therapeutic strategy relies on the dopamine replacement therapy by levodopa or dopamine agonists, which alleviate motor symptoms. However, the dopamine replacement therapy does not change the course of the disease, the progression being hampered by motor complications (motor fluctuations and abnormal movement called dyskinesia), related both to the progression of the neuronal loss and to pre- and post-synaptic plasticity induced by the treatment. In addition, the dopamine replacement therapy has no benefit on non-motor symptoms not related to the loss of dopaminergic neurons.

PD is the most frequent synucleinopathy. Other neurodegenerative diseases share some clinical and pathophysiological features of PD. Multiple system atrophy (MSA) is a rare disease associated with parkinsonism with low response to levodopa, early dysautonomia, and/or cerebellar symptoms [6]. The synucleinopathy affects the substantia nigra, but also the striatum and the cerebellum, and Lewy bodies are also observed in glial cells. There are two variants of MSA: the parkinsonian variant (MSA-P) characterized by parkinsonism and the cerebellar variant (MSA-C) characterized by gait ataxia with cerebellar dysarthria. Dementia with Lewy bodies (DLB), the second most common neurodegenerative dementia after Alzheimer's disease, is characterized by early cognitive decline, hallucinations, and levodopa-responsive motor symptoms [7]. However, whether DLB and PD with dementia are really two distinct entities is still a matter of debate. There are also other rare atypical parkinsonism syndromes, not related to a synucleinopathy. Progressive supranuclear palsy (PSP) is a tauopathy (aggregation of tau protein) characterized by a nonresponsive, axial predominant parkinsonism, early falls, supranuclear gaze palsy, and a frontal

syndrome [8]. The cortico-basal degeneration (CBD) is also a tauopathy with asymmetric parkinsonism with dystonia and cognitive dysfunction. Table 1 summarizes the characteristics of all these disorders.

Considering the complexity of these disorders, the lack of reliable biomarkers, and the overlapping clinical presentation at the early stage, there is a need for more advanced approaches to support differential diagnosis. In addition, the pathophysiology of these disorders results from the complex interplay of multiple mechanisms. One current challenge is to stratify patients according to specific mechanisms and predict individual progression profile in order to move toward a more personalized medicine. Machine learning consists in extracting information from data by computer programs without providing explicit rules on what to extract, in the sense that machines learn by themselves which information to extract. Given the complexity of Parkinson's disease and its related disorders, there still exist many challenges and open questions for which machine learning could help increase knowledge on these disorders, in particular diagnosis, disease understanding, and precision medicine, and create better clinical decision support systems. Table 2 summarizes the potential benefits of machine learning for Parkinson's disease and related disorders.

The rest of this chapter is organized as follows. We first present research works on the diagnosis of Parkinson's disease and the differential diagnosis between parkinsonian syndromes, including disease understanding (Subheading 2). We then focus on the detection and quantification of motor and non-motor symptoms in Parkinson's disease (Subheading 3). Disease progression in Parkinson's disease, with the prediction of individual progression trajectories, is presented in Subheading 4. We then describe research on the monitoring and adjustment of treatment in Parkinson's disease and discuss the limitations of machine learning in terms of causality (Subheading 5). Finally, we conclude on the existing literature and discuss open questions and research works (Subheading 6). Table 3 summarizes the studies described in this chapter.

---

## 2 Diagnosis

Having an automated model being able to accurately diagnose one or several diseases has not only a concrete utility in clinical routine, but interpreting the decision process of the model may also help better understand these diseases. To assist diagnosis, two different classification tasks are usually considered: (i) being able to differentiate PD patients from healthy controls (HC) and (ii) being able to differentiate several parkinsonian syndromes from each other.

**Table 1**  
**Main characteristics of Parkinson's disease and its related disorders**

Disorder	Parkinson's disease	Multiple system atrophy	Progressive supranuclear palsy	Dementia with Lewy bodies	Cortico-basal degeneration
Pathophysiology	Aggregates of $\alpha$ -synuclein in dopaminergic neurons in brainstem	Aggregates of $\alpha$ -synuclein in neurons (cerebellum, pons, basal ganglia) and oligodendrocytes (glial cytoplasmic inclusions)	Aggregates of 4-repeat tau in astrocytes (tufts), oligodendrocytes (coiled bodies), and neurons (neurofibrillary tangles) in basal ganglia and brainstem	Aggregates of $\alpha$ -synuclein in neurons of the neocortex and brainstem	Aggregates of 4-repeat tau in astrocytes (astrocytic plaques), oligodendrocytes (coiled bodies), and neurons (neurofibrils) in neocortex and basal ganglia
Prevalence	100–300/100,000	5/100,000	5–10/100,000	50/100,000	1/100,000
Main symptoms	Levodopa-responsive parkinsonian syndrome Non-motor symptoms: Dysautonomia, anxiety, depression, sleep disorders, late cognitive dysfunction	Parkinsonian syndrome with poor response to levodopa Early and severe dysautonomia or cerebellar syndrome	Axial predominant parkinsonian syndrome nonresponsive to levodopa Early gait disturbance and falls Supranuclear gaze palsy Frontal syndrome	Parkinsonian syndrome responsive to levodopa Early dementia Early hallucinations or illusions Fluctuations in alertness and attention	Asymmetric parkinsonian syndrome nonresponsive to levodopa Apraxia and dystonia Early dementia
Symptomatic treatment	Levodopa, dopamine agonists, C-dopamine metabolism inhibitors, amantadine Deep brain stimulation for advanced cases	Levodopa Droxidopa, ephedrine, midodrine, or fludrocortisone for hypotension	Levodopa	Levodopa Cholinesterase inhibitors	Levodopa Botulinum toxin A for treating focal dystonia

**Table 2****Summary of the potential benefits of machine learning for Parkinson's disease and related disorders**

<b>Disease stage</b>	<b>Potential benefits</b>
Early PD diagnosis	Better clinical decision support systems Higher performance than current diagnostic criteria Better management and improved quality of life Potential preventive therapeutic strategies
Differential diagnosis	Better clinical decision support systems Higher performance than current diagnostic criteria Better management and improved quality of life
Symptom detection and quantification	More frequent, more robust assessment of symptoms with automatic analysis of sensor data Better management and improved quality of life
Disease progression	Identification of disease subtypes Prediction of future symptoms Treatment adjustment for potential prevention
Treatment adjustment	Better clinical decision support systems Personalized therapy Prevention of adverse events Better management and improved quality of life

### **2.1 Parkinson's Disease Diagnosis Compared to Healthy Subjects**

Given the much larger prevalence of Parkinson's disease compared to the atypical parkinsonian syndromes, gathering data from PD patients and HC is naturally easier, especially easy-to-collect data from sensors compared to clinical, imaging, or genetic data.

Digital technologies including wearable sensors, smartphone applications, and smart algorithms receive a strongly increasing interest and begin to move toward medical applications, particularly in PD [9]. Two main types of sensor data are usually considered: voice data and motion data. Given that the cardinal symptoms of PD are motor, motion data is natural, but speech also involves motor muscles. Dysarthria, which is a motor speech disorder in which the muscles involved in producing speech are damaged, paralyzed, or weakened, is a symptom of PD.

#### **2.1.1 PD Diagnosis Using Motion Data**

Several types of sensors have been investigated to collect motion data depending on the movements of interest.

Wahid and colleagues [10] investigated the discrimination between PD patients and healthy controls using gait data collected during self-selected walking. They extracted spatial-temporal features, such as stride length, stance time, swing time, and step length, from the signals and investigated different strategies of data normalization using dimensionless equations and multiple regression and different machine learning algorithms such as naive Bayes (NB), k-nearest neighbors (kNN), support vector

**Table 3**  
**Summary of the studies reviewed in this chapter**

Study	Subheading	Objectives/tasks	Modalities	Data sets	Number of subjects	Methods
Wahid et al. [10]	2.1.1.1	PD diagnosis (PD vs HC)	Motion data (gait)	Local	23 PD, 26 HC	NB, kNN, SVM, RF, KFD
Mirelman et al. [11]	2.1.1.1	PD diagnosis (PD vs HC), PD stage classification	Motion data (gait)	Local	100 HC, 58 PD-HYI, 190 PD-HYII, 84 PD-HYIII	RUSBoost
Kostikis et al. [12]	2.1.1.1	PD diagnosis (PD vs HC)	Motion data (hand movement)	Local	25 PD, 20 HC	NB, LR, SVM, AdaBoost, DT, RF
Kotsavasiloglou et al. [13]	2.1.1.1	PD diagnosis (PD vs HC)	Motion data (hand movement)	Local	24 PD, 20 HC	NB, LR, SVM, AdaBoost
Amato et al. [14]	2.1.1.2	PD diagnosis (PD vs HC)	Voice data	IPVC, local	28 PD, 22 HC; 26 PD, 18 HC	NB, kNN, SVM, RF, AdaBoost, gradient boosting, bagging ensemble
Jeancolas et al. [15]	2.1.1.2	PD diagnosis (PD vs HC)	Voice data	Local	121 PD, 100 HC	MFCC-GMM, X-vector
Jeancolas et al. [16]	2.1.1.2	PD diagnosis (PD vs HC)	Voice data	Local	117 PD, 41 iRBD, 98 HC	SVM
Quan et al. [17]	2.1.1.2	PD diagnosis (PD vs HC)	Voice data	Local	30 PD, 15 HC	DT, kNN, NB, SVM, MLP, LSTM
Adeli et al. [21]	2.1.1.3	PD diagnosis (PD vs HC)	Imaging data (MRI)	PPMI	374 PD, 169 NC	JFSS+robust LDA
Solana-Lavalle and Rosas-Romero [22]	2.1.1.3	PD diagnosis (PD vs HC)	Imaging data (MRI)	PPMI	114 PD, 49 HC; 234 PD; 110 HC	kNN, SVM, RF, NB, LR, MLP, BNN

Mudali et al. [23]	2.1.3	PD diagnosis (PD vs HC), differential diagnosis	Imaging data (FDG-PET)	Local	20 PD, 21 MSA, 17 PSP, 18 HC	DT
Huppertz et al. [26]	2.2	Differential diagnosis	Imaging data (MRI)	Local	204 PD, 106 PSP, 21 MSA-C, 60 MSA-P, 73 HC	SVM
Archer et al. [27]	2.2	Differential diagnosis	Imaging data (MRI)	Local	511 PD, 84 MSA, 129 PSP, 278 HC	SVM
Chougar et al. [28]	2.2	Differential diagnosis	Imaging data (MRI)	Local	63 PD, 21 PSP, 11 MSA-P, 12 MSA-C, 72 HC; 56 PD, 30 PSP, 24 MSA-P, 11 MSA-C, 22 HC	LR, SVM, RF
Shinde et al. [29]	2.2	Differential diagnosis	Imaging data (MRI)	Local	45 PD, 20 APS, 35 HC	CNN, gradient boosting
Jucaite et al. [30]	2.2	Differential diagnosis	Imaging data (PET)	Local	24 PD, 66 MSA	LDA
Khawaldeh et al. [31]	2.3	Parkinson's disease understanding (subthalamic nucleus activity)	Signal data (LFP)	Local	18 PD	NB
Poston et al. [32]	2.3	Parkinson's disease understanding (cognition)	Imaging data (fMRI)	Local	54 PD	SVM
Trezzi et al. [33]	2.3	Parkinson's disease understanding (metabolism)	CSF	Local	44 PD, 43 HC	LR
Vanneste et al. [34]	2.3	Parkinson's disease understanding (thalamocortical dysrhythmia)	Signal data (rs-EEG)	Local	31 PD, 264 HC, 153 tinnitus, 78 CP, 15 MDD	SVM

(continued)

**Table 3**  
(continued)

Study	Subheading	Objectives/tasks	Modalities	Data sets	Number of subjects	Methods
Ahlich et al. [37]	3.1	Symptom detection and quantification (freezing of gait)	Motion data (gait)	Local	20 PD (8 with FOG, 12 without FOG)	SVM
Aich et al. [38]	3.1	Symptom detection and quantification (freezing of gait)	Motion data (gait)	Local	51 PD (36 with FOG, 15 without FOG)	SVM, kNN, DT, NB
Borzì et al. [39]	3.1	Symptom detection and quantification (freezing of gait)	Motion data (gait)	Local	11 PD	SVM, kNN, LDA, LR
Dvorani et al. [40]	3.1	Symptom detection and quantification (freezing of gait)	Motion data (gait)	Local	16 PD	SVM, AdaBoost
Park et al. [42]	3.2	Symptom detection and quantification (bradykinesia and tremor)	Motion data (resting tremor and finger tapping)	Local	55 PD	SVM
Kim et al. [43]	3.2	Symptom detection and quantification (tremor)	Motion data (tremor)	Local	92 PD	CNN, RF, NB, LR, DT, SVM, MLP
Eskofier et al. [44]	3.2	Symptom detection and quantification (bradykinesia)	Motion data (motor tasks)	Local	10 PD	CNN, AdaBoost, kNN, SVM
Abós et al. [46]	3.3	Symptom detection and quantification (cognition)	Imaging data (rs-fMRI)	Local	27 PD-MCI, 43 PD-CN, 38 HC	SVM



Betrouni et al. [47]	3.3	Symptom detection and quantification (cognition)	Signal data (EEG)	Local	118 PD	SVM, kNN
García et al. [48]	3.3	Symptom detection and quantification (cognition)	Voice data	Local	40 PD, 40 HC	SVM, AdaBoost
Morales et al. [49]	3.3	Symptom detection and quantification (cognition)	Imaging data (MRI)	Local	16 PD-CN, 15 PD-MCI, 14 PD-D	NB, SVM
Shibata et al. [50]	3.3	Symptom detection and quantification (cognition)	Imaging data (MRI)	Local	61 PD-MCI, 59 PD-CN	RF, gradient boosting, light gradient boosting
Hannink et al. [51]	3.4	Symptom detection and quantification (gait)	Motion data (gait)	eGaIT	99 geriatric subjects	CNN
Lu et al. [52]	3.4	Symptom detection and quantification (gait)	Motion data (gait)	Local	55 PD	CNN
Gao et al. [53]	3.4	Symptom detection and quantification (falls)	Clinical data (motor exams)	Local	148 PD (45 fallers, 97 non-fallers); 103 PD (41 fallers, 62 non-fallers)	LR, SVM, RF, gradient boosting
Severson et al. [54]	4.1	Disease progression (disease states)	Clinical data	PPMI, local	423 PD, 196 HC; 610 PD	HMM
Salmanpour et al. [55]	4.1	Disease progression (disease trajectories)	Clinical and imaging (MRI) data	PPMI	885 PD	k-means
Oxtoby et al. [56]	4.2	Disease progression (prediction of sequence of events)	Clinical and genetic data	Local, PPMI	100 PD, 33 HC; 350 PD, 127 HC	Event-based model
Latourelle et al. [57]	4.2	Disease progression (prediction of motor progression)	Clinical and genetic data	PPMI, local	312 PD, 117 HC; 317 PD	REFS

(continued)

**Table 3**  
(continued)

Study	Subheading	Objectives/tasks	Modalities	Data sets	Number of subjects	Methods
Ahmadi, Tastegar et al. [58]	4.2	Disease progression (prediction of motor and non-motor symptoms)	Clinical data and inflammatory cytokine measurements	LRRK2 cohort consortium	80 iPD, 80 PD-LRRK2	Elastic-net, RF
Amara et al. [59]	4.2	Disease progression (prediction of excessive daytime sleepiness)	Clinical and imaging (DaTscan) data	PPMI	423 PD, 196 HC	Random survival forest
Couronné et al. [60]	4.2	Disease progression (prediction of motor and non-motor symptoms)	Clinical and imaging (DaTscan) data	PPMI	362 PD	Generative mixed effect model
Fauzi et al. [61]	4.2	Disease progression (prediction of impulse control disorders)	Clinical and genetic data	PPMI, local	380 PD; 388 PD	LR, SVM, RF, gradient boosting, RNN
Yang et al. [63]	5.1	Treatment adjustment (motor symptoms)	Imaging data (fMRI)	Local	38 PD	Ridge, SVM, AdaBoost, gradient boosting
Kim et al. [64]	5.1	Treatment adjustment (motor symptoms)	Clinical and medication data	PPMI	431 PD	Markov decision process
Boutet et al. [66]	5.2	Optimal deep brain stimulation parameters	Imaging data (fMRI)	Local	67 PD	LDA
Geraedts et al. [67]	5.2	Optimal deep brain stimulation parameters	Signal data (EEG)	Local	112 PD	RF

Phokaewvarangkul et al. [68]	5.3	Electrical muscle stimulation (resting tremor)	Motion data (hand tremor)	Local	20 PD	LR, RF, SVM, MLP, LSTM
Panyakaew et al. [69]	5.3	Adverse event prevention (identification of modifiable risk factors of falls)	Clinical data	Local	305 PD	Gradient boosting

*Modalities:* CSF cerebrospinal fluid, *DuTsam* dopamine transporter scan, *EEG* electroencephalography, *FDG-PET* [18F]-fluorodeoxyglucose positron emission tomography, *fMRI* functional magnetic resonance imaging, *LFP* local field potential, *MRI* magnetic resonance imaging, *PET* positron emission tomography, *rs-fMRI* resting-state functional magnetic resonance imaging

*Data sets:* *eGait* embedded Gait analysis using Intelligent Technologies (<https://www.mad.tf.fau.de/research/activitynet/digital-biobank/>), *IPYC* Italian Parkinson's Voice and Speech (<https://iee-dataport.org/open-access/italian-parkinsons-voice-and-speech>), *LRRK2* Cohort Consortium (<https://www.michaeljfox.org/news/lrrk2-cohort-consortium>), *PPMI* Parkinson's Progression Markers Initiative (<https://www.ppmi-info.org>)

*Subjects:* *APS* atypical parkinsonian syndromes, *CP* chronic pain, *FOG* freezing of gait, *HC* healthy controls, *iPD* idiopathic Parkinson's disease, *iRBD* idiopathic rapid eye movement sleep behavior disorder, *MDD* major depressive disorder, *MSA* multiple system atrophy, *MSA-C* cerebellar variant of multiple system atrophy, *MSA-P* parkinsonian variant of multiple system atrophy, *PD* Parkinson's disease, *PD-CN* Parkinson's disease with normal cognition, *PD-D* Parkinson's disease with dementia, *PD-HTI* Parkinson's disease with Hoehn and Yahr stage 1, *PD-HTII* Parkinson's disease with Hoehn and Yahr stage 2, *PD-HTIII* Parkinson's disease with Hoehn and Yahr stage 3, *PD-LRRK2* Parkinson's disease with LRRK2 mutation, *PD-MCI* Parkinson's disease with mild cognitive impairment, *PSP* progressive supranuclear palsy

*Methods:* *BNN* Bayesian neural network, *CNN* convolutional neural network, *DT* decision tree, *HMM* hidden Markov model, *JFSS* joint feature-sample selection, *KFD* kernel Fisher discriminant, *kNN* k-nearest neighbors, *LR* logistic regression, *LDA* linear discriminant analysis, *LSTM* long short-term memory, *MFCC-GMM* Mel-frequency cepstral coefficients-Gaussian mixture model, *MLP* multi-layer perceptron, *NB* naive Bayes, *REFS* reverse engineering and forward simulation, *RF* random forest, *RNN* recurrent neural network, *RUSBoost* random under-sampling boosting, *SVM* support vector machine

machines (SVM), and random forests (RF). They obtained the best predictive performance with the random forest trained on features normalized using multiple regression.

Mirelman and colleagues [11] also investigated gait and mobility measures that are indicative of PD and PD stages. They gathered data from sensors adhered to the participant's lower back, bilateral ankles, and wrists, during short walks, and extracted gait features. They investigated several strategies to perform feature selection and use a random under-sampling boosting classification algorithm to tackle class imbalance. When comparing PD patients with mild PD severity (Hoehn and Yahr stage 1) to healthy controls, they obtained good discriminative performance (84% sensitivity, 80% specificity). Most discriminative features were extracted from the upper limb sensors, with the remaining features extracted from the trunk sensor, while the lower limb sensors did not contribute to discrimination accuracy.

Kostikis and colleagues [12] investigated upper limb tremor using a smartphone-based tool. Signals from the phone's accelerometer and gyroscope were computed, from which features were extracted. They trained several machine learning algorithms, including random forest, naive Bayes, logistic regression (LR), and support vector machine, using these features as input and obtained the highest discriminative performance between PD patients and HC with the random forest model.

Kotsavasiloglou and colleagues [13] investigated the use of a pen-and-tablet device to study the differences in hand movement and muscle coordination between PD patients and HC. Data consisted of the trajectory of the pen's tip and on the pad's surface from drawings of simple horizontal lines, from which they extracted features. They investigated several machine learning algorithms, such as logistic regression, support vector machine, and random forest, and used nested cross-validation to perform feature selection. They obtained the highest discriminative performance with the naive Bayes model.

### 2.1.2 PD Diagnosis Using Voice Data

Voice data is usually recorded from high-quality microphones or from smartphones during specific vocal tasks focused on characteristics such as phonation and speech. Features are then extracted from the corresponding signals and used as input to machine learning classification algorithms.

Amato and colleagues [14] analyzed specific phonetic groups in native Italian speakers, extracted several spectral moments from the signals, and trained a SVM algorithm on these extracted features to distinguish PD patients from HC. They first worked on a public data set called Italian Parkinson's Voice and Speech,<sup>1</sup> with data

<sup>1</sup> <https://iecc-dataport.org/open-access/italian-parkinsons-voice-and-speech>

recorded in ideal publications, and obtained great performance on the validation and test sets. They then merged this public data set with a data set that they collected, with data being recorded in more realistic, suboptimal conditions, and obtained good but lower performance on the validation and test sets of this merged data set. Experiments with training on one single data set and validation on the other data set were not performed, but it would have been interesting to estimate how well a trained model could generalize on other data sets with data being recorded in different conditions.

Jeancolas and colleagues [15] investigated the early diagnosis of PD and possible gender differences in voice data. They used a pre-trained deep neural network focused on speaker recognition system to extract features and obtained a higher performance than with a standard multidimensional Gaussian mixture model, although the increase was more important among men than women. They also investigated the impact of the quality of the recordings (using either a high-quality microphone or a telephone) and obtained the same conclusions in both cases.

In another study, Jeancolas and colleagues [16] investigated the differentiation between early PD patients and patients with idiopathic rapid eye movement sleep behavior disorders (iRBD), which are important risk factors to develop PD in the near future. They extracted features related to prosody, phonation, speech fluency, and rhythm abilities from speech recordings. They once again obtained a higher predictive performance among men than women in the PD vs HC classification task and a better discriminative power for this classification task than for the iRBD vs HC one, suggesting that discriminating iRBD patients from HC using voice data is a much harder task, but it is also probably a most useful one in practice.

Quan and colleagues [17] investigated the extraction of global static features (from the whole signals) and local dynamic features (using a sliding window on the signals) from voice data during articulation tasks. They trained standard machine learning classification algorithms, such as decision trees (DT), k-nearest neighbors, naive Bayes, and support vector machines, using the static features, while they trained a recurrent neural network, more specifically a bidirectional long short-term memory (LSTM), on the dynamic features and obtained a higher predictive performance with the deep learning approach.

Although many studies reported high predictive performances, some results must be taken with caution. Indeed, a recent study reported methodological issues in several studies, including record-wise cross-validation instead of subject-wise cross-validation, high imbalance in ages between PD patients and HC, and performance metrics computed on the validation folds of k-fold cross-validation and not on an independent test set, which may lead to overly optimistic results [18].

### 2.1.3 PD Diagnosis Using Imaging Data

The diagnosis of PD remains based on its clinical presentation [19]. Imaging of dopaminergic terminals loss can be assessed using nuclear imaging, but it is not recommended in clinical routine and does not differentiate PD from other related disorders associated with dopamine neuron loss [20]. Standard brain magnetic resonance imaging (MRI) is normal in PD. However, several new markers have been recently been investigated in several studies, with mixed results.

Adeli and colleagues [21] investigated the use of T1-weighted anatomical MRI data to differentiate PD patients from HC. They developed a joint feature-sample selection algorithm in order to select an optimal subset of both features and samples from a training set, and a robust classification framework that performs denoising of the selected features and samples then learns a classification model. They analyzed data from 374 PD patients and 169 HC from the Parkinson's Progression Markers Initiative<sup>2</sup> (PPMI) cohort and included white matter, gray matter, and cerebrospinal fluid measurements from 98 regions of interest. The combination of the proposed feature selection/extraction method and classifier achieved the highest predictive accuracy (0.819), being significantly better than almost every other combination of a feature selection/extraction method and a classification algorithm.

Solana-Lavalle and Rosas-Romero [22] investigated the use of voxel-based morphometry features extracted from T1-weighted anatomical MRI to perform a PD vs HC classification task. Their pipeline consisted of five stages: (i) identification of regions of interest using voxel-based morphometry, (ii) analysis of these regions for PD detection, (iii) feature extraction based on first- and second-order statistics, (iv) feature selection based on principal component analysis, and (v) classification with tenfold cross-validation based on seven different algorithms (including k-nearest neighbors, support vector machine, random forest, naive Bayes, and logistic regression). They obtained excellent predictive performance for both male and female genders and for both 1.5 T and 3 T MRI scans (accuracy scores ranging from 0.93 to 0.99 for the best classification algorithms). However, cross-validation was performed very late in their pipeline (after the feature subset selection), which could lead to biased models and overly optimistic predictive performances.

Mudali and colleagues [23] investigated another modality, [18F]-fluorodeoxyglucose positron emission tomography (FDG-PET), to compare 20 PD patients and 18 HC. They applied the subprofile model/principal component analysis method to extract features from the images. They considered a DT algorithm and used leave-one-out cross-validation to evaluate the predictive

---

<sup>2</sup> <https://www.ppmi-info.org>

performance of the models. They obtained really low predictive performance (50% sensitivity, 45% specificity), close to chance level.

Overall, it is unclear if machine learning applied to anatomical MRI or FDG-PET can bring added value for the diagnosis of PD. However, advanced MRI sequences have the potential to bring much more valuable information [24].

## **2.2 Differential Diagnosis**

The PD vs HC binary classification task has limited utility as, even at the early stage of PD, patients have clinical symptoms strongly suggesting that they suffer from a movement disorder and thus are not healthy subjects. However, the accurate early diagnosis of parkinsonian syndromes is difficult but needed due to the different pathologies and thus the different care. Although one study investigated the differential diagnosis using sensor-based gait analysis [25], most studies investigated it using imaging data, particularly diffusion MRI.

Huppertz and colleagues [26] investigated the differential diagnosis with data from a relatively large cohort (73 HC, 204 PD, 106 PSP, 20 MSA-C, and 60 MSA-P). Using atlas-based volumetry of brain MRI data, they extracted volumes in several regions of interest and trained and evaluated a linear SVM algorithm using leave-one-out cross-validation. They obtained good predictive performance in most binary classification tasks and showed that midbrain, basal ganglia, and cerebellar peduncles were the most relevant regions.

A landmark study on this topic was published in 2019 by Archer and colleagues [27], with diffusion-weighted MRI data being collected for 1002 subjects from 17 MRI centers in Austria, Germany, and the USA. They extracted 60 free-water and 60 free-water-corrected fractional anisotropy values from diffusion-weighted MRI data, and the other features consisted of the third part of the Movement Disorder Society-Sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS III), sex, and age. They trained several SVM models and showed that the model trained using MDS-UPDRS III (with sex and age also, for all the models) performed poorly in most classification tasks, whereas the model trained using DWI features had much higher predictive performance (particularly for the MSA vs PSP task), and adding MDS-UPDRS III to this model did not improve the performance.

More recently, Chougar and colleagues [28] investigated the replication of such differential diagnosis models in clinical practice on different MRI systems. Using MRI data from 119 PD, 51 PSP, 35 MSA-P, 23 MSA-C, and 94 HC, split into a training cohort ( $n = 179$ ) and a replication cohort ( $n = 143$ ), they extracted volumes and diffusion tensor imaging (DTI) features (fractional anisotropy, mean diffusivity, axial diffusivity, and radial diffusivity) in 13 regions of interest. They investigated two feature normalization strategies (one based on the data of all subjects in the training

set and one based on the data of HC for each MRI system to tackle the different feature distributions, in particular for DTI features, because of the use of different MRI systems) and four standard machine learning algorithms, including logistic regression, support vector machines, and random forest. They obtained high performances in the replication cohort for many binary classification tasks (PD vs PSP, PD vs MSA-C, PSP vs MSA-C, PD vs atypical parkinsonism), but lower performances for other classification tasks involving MSA-P patients (PD vs MSA-P, MSA-C vs MSA-P). They showed that adding DTI features did not improve performance compared to using volumes only and that the usual normalization strategy worked best in this case.

Shinde and colleagues [29] investigated the automatic extraction of contrast ratios of the substantia nigra pars compacta from neuromelanin-sensitive MRI using a convolutional neural network. Based on the class activation maps, they identified that the left side of substantia nigra pars compacta played a more important role in the decision of the model compared to the right side, in agreement with the concept of asymmetry in PD.

A recent study [30] investigated the use of positron emission tomography of the translocator protein, expressed by glial cells, and extracted normalized standardized uptake value images and normalized total distribution volume images. Using a linear discriminant analysis algorithm with leave-one-subject-out cross-validation, they obtained great discriminative power between MSA and PD patients, with better performance with normalized total distribution volume images.

### **2.3 Disease Understanding**

Rather than focusing on the diagnosis of Parkinson's disease itself, several studies were more focused on interpreting the trained machine learning models in order to better understand the mechanisms of Parkinson's disease.

Khawaldeh and colleagues [31] investigated the task-related modulation of local field potentials of the subthalamic nucleus before and during voluntary upper and lower limb movements in 18 consecutive Parkinson's disease patients undergoing deep brain stimulation (DBS) surgery of the subthalamic nucleus in order to improve motor symptoms. Using a naive Bayes classification algorithm, they obtained chance-level performance at rest, but much higher performance during the pre-cue, pre-movement onset, and post-movement onset tasks. They showed that the presence of bursts of local field potential activity in the alpha and, even more so, in the beta frequency band significantly compromised the prediction of the limb to be moved, concluding that low-frequency bursts restrict the capacity of the basal ganglia system to encode physiologically relevant information about intended actions.



Poston and colleagues [32] investigated brain mechanisms that allow some PD patients with severe dopamine neuron loss to remain cognitively normal. Using functional MRI data from PD patients without cognitive impairment and from HC collected during a working memory task, they trained a support vector machine classifier and identified robust differences in putamen activation patterns, providing novel evidence that PD patients maintain normal cognitive performance through compensatory hyperactivation of the putamen.

Trezzi and colleagues [33] investigated cerebrospinal fluid biomarkers, and more precisely the metabolome, in early-stage PD. The logistic regression model trained on such data provided good discriminative power, and the most associated biomarkers were mannose, threonic acid, and fructose. These biomarkers were associated with antioxidative stress response, glycation, and inflammation and may help better understand PD pathogenesis.

Vanneste and colleagues [34] investigated thalamocortical dysrhythmia, which is a model proposed to explain divergent neurological disorders and is characterized by a common oscillatory pattern in which resting-state alpha activity is replaced by cross-frequency coupling of low- and high-frequency oscillations. The trained support vector machine model identified specific brain regions that provided good discriminative power between PD patients and HC, including subgenual anterior cingulate cortex, posterior cingulate cortex, parahippocampus, dorsal anterior cingulate cortex, and motor cortex. Another model also identified brain areas that are common to the pathology of Parkinson's disease, pain, tinnitus, and depression, including dorsal anterior cingulate cortex and parahippocampal area.

---

### 3 Symptom Detection and Quantification

Given the complexity and heterogeneity of Parkinson's disease, prompt accurate assessment of symptoms is needed. A detailed scale, called the Movement Disorder Society-Sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [35], is currently the gold standard to assess motor (and non-motor) features of PD patients by movement disorder specialists. The scale is divided into four sections. The first two sections allow for assessing the non-motor and motor activities of daily living, respectively, while the third section consists of a motor exam, and the fourth section allows for assessing motor complications.

Nonetheless, the MDS-UPDRS has several limitations. First, it requires time (30–45 minutes for the full scale) and a trained movement disorder specialist to fill it, limiting its use during clinical routine visits. Second, part of subjectivity from a human evaluation,

and thus variance in the MDS-UPDRS scores, cannot be excluded, with a recent study suggesting that MDS-UPDRS scores contain a substantial amount of variance [36]. Moreover, other scales are typically used to more precisely assess non-motor symptoms such as depression, anxiety, and cognition. Finally, scales are addressed during a visit at the hospital and may not reflect the symptoms in a more ecological setting, at home, during the daily life of the patient. Automatic detection and quantification of symptoms using machine learning may help tackle these limitations, and several studies investigated this topic. In the remaining of this section, we group these studies based on the symptoms investigated.

### 3.1 Freezing of Gait

Freezing of gait (FOG) is a common motor symptom and is associated with life-threatening accidents such as falls. Prompt identification or prediction of freezing of gait episodes is thus needed.

Ahlich and colleagues [37] investigated freezing of gait in 20 PD patients (8 with FOG, 12 without FOG), split into a training set (15 patients) and a test set (5 patients). They collected sensor (accelerometer, gyroscope, and magnetometer) data during scripted activities (e.g., walking around the apartment, carrying a full glass of water from the kitchen to another room) and non-scripted activities (e.g., answering the phone). Two recording sessions were considered, one in “OFF” motor state and one in “ON” motor state, and the data was labeled by experienced clinicians based on the corresponding video recordings. The task was a binary classification task (FOG vs no FOG) for each window. They extracted sub-signals from the whole signals using a sliding window and then extracted features and in the time and frequency domains for each sub-signal. They trained two SVM algorithms (one with a linear kernel, one with a Gaussian kernel) and obtained high and better results with the linear kernel.

Aich and colleagues [38] gathered sensor data for 36 PD patients with FOG and 15 PD patients without FOG from 2 wearable triaxial accelerometers during clinical experiments. They extracted features, such step time, stride time, step length, stride length, and walking speed, from the signals. They trained several classic machine learning classification algorithms (SVM, kNN, DT, NB) and obtained good predictive performances with all of them, although the SVM model had the highest mean accuracy on the test sets of the cross-validation procedure.

Borzi and colleagues [39] collected data from 2 inertial sensors placed on each shin of the 11 PD patients during the “timed up and go” test in order to investigate FOG and pre-FOG detection. They extracted features in the time and frequency domains and trained decision tree algorithms. They obtained great predictive performance to detect FOG episodes, but lower performance to predict pre-FOG episodes, with the performance decreasing even more as the window length increased.

Dvorani and colleagues [40] were interested in detecting foot motion phases using a shoe-placed inertial sensor in order to detect FOG episodes. They extracted ten features, including stride length, maximum gait velocity, and step duration, from each motion phase and trained a SVM algorithm to detect FOG episodes. They obtained great performance when using features from the current and two preceding motion phases, but lower performance when using only features from the two preceding motion phases, highlighting the higher difficulty to predict FOG episodes in advance. Shalin and colleagues [41] reached the same conclusion using plantar pressure data and a long short-term memory neural network.

### **3.2 *Bradykinesia and Tremor***

Bradykinesia and tremor are two other motor symptoms that are frequently investigated for automatic assessment.

Park and colleagues [42] investigated automated rating for resting tremor and bradykinesia from video clips of resting tremor and finger tapping of the bilateral upper limbs. They extracted several features from the video clips, including resting tremor amplitude and finger tapping speed, amplitude, and fatigue, using a pre-trained deep learning model. These features were used as input of a SVM algorithm to predict the corresponding scores from the MDS-UPDRS scale. For resting tremors, the automated approach had excellent reliability range with the gold standard rating and higher performance than that of non-trained human rater. For finger tapping, the automated approach had good reliability range with the gold standard rating and similar performance than that of non-trained human rater.

Kim and colleagues [43] performed a study in which they investigated tremor severity using three-dimensional acceleration and gyroscope data obtained from wearable device. They investigated a convolutional neural network to automatically extract features and perform classification, compared to extracting defined features from the time and frequency domains and training standard machine learning algorithms (random forest, naive Bayes, linear regression, support vector machines) using these features. They obtained better higher predictive performance with the deep learning approach than the standard machine learning approach. Eskofier and colleagues [44] obtained similar results using inertial measurement units collected during motor tasks.

### **3.3 *Cognition***

Cognitive impairment is frequent in PD, with the point prevalence of PD dementia being around 30% and the cumulative prevalence for patients surviving more than 10 years being at least 75% [45]. Due to its high negative impact on the quality of life of PD patients and their caregivers, it is important to identify and quantify cognitive impairment. Several scales to assess cognition already

exist, such as the Mini-Mental State Examination and the Montreal Cognitive Assessment, but automatic assessment of cognition could be helpful.

Abós and colleagues [46] investigated discriminating cognitive status in PD through functional connectomics. Using resting-state functional MRI data, they extracted features consisting of connection-wise pattern of functional connectivity. They performed feature selection using randomized logistic regression with leave-one-out cross-validation and then trained a SVM algorithm. They obtained good discriminative performance between PD patients with mild cognitive impairment and with no cognitive impairment, but could not report significant connectivity reductions between both groups.

Betrouni and colleagues [47] investigated the use of electroencephalograms to automatically assess their cognitive status. A cluster analysis of the neuropsychological assessments of 118 PD patients revealed 5 cognition clusters. They extracted quantitative features from the electroencephalograms and performed feature selection based on Pearson correlation tests. They trained two machine learning algorithms (kNN and SVM), using a fivefold cross-validation procedure that was repeated five times, and obtained good similar predictive performances for the five-class classification task with both models.

García and colleagues [48] investigated cognitive decline using dysarthric symptoms. They extracted prosodic, articulatory, and phonemic identifiability features from speech signals recorded during the reading of two narratives. Using a SVM algorithm and nested cross-validation, they obtained correct discriminative performance (area under the receiver operating characteristics curve of 0.76), with the highest performance being obtained using phonemic identifiability features.

Morales and colleagues [49] investigated the classification of PD patients with no cognitive impairment ( $n = 16$ ), with mild cognitive impairment ( $n = 15$ ), and with dementia ( $n = 14$ ). They trained several variants of the naive Bayes algorithm and 1 SVM algorithm on 112 MRI features consisting of volumes of subcortical structures and thickness of cortical parcels and obtained good discriminative performance in the 3 binary classification tasks, the lower performance corresponding to the differentiation between PD patients with no cognitive impairment with mild cognitive impairment. The most important features involved the following brain regions: left cerebral cortex, left caudate, left entorhinal, right inferior left hippocampus, and brainstem.

A recent study [50] also investigated MRI data, more specifically quantitative susceptibility mapping images parcellated into 20 regions of interest, for the early detection of cognitive impairment in PD. Using tree-based ensemble machine learning algorithms, such as random forest and extreme gradient boosting,

they obtained acceptable predictive performance and showed that the features corresponding to the caudate nucleus were important for classification and also inversely correlated with Montreal Cognitive Assessment scores.

### 3.4 Other Symptoms

Although less prevalent in the literature, studies also investigated other PD symptoms such as falls and motor severity.

An early study by Hannink and colleagues [51] was performed to investigate gait parameter extraction from sensor data using convolutional neural networks. Using 3d-accelerometer and 3d-gyroscope data from 99 geriatric patients, the objective was to predict the stride length and width, the foot angle, and the heel and toe contact times. They investigated two approaches to tackle this multi-output regression task, either training a single convolutional neural network to predict the five outcomes or training a convolutional neural network for each outcome, and obtained better performance on an independent test set with the latter approach. Although the considered population was not parkinsonian, the prevalence of gait symptoms in this population and the obtained results might be relevant to better understand gait in this population. Lu and colleagues [52] investigated gait in PD, as measured by MDS-UPDRS item 3.10, which does not include freezing of gait. They collected video recordings of MDS-UPDRS exams from 55 participants which were scored by 3 different trained movement disorder neurologists, and the ground truth score was defined using majority voting among the 3 raters. They performed skeleton extraction from the videos and trained a convolutional neural network, with regularization using rater confusion estimation to tackle noise in labels, to predict gait severity. They obtained correct performance on the test set (72% accuracy with majority voting, 84% accuracy with the model predicting at least one of the raters' scores).

Gao and colleagues [53] investigated falls in two data sets independently collected at two different sites. Using clinical scores as input, they trained several classic machine learning classification algorithms to differentiate fallers from non-fallers. They obtained acceptable predictive performance in both data sets when training and evaluating (using cross-validation) a model in each data set independently. They also showed that the predictive performance was lower when training the model on one data set and evaluating it on the other data set, which is not surprising, but it is important to have this possible issue in mind when a model is not evaluated on a different cohort.

---

## 4 Disease Progression

Given the complexity and heterogeneity of Parkinson's, prediction of disease progression with individual trajectories is challenging. Two subtypes of PD, one with more postural instability and gait difficulty and the other one with more tremor symptoms, are already known. Nonetheless, there are other motor symptoms in PD, and many PD symptoms are non-motor; thus, deeper knowledge is required to understand disease progression.

### 4.1 Disease Subtypes

Several studies focused on the identification of more specific disease subtypes than the two aforementioned well-known ones characterized by postural instability and gait difficulty for one and tremor-predominant for the other.

Severson and colleagues [54] worked on the development of a statistical progression model of Parkinson's disease accounting for intra-individual and inter-individual variability, as well as medication effects. They built a contrastive latent variable model followed by a personalized input-output hidden Markov model to define disease states and assessed the clinical significance of the states on seven key motor or cognitive outcomes (mild cognitive impairment, dementia, dyskinesia, presence of motor fluctuations, functional impairment from motor fluctuations, Hoehn and Yahr score, and death). They identified eight disease states that were primarily differentiated by functional impairment, tremor, bradykinesia, and neuropsychiatric measures. The terminal state had the highest prevalence of key clinical outcomes, including almost every recorded instance of dementia. The discovered states were non-sequential, with overlapping disease progression trajectories, supporting the use of non-deterministic disease progression models, and suggesting that static subtype assignment might be ineffective at capturing the full spectrum of PD progression.

Salmanpour and colleagues [55] performed a longitudinal clustering analysis and prediction PD progression. They extracted almost a thousand features, including motor, non-motor, and radiomics features. They performed a cross-sectional clustering analysis and identified three distinct progression trajectories, with two trajectories being characterized by disease escalation and the other trajectory by disease stability. They also investigated the prediction of progression trajectories from early stage (baseline and year 1) data and obtained the highest predictive performance with a probabilistic neural network.

### 4.2 Prediction of Future Motor and Non-motor Symptoms

Prediction of future symptoms and individual disease trajectories was the main focus of several studies.

Oxtoby and colleagues [56] aimed at estimating the sequence of clinical and neurodegeneration events, and variability in this

sequence, using data-driven disease progression modelling, with a focus on PD patients with higher risk of developing dementia (defined as PD patients being diagnosed at age 65 or later). They analyzed baseline visit data from two separate cohorts: a local discovery cohort (100 PD patients and 33 HC) and a replication cohort (PPMI study, 350 PD patients and 127 HC). They considered 42 features, including 8 clinical/cognitive measures, 6 vision measures, 4 retinal measures, 8 regional measures of cortical thickness, 4 measures of white matter neurodegeneration in the substantia nigra, and 12 regional measures of brain iron content. They trained event-based models that incorporate non-parametric mixture modelling using ten fivefold cross-validation procedures to estimate the robustness of the models. The authors showed that Parkinson's progression in patients at higher risk of developing dementia starts with classic prodromal features of PD (sleep and olfactory disorders), followed by early deficits in visual abilities and increased brain iron content, followed later by a less certain ordering of neurodegeneration in the substantia nigra and cortex, neuropsychological cognitive deficits, retinal thinning in dopamine layers, and further visual deficits. Their results support the growing piece of evidence that visual processing specifically is affected early in PD patients with high risk of developing dementia.

Latourelle and colleagues [57] investigated the development of predictive models of motor progression using longitudinal clinical, molecular, and genetic data. More specifically, the objective was to predict the annual rate of changes in combined scores from the second and third parts of the MDS-UPDRS. The trained model showed strong performance in the training cohort (using fivefold cross-validation) and lower but still significant performance in an independent replication cohort. The most relevant features included baseline MDS-UPDRS motor score, sex, and age, as well as a novel PD-specific epistatic interaction. Genetic variation was the most useful prediction of motor progression, and baseline CSF biomarkers had a lower but still significant effect on predicting motor progression. They also performed simulations with the trained model and concluded that incorporating the predicted rates of motor progression into the final models of treatment effect reduced the variability in the study outcome, allowing significant differences to be detected at sample sizes up to 20% smaller than in naive trials.

Ahmadi Rastegar and colleagues [58] investigated the prediction of longitudinal clinical outcomes after 2-year follow-up from baseline and 1-year follow-up data. They also measured 27 inflammatory cytokines and chemokines in serum at baseline and after 1 year to investigate cytokine stability. Training random forest algorithms, the best prediction models were for motor symptom severity scales (Hoehn and Yahr stage and MDS-UPDRS III total score), and several inflammatory cytokine and chemokine features



were among the most relevant features to predict Hoehn and Yahr stage and MDS-UPDRS III total score, giving evidence that peripheral cytokines may have utility for aiding prediction of PD progression using machine learning models.

Amara and colleagues [59] investigated the prediction of future incidents of excessive daytime sleepiness. They trained a random survival forest using 33 baseline variables, including anxiety, depression, rapid eye movement sleep, cognitive scores,  $\alpha$ -synuclein, p-tau, t-tau, and ApoE  $\epsilon 4$  status. The performance of the model was only marginally better than random guess, but the strongest predictive features were p-tau and t-tau.

Couronné and colleagues [60] performed longitudinal data analysis to predict patient-specific trajectories. They proposed to use a generative mixed effect model that considers the progression trajectories as curves on a Riemannian manifold and that can handle missing values. They applied their model to PD progression with joint modelling of two features (MDS-UPDRS III total score and striatal binding ratio in right caudate). Interpretation of the model revealed that patients with later onset progress significantly faster and that  $\alpha$ -synuclein mean level was correlated with PD onset.

Faouzi and colleagues [61] investigated the prediction of future impulse control disorders (psychiatric disorders characterized by the inability to resist an urge or an impulse and which include a wide range of types including compulsive shopping, internet addiction, and hypersexuality, for instance) in Parkinson's disease. The objective of their study was to predict the presence or absence of these disorders at the next clinical visit of a given patient. Using clinical and genetic data, they trained several machine learning models on a training cohort and evaluated the models on the training cohort (using cross-validation) and on an independent replication cohort. They showed that a recurrent neural network model achieved significantly better performance than a trivial model (predicting the status at the next visit with the status at the most recent visit), but the increase in performance was too small to be deemed clinically relevant. Nevertheless, this proof-of-concept study highlights the potential of machine learning for such prediction.

---

## 5 Treatment Adjustment and Adverse Event Prevention

Being able to predict future adverse events in Parkinson's disease is useful, but being able to prevent them would be even more useful. Parkinson's disease is one of the few neurodegenerative diseases where current therapies can greatly improve the quality of life of the patients, but these therapies also have adverse effects. Providing personalized adapted therapies to every patient is of high importance.



Machine learning allows for unveiling complex correlations or patterns from data. However, correlation does not imply causality: if two variables are correlated, one variable does not necessarily cause an effect on the other. Therefore, standard machine learning is not always well adapted to draw conclusions for personalized therapies. Ultimately, clinical trials with a specific hypothesis tested are the best solution to draw causality effect conclusions. Nonetheless, several machine learning approaches can investigate causality effects. Causal inference, that is, being able to discover which variables have which impacts on which other variables, is an open research topic in machine learning, but usually requires a lot of data, limiting its use in Parkinson's disease. Nonetheless, exploratory studies suggesting potential options for personalized therapies and adverse event prevention have been published.

### **5.1 Dopamine Replacement Therapy**

Dopamine replacement therapy, as a way to compensate the loss of dopamine neurons in the brain, is the most common therapy due to its efficacy and simplicity (drug intake). Nonetheless, it also comes with adverse effects and long-term motor complications such as motor fluctuations (worsening or reappearance of motor symptoms before the next drug intake) and dyskinesia (involuntary muscle movements) [62].

Yang and colleagues [63] investigated the utility of amplitude of low-frequency fluctuation computed from functional MRI data of 38 PD patients in order to predict individual patient's response to levodopa treatment. They applied principal component analysis to perform dimensionality reduction and trained gradient tree boosting algorithms to discriminate between moderate and superior responders to levodopa treatment. Treatment efficacy was defined based on motor symptom improvement from the state of medication off to medication on, as assessed by MDS-UPDRS III total score. They obtained great discriminative performance between both groups, even though no significant difference in clinical data was observed between both groups. The mainly contributed regions for both models included the bilateral primary motor cortex, the occipital cortex, the cerebellum, and the basal ganglia. These results suggest the potential utility of amplitude of low-frequency fluctuation as promising predictive markers of dopaminergic therapy response in PD patients.

Kim and colleagues [64] investigated the use of reinforcement learning to predict optimal treatment for reducing motor symptoms. They derived clinically relevant disease states and an optimal combination of medications for each of them by using policy iteration of the Markov decision process. Their model achieved a lower level of motor symptom severity scores than what clinicians did, whereas the clinicians' medication rules were more consistent than their model. Their model followed the clinician's medication rules in most cases but also suggested some changes, which leads to the

difference in lowering symptom severity. This proof of concept showed the potential utility of reinforcement learning to derive optimal treatment strategies.

## 5.2 Deep Brain Stimulation

Deep brain stimulation is a neurosurgical procedure that uses implanted electrodes and electrical stimulation and has proven efficacy in advanced Parkinson's disease by decreasing motor fluctuations and dyskinesia and improving quality of life [65]. The most commonly stimulated region is the subthalamic nucleus, but the globus pallidus is sometimes preferred. Although DBS usually greatly improves the motor symptoms, it also has downsides, such as requiring personalized parameters and potential adverse events such as postoperative cognitive decline.

Boutet and colleagues [66] investigated the prediction of optimal deep brain stimulation parameters from functional MRI data. They extracted blood-oxygen-level-dependent (BOLD) signals in 16 motor and non-motor regions of interest for 67 PD patients, from which 62 underwent DBS of the subthalamic nucleus and 5 underwent DBS of the globus pallidus. They trained a linear discriminant analysis algorithm on normalized BOLD changes using fivefold cross-validation and obtained great performance in classifying optimal vs non-optimal parameter settings, although the performance was lower on two additional (a priori clinically optimized and in stimulation-naïve patients) unseen data sets.

Geraedts and colleagues [67] also investigated deep brain stimulation in the context of cognitive function, as a downside of DBS for PD is the potential deterioration of cognition postoperatively. They extracted features from electroencephalograms, trained random forest algorithms using tenfold cross-validation, and obtained great discrimination between PD patients with the best and worst cognitive performances. However, it should be noted that they only included the best and worst cognitive performers ( $n = 20$  per group from 112 PD patients), making the classification task probably much easier than if it was performed on the 112 PD patients, thus requiring their model to be evaluated on PD patients independently on their cognitive performance. Nonetheless, their results suggest the potential utility of electroencephalography for cognitive profiling in DBS.

## 5.3 Others

Phokaewvarangkul and colleagues [68] explored the effect of electrical muscle stimulation as an adjunctive treatment for resting tremor during "ON" period, with machine learning used to predict the optimal stimulation level that will yield the longest period of tremor reduction or tremor reset time. They used sensor data from a glove incorporating a three-axis gyroscope to measure tremor signals. The stimulation levels were discretized into five ordinary classes, with the objective to predict the accurate class from the sensor data. They observed a significant reduction in tremor

parameters during stimulation. The best performing machine learning model was a LSTM neural network in comparison to classic algorithms such as logistic regression, support vector machine, and random forest. The high predictive performance of the LSTM model confirmed the potential utility of electrical muscle stimulation for the reduction of resting tremors in PD.

Panyakaew and colleagues [69] investigated the identification of modifiable risk factors of falls. The input data consisted of clinical demographics, medications, and balanced confidence scaled by the 16-item Activities-Specific Balance Confidence (ABC-16) scale, from 305 PD patients (99 fallers, 58 recurrent fallers, and 148 non-fallers). They trained two gradient tree boosting algorithms using sevenfold cross-validation. They obtained good predictive performance at differentiating fallers from non-fallers, the most relevant features being item 7 (sweeping the floor), item 5 (reaching on tiptoes), and item 12 (walking in a crowded mall) from the ABC-16 scale, followed by disease stage and duration. They obtained even better performance at differentiating recurrent fallers from non-fallers, the most relevant features being items 12, 5, and 10 (walking across a parking lot) from the ABC-16 scale, followed by disease stage and current age.

---

## 6 Conclusion

Many research works on Parkinson's disease and related disorders using machine learning have been published in the literature, investigating diagnosis, symptom severity, disease progression, and personalized therapies. These studies provide new insights to better understand these neurodegenerative disorders.

However, many questions and challenges are still open. The early-stage, and even more so the prodromal-stage, classification of Parkinson's disease is still very challenging. The early differential diagnosis of parkinsonian syndromes is another topic for which higher performance is needed at an early stage. More highly personalized therapies are also needed to better improve the quality of life of the PD patients. All the research works on these topics also need to be evaluated in non-research environments in order to be translated to the clinics.

Right usage of machine learning is required to try to answer these questions and challenges. The most common methodological issues are usually related to the cross-validation procedure used, which can lead to biased, overly optimistic, reported predictive performance. Nonetheless, our anecdotal experience after performing this literature review is that these issues are less and less frequent over time. Nonetheless, many studies use small data sets and leave-one-out cross-validation, which provides an unbiased estimation of the predictive performance, but with high variance.

The few studies that investigated replicating their results in another independent data set all reported (much) lower predictive performance, highlighting the critical need of replication. Using artificial intelligence algorithms also rises ethical and legal issues regarding the collection, processing, storage, and reuse of potentially sensitive patient data, particularly coming from sensor-based digital data [9]. These aspects will have to be taken into consideration when transferring results obtained from clinical research to clinical routine use.

To conclude, the use of machine learning has allowed researchers to better understand Parkinson's disease and related disorders and suggested potential to better diagnose these disorders as well as to provide better care for the patients, but more research works and replication studies are required to translate these results into the clinics through clinical decision support systems.

---

## Acknowledgments

The authors would like to thank Jochen Klucken for his fruitful remarks. This work was supported by the French government under the management of Agence Nationale de la Recherche as part of the "Investissements d'avenir" program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute) and reference ANR-10-IAIHU-06 (Agence Nationale de la Recherche-10-IA Institut Hospitalo-Universitaire-6), by the European Union H2020 Programme (grant number 826421, project TVB-Cloud), and by the ERA PerMed EU-wide project DIGIPD (01KU2110).

## References

1. GBD (2016) Parkinson's disease collaborators (2018) global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol* 17:939–953
2. Ascherio A, Schwarzschild MA (2016) The epidemiology of Parkinson's disease: risk factors and prevention. *Lancet Neurol* 15:1257–1272
3. Blauwendraat C, Nalls MA, Singleton AB (2020) The genetic architecture of Parkinson's disease. *Lancet Neurol* 19:170–178
4. Corti O, Lesage S, Brice A (2011) What genetics tells us about the causes and mechanisms of Parkinson's disease. *Physiol Rev* 91:1161–1218
5. Sambin S, Lavisse S, Decaix C et al (2022) Compensatory mechanisms nine years before Parkinson's disease conversion in a LRRK2 R1441H family. *Mov Disord* 37:428–430
6. Wenning GK, Stankovic I, Vignatelli L et al (2022) The Movement Disorder Society criteria for the diagnosis of multiple system atrophy. *Mov Disord* 37:1131
7. Tolosa E, Garrido A, Scholz SW et al (2021) Challenges in the diagnosis of Parkinson's disease. *Lancet Neurol* 20:385–397
8. Boxer AL, Yu J-T, Golbe LI et al (2017) Advances in progressive supranuclear palsy: new diagnostic criteria, biomarkers, and therapeutic approaches. *Lancet Neurol* 16:552–563
9. Fröhlich H, Bontridder N, Petrovska-Dela-créta D et al (2022) Leveraging the potential of digital technology for better individualized treatment of Parkinson's disease. *Front Neurol* 13:788427
10. Wahid F, Begg RK, Hass CJ et al (2015) Classification of Parkinson's disease gait using spatial-temporal gait features. *IEEE J Biomed Health Inform* 19:1794–1802

11. Mirelman A, Ben Or Frank M, Melamed M et al (2021) Detecting sensitive mobility features for Parkinson's disease stages via machine learning. *Mov Disord* 36:2144–2155
12. Kostikis N, Hristu-Varsakelis D, Arnaoutoglou M et al (2015) A smartphone-based tool for assessing parkinsonian hand tremor. *IEEE J Biomed Health Inform* 19:1835–1842
13. Kotsavasiloglou C, Kostikis N, Hristu-Varsakelis D et al (2017) Machine learning-based classification of simple drawing movements in Parkinson's disease. *Biomed Signal Process Control* 31:174–180
14. Amato F, Borzi L, Olmo G et al (2021) Speech impairment in Parkinson's disease: acoustic analysis of unvoiced consonants in Italian native speakers. *IEEE Access* 9:166370–166381
15. Jeancolas L, Petrovska-Delacrétaz D, Mangone G et al (2021) X-vectors: new quantitative biomarkers for early Parkinson's disease detection from speech. *Front Neuroinform* 15:578369
16. Jeancolas L, Mangone G, Petrovska-Delacrétaz D et al (2022) Voice characteristics from isolated rapid eye movement sleep behavior disorder to early Parkinson's disease. *Parkinsonism Relat Disord* 95:86–91
17. Quan C, Ren K, Luo Z (2021) A deep learning based method for Parkinson's disease detection using dynamic features of speech. *IEEE Access* 9:10239–10252
18. Ozbolt AS, Moro-Velazquez L, Lina I et al (2022) Things to consider when automatically detecting Parkinson's disease using the phonation of sustained vowels: analysis of methodological issues. *Appl Sci* 12:991
19. Berg D, Adler CH, Bloem BR et al (2018) Movement disorder society criteria for clinically established early Parkinson's disease. *Mov Disord* 33:1643–1646
20. de la Fuente-Fernández R (2012) Role of DaTSCAN and clinical diagnosis in Parkinson disease. *Neurology* 78:696–701
21. Adeli E, Shi F, An L et al (2016) Joint feature-sample selection and robust diagnosis of Parkinson's disease from MRI data. *NeuroImage* 141:206–219
22. Solana-Lavalle G, Rosas-Romero R (2021) Classification of PPMI MRI scans with voxel-based morphometry and machine learning to assist in the diagnosis of Parkinson's disease. *Comput Methods Prog Biomed* 198:105793
23. Mudali D, Teune LK, Renken RJ et al (2015) Classification of parkinsonian syndromes from FDG-PET brain data using decision trees with SSM/PCA features. *Comput Math Methods Med* 2015:136921
24. Mitchell T, Lehericy S, Chiu SY et al (2021) Emerging neuroimaging biomarkers across disease stage in Parkinson disease: a review. *JAMA Neurol* 78:1262–1272
25. Gaßner H, Raccagni C, Eskofier BM et al (2019) The diagnostic scope of sensor-based gait analysis in atypical parkinsonism: further observations. *Front Neurol* 10:5
26. Huppertz H-J, Möller L, Südmeyer M et al (2016) Differentiation of neurodegenerative parkinsonian syndromes by volumetric magnetic resonance imaging analysis and support vector machine classification. *Mov Disord* 31:1506–1517
27. Archer DB, Bricker JT, Chu WT et al (2019) Development and validation of the automated imaging differentiation in parkinsonism (AID-P): a multi-site machine learning study. *Lancet Digit Health* 1:e222–e231
28. Chougar L, Faouzi J, Pyatigorskaya N et al (2021) Automated categorization of parkinsonian syndromes using magnetic resonance imaging in a clinical setting. *Mov Disord* 36:460–470
29. Shinde S, Prasad S, Saboo Y et al (2019) Predictive markers for Parkinson's disease using deep neural nets on neuromelanin sensitive MRI. *Neuroimage Clin* 22:101748
30. Jucaite A, Cselényi Z, Kreisl WC et al (2022) Glia imaging differentiates multiple system atrophy from Parkinson's disease: a positron emission tomography study with [<sup>11</sup>C] PBR28 and machine learning analysis. *Mov Disord* 37:119–129
31. Khawaldeh S, Tinkhauser G, Shah SA et al (2020) Subthalamic nucleus activity dynamics and limb movement prediction in Parkinson's disease. *Brain* 143:582–596
32. Poston KL, YorkWilliams S, Zhang K et al (2016) Compensatory neural mechanisms in cognitively unimpaired Parkinson disease. *Ann Neurol* 79:448–463
33. Trezzi J-P, Galozzi S, Jaeger C et al (2017) Distinct metabolomic signature in cerebrospinal fluid in early parkinson's disease. *Mov Disord* 32:1401–1408
34. Vanneste S, Song J-J, De Ridder D (2018) Thalamocortical dysrhythmia detected by machine learning. *Nat Commun* 9:1103
35. Goetz CG, Fahn S, Martinez-Martin P et al (2007) Movement Disorder Society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): process, format, and clinimetric testing plan. *Mov Disord* 22:41–47
36. Evers LJW, Krijthe JH, Meinders MJ et al (2019) Measuring Parkinson's disease

- over time: the real-world within-subject reliability of the MDS-UPDRS. *Mov Disord* 34: 1480–1487
37. Ahlrichs C, Samà A, Lawo M et al (2016) Detecting freezing of gait with a tri-axial accelerometer in Parkinson's disease patients. *Med Biol Eng Comput* 54:223–233
  38. Aich S, Pradhan PM, Park J et al (2018) A validation study of freezing of gait (FoG) detection and machine-learning-based FoG prediction using estimated gait characteristics with a wearable accelerometer. *Sensors* 18: 3287
  39. Borzi L, Mazzetta I, Zampogna A et al (2021) Prediction of freezing of gait in Parkinson's disease using wearables and machine learning. *Sensors* 21:614
  40. Dvorani A, Waldheim V, Jochner MCE et al (2021) Real-time detection of freezing motions in Parkinson's patients for adaptive gait phase synchronous cueing. *Front Neurol* 12:720516
  41. Shalin G, Pardoel S, Lemaire ED et al (2021) Prediction and detection of freezing of gait in Parkinson's disease from plantar pressure data using long short-term memory neural networks. *J Neuroeng Rehabil* 18:1–15
  42. Park KW, Lee E-J, Lee JS et al (2021) Machine learning-based automatic rating for cardinal symptoms of Parkinson disease. *Neurology* 96:e1761–e1769
  43. Kim HB, Lee WW, Kim A et al (2018) Wrist sensor-based tremor severity quantification in Parkinson's disease using convolutional neural network. *Comput Biol Med* 95:140–146
  44. Eskofier BM, Lee SI, Daneault J-F, et al (2016) Recent machine learning advancements in sensor-based mobility analysis: deep learning for Parkinson's disease assessment. In: 2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 655–658
  45. Litvan I, Aarsland D, Adler CH et al (2011) MDS task force on mild cognitive impairment in Parkinson's disease: critical review of PD-MCI. *Mov Disord* 26:1814–1824
  46. Abós A, Baggio HC, Segura B et al (2017) Discriminating cognitive status in Parkinson's disease through functional connectomics and machine learning. *Sci Rep* 7:45347
  47. Betrouni N, Delval A, Chaton L et al (2019) Electroencephalography-based machine learning for cognitive profiling in Parkinson's disease: preliminary results. *Mov Disord* 34: 210–217
  48. García AM, Arias-Vergara TC, Vasquez-Correa J et al (2021) Cognitive determinants of dysarthria in Parkinson's disease: an automated machine learning approach. *Mov Disord* 36: 2862–2873
  49. Morales DA, Vives-Gilabert Y, Gómez-Ansón B et al (2013) Predicting dementia development in Parkinson's disease using Bayesian network classifiers. *Psychiatry Res* 213:92–98
  50. Shibata H, Uchida Y, Inui S et al (2022) Machine learning trained with quantitative susceptibility mapping to detect mild cognitive impairment in Parkinson's disease. *Parkinsonism Relat Disord* 94:104–110
  51. Hannink J, Kautz T, Pasluosta CF et al (2017) Sensor-based gait parameter extraction with deep convolutional neural networks. *IEEE J Biomed Health Inform* 21:85–93
  52. Lu M, Zhao Q, Poston KL et al (2021) Quantifying Parkinson's disease motor severity under uncertainty using MDS-UPDRS videos. *Med Image Anal* 73:102179
  53. Gao C, Sun H, Wang T et al (2018) Model-based and model-free machine learning techniques for diagnostic prediction and classification of clinical outcomes in Parkinson's disease. *Sci Rep* 8:7129
  54. Severson KA, Chahine LM, Smolensky LA et al (2021) Discovery of Parkinson's disease states and disease progression modelling: a longitudinal data study using machine learning. *Lancet Digit Health* 3:e555–e564
  55. Salmanpour MR, Shamsaei M, Hajianfar G et al (2022) Longitudinal clustering analysis and prediction of Parkinson's disease progression using radiomics and hybrid machine learning. *Quant Imaging Med Surg* 12:90619–90919
  56. Oxtoby NP, Leyland L-A, Aksman LM et al (2021) Sequence of clinical and neurodegeneration events in Parkinson's disease progression. *Brain J Neurol* 144:975–988
  57. Latourelle JC, Beste MT, Hadzi TC et al (2017) Large-scale identification of clinical and genetic predictors of motor progression in patients with newly diagnosed Parkinson's disease: a longitudinal cohort study and validation. *Lancet Neurol* 16:908–916
  58. Ahmadi Rastegar D, Ho N, Halliday GM et al (2019) Parkinson's progression prediction using machine learning and serum cytokines. *NPJ Park Dis* 5:1–8
  59. Amara AW, Chahine LM, Caspell-Garcia C et al (2017) Longitudinal assessment of excessive daytime sleepiness in early Parkinson's disease. *J Neurol Neurosurg Psychiatry* 88:653–662
  60. Couronne R, Vidailhet M, Corvol JC, et al (2019) Learning disease progression models with longitudinal data and missing values. In: 2019 IEEE 16th international symposium on

- biomedical imaging (ISBI 2019), pp 1033–1037
61. Faouzi J, Bekadar S, Artaud F et al (2022) Machine learning-based prediction of impulse control disorders in Parkinson's disease from clinical and genetic data. *IEEE Open J Eng Med Biol* 3:96–107
  62. You H, Mariani L-L, Mangone G et al (2018) Molecular basis of dopamine replacement therapy and its side effects in Parkinson's disease. *Cell Tissue Res* 373:111–135
  63. Yang B, Wang X, Mo J et al (2021) The amplitude of low-frequency fluctuation predicts levodopa treatment response in patients with Parkinson's disease. *Parkinsonism Relat Disord* 92:26–32
  64. Kim Y, Suescun J, Schiess MC et al (2021) Computational medication regimen for Parkinson's disease using reinforcement learning. *Sci Rep* 11:9313
  65. Perestelo-Pérez L, Rivero-Santana A, Pérez-Ramos J et al (2014) Deep brain stimulation in Parkinson's disease: meta-analysis of randomized controlled trials. *J Neurol* 261: 2051–2060
  66. Boutet A, Madhavan R, Elias GJB et al (2021) Predicting optimal deep brain stimulation parameters for Parkinson's disease using functional MRI and machine learning. *Nat Commun* 12: 3043
  67. Geraedts VJ, Koch M, Contarino MF et al (2021) Machine learning for automated EEG-based biomarkers of cognitive impairment during deep brain stimulation screening in patients with Parkinson's disease. *Clin Neurophysiol* 132:1041–1048
  68. Phokaewvarangkul O, Vateekul P, Wichakam I et al (2021) Using machine learning for predicting the best outcomes with electrical muscle stimulation for tremors in Parkinson's disease. *Front Aging Neurosci* 13:727654
  69. Panyakaew P, Pornputtpong N, Bhidayasiri R (2021) Using machine learning-based analytics of daily activities to identify modifiable risk factors for falling in Parkinson's disease. *Parkinsonism Relat Disord* 82:77–83

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

