Master's Degree in Biomedical Engineering

Master's Degree Thesis

A retrospective study on deep brain stimulation effects on impulsivity in Parkinsonian patients

Politecnico di Torino

Supervisor

Prof. Molinari Filippo

Co-Supervisor

Prof. Mazzoni Alberto

Candidate

Cubeddu Laura

July 2023

Acknowledgment

It has been three intense years, at the end of which I feel I have earned a lot of knowledge and important skills, but above all I have grown as a person. I have faced many difficulties that taught me how to give the right weight to things, to be proud of myself and to be grateful for the people around me, their help and their affection.

I would like to thank Professor Alberto Mazzoni for this opportunity, for his enthusiasm and humanity, he has been an inspiration during this journey. An important thank goes to Ahmet, who followed me closely, reassured me often and believed in me. Thanks also to the team of doctors and residents at the Careggi Hospital in Florence, especially Dr. Silvia Ramat, with whom I was able to have some interesting discussions and who gave me a closer insight into the reality that is the subject of my thesis.

First and foremost, I thank my family, who always supported me even when I thought I couldn't do it, and with whom I am happy to share this achievement, which is also a bit of theirs. To my friends, who are always there, even those who are physically far away, ready to give me their support and to bring a smile to my face, making everything more colorful and lighter, I am grateful to have you in my life. I thank all those who have joined me even for a small part of the journey, you are part of it too. And why not, I want to thank myself too because, although there were many times when I did not believe in it, I managed to get this far and be satisfied.

Abstract

Parkinson's disease (PD) is a neurodegenerative disorder characterized by motor symptoms such as tremors, rigidity and bradykinesia and non-motor symptoms such as mood disorders, cognitive issues, and impulse control disorders (ICDs). ICDs are often associated with dopamine agonists (DAs), although this association is still debated. DAs are prescribed as an alternative or in combination with levodopa, to reduce levodopa-induced complications such as dyskinesia. Deep brain stimulation (DBS) proves to be a viable alternative: although it is invasive, being localized, it provides a more targeted effect. Moreover, some studies have shown that DBS has an improvement effect not only on motor symptoms but also on impulsive behavior. However, other studies have found no positive impact or even the appearance of ICD after surgery (de novo ICD). This study tries to shed light on this issue by analyzing the effect of DBS in PD patients with ICD. To realize this purpose, the clinical and neural features of 24 PD patients from Careggi Hospital (Florence, Italy) have been analyzed. 12 of them were diagnosed with ICD, and all were treated with STN-DBS. Statistical analyses were conducted to compare patients, those who recovered from ICD after surgery (improved group) and other ICD+ patients (stable group), before, during and after surgery (1-month and 1-year follow-ups). The results obtained were exploited to cluster patients' profiles before surgery and to classify patients with neural features (during surgery). Analysis before surgery pointed out that the improved group displayed significantly lower values in UPDRS III off (medication off), UPDRS III off-on (difference between off and on) and in bradykinesia subscore, and that clustering with high effect size features (Hoehn Yahr off and rigidity subscore together with the previous ones) allowed to differentiate the two groups with an accuracy of 83%. The significant difference in UPDRS III off remained the same at the 1-year follow-up. Another interesting finding, in contrast to many studies, was the lack of a relationship between dopamine replacement therapy (DRT) and recovery from ICD. After 1 month, DRT was decreased due to the positive effect of DBS, but this decrease was not significantly different between stable and improved groups, nor a DRT difference was found at the baseline condition. When data were available, comparisons between ICD+ and ICD- were conducted, and a significant difference in UPDRS III off and in the Barratt impulsiveness Scale (BIS) was found before surgery. Among neural features retrieved through microelectrode recordings (MERs) during surgery, beta oscillation amplitude, theta oscillation frequency and interburst interval (IBI) differed significantly between improved and stable groups. Beta oscillation amplitude was lower in the improved group, an evidence that was confirmed by literature since a greater beta activity is considered as a pathological oscillation. IBI and theta oscillation frequency showed significantly higher values in the improved group. Beta oscillation amplitude and IBI were selected to train a classifier to recognize individual improved or stable neurons and subsequently, by majority voting, to define the patient's class. The prediction of patients' conditions reached an accuracy value of 83%, as did clustering performance. This may have clinical relevance, as it allows to conduct an analysis of the patient's profile and future outcomes before surgery, better assessing the procedure to be followed.

Contents

List of Figures

List of Tables

1 | Introduction

1.1 Basal ganglia

1.1.1 Basal ganglia structures

The basal ganglia are situated in the forebrain circuitry and are involved in the selection of actions [1, 2] in order to achieve reward, avoid suffering and minimize dangerous events. Most of these behaviors are automatic, and often, even if awareness is present, it is achieved only once an action is completed [3]. In the past, the basal ganglia were considered part of the extrapyramidal motor system, along with the cerebellum. Together with the pyramidal motor system, they connect neurons in the motor cortices to motor neurons in the spinal cord. For this reason, diseases of the basal ganglia are classified as "extrapyramidal movement disorders" [3]. They can range from those characterized by hypokinesia, such as Parkinson's disease, to those characterized by hyperkinesia, such as chorea and dystonia [4]. Before analyzing brain disorders related to the basal ganglia, their anatomy, nuclei and structures will be described and shown in Figure 1.1 [3]:

- Striatum, composed of:
  - Caudate nucleus
  - Putamen, which, together with the caudate, forms the dorsal striatum
  - Nucleus accumbens (NAc), which is the main structure of the ventral striatum, and is not shown in Figure 1.1.
- Globus pallidus (GP), divided into:
  - External globus pallidus (GPe)
  - Internal globus pallidus (GPi)
  - Ventral pallidum (VP), which is not depicted in Figure 1.1.
- Substantia nigra (SN), composed of:
  - Substantia nigra pars reticulata (SNr)
  - Substantia nigra pars compacta (SNc)
  - Ventral tegmental area (VTA), which is not shown in Figure 1.1.
- Subthalamic nucleus (STN)

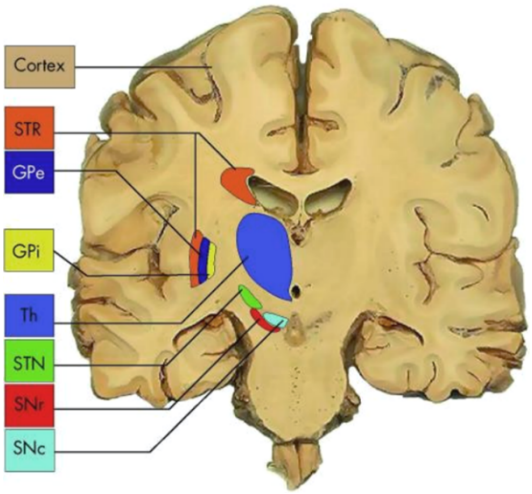


Figure 1.1: A coronal section showing structures of basal ganglia [5].

Two other structures, the cortex and the thalamus need to be considered, because of their strong connectivity with the basal ganglia. In fact, the striatum receives excitatory afferents from the cerebral cortex and the thalamus. Each region of the cortex projects to a specific region of the striatum. For example, the frontal lobe projects to the caudate head and the putamen, the parietal and occipital lobes project to the caudate body and the temporal lobe projects to the caudate tail. Moreover, the NAc receives projections from the limbic cortex [6]. The STN receives input from the ipsilateral motor and the somatosensory cortices. The GPi and the SNr are the output nuclei of the basal ganglia and project to the thalamus, which is connected to the frontal cortex [3].

1.1.2 Direct and indirect pathways

Three major dopaminergic pathways involve the basal ganglia: the mesolimbic pathway, in which the projections from the VTA reach the NAc and which plays a central role in reward and pleasure mechanisms [7]; the mesocortical pathway, which has neuron bodies situated in the VTA and has axons' projections to the frontal and prefrontal cortex, and it is involved in executive functions [8]; the nigrostriatal pathway, which leads dopaminergic neurons from the SNc to the caudate nucleus and putamen (dorsal striatum) [9]. Dopaminergic neurons of the nigrostriatal pathway mainly project to GABAergic spinal neurons (SPNs) of the striatum, which constitute about 90% of the total population of striatal neurons. SPNs can be divided into two populations, depending on their axonal projections: the direct pathway and the indirect one. The former has an excitatory effect, whereas the latter leads to inhibition activity [10]. Before delving into these mechanisms, it is important to dwell on dopamine (DA) receptors: D1 and D2 are the majority of the five DA receptors in the striatum. In particular, D1 receptors are expressed by the direct pathway, whereas D2 receptors are expressed by the indirect pathway [10]. Many models exist to explain the direct and indirect pathways, including the D1/D2 direct/indirect pathways box and arrow model, shown in Figure 1.2. The underlying theory states that [3]:

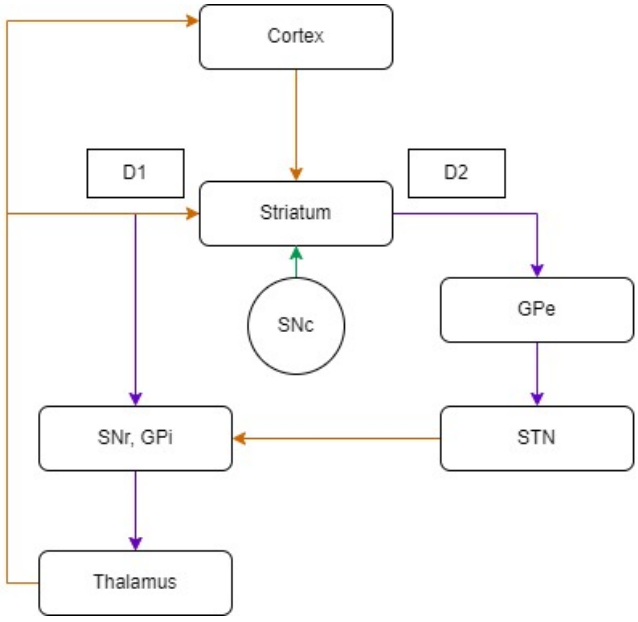


Figure 1.2: The D1/D2 direct/indirect pathways box and arrow model.

- Each box represents neurons in the specific structure.
- Arrows reveal the projections of neurons that can excite or inhibit the target structure, specifically:
  - Orange arrows represent the excitatory action of glutamate.
  - Purple arrows represent GABA, an inhibitory neurotransmitter.
  - The green arrow depicts dopamine neurotransmitters, which can have both functions depending on the postsynaptic receptor in the striatum: inhibitory with D2 dopamine receptors and excitatory with D1 receptors.

The direct pathway has the net effect of exciting the thalamus, which in turn excites cortical neurons. The indirect pathway, instead, inhibits the thalamic neurons, which become unable to excite the cortex. The abnormal functioning of this mechanism is due to an incorrect balance between these two pathways [10]. For example, an excess of dopamine leads to an overstimulation of the direct pathway and causes hyperkinetic motor symptoms, whereas a lack of dopamine leads to overstimulation of the indirect pathway and, consequently, hypokinetic motor symptoms [3]. Other box and arrow models have been designed to explain better the complex connectivity of the basal ganglia:

- The three-layer model in which the striatum and the STN are target structures of the cortex and thalamus: the striatum and the STN project to the GPe and GPi/SNr. The GPe projects back to the basal ganglia input (the striatum and the STN) and forward to output structures.

- The action-selection model states that the direct pathway facilitates the execution of a selected action that is suitable for the present task, whereas the indirect pathway simultaneously inhibits the execution of other competitive actions. Therefore, an excessive amount of dopamine leads to multiple action selection and hyperkinetic disturbances, whereas dopamine depletion leads to an inability to focus and no action selection.

The objective of the models described above is a static regulation of motor vigor. Therefore, different types of models need to be introduced to describe the behavioral and fast dynamics of motor learning. One of them is based on the theory of reinforcement learning, shown in Figure 1.3. The large rectangle represents the agent interacting with the world, receiving negative or positive rewards and facing a new state. It is composed of the actor and the critic: the actor connects possible states and actions, based on the current behavioral policy; the critic calculates the prediction error (PE), the difference between the predicted and the actual value of the previous state. The critic uses the PE to modify the estimate of the state value (which will represent the predicted state for the next cycle). If the PE is positive, the association between the previous state-to-action is reinforced (hence reinforcement learning), whereas if it is negative, the association is decreased, and if it is equal, the policy is not changed. The main axis of the basal ganglia acts as the actor, and SNc dopaminergic neurons as the critic [3].

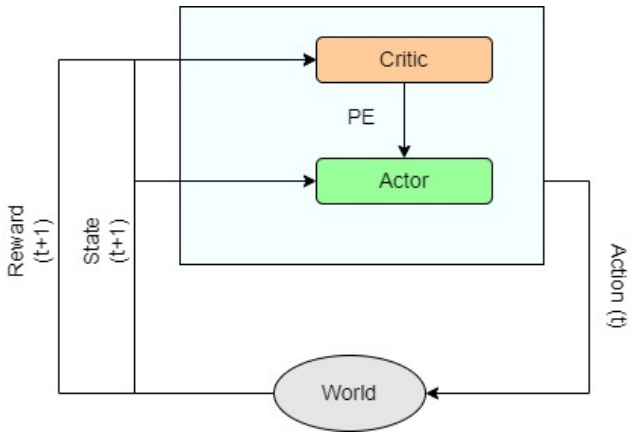


Figure 1.3: Reinforcement learning.

A more complex, but biologically more complete model is the multiple-critics and multi-objective optimization model, in which the basal ganglia doesn't have dopaminergic neurons as the only critic but also another three neuromodulators: acetylcholine, serotonin and histamine. An overview of the model is shown in Figure 1.4.

Broadening the focus used to study the basal ganglia, it is clear that they are only a part of the brain, and therefore it is also important to analyze the other structures that constitute the state-to-action loop. The corticocortical network begins with afferent neurons that project their axons from the sensory organs to the primary sensory cortex, which is connected to the motor cortices. The motor cortices, in turn, have efferent neurons that reach the spinal cord and, finally, the muscles that perform the movement.

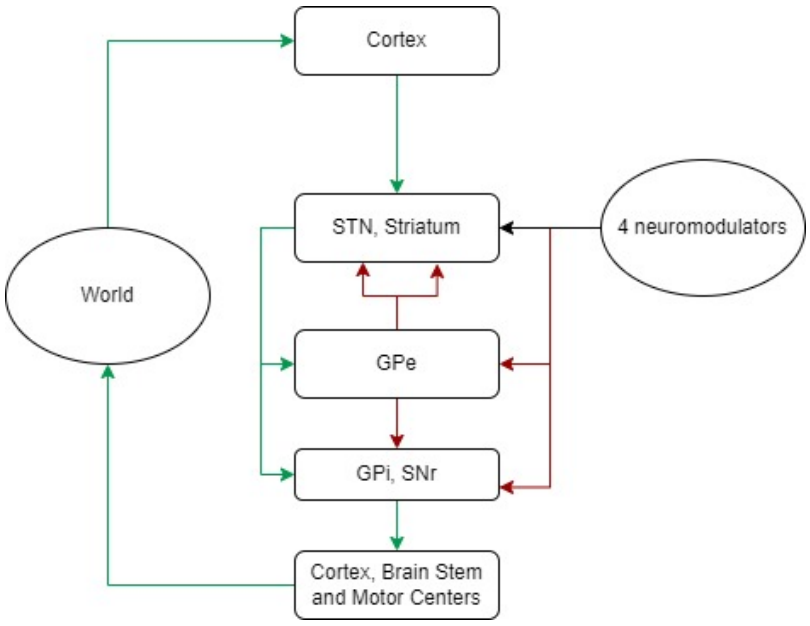


Figure 1.4: Multiple critics with a three-layer box and arrow model.

In 2011, David Kahneman claimed that two different brain systems are present in the state-to-action loop [11]:

- System 1: automatic, subconscious network composed of the basal ganglia and the cerebellum.

- System 2: logical, calculating, conscious network, that is, the corticocortical network.

As mentioned above, the basal ganglia are implicated in a wide variety of disorders, for which the altered regulation of DA in the direct and indirect pathways seems to be determinant [10]. PD studies show that loss of DA drives the excitability of the direct and indirect pathways in the opposite direction, thus creating a net effect of movement suppression [12]. Another study reveals that loss of DA alters the establishment of long-term plasticity induction in glutamatergic synapses [13]. In direct-pathway SPNs, loss of D1 receptors orients glutamatergic synapses toward long-term depression (LTD). In contrast, in indirect-pathway SPNs, loss of D2 receptor signalling promotes induction of long-term potentiation (LTP) [10].

One question that might arise is why other structures do not compensate for basal ganglia disorders, and one possible answer is that the basal ganglia failed to transfer the command to Kahneman's system 2 [3].

1.2 Parkinson's disease

Parkinson's disease (PD) is the second most common neurodegenerative disorder after Alzheimer, described for the first time in 1817 by James Parkinson, in his publication "Essay on the Shaking Palsy" [14]. PD was originally described as a movement disorder with three cardinal motor symptoms: tremors, rigidity and bradykinesia. Over the year, postural instability was considered the fourth motor symptom, and non-motor symptoms (NMSs) were included in the clinical picture [15, 16]. Among the heterogeneity, different motor subtypes have been defined, in order to associate them with aetiological aspects and a different response to treatment: tremor-dominant Parkinson's disease, with a lower presence of other symptoms and a slower disease progression and non-tremor-dominant Parkinson's disease, with a higher functional disability [17].

1.2.1 Epidemiology and risk factors

An estimated 7 to 10 million people are affected by PD worldwide [18], with a prevalence in Europe, North America and South America [19]. Age is a risk factor, with a peak after 80 years old: in industrialized countries, PD affects 0.3% of the total population, 1% of people over 60 years of age, and 3% of people over 80 years of age [20]. Male gender represents a risk factor: men are 1.5 times more likely to develop the disease than women [18]. One of the main risk factors that increase the likelihood of two- or three-fold is a first-degree relative with a history of PD [21]. Several genes have been identified as causing familial forms of PD, and their study can help uncover the pathogenesis of PD [22]. Exposure to certain environmental substances has been associated with increased risk of PD, such as certain pesticides, herbicides or insecticides [23]. Conversely, smoking cigarettes, drinking coffee, and practicing physical activity, are inversely related to PD risk [24].

1.2.2 Pathology

PD originated from a loss of dopaminergic neurons collocated in the substantia nigra pars compacta (SNc), but the disorder reaches other brain regions and also affects non-dopaminergic neurons [24]. A pathologic hallmark of PD is the presence of Lewy Bodies (LBs), composed of more than 90 proteins [25], their main components are alpha-synuclein and ubiquitin [26]. Alpha-synuclein, due to a genetic mutation, becomes insoluble and self-aggregate in LBs [18]. By the time motor symptoms occur, it is estimated that 50-70 percent of the neuron in the SNc have already died [27]. This is one of the studied causes of neuronal death. The other four main ones are:

- Autophagy dysfunction. Autophagy is a mechanism by which a cell's interior components are broken down and recycled. Since autophagy has been proven to contribute to brain functioning, disruption of the autophagy system can result in a variety of illnesses, including Parkinson's disease [28].

- Mitochondrial dysfunction. The hypothesis is that PINK1, a protein normally transported into the mitochondrion, accumulates on the surface of impaired mitochondria and then recruits the Parkin complex, which begins to break down defective mitochondria. In Parkinson's disease, this mechanism of rupture may be flawed: the genes coding PINK1 and Parkin are thought to be mutated, therefore causing mitochondrial dysfunction and cell death will ultimately occur [29, 30].

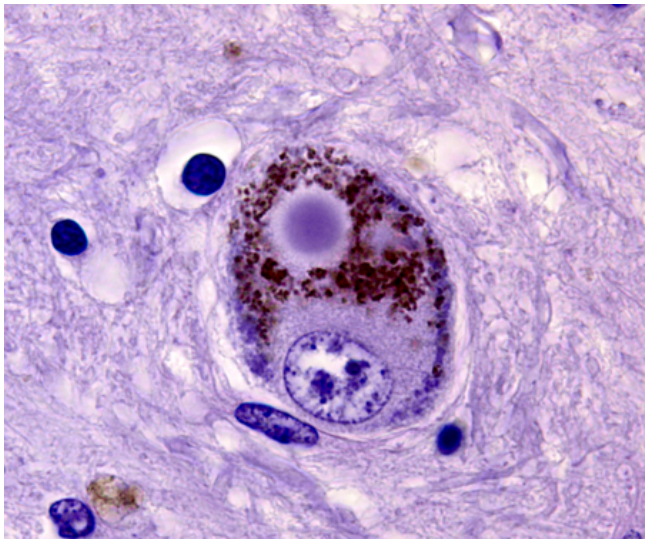


Figure 1.5: Lewy body [34].

- Neuroinflammation, which mostly involves microglia. Microglia are the immune cells of the central nervous system: they can change their morphology in case of neural injury. Usually, microglia are in a resting state, but in Parkinson's disease, alpha-synuclein aggregates can activate them. In this state microglia secrete pro-inflammatory factors, which can cause neurons' death, leading in turn to the activation of microglia. Thereby a positive loop is established that leads to the death of more and more neurons [31].
- Breakdown of the blood-brain barrier (BBB). The function of the BBB can be disrupted by protein aggregates which alter the function of cellular receptors in the BBB [32, 33].

1.2.3 Clinical features

Parkinsonian patients suffer from both motor and non-motor symptoms. Motor symptoms were described first since Parkinson's was originally discovered as a movement disorder. The three main motor features are tremor, bradykinesia, and rigidity, followed by postural instability [21]. Other features of the common PD form are flexed posture and freezing of gait [35]. Table 1.1 shows the motor and non-motor symptoms of PD and the main ones are described below.

Bradykinesia is the slowness of the movement, reflecting the slowness of the signals in the brain [36]. It is a typical feature of basal ganglia disorders, with great difficulty in initiating and executing the movement. The difficulty is less in the presence of external triggers, such as auditory or visual cues, and depends on the patient's emotional state. An example is the phenomenon of kinesia paradoxica, whereby a patient would be able to run if faced with danger, which suggests that patients keep motor programs intact [35]. Bradykinesia is the feature of PD that seems to correlate best with dopamine decrease [37]. The deficit, which causes a decrease in force at the onset of movement, is localized in the putamen and globus pallidus [38]. One of the first manifestations

Motor symptoms	Non-motor symptoms
Tremor, bradykinesia, rigidity,postural instability	Cognitive impairment, bradyphrenia, tip-of-the-tongue (word finding) phenomenon
Hypomimia, dysarthria, dysphagia, sialorrhoea	Depression, apathy, anhedonia, fatigue, other behavioral and psychiatric problems
Decreased arm swing, shuffling gait, festination difficulty arising from a chair, turning in bad	Sensory symptoms: anosmia, ageusia, pain (shoulder, back), paresthesias
Micrographia, cutting food, feeding, hygiene, slow activities of daily living	Dysautonomia (orthostatic hypotension, constipation, urinary and sexual dysfunction, abnormal sweating, seborrhoea), weight loss
Glabellar reflex, blepharospasm, dystonia, striatal deformity, scoliosis, camptocormia	Sleep disorders (REM behavior disorder, vivid dreams, daytime drowsiness, sleep fragmentation, restless legs syndrome)

Table 1.1: PD symptoms [35].

is a general slowness in daily activities and a lengthening of reaction time [39]. Other expressions of bradykinesia are dysarthria, hypomimia (loss of facial expression), decreased blinking, difficulty in swallowing, and reduced arm swing in walking [35].

Resting tremor appears in more than 70% of PD patients, causing an evident shaking of the body extremities in a non-movement condition, at a frequency between 4-6 Hz [36]. It can also affect the chin, jaw, lips and legs and usually subsides during sleep or with movement. Many patients suffer also from postural tremor, which is more substantial and debilitating than resting tremor, nevertheless, it is considered a variant of resting tremor, because of the same frequency and responsiveness to dopaminergic treatment. Some studies have shown degeneration of A8 neurons (a subgroup of the midbrain) in PD patients with resting tremors, rather than in those without tremors [35]. Differently from the essential tremor, resting tremor does not affect the head or voice, and postural tremor appears when the patient has assumed a lying horizontal position for a whereas [40].

Rigidity is characterized by improper relaxation of antagonistic muscles, leading to constant contraction even when the patient is tired [36]. Froment's maneuver [41], voluntary movements of the contralateral limb, are used to detect mild rigidity cases. Pain caused by rigidity is one of the most common signs: for example, it affects the shoulders, often in the early stage of the disease [35].

Postural instability is the fourth cardinal motor symptom and consists in the difficulty in maintaining the standing position, either during movement or in a stationary position. This leads to instability during many daily actions such as walking, getting up from a chair, and so on [36], to the point that it is one of the most common causes of falling in patients with PD [42]. It usually appears after other symptoms, with the progress of the disease. Postural instability is evaluated by the pull test, which consists of pulling the patient by the shoulders forward or backward. If the patient takes more than two steps or there is any postural adjustment, postural instability is considered [35].

Freezing is another typical characteristic of PD, albeit is not always present. It is a form of akinesia, a temporary inability to move [35]. Its most common manifestation is whereas walking, in the beginning, at narrow passages, and in other situations where there is some change in direction or there are obstacles. Five types of freezing can be distinguished: start hesitation, turn hesitation, hesitation in tight quarters, destination hesitation and open space hesitation [43].

Secondary motor symptoms may occur, some of them as a consequence of primary motor symp-

toms but equally or even more disabling. One example is primitive reflexes, such as the glabellar reflex or the palmomental reflex. Dysarthria, hypophonia, dysphagia, and sialorrhea are often associated with PD and are due to bulbar dysfunction. Patients with PD may experience neuro-ophthalmological impairments, such as blurred vision, impaired upward gaze, decreased blink rate, visual hallucinations, and oculogyric crises [35]. Other complications may be respiratory, either restrictive or obstructive [44].

Non-motor symptoms are often present in PD and can be divided into autonomic dysfunction, sensory abnormalities, cognitive problems, and sleep disorders. In the former group, symptoms may manifest as, for example, orthostatic hypotension and sweating dysfunction, whereas in the latter group, typical symptoms are olfactory dysfunction, pain, paresthesia, and akathisia. Rapid eye movement sleep behavior disorder, present in a pre-parkinsonian state, and insomnia are common sleep disorders in PD [35]. Among neuropsychiatric disorders, depression affects more than 40% of patients [36]. Other mood disorders such as anxiety, which affects 1/3 of patients, apathy and abulia are also present in the early stages of the disease. Cognitive problems and dementia usually appear at a later stage of the disease [19].

Many patients may develop impulsive behaviors such as binge eating, hypersexuality, pathological gambling, etc., that are called ICDs (impulse control disorders) and will be described in section 1.4. Under pharmacological treatment, certain clinical features develop over time: dyskinesias (involuntary movements); visual (more frequent) and non-visual (auditory, tactile, olfactory) hallucinations; illusion and delusions [19].

Pathological oscillations

The local field potential (LFP) registered in PD patients at rest without any treatment is characterized by an excess of beta activity (13–35 Hz). Brain oscillations are balanced by an alternation between excitation and inhibition, which in PD patients is lost due to decreased dopaminergic excitation at the basal ganglia [45]. Beta oscillations decrease with dopaminergic treatment and deep brain stimulation (DBS)[46, 47]. Beta oscillations are related to motor symptoms, particularly rigidity and bradykinesia [48] whereas their suppression means an improvement in motor symptoms [49]. Low beta activity (13-20 Hz) is considered a pathological oscillation [50], it is higher in patients in the off state (without dopaminergic treatment)[51]. LFP may also be useful in classifying PD patients: akinetic rigid patients show higher beta power with respect to tremor-dominant patients [52]. This may be due to more severe motor symptoms in akinetic rigid patients than in tremor-dominant patients[53]. Another indicator of pathological beta activity is the presence of longer beta bursts, which is associated with disease severity in terms of motor symptoms [54, 55]. The explanation may be in the over-synchronization of neurons in motor circuits. Levodopa treatment and DBS stimulation lead to shorter beta bursts [54], a condition also achieved with movement[56]. Not only do beta bursts shorten during movement, but the entire beta activity is attenuated, and this also occurs with movement imagination or movement observation [57, 58, 59] Another study [60] shows that beta activity in the STN is also related to the termination of the movement.

The other bands studied in the LFP activity of PD patients are theta (4–7 Hz) and alpha (8–12 Hz), often considered together as low-frequency band (LF); gamma and high-frequency bands. The

LF activity of the STN [61, 47] shows an increase with DBS or pharmacological treatments. In the on state (with the effects of dopaminergic treatments) LF has been associated with the presence of dyskinesia and impulse control disorders. In the on state, the STN is characterized by power peaks in the gamma and theta domains [62, 63]. Gamma oscillations are attenuated in PD, as shown in Tan et al. (2013) study [60]: during sustained contraction, PD patients showed, more than the control group, gamma power decreases, in association with reduced grip strength.

1.2.4 Diagnosis and clinical scales

Clinical factors are used to diagnose Parkinson's disease, but there is no reliable test except for pathological confirmation of Lewy bodies at autopsy, which is still considered the gold standard. The most widely accepted criteria come from the UK Queen Square Brain Bank for Neurological Disorders and the U.S. National Institute of Neurological Disorders and Stroke. The first criteria require bradykinesia and one of rigidity, resting tremor, or postural instability. Moreover, other possible causes of these symptoms must be excluded. Finally, associated symptoms, such as unilateral onset, tremor at rest, progression in time, asymmetry of motor symptoms and response to levodopa therapy during disease evolution may be required [64].

The major difficulty lies in differentiating PD from other forms of parkinsonism in the early stages of the disease course, when signs and symptoms overlap with other disorders [35]. The aforementioned reasons explain the difficulty in PD clinimetrics. The oldest scale was introduced by M. Hoehn and M. Yahr [65]. It consists of five stages, representing the disease progression, from the first stage, which has unilateral involvement, to the last one, when the patient is confined to bed or in a wheelchair [66]. As shown in 1.2, Hoehn and Yahr scale was modified by adding stages 1.5 and 2.5 to describe better the intermediate course of the disease [67].

Stage	Hoehn and Yahr Scale	Modified Hoehn and Yahr Scale
1	Unilateral involvement only usually with minimal or no functional disability	Unilateral involvement only
1.5		Unilateral and axial involvment
2	Bilateral or midline involvement without impairment of balance	Bilateral involvement without impairment balance
2.5		Mild bilateral disease with recovery on pull test
3	Bilateral disease: mild to moderate disability with impaired postural reflexes; physically independent	Mild to moderate bilateral disease; some postural; instability; physically independent
4	Severely disabling disease; still able to walk or stand unassisted	Severe disability; still able to walk or stand unassisted
5	Confinement to bed or wheelchair unless aided	Wheelchair bound or bedridden unless aided balance

Table 1.2: Hoehn and Yahr scale and modified Hoehn and Yahr scale [67].

In 1987, was developed the Unified Parkinson's Disease Rating Scale (UPDRS), consisting of six integral parts [68]:

- Part I: mentation, behavior and mood

- Part II: activities of daily life (ADL)
- Part III: motor examination
- Part IV: complications of therapy
- Part V: Hoehn and Yahr scale
- Part VI: Schwab and England ADL scale [69]

The first three areas and part of the fourth are organized on a five-point scale (0–4 points) [66], in which:

- 0 = no involvement
- 1 = detectable disorders
- 2 = moderate disorders
- 3 = considerable disorders
- 4 = no function or severe disorders

UPDRS III (motor examination) consists of the following 14 items: speech, facial expression, tremor at rest, action or postural tremor of hands, rigidity, finger taps, hand movements, rapid alternating movements of hands, leg agility, arising from a chair, posture, gait, postural stability, body bradykinesia, and hypokinesia [68].

In 2008, Goetz et al. published a modified UPDRS, known as MDS-UPDRS, where MDS stands for Movement Disorder Society (MDS). Two limitations of the original UPDRS are the low emphasis on non-motor symptoms and the lack of instructions for standardized application. The MDS-UPDRS consists of 4 subscales [70]:

- Part I: Nonmotor aspects of experiences of daily living
- Part II: Motor aspects of experiences of daily living
- Part III: Motor examination
- Part IV: Motor complications

There are other specific scales employed for motor and functional assessment, motor complications, and non-motor symptoms [71]. They will only be mentioned as they are beyond the scope of this thesis:

- Scales for Outcomes in Parkinson's Disease-Motor (SCOPA-Motor)
- Schwab & England Activities of Daily Living Scale (included in UPDRS)
- Self-Assessment Parkinson's Disease Disabilities Scale (SPDDS)
- The Postural Instability and Gait Difficulty Score (PIGD)
- The freezing of gait questionnaire (FOGQ)
- The nonmotor symptoms questionnaire (NMSQuest)
- Nonmotor Symptoms Scale (NMSS)
- Unified Dyskinesia Rating Scale (UDysRS)
- Wearing-Off Questionnaires (WOQ)

1.3 Treatments

1.3.1 Pharmacological treatments

Nowadays dopamine replacement therapy (DRT) is a good pharmacological treatment for the first 5-10 years of the disease. It is important to underline that it is a symptomatologic treatment: it does not slow disease progression, and for advanced disease no good drug treatment is available. In the beginning, anticholinergic treatment was adopted, but it had only mild effects and it is now mainly used in tremor-dominant patients [3].

The discovery of the levodopa (L-dopa) as the actual gold-standard pharmacological treatment began with Arvind Carlson who considered that dopamine might have been a neurotransmitter and not only the precursor to adrenaline and noradrenaline. Today it is known that dopamine is part of the family of neurotransmitters called catecholamine, and its synthesis begins with phenylalanine, an amino acid. Phenylalanine is converted to l-tyrosine and l-tyrosine is then converted to L-dopa. Finally, L-dopa is converted to dopamine, and when necessary, dopamine is converted to adrenaline and noradrenaline. With some experiments conducted in 1957, Arvind Carlson showed that in rabbits reserpine led to dopamine depletion, whereas dopa awakened them from immobility [72]. Oleh Hornykiewicz discovered that Parkinsonian patients have dopamine depletion in the striatum [73] and George Cotzias used high dopa dosage in patients with PD, and observed, notwithstanding the presence of nausea and vomiting, a significant improvement in symptoms [74]. The side effects of nausea and vomiting are caused by dopamine activation in the area postrema, which is collocated outside the blood-brain barrier (BBB). Therefore, to avoid these side effects, L-dopa is prescribed with carbidopa, an inhibitor of the enzyme that converts L-dopa to dopamine. Considering that carbidopa does not pass through the BBB, carbidopa inhibits the conversion of L-dopa to dopamine only outside the BBB, thus preventing vomiting but still allowing the release of dopamine into the striatum [3].

Until the first 10 years of the disease, L-dopa is buffered in the dopaminergic terminals and released as needed. As the disease progresses this buffer deteriorates, leading to rapid fluctuations in dopamine levels: dopamine is released freely causing a peak in the striatum followed by a decline until the next pill. These fluctuations probably induced abnormal dopamine receptor sensitivity, and after 5-10 years, many Parkinsonian patients begin to suffer from on-off fluctuations and levodopa-induced dyskinesias. Nowadays amantadine is used to reduce dyskinesias and it is prescribed together with levodopa. Acting as an NMDA (N-methyl-D-aspartate) antagonist, it can achieve the same purpose as the STN inactivation: it can reduce glutamate excitation from the STN to the GPi/SNr and thus lead to symptom improvement.

Postsynaptic dopamine agonists are prescribed as an alternative or in combination with levodopa to reduce complications such as dyskinesias, although they are less effective in controlling motor symptoms, particularly rigidity and bradykinesia [26, 36]. Unlike levodopa, they act directly on dopamine receptors in the striatum. For this reason, they show a great advantage with respect to levodopa: their effects are more long-lasting and dopamine fluctuations are smaller. Conversely, they produce side effects, like hallucinations and dopamine dysregulation disorders [3].

As mentioned at the beginning of this section, pharmacological treatments do not slow disease progression, and many patients may develop PDD (Parkinson's disease dementia). One reason may be that neuromodulators used in treatments (such as dopamine) are widespread in many structures and pathways: this results in non-selective treatment and the difficulty of adapting therapy without

causing imbalances and side effects. Deep brain stimulation (DBS) could be an option because, although it is invasive, it is localized: it might overcome the issues of pharmacotherapy [3].

1.3.2 Deep brain stimulation (DBS)

Before the introduction of levodopa as the gold standard treatment in 1960, thalamotomy was the most popular surgical method, used to treat tremors. Tremors and rigidity, indeed, originated in the 'extrapyramidal system', composed of the globus pallidus (GP) and the ventrolateral motor thalamus [75]. DBS was deployed for the first time in psychiatric illness and pain. However, DBS for pain was never approved by the FDA (Food and Drug Administration) for clinical use [76]. Neurosurgeon Lawrence Pool, in 1948, implanted an electrode in a 60-year-old woman who suffered from depression and had an advanced form of Parkinson's disease. The results were positive, and stimulation was interrupted only because a wire ruptured [77]. Initially, the aim was to stimulate brain structures in order to identify the target for ablation. Sem-Jacobsen extended the treatment from patients with psychiatric disorders to PD patients [78].

The introduction of stimulation for therapeutic purposes became possible through the development of implantable pacemakers. If it was possible to stimulate the heart continuously, it may also have been possible to stimulate the brain in the same way, and the term "deep" was coined to distinguish this stimulation of subcortical areas (basal ganglia structures) from cortical stimulation of the brain [75]. The hardware consisted of an electrode and an extension cable, whereas the battery, which charged the electrode through radiofrequency, was external and maintained by the patient [76]. It was a group from Grenoble that brought DBS to the modern era in 1987: neurosurgeon Benabid and neurologist Pollak applied a high-frequency stimulation (130 Hz) at VIM of the thalamus (ventral intermediate nucleus) to treat tremor and saw that tremor ceased as soon as stimulation occurred. As long as the current was injected the tremor disappeared, thus they decided to implant the electrode permanently in lieu of proceeding with thalamotomy [79]. Since that time, many technical innovations were introduced, and DBS was tested to treat various disorders in different brain structures, some of which were approved by the FDA. In Figure 1.6 below are shown the major innovations (with red background) and the FDA approvals (with orange background).

DBS technology

Modern DBS hardware is composed of cylindrical leads of 1.27 mm in diameter, with 4-8 contacts of 1.5 mm each. The contacts are separated by 0.5 or 1.5 mm (Figure 1.3.2). The actual DBS electrodes are wires in platinum-iridium and nickel alloy connectors embedded in a polyurethane sheath[76]. The electrodes are connected through cables under the skin to an internal pulse generator (IPG), which delivers stimulation to the target brain structure. It is programmed by an external computer [75], and the parameters that can be programmed are the following [3]:

- Geometry: monopolar stimulation or bipolar stimulation. The first is between one contact (the cathode) and the IPG case (the anode). The second is between two contacts and, although less powerful, it is more targeted, preventing surrounding structures from being involved in the stimulation.

- The stimulation can be set as current or voltage stimulation. Typical values are 0.25 to 6 V and 0.25 to 6 mA; this is true if the contact impedance is 1 $k\Omega$; however, if, for example, the impedance is lower than 1 $k\Omega$, the current is expected to be at a higher range.

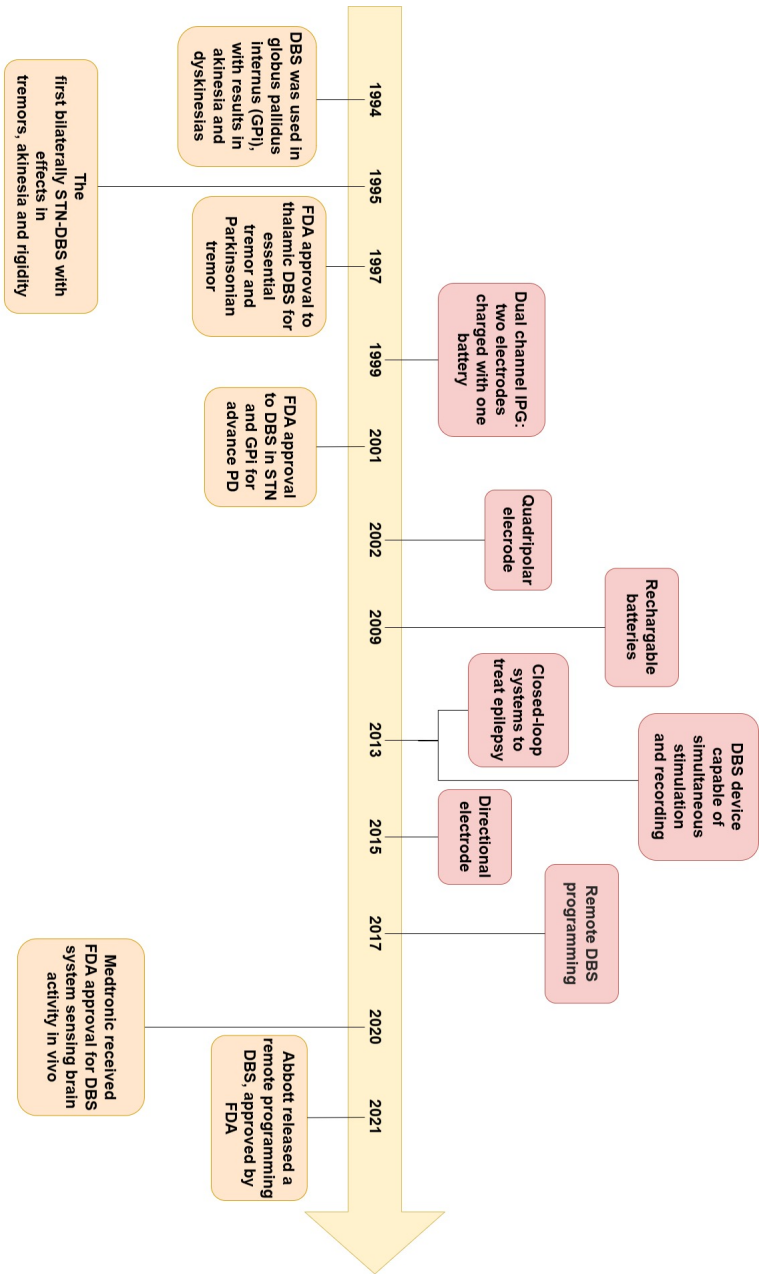
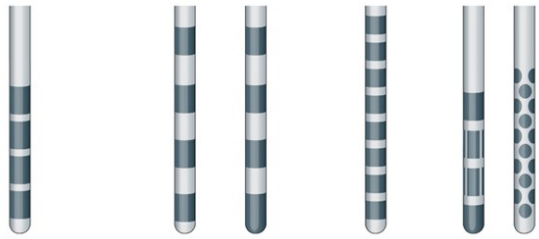
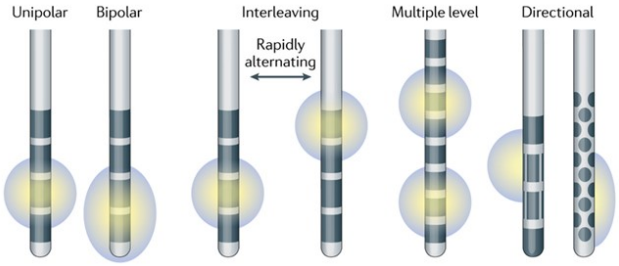


Figure 1.6: DBS history [76, 80, 81, 82, 83, 84, 85].



(a) Common DBS electrode configurations



(b) Types of stimulation

Figure 1.7: DBS configurations [76].

- Frequency is chosen between 80 and 180 Hz; typical values are in the middle, around 130 Hz.

- Pulse width typical range is between 60 and 120 μs, but higher pulse widths, up to 180 μs, can be adopted.

In addition to monopolar and bipolar configurations, there are other more complex configurations, shown in Figure 1.3.2, such as: interleaving, multiple level, and directional. The advantage of the last configuration is the stimulation field on the horizontal plane or in a particular direction, in order to minimize the stimulation in non-interested areas and thus minimize side effects [76]. The volume of tissue activated (VTA) is a modeling technique used to evaluate brain areas that may be stimulated and led to the development of directional stimulation (current steering) mentioned above [80].

The efficacy and adverse effects depend on the location of the electrodes and the correctness of their placement in the target structure. To find the correct target for DBS electrodes, stereotaxic and image guidance are adopted during surgery. In addition, microelectrode recordings (MERs) or anatomical atlas, imagining sequences, and connectomics (map of the neural connections within the brain) are applied as refining methods [80]. If the electrodes are correctly placed in the STN, a reduction in rigidity and bradykinesia may be visible with a low amplitude of stimulation, whereas tremor arrest requires a longer stimulation time. Incrementing the stimulation amplitude, stimulation-induced dyskinesia may appear: this indicates good positioning and will not result in future side effects. If the position is not correct, for example, if it is too lateral, near the internal capsule, it can cause dysarthria [3].

Neural mechanism

The internal mechanism by which the DBS works is still a matter of debate; the net effect is the mimics of ablation [75]. One of the first hypotheses was the depolarization block: a continuous depolarization provided by DBS inactivates sodium channels, preventing the generation of subsequent action potentials [3]. Another hypothesis is called "information jam" and claims that the transmission of pathological information is disrupted by high-frequency stimulation induced by DBS [86]. DBS can lead to orthodromic activations of afferent and efferent neurons in the target structure and to antidromic activation of the afferent neurons to the structure, thereby blocking intrinsic action potentials traveling orthodromically. This results in the replacement of the pathological frequency pattern of the parkinsonian activity with a stimulation-induced high-frequency pattern. Moreover, DBS can lead to synaptic failure: most synapsis would be depleted of their neurotransmitter molecules and vesicles, due to prolonged high-frequency stimulation. Inactivation of the basal ganglia, which previously acted as the default system, allows other systems (corticocortical network) to replace them and compensate for abnormal activity [3].

DBS target

The two most common targets for DBS in PD patients are the STN and the GPi [87]. The GPi receives glutamatergic afferent neurons only from the STN, therefore it is possible to inactivate the STN or directly the GPi. The STN is considered a better target, and it is more common in most centers, although it is riskier. Therefore, the GPi is used mostly when the criteria for excluding the STN are met [3]. Moreover, STN-DBS has the best results in terms of improvement in PD symptoms and allows for lower doses of dopaminergic drugs. On the other hand, it has strict eligibility criteria. The first is responsiveness to levodopa; the second is age: in some centers, the age limit is 69-70 years old, because of reduced improvements and increased side effects; the third is to not have any cognition disorder because it has been found that STN-DBS could worsen them; and the last is to not have mood disorders, because they could worsen in postoperative [75].

Local field potential (LFP)

DBS has enabled electrophysiological findings on the oscillation frequency of neurons in various structures, in both physiological and pathological states. In fact, the DBS electrode can record the subcortical electrophysiological activity, allowing a deeper insight than EEG, ECoG, or MEG. The activity recorded by DBS is the local field potential (LFP) around the electrode, which is the electric potential in the extracellular space originated by neurons near the electrode. There are 3 different ways to record LFP with a DBS electrode, depending on the time window. During surgery, microelectrode recordings are used to optimize DBS localization and recorded activities from the target and surrounding structures. Secondly, LFP can be recorded for some days after surgery, before connecting the leads to the IPG. Finally, in the recent generations of DBS, it is possible to record LFP during treatments. The most recent device with chronic sensing is Medtronic's FDA-approved device [88].

Telemedicine, automated and closed-loop DBS

This paragraph briefly describes recent innovations and developments to give an idea of the current state of the art and future directions. Remote DBS programming has many advantages,

such as a reduction of costs and time spent traveling, more frequent adjustment and monitoring that can avoid some of the stimulation side effects [80]. One of the major difficulties is quantifying symptoms, an issue that has seen a first resolution during the COVID-19 pandemic, with the spread of telemedicine and especially the introduction of new technologies or symptom quantification such as wearable sensors [89] and advanced video recognition software [90]. In 2021, FDA approved Abbott's technology for remote DBS programming [91]. Automated programming relies on objective measurement of symptoms and automated treatment therapy [92]. In addition to wearable sensors, LFP is beginning to be adopted as a biomarker of disease states [93]. LFP recorded by DBS can be used also in closed-loop DBS. In particular, the ratio between the power of the LF (alpha and theta) and the beta band can be exploited to adjust stimulation in real-time: in an off state, the major power is concentrated in the beta band, whereas if stimulation is on, the power shifts to the LF band. [94]. In 2020, FDA approved Medtronic Percept PC, a system with chronic sensing during daily brain activities [95].

Clinical effects

Deep brain stimulation is widely regarded as the most successful surgical treatment for Parkinson's disease motor symptoms: multiple clinical trials [96] discovered that both STN-DBS and GPi-DBS are efficacious in mild and advanced Parkinson's disease and in a randomized controlled trial STN-DBS [97]. STN-DBS provided evidence of better results than pharmacological treatment. Furthermore, STN-DBS allows patients with advanced Parkinson's disease to reduce or even discontinue dopaminergic treatment [75]. The reduction of dopamine treatment due to STN-DBS may be from 30 to 50 percent of the preoperative dose [3]. In addition to motor improvements, DBS can lead to improvement in impulse control disorder, whereas one of the side effects is the development or worsening of apathy [98, 97]. The study of DBS effects on impulse control disorders (ICDs) is still debated and controversial, but before delving into DBS effects on ICDs, impulse control disorders in PD will be described in the next section.

1.4 Impulse control disorders (ICDs)

Impulse control disorders (ICDs) are characterized by excessive behavior or damage towards oneself or others and impair normal functioning in various areas of life [99]. The criteria through which classify ICDs are difficult to define: the DSM-IV (Diagnostic and Statistical Manual of Mental Disorders) classifies pathological gambling, kleptomania, trichotillomania, intermittent explosive disorder, and pyromania as ICDs, whereas pathological skin picking, compulsive sexual behavior, and compulsive buying, are considered ICDs unless otherwise specified [100]. The difficulty in classification lies in the wide variety of ICD types and the partial overlap within them. The following common characteristics can be identified: repetition of behaviors leading to dangerous consequences; lack of control of the impulse behavior; an attractive stimulus before the problematic attitude; the continuation of the pleasant state during the application of the harmful behavior [101]. The description of these behaviors is very similar to those assumed in drug addiction, thus the DSM-V (the latest version of the DSM) classifies gambling disorder as "Substance-Related and Addictive Disorders" [102]. Patients with PD seem to have an increased likelihood of developing ICDs, particularly the following: gambling, buying, sexual, and eating impulsive disorders[103]. Dopamine dysregulation syndrome (DDS) is another type of ICD affecting PD patients and regards overuse of dopaminergic drugs, analogous to drug addiction [104]. Other behaviors such as punding, hobbyism,

walkabouts, and hoarding are present in PD patients although, except for hoarding disorder, are not classified as ICDs, and it is still unclear whether neural mechanisms are similar. As mentioned before, compulsive buying or sexual behaviors are not present in DSM-V. However, the four major impulsive and compulsive behaviors present in PD patients (gambling, buying, sexual, and eating) are commonly considered ICDs [105]. There are few studies on the prevalence rates of ICDs worldwide, which seem to vary between nations and cultures. Some studies find that ICDs were more present in PD patients than control group [106, 107] but another study was of the opposite advise [108]. Due to multiple factors, such as lack of routine screening and general patient embarrassment, ICDs often go unrecognized and untreated properly. Despite this, numerous screening tests exist, such as BIS (Barratt Impulsiveness Scale), QUIP (Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease) [109], the Movement Disorder Society–Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [110], the Structured Clinical Interview for Obsessive-Compulsive Spectrum Disorders (SCID-OCSD) [111], the Ardouin scale [112] and others.

1.4.1 The Barratt Impulsiveness Scale (BIS)

Among these, the Barratt Impulsiveness Scale Version 11 [113] will be briefly described since it is the impulsivity scale used in this study. The original idea behind this scale was to create a unidimensional measure, not closely related to specific self-reported symptoms [114]. In the BIS-10, however, the scale items can be grouped into three main subtypes of impulsivity: motor, attentional and non-planning [115]. The BIS 11 consists of 30 items with a rating range of 1 to 4: 1) rarely/never, 2) occasionally, 3) often, and 4) almost always. It was derived from psychometric analyses of the 34-item BIS-10 [113], 4 of which were not considered due to poor psychometric properties. Applying principal component analysis (PCA), six correlated first-order components were discovered [116]. The following table shows all 30 items and the subdomains of the first-order factor and the higher-order factor (second-order factor).

1.4.2 Risk factors and other psychiatric symptoms

DRT (dopamine replacement therapy), particularly dopamine agonists (DAs), is believed to play a role in the development of ICDs, although there are limited prospective studies and controversial results. In the DOMINION study, patients on DA treatment were the majority of ICD patients. Both men and women were affected by ICD, although of different types. Compulsive sexual behavior (CSB) was prevalent in males, whereas compulsive shopping and binge eating were more common in women[105]. Psychiatric disorders associated with ICD in PD are multiple: affective and anxiety symptoms, elevated obsessionality, novelty seeking, impulsivity, and sleep disorders; obsessive-compulsive disorder (OCD), which may appear with greater frequency in PD [105]; cyclical mood disorders, which are associated with DDS (dopamine dysregulation syndrome) [118].

Apathy is considered the opposite side of ICD. A study on DBS found that patients had an improvement in ICD symptoms, but apathy aggravated post-DBS [119]. The relationship between ICDs and DBS is still debated: some studies found an improvement in ICD symptoms, other studies found a worsening or no effect and in some cases, patients developed de novo ICD after DBS [120, 121]. The answer to this question, the study of the neural mechanism of DBS, which neural and clinical features are important to consider, and whether it is possible to predict the postoperative outcome of ICD, are all topics of the present study and the background works (Section 1.5).

Item	Item content	First-order factor	Second-order factor
5*	don't pay attention	Attention	Attentional
9*	concentrate easily		
11	squirm at plays or lectures		
20	am steady thinker		
28	am restless at the theater		
6	have racing thoughts	Cognitive instability	
24	change hobbies		
26	have extraneous thoughts		
2*	do things without thinking	Motor	Motor
3	make up my mind quickly		
4	am happy-go-lucky		
17	act on impulse		
19*	act on spur of the moment		
22	buy things on impulse		
25	spend more than earn		
16	change jobs	Perseverance	
21	change residences		
23	think about only one thing		
30	am future-oriented		
10	save regularly	Cognitive complexity	Non-Planning
15	like to think about problems		
18	bored solving problems		
27	interested in present		
29	like puzzles		
1*	plan task carefully	Self-Control	
7	plan trips ahead of time		
8*	am self-controlled		
12*	am a careful thinker		
13	plan for job security		
14*	say things without thinking		

Table 1.3: The BIS-11 [113]. Asterisks indicate items from Brief-BIS [117].

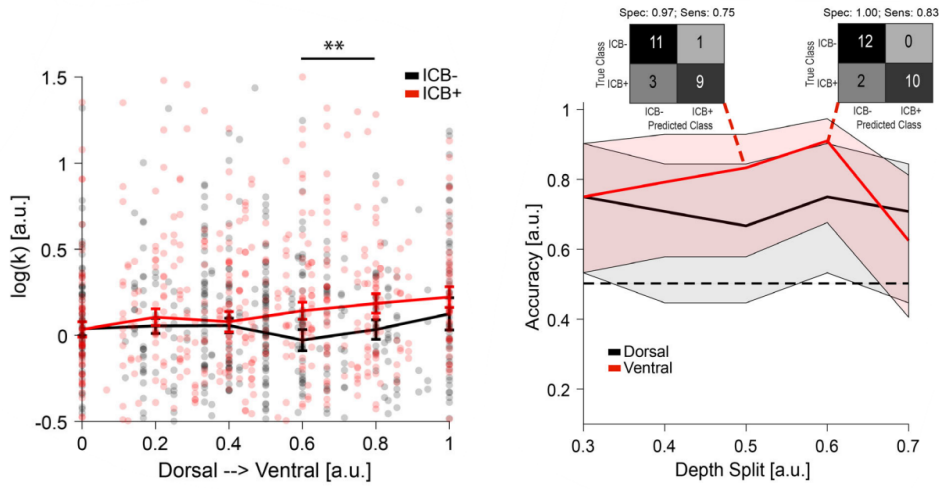
1.5 Background works

This study is the third in a series of studies involving Parkinsonian patients at Careggi Hospital (Florence, Italy). In all these studies, 24 patients were retrospectively selected, 12 of them with ICDs, and all patients were operated on with STN-DBS. In the first study, 742 single-unit activity (SUA) were analyzed from electrophysiological recordings during surgical procedures (described in 2.1). Neural features such as firing rate, fraction of tonic neurons, intraburst frequency (IBF), and all other neural features described in 2.2.1 were extracted from the SUA. SUA-BUA coherence in the theta and beta bands was calculated from the SUA and the background unit activity (BUA), which is the signal without action potentials. The aim was to find out whether certain neural features of the STN were associated with the impulsivity condition. Applying a Support Vector Machine (SVM) classifier, the discriminatory power between patients with ICDs (defined as ICD+) and patients without ICDs (defined as ICD-) of each neural feature was evaluated. IBF (intraburst frequency) and the combination between the fraction of tonic neurons and SUA-BUA beta coherence resulted in the highest accuracy in identifying ICD+ patients. This analysis, conducted in an off state (without medication treatment), shows that patients with ICD have longer intervals between beta bursts and a higher number of tonic neurons in combination with weaker beta coherence [122].

In the second study, ventral and dorsal subthalamic nucleus (STN) neurons were compared to assess which STN region is more involved in impulsivity. The interaction between impulsivity and depth was analyzed through a linear mixed model (LMM), and an SVM classifier was implemented to discriminate between the ICD+ and ICD- groups. The classifier was based on the regularity of firing of ventral and dorsal neurons separately. In ICD+ patients, ventral neurons are more regular than those in ICD- patients, whereas in the dorsal part both patient populations have irregular neurons. This difference was also evident with SVM implementation: classification based on the ventral neurons achieved significant accuracy, as shown in Figure 1.8. The greater regularity of ventral neurons in ICD+ patients may be a possible explanation for the role of dopamine agonists in ICD. Indeed, in patients with a preserved ventral part, dopamine agonists, because of their mechanism of action, could overcompensate and increase the probability of ICD [123].

1.6 Aim of the Master Thesis

This thesis is the third work of those just described. The purpose is to analyze the 12 ICD+ patients knowing that 5 of them improved one month after the intervention (defined as the improved group), as opposed to the other 7 (defined as the stable group), trying to understand whether there were differences already visible before or during the intervention and what the differences were at one-month and one-year follow-ups. To do this, I conducted statistical tests on clinical features before surgery, provided by physicians at the Florence hospital, and on neural features during surgery, obtained in the previous works. I used the results obtained from the pre-surgical analysis to test different clustering methods to assess whether the clinical features before surgery could lead to separability between the two groups. On the other hand, statistical analysis of neural features allowed me to select features to be given as input to a classifier that could label individual neurons as improved or stable and, from this classification, predict the class of patients. I then conducted further statistical analyses to compare the clinical features of the two groups at follow-ups. Finally, I also conducted statistical analyses to compare the ICD+ and ICD- groups before the intervention (defined as baseline condition) and at the 1-year follow-up.



(a) (b)

Figure 1.8: a) Topographical distribution of the neural irregularity based on the parameter log(k) in ICD+ (black) and ICD- patients (red) in normalized relative coordinates (STN entry point: 0 and STN exit point: 1). b) Decoding performance of neural irregularity in discriminating ICD- and ICD+ patients based on various depths, from the dorsal part [0 depth] (black line) to the ventral part [depth 1] (red line) [123].

2 | Methods

2.1 Data acquisition procedure

The description of the data acquisition procedure is intended to explain the background of this study and how data were collected from the previously analyzed studies. This study was conducted in accordance with the Declaration of Helsinki and after ethical approval (Careggi Hospital, Florence, Italy). Among the 24 PD patients considered retrospectively, 12 were diagnosed with at least one ongoing ICD. All patients were treated with bilateral STN-DBS and the surgical procedure was conducted at Careggi Hospital in Florence. Frame-based stereotaxy was applied for the procedure, together with microelectrode recordings (MERs) and macrostimulation of the STN. For STN target localization, T2, SWI and T1 sequences from the preoperative MRI (1.5 T) were fused with the preoperative stereotactic CT. The quadripolar DBS electrodes (Medtronic model 3389) was implanted bilaterally, with local anesthesia. Electrophysiological signals were recorded through a Medtronic lead-point system. The recordings, lasting at least 10 seconds each, began 10 mm above the previously identified target position (by stereotactic imaging) and were made in 0.5-mm steps along three parallel traces (anterior, middle, and lateral) with a lateral distance of 2 mm. Subsequently, under general anesthesia, an implanted pulse generator (IPG) was connected to the DBS electrodes. Finally, postoperative imaging was performed to verify the correct positioning of the electrodes in the STN. [122, 123].

In the previous studies [122, 123], single-unit activity (SUA) and background unit activity (BUA) were analyzed from electrophysiological recordings during surgery, and the data obtained are the neural characteristics described in the following section.

2.2 Dataset

Table 2.1 shows an overview of the patients: before the surgery, 12 of the 24 patients had a diagnosis of at least one ongoing ICD (ICD+) whereas the remaining 12 didn't (ICD-). At the 1-month follow-up, among the 12 patients with ICDs, 7 patients still had them, whereas the other 5 had recovered. In the ICD- group, however, at the 1-year follow-up, one had not tolerated DBS, and another had developed a de novo ICD after the 3 years follow-up.

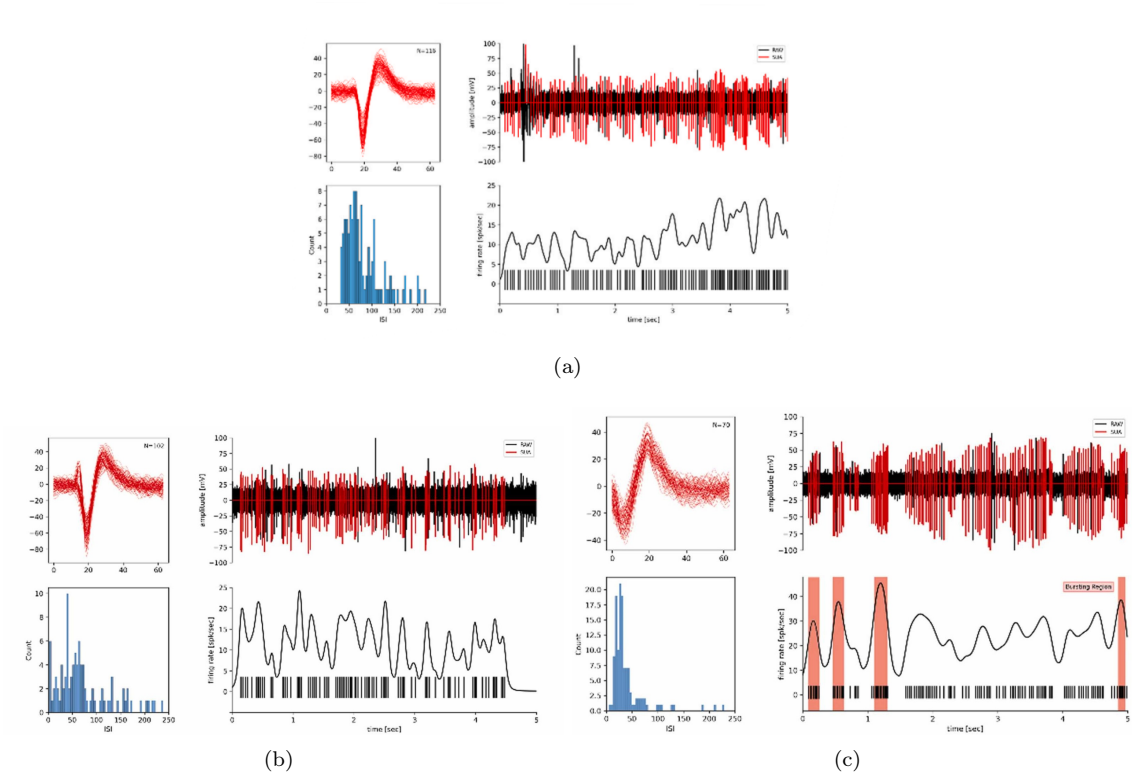
2.2.1 Neural features

A total of 412 neurons from ICD+ patients were provided. The continuous and binary neural features, extracted from the SUA, are as follows:

Timeline	
Before surgery	Follow-up
12 ICD+	7 ICD+
	5 ICD-
12 ICD-	1 ICD+
	10 ICD-
	1 unavailable data

Table 2.1: Dataset.

- Coefficient of variation (CV): the measure of spike train irregularity. It is evaluated to quantify the width of the interspike interval (ISI) distribution.
- Firing rate: the number of spikes/s.
- Firing regularity: the parameter log(k) [66] evaluates the regularity of neurons' spikes and it allows to classify neurons into subcategories (bursting, tonic and irregular).
- Depending on the firing regularity, three binary features are obtained: tonic, bursting and irregular. A value of 1 in a binary feature indicates the presence of that firing feature in the neuron, otherwise the value is set to 0. Thus, the sum of "1s" represents the total number of a certain type of neuron.
- Frequency bands of neural oscillations are the following: delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12–30 Hz), and gamma (30-100 Hz). Only the beta and theta bands were considered in this study because they are the ones of greatest interest in Parkinson's disease. Six features were extracted from them:
  - Two binary features (beta oscillatory and theta oscillatory) indicate whether the neuron is oscillating in the beta and theta band, respectively.
  - Four continuous features: amplitude and frequency of the beta and theta bands.
- Ccbeta and cctheta: cross-coherence between SUA-BUA in the beta and theta bands, respectively.
- Intraburst frequency (IBF): the frequency of the spikes within one burst event, defined only for bursting neurons (Hz).
- Interburst interval (IBI): the duration between two successive bursts (sec).
- Total bursting duration: the duration of one burst event (sec).



(a)

(b)

(c)

Figure 2.1: a) Tonic neuron. b) Irregular neuron. c) Bursting neuron. For all, in the top left is represented the action potential of single-unit activity (SUA); in the top right: BUA (black) and SUA (red) components of MERs; in the bottom left: the interspike-interval (ISI) distribution of spike train; in the bottom right: instantaneous FR (firing rate) and raster plot (bottom). Bars in c) indicate bursts [124].

2.2.2 Clinical features

In addition to neural data, the following clinical characteristics were provided before surgery, at the 1-month and 1-year follow-ups:

- Generic information: age at DBS, PD duration and gender.
- Fenotype: refers to the classification of the patient as tremor dominant (TR) or non-tremor dominant (NTR).
- The Hoehn Yahr and UPDRS III scales, described in 1.2.4, were measured in the off condition (without dopamine treatment) and the on condition. In addition, the difference between the off and on conditions was calculated to measure the effectiveness of medication in improving symptoms.
- LD: levodopa doses.

- DA LEDD: dopamine agonist doses converted in levodopa equivalent daily doses (LEDD).
- LEDD: total levodopa equivalent daily doses (the sum of LD and DA LEDD).
- DA: binary variable where "0" indicates levodopa monotherapy, and "1" indicates treatment with levodopa and dopamine agonists.
- BIS (Barratt impulsiveness scale): the score of each item (31 in total, described in 1.4.1) and the total score.

Not all features were available for all patients at the various stages of the analysis. The Barratt impulsiveness scale (BIS) was provided for all 24 patients, but only before surgery (defined as baseline condition). At the 1-month follow-up only pharmacological treatment information was available (DA, LEDD and DA LEDD) for ICD+ patients (12), whereas at the 1-year follow-up, data were available for 22 patients, 12 ICD+ and 10 ICD- patients. Of the two missing ICD- patients, one did not tolerate DBS, whereas the other became de novo ICD, and it was decided not to consider it for analysis.

2.3 Statistical analysis

The main purpose of this part is to test whether the group that recovered from ICD after DBS (improved group) differs significantly in some characteristics from the group of patients who instead continues to be affected by ICD even postsurgery. Statistical analyses were conducted in chronological order:

- Continuous and categorical clinical features collected before DBS surgery were tested to assess if there was already any difference between healed and non-healed patients, before surgery and before recovery, thus when they were all still suffering from ICD.
- Continuous and categorical neural features during surgery were tested to check whether there were differences in neurons' firing regularity, or frequency of neural oscillations, therefore whether there were already indicative signs to be able to distinguish future improvement or not.
- Continuous pharmacological features at the 1-month follow-up, to check whether there were different drug doses between stable and improved groups in the month of recovery.
- Continuous clinical features at the 1-year follow-up, to test whether there were the same differences remained constant.

Additional tests were conducted to complete the analysis:

- It was tested whether the difference in each feature between follow-ups and baseline conditions was different between the two groups.
- Within-group differences, between before and after surgery, were calculated to test for changes in features over time.
- Clinical features were compared between the ICD+ and ICD- groups, and within each one, before surgery and at the 1-year follow-up, also calculating their difference.

2.3.1 Statistical tests

Two types of tests can be defined in statistical analysis: parametric tests and non-parametric tests. Parametric tests can be applied only if the distribution of test populations is approximately normal. To check the normality distribution, for sample sizes lower than 50, the Shapiro-Wilk test is usually used [125]. In our case, comparisons between the improved and stable groups were conducted directly with non-parametric tests, due to small sample sizes, whereas for comparison between the ICD+ and ICD- groups, the Shapiro-Wilk test was performed first. Since the number of features with normal distribution was low and the sample sizes were small, it was decided to use non-parametric tests to compare the ICD+ and ICD- groups as well.
The tests used are the following:

- Mann-Whitney U test, to test continuous variables between groups (unpaired tests).
- Wilcoxon signed-rank test, to test continuous variables within groups (paired tests).
- Fisher exact test, to test categorical variables between groups.

Each test determines whether to accept or reject the null hypothesis. The null hypothesis assumes that the means of the two groups (unpaired tests) or within the same group, before and after (paired tests), are equal. The alternative hypothesis, instead, assumes that the means are different. To measure the strength of H0 a p-value is calculated: the probability that a given result would occur under the null hypothesis. If the p-value is below the level of significance α (conventionally $\alpha = 0.05$), the H0 hypothesis is rejected.
To deal with the issue of multiple comparisons a p-value correction was applied using the Holm-Bonferroni method, described in 2.3.2. Moreover, the effect size was calculated for each comparison to quantify the magnitude of the difference between the groups, using Coehn's d measure, described in 2.3.2.

Continuous variables

- **Paired: Wilcoxon signed-rank test.** It is a non-parametric version of the paired t-test: Wilcoxon signed-rank test [126] checks whether the distribution of differences of two matched samples is symmetric with respect to zero. It computes the difference between matched data and ranks the absolute values of the differences. The rank of the smallest difference is one, the rank of the next smallest is two, and so on. Then the test sums the ranks of negative and positive differences separately and preserves the absolute value of the smallest, which is the test statistic w_s. Regarding zeros, many methods can be used, in this case, "zsplit" was chosen: it includes zero differences in the ranking process and splits the ranks of zeros between positive and negative ones. At this point, the algorithm, by using a table of critical values, compares w_s with its corresponding critical value, and if it is greater, H0 is accepted.

- **Unpaired: Mann-Whitney U test.** The Mann-Whitney U test [127] is the non-parametric equivalent of the t-test: it tests that two independent samples belong to the same distribution. The test statistic is U, the smaller of:

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \qquad \text{and} \qquad U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

where R_1 and R_2 are the sums of the ranks in the two groups, and n_i is the number of samples. If U is less than or equal to the critical value in the reference table, the null hypothesis is rejected.

Categorical variables

- **Fisher exact test**. It tests the null hypothesis that two categorical variables are independent. It receives as input a 2x2 contingency table, i.e., a table that shows the distribution of one variable in rows and another in columns (two categorical variables). It calculates the probability that this or any other table with more extreme joint values and the same marginal totals (the totals across rows and across columns) would occur under the null hypothesis. If the probability is small, the null hypothesis of unassociated variables is rejected in favor of the alternative hypothesis. [128].

2.3.2 P-value correction and effect size

Holm–Bonferroni method

Holm–Bonferroni method [129] is used to deal with the problem of multiple comparisons. Its main objective is to control the family-wise error rate and the probability of making type I errors (false positive) when performing multiple hypotheses tests. It is based on the following steps:

- Reorder p-values in an ascendant way and mark them with a rank (the smallest has the first rank).

- Starting with the first one, calculate the adjusted alpha level:

$$\alpha_{corr} = \frac{\alpha}{n - rank + 1}$$

  Where α is the level of significance established in advance, n is the number of comparisons and rank refers to the first step.

- Compare the p-value with α_{corr}: if the p-value is lower, the first null hypothesis is rejected.

- Repeat until a null hypothesis is accepted, all subsequent hypotheses will also be accepted. If it happens with the first p-value, then no p-value is low enough for rejection (all the null hypotheses are accepted).

Tables in Chapter 3 show adjusted p-values, thus the significance level to compare them with is 0.05.

Cohen's d

Cohen's d [130] is one of the most common ways to measure effect size which, as written above, quantifies the magnitude of the difference between two samples. It is defined as:

$$d = \frac{\overline{x_1} - \overline{x_2}}{s}$$

Where at the numerator there is the difference between the mean of the two samples and s is the pooled standard deviation:

$$s = \sqrt{\frac{(n_1 - 1) \cdot s_1 + (n_2 - 1) \cdot s_2}{n_1 + n_2 - 2}}$$

Where n_i is the number of samples of the respective group and s_i is its standard deviation. Cohen's d results are commonly interpreted in this way:

- A value of 0.2 represents a small effect size.
- A value of 0.5 represents a medium effect size.
- A value of 0.8 represents a large effect size.

2.4 Clustering

The purpose of this part is to cluster the improved group and the stable group through clinical features before surgery in order to find out, if the two groups are well separable, which features are most important in predicting outcome after STN-DBS intervention.

2.4.1 Feature selection

Exploiting the previous statistical analysis, significant features and those with a large effect size were selected, representing the first two subsets of features as input to the various clustering methods. Principal Component Analysis (PCA), described below, was another method adopted to select features.

Principal component analysis (PCA)

Principal Components Analysis is a method of dimensionality reduction. PCA searches for k n-dimensional orthogonal vectors that can best be used to represent the data, where n is the number of attributes and k≤n. In other words, it projects data onto a smaller space. One of the major advantages of PCA is that it does not select only a subset of features but combines old features revealing relationships that before were not clear [131]. The PCA steps are the following:

1. Input data are normalized, with standard normalization, to prevent features with larger domains from weighing more than those with smaller domains.
2. The PCA calculates an orthonormal basis composed of k orthonormal vectors (unit vectors perpendicular to each other at each point). The data are a linear combination of these principal components.
3. The principal components are sorted in a descending manner, where the first is the component that most explains the variance of the data. Data, re-mapped into a new set of axes, provide information about variance and allow identification of patterns within them.
4. At this point, dimensionality reduction can be applied, eliminating components with low variance. The original data can be reconstructed to a good approximation, using only the first principal components.

5. Several methods can be used to choose the number of components. In this case, a scree plot was used and Kaiser's rule [132] was followed: only principal components with eigenvalues greater than one were retained.

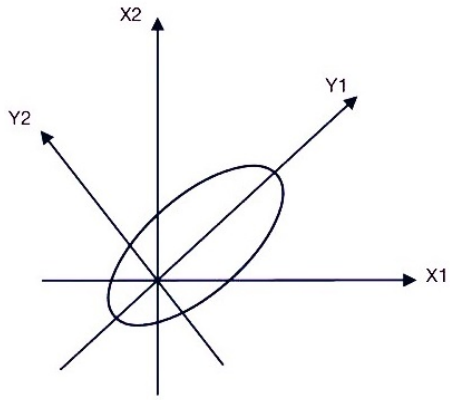


Figure 2.2: Example of PCA. Y1 and Y2 are the first 2 principal components, X1 and X2 are the original axes.

2.4.2 Clustering methods

The Scikit-learn library, a machine learning library for the Python programming language, was used to test different types of methods: k-means, OPTICS, spectral clustering, affinity propagation, mean shift, and Gaussian mixture model. Due to low performances, only Spectral Clustering and Gaussian Mixture Model were selected, for which hyperparameter optimization was performed. The following paragraphs describe the selected methods and hyperparameter tuning. The search for optimal hyperparameters was repeated for different types of feature selection, and each feature was normalized with standard normalization.

Spectral clustering

Spectral clustering is a multivariate statistics technique that performs dimensionality reduction using the eigenvalues (the spectrum) of the data affinity matrix. It then performs a traditional clustering algorithm, such as k-means, in the lower dimension. An affinity matrix, also called a similarity matrix, is a technique used to organize the mutual similarities between a set of data points. Similarity is a measure of distance, however, it does not satisfy the properties of a metric: two points that are the same will have a similarity score of 1, whereas computing the metric would result in 0.

In the case of two clusters, there is a link to spectral graph partitioning, where the second eigenvector of a Laplacian graph is used to define a semi-optimal cut. A cut based on the second eigenvector, in fact, provides an approximation to the optimal cut [133]. This analysis can be extended to clustering, which can be seen as a weighted graph where nodes correspond to data points, edges are related to the distance between the points and the objective is to find the optimal cut, that is, the optimal separation between points.

The main steps of the algorithm are the following [133, 134]:

1) Form a similarity graph.

2) Calculate the weighted adjacency matrix of the graph.

3) Compute the normalized Laplacian:

$$L = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$$

Where I is the identity matrix, A is the adjacency matrix and D is the diagonal matrix $D_{ii} = \sum_j A_{ij}$

4) Calculate the eigenvalues of the normalized Laplacian.

5) Cluster points into two sets based on the eigenvector corresponding to the second-smallest eigenvalue.

The scikit-learn algorithm requires that the number of clusters is specified in advance and works well for a small number of clusters.

Hyperparameter tuning

- Affinity, the method used to construct the affinity matrix:
  - Nearest neighbors: constructs the affinity matrix by computing a graph of the nearest neighbors.
  - RBF: constructs the affinity matrix using a radial basis function (RBF) kernel.
- Eigen solver, the eigenvalue decomposition strategy:
  - Arpack.
  - Amg (AMG requires pyamg installation).
- Assign labels, the strategy for assigning labels after the Laplacian embedding:
  - K-means: common, but can be sensitive to initialization.
  - Discretize: less sensitive to random initialization [135].
  - Cluster qr: it has no tuning parameters and runs no iterations, but it may be better than other methods in terms of both quality and speed [136].

Fixed hyperparameters

- N°clusters = 2, the dimension of the projection subspace.

- Random state = 0, an int to make results deterministic across multiple function calls.

- N°neighbors = 3, the number of neighbors to use when the affinity matrix is the nearest neighbors.

Gaussian Mixture Model

A mixture model assumes a dataset is a mixture of instances from multiple probabilistic clusters. A probabilistic cluster is characterized by a probability density function f, which gives the relative likelihood that an instance of the cluster appears at one point o in the dataspace ($f(o)$). Given dataset, D, and a k number of clusters, the task of cluster analysis is to infer a set of k probabilistic clusters that is most likely to generate D. In other words, the objective is to infer a set of parameters that maximizes the equation:

$$P(O|\Theta) = \prod_{i=1}^{n} \sum_{j=1}^{k} w_j P_j(o_i|\Theta_j)$$

where $P_j(o_i|\Theta_j)$ is the probability that o_i is generated from the jth distribution by using parameter Θ_j and w_j is the probability of the cluster C_j with $j = 1,...,k$ number of clusters.

To estimate the maximum likelihood the algorithm uses an iterative process called Expectation-maximization. It is used by k-means and generalized to mixture models clustering. It consists of two steps:

- The expectation step (E-step): assigns data points to clusters according to the current parameters of the probabilistic clusters.

- The maximization step (M-step): finds the new parameters that maximize the expected likelihood found in the previous step.

The algorithm starts with an initial set of parameters and repeats the two steps until the clustering converges or when the improvement is less than a certain threshold. A Gaussian mixture model is one in which the probability density functions are Gaussian distributions with unknown parameters: mean, covariance and weight parameters. [131].

Hyperparameter tuning

- Covariance type, the type of covariance parameters: full, tied, diagonal, or spherical.
- Init params, used to initialize weights, means, and precisions:
  - K-means: the default method, may be more expensive than the others.
  - K-means++: an initialization method of k-means, which is faster than the standard one. Initially, the center is picked randomly from data, whereas subsequent centers are chosen by prioritizing the points farthest from the existing centers.
  - Random: centers are chosen by adding a small perturbation from the average of all data.
  - Random from data: centers are chosen randomly from the input data.

Fixed hyperparameters

- N°components = 2, the number of mixture components.

- Random state = 0, an int to have a reproducible output across multiple function calls.

2.4.3 Performance evaluation

Both supervised and unsupervised metrics have been used to evaluate the goodness of clustering. Usually, the main goal of clustering is to discover the existence of distinct groups within a dataset, therefore unsupervised metric is used. Since true labels are also available, the idea is to use also supervised metrics to obtain a more complete result.

Unsupervised metric

- Silhouette coefficient (given for a single sample):

$$s = \frac{b - a}{max(a,b)}$$

  Where:

  - a: the mean distance between a sample and all other points in the same class.
  - b: the mean distance between a sample and all other points in the next nearest cluster.

  The score is between -1 and 1: a higher score indicates that clusters are well separated; scores around 0 indicate overlapping clusters whereas a score near -1 indicates incorrect clustering.

Supervised metrics:

- Homogeneity: to assess whether each cluster contains only members of a single class.
- Completeness: to assess if all members of a given class are assigned to the same cluster.

  Both metrics are between 0 and 1. Higher values indicate a clustering closer to the ground truth.

- V-measure is their harmonic mean:

$$v = \frac{(1 + \beta) \cdot homogeneity \cdot completeness}{\beta \cdot homogeneity + completeness}$$

  where β is a weight coefficient: a value less than 1 gives more weight to homogeneity, whereas a value greater than 1 gives more weight to completeness. In all evaluations, the default value of 1 was chosen.

- Misclassified points (how many points are misclassified):

$$Ms = FP + FN$$

  Where FP = false positive and FN = false negative.

- Confusion matrix, a 2x2 table displaying the performance of supervised learning, by directly comparing true labels and predicted labels, as in Table 2.2:

		Predicted labels	
		Class 0	Class 1
True labels	Class 0	True 0	False 1
	Class 1	False 0	True 1

Table 2.2: Confusion matrix.

The best methods, which depend on the type of method, hyperparameters, and features selection, were chosen by first considering the misclassified points and then the other metrics. Moreover, multidimensional scatter plots with true labels were displayed. Multidimensional scatter plots are a way to visualize features of more than 2 dimensions. The idea is to display as many 2D scatter plots as n-1 where n is the number of features, into a giant matrix in which both the x-axis and y-axis represent all features. In the diagonal of the matrix, the distributions of each feature are plotted. The third dimension, represented by color, provides information about the class label.

2.5 Classification

Neural features of 412 SUA, extracted from microelectrode recordings (MERs) during surgery, were used in input for classification methods. The aim of this part is to classify correctly the improved and stable groups, to have a prediction of their future outcomes, and to better understand which neural features may be more important to distinguish the two groups.

2.5.1 Feature selection

Initially, all neurons were used as inputs for the classification algorithms. Then, based on previous findings [123], that is, the correlation between the ventral part of the STN and ICDs, the neurons were divided into dorsal and ventral neurons with a threshold of 0.5 on the normalized depth. All subsequent steps were performed for these 3 datasets: all neurons, dorsal neurons, and ventral neurons. As with the clustering part, feature selection depended on the results of statistical analysis and PCA, which was performed with all non-binary features; Kaisers' rule was applied to choose the number of principal components.

2.5.2 Training

Using the scikit-learn library, the following classifiers were tried: decision tree, random forest, k-nearest-neighbour, Gaussian process, support vector, adaptive boosting, and gradient boosting. A hyperparameter tuning was performed for each of them with a function provided by scikit-learn: GridSearchCV. GridSearchCV considers all parameter combinations and returns the best estimator and parameters according to a scoring function and a cross-validation scheme provided as input. For all classifiers, balanced accuracy was chosen as the scoring function and a 5-cross validation was performed. Balanced accuracy is defined as:

$$Balanced\ accuracy = \frac{Sensititivty\ +\ Specificity}{2}$$

Where:

$$Sensitivity = \frac{True\ positive\ (TP)}{True\ positive\ (TP)\ +\ False\ negative\ (FN)}$$

$$Specificity = \frac{True\ negative\ (TN)}{True\ negative\ (TN)\ +\ True\ positive\ (TP)}$$

K-fold cross-validation is a type of cross-validation technique. The idea behind this technique is to split the dataset into two parts: one to train the model and the other to perform the evaluation. K-fold cross-validation is mostly used when the dataset is limited because allows each sample to be used. After defining a k parameter, the dataset is randomly divided into k folds. For each k fold the model is trained on k-1 folds and validated on the remaining fold. This process is repeated k times until each fold has been a test set. The performance is calculated as the average of all folds [137].

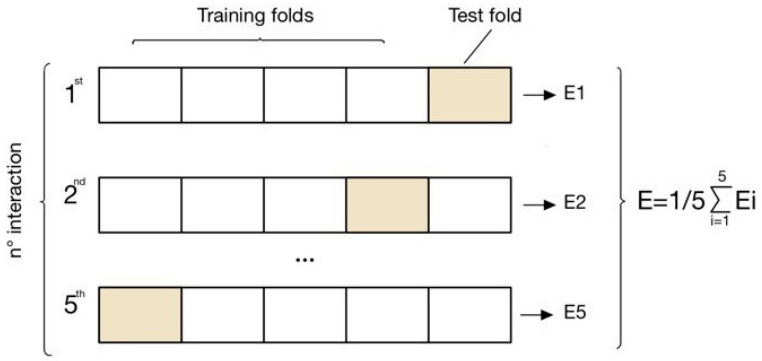


Figure 2.3: 5-cross validation.

After hyperparameters were selected, each model was trained and tested. The division between the training set and the test set was performed by stratified k-fold validation (with k=5), a special type of k-fold cross-validation that maintains the original distribution of classes. In addition, the samples of each class were shuffled before division. The training set samples were augmented by implementing Synthetic Minority Over-sampling Technique (SMOTE) [138] provided by the imbalanced-learn library, to overcome the problem of data set imbalance. The idea is to generate new samples belonging to the smallest class. SMOTE generates new samples by interpolation. Considering a sample x_i and a point of its nearest neighbor, x_{zi} a new sample x_{new} is generated as follows:

$$x_{new} = x_i + \lambda \cdot (x_{zi} - x_i)$$

Where λ is a random number in the interval [0,1]. This interpolation will create a sample on the line between x_i and x_{zi}.

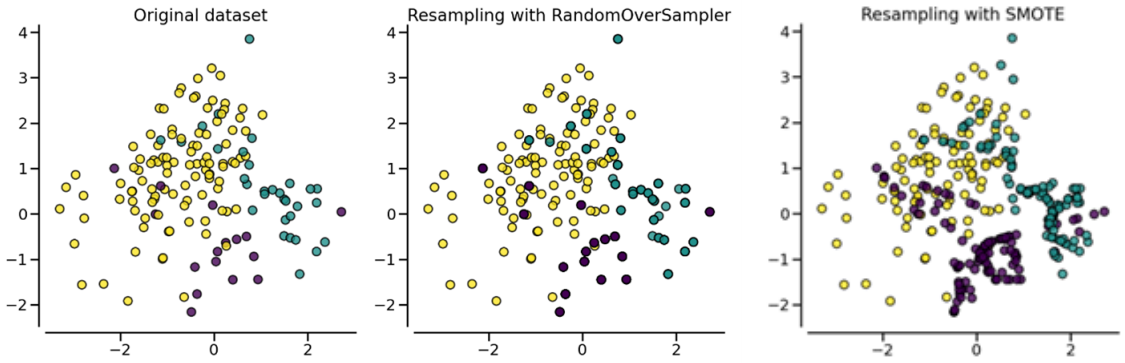


Figure 2.4: Example of RandomOverSampler and SMOTE [139].

If the SMOTE algorithm does not work, a simpler algorithm is tried: RandomOverSampler, which generates new samples by randomly drawing with replacement from available samples.

In addition to the classifiers listed above, ensemble methods were used by applying VotingClassifier, a function provided by the scikit-learn library. It combines multiple classifiers, predicting the class of an object in two possible ways: by majority voting (hard vote) or by averaging the predicted probabilities of the selected classifiers (soft vote). The input classifiers for the ensemble method were chosen through performance evaluation, and the prediction of neurons was obtained by soft voting.

2.5.3 Classifiers

Decision tree

A decision tree is a tree-like structure composed of branches, nodes, and leaves (terminal nodes). Each internal node (non-leaf node) defines a test on a feature, each branch represents a test result, and each leaf node (or terminal node) returns a class label. Internal nodes are denoted by rectangles, whereas leaf nodes are represented by ovals, as can be seen in Figure 3.18. Suppose that X is a tuple with an unknown class label, to be tested with a decision tree. Each attribute value is tested by defining a path from the root node to a leaf node, where a class prediction is provided.

Training a decision tree starts with a single node, with all the training tuples in D, which is the set of training tuples and labels. Then the training set is recursively partitioned into smaller subsets, in a way that depends on the algorithm selected. Examples of algorithms are ID3, C4.5, and CART, whose main difference is the attribute selection measure. This is a heuristic that defines the splitting criterion, which selects the attribute to be tested at a given node, and the split point. The splitting criterion is determined in order to have partitions that are as pure as possible in each branch, which means that, ideally, all tuples in the same partition belong to one class. The recursive partitioning stops in three cases: all the tuples in partition D belong to the same class; if a partition D_j is empty, a leaf node is created with the majority class in D; if there are no more attributes, majority voting is used to define the common class in D. An attribute A could be continuous and in this case, the test node has two possible outcomes: $A \leq split\_point$ and

$A > split\_point$, where the split point is returned by the attribute selection method. If A is a discrete value, the test outcomes correspond directly to each of the possible values of A. Finally, there is a third possibility, A with discrete value and the tree is binary: outcomes usually are "yes" or "no" answers. Some decision tree algorithms construct only binary trees, whereas others can produce non-binary trees [131].

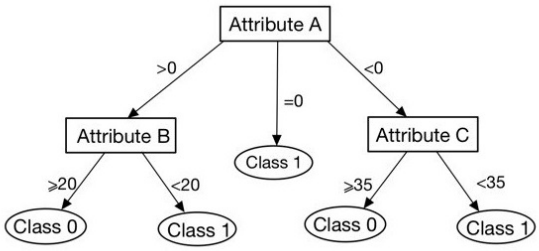


Figure 2.5: Decision tree example.

Hyperparameter tuning

- Max depth, the maximum size of the tree: [2, 7, 12, 17, 22, 27, 32, 37].
- Min samples leaf, the minimum number of samples needed to be at a leaf node: [2, 7, 12, 17, 22, 27, 32, 37].
- Min samples split, the minimum number of samples needed to split an internal node: [2, 7, 12, 17, 22, 27, 32, 37].
- Max leaf nodes, the number of best nodes, defined as relative reduction in impurity: [2, 7, 12, 17, 22, 27, 32, 37].

Fixed hyperparameters

- Random state = 0, an int to have determined results across many calls.
- Criterion, the function to measure the quality of a split: entropy was selected. It is defined as:

$$H(Q_m) = -\sum_k p_{mk} \log(p_{mk})$$

Where $p_{mk} = \frac{1}{n_m} \sum_{y \in Q_m} I(y = k)$ is the proportion of class k observations in node m and Q_m is the data representation at node m with n_m samples. If m is a terminal node, the prediction probability is set to p_{mk}. Note that the entropy criterion calculates the Shannon entropy of the possible classes.

Random forest

The random forest is an ensemble method in which each classifier is a decision tree. To create a random forest, the bagging method is applied, which works as follows: given a set of attributes D, of d tuples, and suppose an ensemble method consisting of k classifiers. For each classifier, a training set D_i (composed of d tuples) is sampled with replacement from the original set. Some elements of D may be included two times or more, whereas others may not be sampled. Each classifier has its own training set and returns its prediction [131]. Usually, the final class is determined using majority voting. The scikit-learn implementation, instead of having each classifier vote for a single class, averages the probabilistic predictions of the classifiers. When a tree is constructed, the split of each node is found using all the input features or a randomly chosen subset (the default number is the square root of the input features). The idea behind this method of constructing each decision tree and the entire forest is to reduce the variance and try to overcome the overfitting issue [140]. The algorithm used is CART, introduced in 1984 by the book "Classification and Regression Trees" [141]. CART constructs binary trees with a criterion based on the largest information gain at each node.

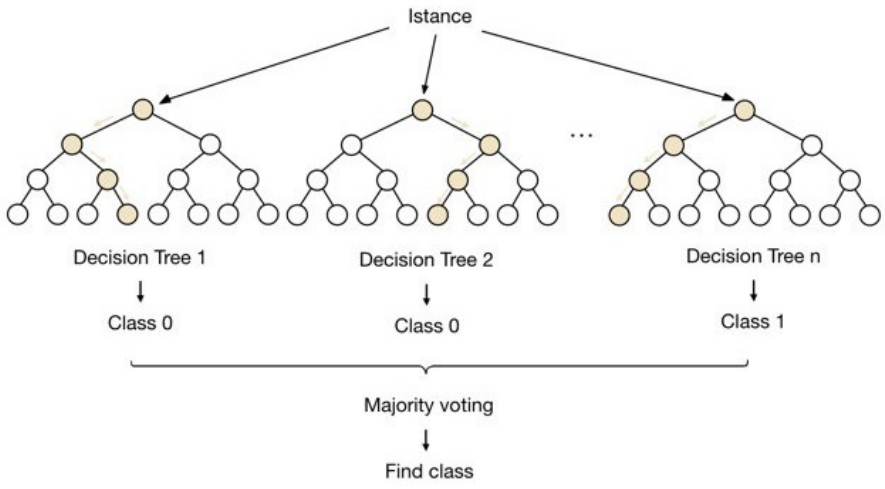


Figure 2.6: Random forest example.

Hyperparameter tuning: the same as the decision tree, except for the number of estimators, set at 50.

K-nearest neighbors (K-NN)

Nearest-neighbour is a classifier that stores training tuples of n attributes and, in order to classify an unknown tuple, it searches the pattern space for the k training tuples closest to the new tuple (k-nearest neighbors). A distance metric, such as Euclidean distance, is used to calculate this closeness. Then the unknown tuple is classified by considering which class is most common

among its k-nearest neighbors. For example, if k=1, the unknown tuple will be assigned to the class of the nearest tuple. The drawback of the k-NN is that k is decided a priori, but it is possible to try different k values, starting with k=1 and using a test set to estimate prediction error. K is then increased until it provides the minimum error rate or an error below a certain threshold. Usually, larger numbers of training tuples require a higher k to have a lower prediction error. To compute distance-based comparison, k-NN uses uniform weights: all the k-nearest neighbors contribute equally to the majority voting to assign the unknown class. In some cases, it is more worthfully to assign neighbors weights inversely proportional to the distance, such that the nearest neighbour has the greatest contribution [131].

Nearest-neighbour algorithms provided by the scikit-learn module are the following:

- Brute force. It computes distances between all pairs of points in the dataset. For small dataset might be a good choice, but since its complexity is $O[DN^2]$, with N number of samples and D dimensions, it grows rapidly as N increases.

  Tree-based algorithms overcome the infeasibility of brute force: aggregation of information avoids the calculation of all distances, thus reducing the number of total operations. For example, if point P is distant from D and point D is near to point N, it is clear that P and N are distant, without having to compute it [142]. The complexity decreases to at most $O[DN\log(N)]$. The available tree-based algorithms are:

- K-D tree. It performs a recursive partition along data axes, dividing the space into nested orthotropic regions. The nearest neighbour of a point can be obtained with $O[\log(N)]$ distance computations. For low dimensionality ($D < 20$) it works quickly, otherwise, another approach such as ball tree [142] is preferable.

- Ball tree. It adopts a recursive partition as the K-D tree algorithm but in nested hyperspheres. This results in an increased construction tree cost to achieve better efficiency on high dimensions. The growing time is approximately $O[D\log(N)]$ [143].

Hyperparameter tuning

- N°neighbors, the number of neighbors (k): a vector from 2 to (n°neurons $\cdot$ 0.2+2) in steps of 2.

- Weights, a method for assigning weights to the nearest neighbors:
  - Uniform: assigns the same weights to all the k-nearest neighbors.
  - Distance: assigns different weights inversely proportional to the distance.

- Algorithm: ball tree, k-D tree, brute force and auto (automatically chooses among these three).

Support vector machine (SVM)

Support vector machines are a set of methods that can be used for both classification and regression. In the former case, the method is called support vector classifier (SVC) and allows the classification of both linear and non-linear data. The idea behind it is that with a non-linear data transformation in a higher dimensional space, two classes of data can be always separated by a hyperplane. SVC finds the optimal hyperplane to separate the two classes. The following figure

shows an example of linearly separable 2D data:

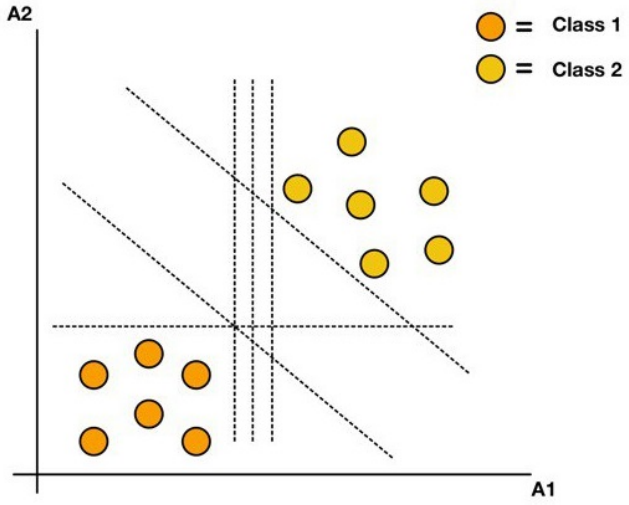


Figure 2.7: Example of 2 data classes linearly separable.

As can be seen in the figure above, there are many lines that can separate the two groups, but the purpose of the SVC is to find the line for the best separation in order to reduce classification errors. Considering higher dimensions, the lines become hyperplanes: decision boundaries of k dimensions [131].
Given training vectors $x_i \in \mathbb{R}^p$ with $i = 1,..,n$ and a vector $y \in \{1, -1\}^n$ denoting the two possible classes , the aim of the SVC is to find $w \in \mathbb{R}^p$ and $b \in \mathbb{R}$ to correctly predict most samples using:

$$sign(w^T \Phi(x_i) + b)$$

Where:

- $w^T \Phi(x_i) + b$ is the equation of the hyperplane and the $sign$ defines on which side the sample will be classified.
- w is a weight vector.
- b is a scalar (bias).
- Φ defines a feature-space transformation.

SVC solves a primal problem:

$$\min_{w,\xi,b} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i$$

$$\text{Subject to} \quad y_i(w^T \Phi(x_i) + b) \geq 1 - \xi_i,$$

$$\xi_i \geq 0 \quad i = 1,...,n$$

SVC tries to minimize $w^t w$, that is, to maximize the margin: it is the shortest distance from the hyperplane to the closest training tuple of either class. Minimization has a penalty term if a sample is misclassified or within the margin boundary, and it is controlled by C, an inverse regularization term. In the ideal case, indeed, the value of the equation would always be >=1, meaning that the two classes are perfectly separable by a hyperplane, but in real cases, ξ_i is the distance allowed from the correct margin boundary. The term C, as mentioned above, controls the strength of this penalty [144, 145]. Due to the possible high dimensionality of the vector variable w, usually, SVC solved the dual problem, but it will not be described as it is beyond the scope of this thesis.

Hyperparameter tuning

- C, the inverse regularization term: [1:11]/10.
- Degree, indicates the degree of the polynomial kernel: [1:11].
- Probability, a boolean to enable or not probability estimates (True, False).
- Kernel types used for data transformation:
  - Common kernels provided by SVC function: linear, polynomial, radial basis function (RBF) and sigmoid.
  - Kernels added from sklearn.gaussian-process.kernels: DotProduct(), Matern(), RationalQuadratic(), WhiteKernel() and RBF().

Fixed hyperparameters: random state = 0, an int to have determined results across many function calls.

Gaussian process classifier (GPC)

Gaussian Processes (GPs) are machine learning methods for supervised probabilistic classification and regression problems. In the former case, they are called Gaussian process classifiers (GPCs). Probabilistic classification means that prediction involves the identification of class probabilities. A latent function (nuisance function) is selected, with the aim of simplifying the model formulation. A Gaussian process prior is applied to the latent function, which is then flattened through a link function, to calculate the probabilistic classification. A Gaussian link function is used for regression, whereas a non-Gaussian likelihood, corresponding to the logistic link function (logit), obtained through the Laplace approximation, is used for classification because the labels are discrete. The mean of the prior is assumed to be zero, whereas its covariance depends on the selected kernel. Hyperparameters are optimized by maximizing the log-marginal-likelihood (LML). Initial hyperparameters come from the kernel values, whereas subsequent hyperparameters are randomly chosen from the range of allowed values. Kernels are an important part of this type of method: they define the shape of the prior and posterior of the Gaussian process. They define the similarity between two data, considering that two similar samples have similar outputs. There are two types of kernels: stationary and non-stationary. The first group is based on the distance of two data points and it is further divisible into isotropic and anisotropic kernels. The second group, instead, depends also on the specific values of the data points [146].

Hyperparameter tuning

- Kernel, indicates the covariance function: DotProduct(), Matern(), RationalQuadratic(), RBF(), DotProduct()+Matern(), DotProduct()+RBF(), Matern()+RBF().

Fixed hyperparameters

- Random state = 0, an int to have determined results among multiple function calls.
- N°jobs = 100, indicates the maximum number of concurrent processes.
- Warm start = True, allows aspects of the model learned in the previous iteration to be reused to initialize the new model in a subsequent call, saving time.

Adaptive Boosting (AdaBoost)

Adaptive Boosting is a boosting algorithm, a type of ensemble method. The idea behind boosting is to train k classifiers in succession. After a classifier C_i has been trained, the subsequent classifier learns from the previous one by paying more attention to tuples misclassified by C_i. The final classifier is an ensemble that combines the vote of each classifier by weighting it according to the accuracy of that classifier. Suppose dataset D consists of d tuples with labels, AdaBoost at the beginning sets for each tuple a weight value equal to 1/d. With k classifiers, there are k rounds to train the model, and for each round, the tuples from D are sampled with replacement to form the training set D_i. The chance of being sampled depends on the weight of the tuple. Each model C_i is trained with its dataset D_i. D_i is used as a test set to calculate the prediction error, then the weight of each tuple is updated according to the difficulty of its classification: the weight increases if the tuple has been misclassified. The next classifier is trained with this new set of weights. In this way, the ensemble classifier is composed of complementary classifiers: if a classifier misclassifies some points, the next classifier may be better at classifying these points but may make mistakes in the points correctly classified by the previous one.
The error rate of a model C_i is calculated as follows:

$$error(C_i) = \sum_{j=1}^{d} w_j \cdot err(X_j)$$

Where $err(X_j)$ is the misclassification error of X_j, which is equal to 0 if it is classified correctly, otherwise is 1. C_i is discarded if its error is higher than 0.5. The weights of correctly classified tuples are adjusted by multiplying them by:

$$\frac{error(C_i)}{1 - error(C_i)}$$

Then all weights are normalized so that their sum is equal to that of the previous classifier (as the weights of correct classifiers decrease, those of incorrect classifiers increase). Each classifier, contributes to classification according to its performance, with a weight defined as:

$$\log \frac{1 - error(C_i)}{error(C_i)}$$

Finally, a tuple is assigned to the class with the highest sum of the weights of the classifiers [131].

Hyperparameter tuning

- Learning rate, a weight assigned to each classifier at each iteration: [0.001, 0.0025, 0.005, 0.0075, 0.01, 0.025, 0.05, 0.075, 0.1, 0.25, 0.5].

Fixed hyperparameters

- N°estimators = 1000, the maximum number of estimators at which AdaBoost stops.
- Random state = 0, an int to have determinate outcomes among multiple function calls.

Gradient Boosting Classifier (GBC)

Gradient Boosting (GB) was introduced by Jerome H. Friedman [147], and the idea behind it was that with small steps it is possible to better predict the outcome of a dataset. GB is a boosting method used in both regression and classification. In the second case, it is called Gradient Boosting Classifier (GBC). It is an ensemble method composed of a number of weak learners, which combine into a single stronger model sequentially with each model that learns from the previous one. GBC usually uses decision trees to construct classification trees. To construct the first leaf, an initial prediction of the outcome is made, then other trees are constructed using the residuals of the previous weak learner (residual = difference between the predicted and observed outcome). In GB classification, the logarithmic loss is used to calculate residuals. For each tree, it is possible to define the depth of the tree, the number of maximum nodes, the minimum number of samples to split, and so on, as in the random forest. To avoid overfitting, the learning rate is chosen between 0 and 1. It is a hyperparameter that shrinkages the contribution of each tree to the classification.

Hyperparameter tuning

- Learning rate, the contribution of each tree: [0.001, 0.0025, 0.005, 0.0075, 0.01, 0.025, 0.05, 0.075, 0.1, 0.25].
- Max depth, min samples leaf, and min samples split are all hyperparameters derived from the random forest.

Fixed hyperparameters

- N° estimators = 2000, the number of boosting stages to be performed.
- Random state = 0, an int to have determinate outcomes among multiple function calls.
- Max leaf nodes = 32, keeps only the first 32 best nodes, defined as relative reduction in impurity.

2.5.4 Performance evaluation

Once the hyperparameters were selected for each combination of classifier and feature selection, the following metrics were used to evaluate the performance of each combination:

- A normalized confusion matrix, calculated by normalizing over the true condition (the sum of each row is 1).
- Balanced accuracy (defined in 2.5.2).
- F1 score, defined as:

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision \ + \ recall}$$

  Where:

$$precision = \frac{TP}{TP + FP} \qquad recall = \frac{TP}{TP + FN}$$

- AUC: area under the ROC curve. ROC (receiver operating characteristics) curve is a graphical plot that displays the true positive rate (TPR) against the false positive rate (FPR), at various threshold settings. TPR and FPR are defined as:

$$TPR = \frac{TP}{TP + FN} \qquad FPR = \frac{FP}{FP + TN}$$

  In Figure 2.8 is shown the ROC curve: higher AUC values indicate a better classifier.

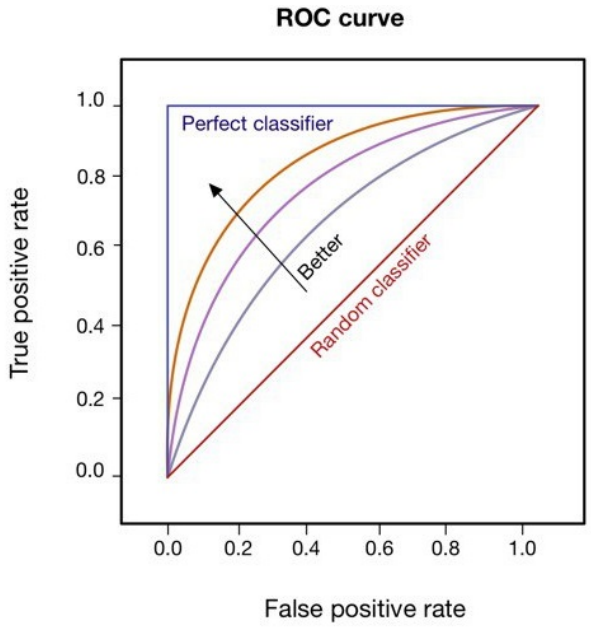


Figure 2.8: ROC curve.

2.5.5 Post-classification analysis and patient prediction

It is possible to get an inside view of the decision-making mechanisms of some classifiers, such as those based on decision trees. The Dtreeviz library enables the graphical visualization of a decision tree. It also allows a selection of individual trees that are part of ensemble methods, such as random forest or gradient boosting. In this case, it can be seen that different trees have different subsets of input features. For each tree, the features selected as input and along the path, the thresholds used in each partition, and the results of the leaf nodes are shown.

The results of the best-performing classifier are applied to classify the improved and stable patients. The prediction class of an individual patient is obtained by majority voting, using the classification of all neurons belonging to the patient. The prediction results are reported together with true labels in a confusion matrix, next to which a table with the percentage of winning neurons is reproduced to show the ratio of winning neurons to losing ones (since a class can be selected with only 51% of the neurons belonging to that class).

3 | Results

The following tables show the clinical and neural features described in 2.2, in terms of mean, standard deviation, minimum, and maximum for continuous variables and in terms of count (total number) and percentage for categorical variables. The upper part of each table shows the time of data gathering, and the lower part shows the number of patients for whom data were collected. Table 3.1 and Table 3.2 report information on the regularity of neuron firing, frequency of neural oscillations, SUA-BUA cross-coherence, and all neural information recorded during surgery. Since the differences between the ICD+ and ICD- groups at the neural level have already been addressed in background works [122, 123], this study compared only the improved and stable groups on neural features. Clinical features are shown in Table 3.3, Table 3.4 and Table 3.5. Table 3.5 shows the BIS second-order factor, collected for all patients at baseline condition (indicated as T0 in the tables). Categorical and continuous clinical features, instead, were provided before surgery and at the follow-ups (indicated as T1-month and T1-year in the tables), and they were not all always available for all patients, as shown in Table 3.4 and Table 3.3.

	T-surgery			
Feature	mean	std	min	max
CV	1.158	0.283	0.401	1.902
Firing rate	17.386	15.896	1.606	149.570
Firing regularity	0.189	0.466	-0.752	2.142
Beta oscillation frequency	20.925	5.740	12.000	30.000
Beta oscillation amplitude	0.513	0.407	0.000	2.382
Ccbeta	0.081	0.152	0.000	0.702
IBF	138.922	75.855	0.000	375.557
Theta oscillation frequency	5.653	1.454	4.000	8.000
Theta oscillation amplitude	0.448	0.494	0.000	3.330
Cctheta	0.059	0.127	0.000	0.683
Burst duration	0.109	0.213	0.000	2.970
IBI	1.512	2.043	0.000	17.150
Patients	12			

Table 3.1: Neural features.

	T-surgery	
Feature	count	percentage (%)
Is tonic	127	30.83
Is bursting	112	27.18
Is irregular	173	41.99
Is beta oscillatory	123	29.85
Is theta oscillatory	94	22.82
Patients	12	

Table 3.2: Neural binary features.

	T0		T1-month		T1-year	
Feature	mean	std	mean	std	mean	std
Age at DBS	61.4	6.7				
PD duration	12.8	5.5				
Hoehn-Yahr off	3.3	0.9			2.3	0.8
Hoehn-Yahr on	2.0	1.0			1.8	0.7
Hoehn-Yahr off-on	1.3	0.8			0.5	0.4
UPDRS III off	31.8	11.6			23.5	13.1
UPDRS III on	10.2	6.9			10.9	5.8
UPDRS III off-on	21.7	7.9			12.6	10.8
LEDD	1266.1	502.8	825.3	381.3	741.0	352.4
DA LEDD	302.5	255.2	95.1	108.3	98.4	106.8
LD	963.6	356.8	730.2	342.6	642.6	360.9
Tremor subscore	1.3	1.7				
Bradykinesia subscore	12.6	5.7				
Rigidity subscore	5.3	3.5				
Patients	24		12		22	

Table 3.3: Clinical features.

	T0	T1-month	T1-year
Feature			
Gender	17M - 7F	8M/ 4M	15M / 7F
Fenotype	13 TR - 11 NTR	6 TR - 6 NTR	12 TR - 10 NTR
DA (yes)	18 (75%)	7 (58.3%)	12 (54.5%)
Patients	24	12	22

Table 3.4: Clinical categorical features. TR=tremor dominant; NTR=non tremor dominant.

	T0			
Feature	mean	std	min	max
Attentional	15.7	2.4	12	21
Motor	19.3	3.6	13	26
Non-planning	24.7	3.4	19	30
BISTOT	59.7	7.3	50	72
Patients	24			

Table 3.5: Barrat Impulsiveness Scale Second-Order Factor.

3.1 Predicting outcome from presurgical evaluation

3.1.1 Statistical differences across outcomes

			ICD+ vs ICD-			ICD stable vs ICD improved	
	Feature	p-value	Cohen's d	Cohen's d interpretation	p-value	Cohen's d	Cohen's d interpretation
Mann-Whitney U test with Holm-Bonferroni Correction	Age at DBS	0.173	0.576	medium	0.29	0.64	medium
	PD Duration	0.505	0.274	small	0.371	0.534	medium
	Hoehn-Yahr Off condition	0.277	0.434	small	0.089	1.061	large
	Hoehn-Yahr On Condition	0.492	0.274	small	0.458	0.432	small
	Hoehn-Yahr Off-On Difference	0.470	0.286	small	0.802	0.141	small
	UPDRS III Off	0.049	0.875	large	0.034	1.537	large
	UPDRS III On	0.083	0.756	medium	0.625	0.284	small
	UPDRS III Off-On Difference	0.087	0.742	medium	0.034	1.537	large
	Levodopa Dose	0.453	0.31	small	0.343	0.586	medium
	LEDD	0.506	0.274	small	0.53	0.382	small
	DA LEDD	0.793	0.106	small	0.624	0.284	small
	Tremor Subscore	0.902	0.047	small	0.434	0.432	small
	Bradykinesia Subscore	0.099	0.713	medium	0.050	1.361	large
	Rigidity Subscore	0.055	0.844	large	0.06	1.280	large
	BISFOF Attention	0.216	0.511	medium	0.408	0.482	small
	BISFOF Motor	***	3.13	large	0.182	0.789	medium
	BISFOF Self Control	0.069	0.785	medium	0.868	0.094	small
	BISFOF Cognitive Complexity	0.001	1.904	large	0.180	0.789	medium
	BISFOF Perseverance	0.460	0.274	small	0.640	0.236	small
	BISFOF Cognitive Instability	***	2.919	large	0.381	0.482	small
	BISSOF Attentional	0.079	0.756	medium	0.406	0.482	small
	BISSOF Motor	***	3.056	large	0.364	0.534	medium
	BISSOF Non-Planning	0.003	1.481	large	0.368	0.534	medium
	BISSOF Total	***	3.056	large	0.625	0.284	small
Fisher Exact Test	Sex	1.000			0.576		
	Dopamine Agonist Use	1.000			0.205		
	Phenotype (Tremor Dominant)	1.000			1.000		

Table 3.6: Pre-surgery population-wise comparisons. BISFOF: Barrat Impulsiveness Scale First Order Factor; BISSOF: Barrat Impulsiveness Scale Second-Order Factor.

*** $p<0.001$

Table 3.6 shows the comparison between the ICD+ and ICD- groups and between the stable and improved groups, performed with the Mann-Whitney U test (with Holm-Bonferroni correction) for continuous variables and the Fisher Exact test for categorical variables. The compared features are the clinical data collected presurgery, and the line under the rigidity subscore separates the continuous clinical variables from those belonging to the BIS. The BIS subscore of the first-order factor (BISFOF), second-order factor (BISSOF), and total BIS scores (Total BISSOF) were compared. Significant differences are highlighted in orange.

Comparisons between the ICD and non-ICD groups showed a significant difference in almost all BIS scores, it is not present between the improved and stable groups. However, they differed significantly in some clinical features: UPDRS III off, UPDRS III off-on difference and bradykinesia subscore. To obtain a clearer view of significant differences between the two groups, categorical scatter plots are shown with an adjustment to avoid overlap between points. The improved group is represented by orange dots, whereas the stable group is colored dark red. Each scatter plot shows one of the significantly different features between the two populations, also reporting p-values and Cohen's d measure of effect size. The rigidity subscore is displayed even if its p-value is just above 0.05 because it has a high effect size and the p-value is affected by the presence of an outlier.

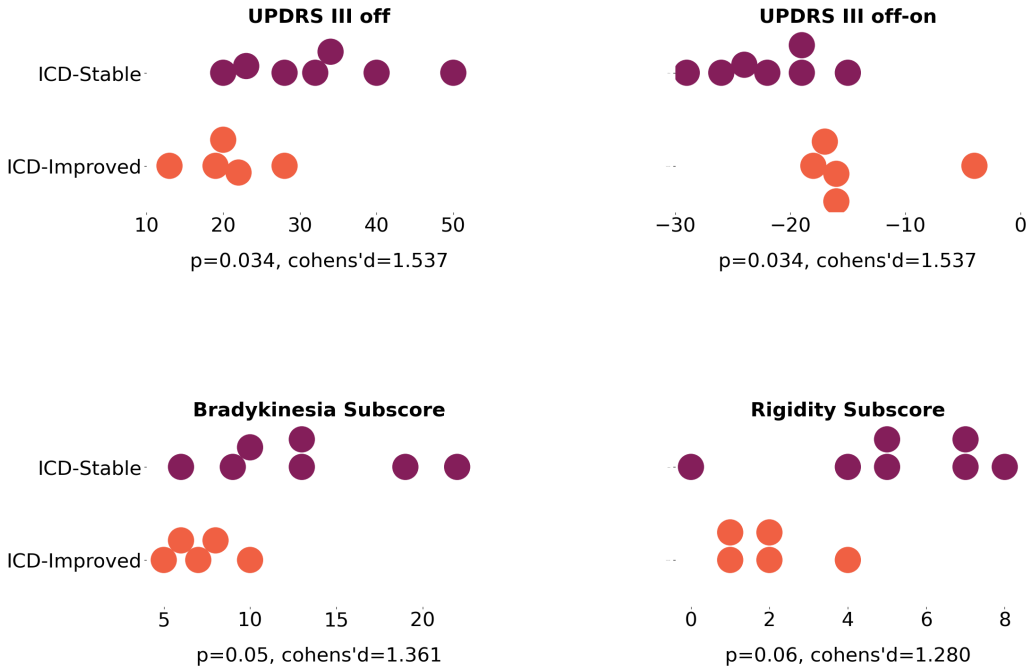


Figure 3.1: Categorical scatter plot of significantly different features between the improved and stable groups.

The UPDRS III off and its subscores of bradykinesia and rigidity are lower on average in improved patients, whereas the UPDRS III off-on has a reverse trend: improved patients have greater progress in UPDRS III score with medication intake.

3.1.2 Clustering of patient's profiles

The purpose of separating the improved group from the stable group by clustering methods was pursued using pre-surgical clinical characteristics; three types of features selection were performed:

- Significant features: UPDRS III off, UPDRS III off-on and bradykinesia subscore.
- Large effect size features: UPDRS III off, UPDRS III off-on, bradykinesia subscore, rigidity subscore and Hoehn-Yahr off.
- Principal component analysis (PCA) with all presurgical features. A scree plot was displayed to select the number of components to be retained. The Y axis shows the eigenvalues, from largest to smallest, whereas the X axis shows the principal components (PCs). Following Kaiser's rule, components with eigenvalues greater than 1 were kept, as shown in Figure 3.2:

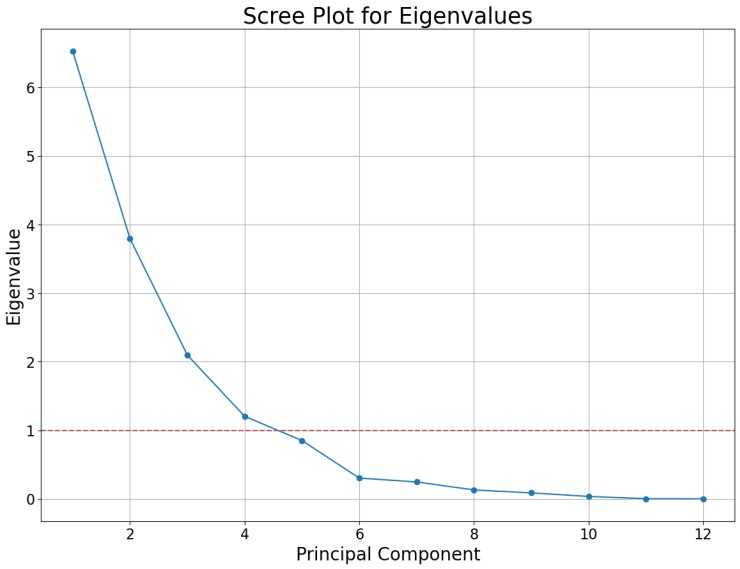


Figure 3.2: Scree plot for eigenvalues. The first four PCs have an eigenvalue greater than 1.

The first four principal components were retained and to understand the contribution of the original variables to the principal components, the loadings matrix, that is, the correlations between the original variables and the principal components, was computed and displayed in the following heatmap (Figure 3.3).

The Loading matrix, using the scikit-learn library, was calculated as follows:

$$Loadings\ matrix = pca.components\_.T \cdot \sqrt{pca.explained\_variance\_}$$

Where *pca.components* calculates the principal axes in feature space, which represent the directions of maximum variance in the data, and *pca.explained_variance_* calculates the amount of variance explained by each component. Transposition (.T) was used because *pca.components* has shape

(n°components, n°features) whereas *pca.explained_variace_* is a vector of rows of n°components.

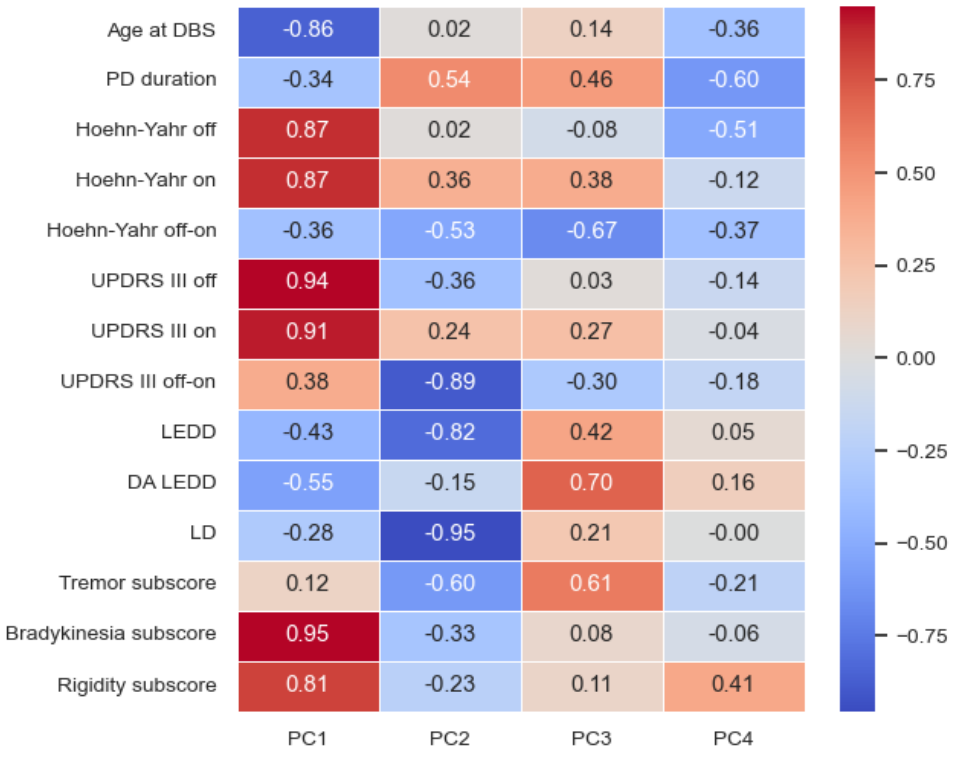


Figure 3.3: Heatmap representing correlations between clinical features and the first four PCs.

It can be seen that the features that were significant or had a large effect size from the statistical analysis were those that correlated strongly with the first of the principal components: UPDRS III off, Hoehn-Yahr off, bradykinesia and rigidity subscores. UPDRS III off-on, instead, had high negative correlation with PC 2, along with LD and LEDD. In addition, UPDRS III on and Hoehn-Yahr on also have positive correlations with PC 1.

Large effect size features

Figure 3.4 represents a pairwise plot, in which the diagonal plots show the kernel density estimation (KDE) of the data, separately for the improved (light blue) and stable (orange) groups, whereas the other plots are scatter plots of the data in multiple 2D features of large effect size.

The selection of large effect size features led to the best performances for both selected clustering methods: spectral clustering and Gaussian mixture models.

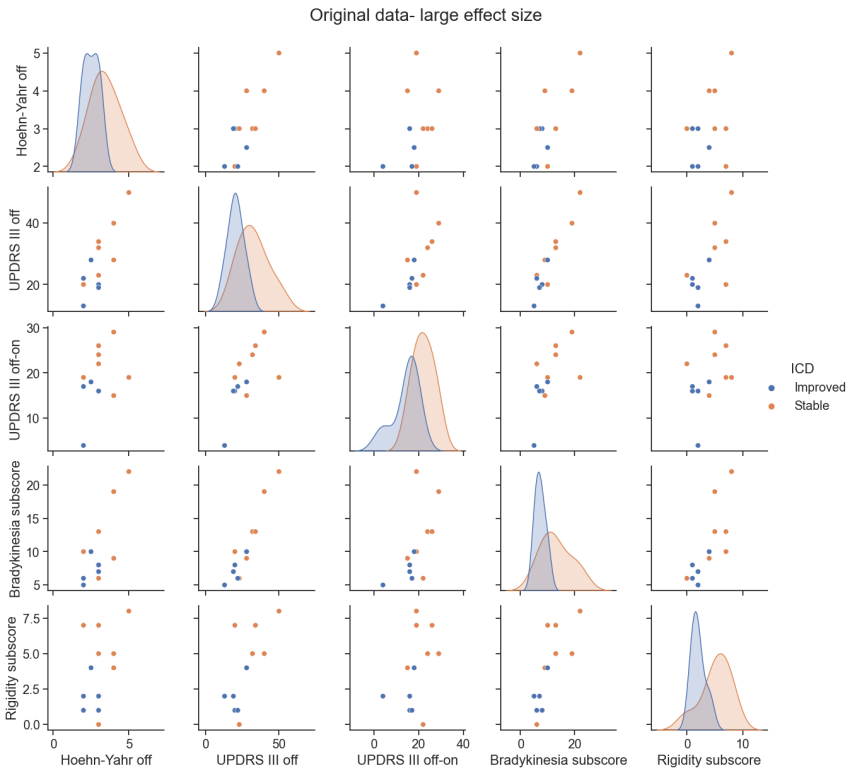


Figure 3.4: Pairwise plot with large effect size features.

Large effect size features - Spectral clustering

Table 3.7 shows the hyperparameters chosen during optimization and the unsupervised and supervised scores. Below the table, the confusion matrix shows the predicted versus true labels, and next to it are the number of misclassified patients and the accuracy of the clustering method. The way the results are presented will be the same for the Gaussian mixture model.

Hypermarameters					
Affinity	Eigen_solver	Assign_labels	n_clusters	random_state	n_neighbors
rbf	Arpack	discretize	2	0	3
Unsupervised and supervised metrics					
Silhouette score		Homogeneity	Completeness	V_measure	
0.343		0.486	0.486	0.486	

Table 3.7: Hyperparameters and performance metrics of spectral clustering with large effect size features.

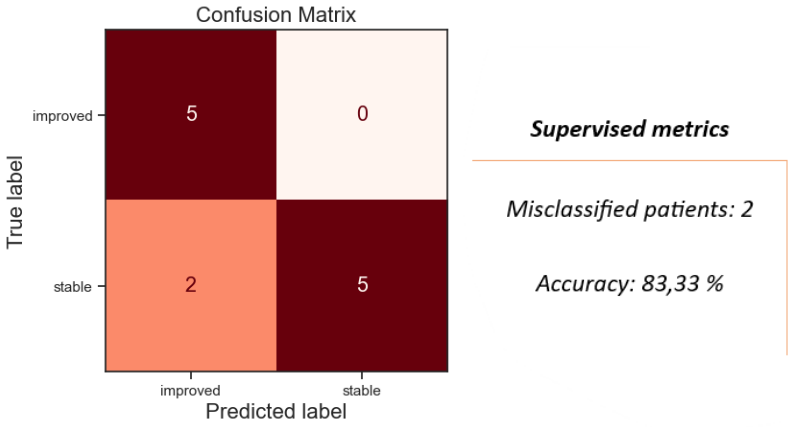


Figure 3.5: Large effect size features - Spectral clustering. a) Confusion matrix; b) supervised metrics.

Large effect size features - Gaussian mixture model

Hypermarameters			
Covariance_type	init_params	n_components	random_state
full	K-means++	2	0
Unsupervised and supervised metrics			
Silhouette score	Homogeneity	Completeness	V_measure
0.340	0.486	0.486	0.486

Table 3.8: Hyperparameters and performance metrics of Gaussian mixture model with large effect size features.

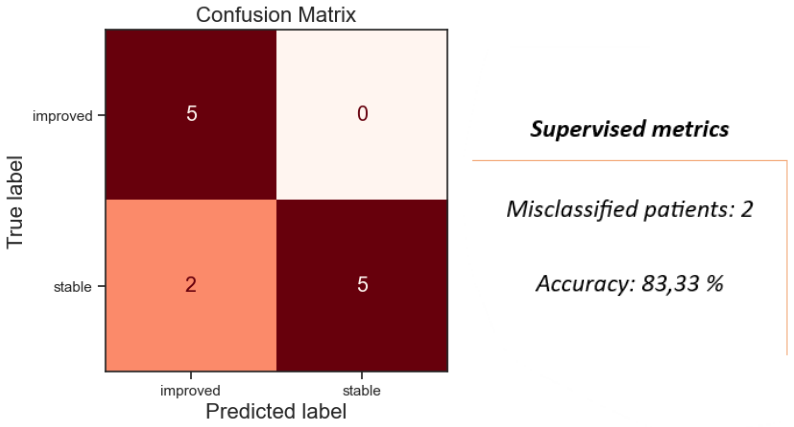


Figure 3.6: Large effect size features - Gaussian mixture model. a) Confusion matrix; b) supervised metrics.

Spectral clustering and the Gaussian mixture model achieved an accuracy of 83.33%, and both misclassified two stable patients. The silhouette score does not show high values: a separation between the two groups can be seen in the pairwise graph, but it is not very pronounced. This may be because the selection of the best methods was based on the supervised metrics of misclassified patients.

The results with significant features and the first four principal components for both clustering methods are presented below with the same previous pattern. In both pairwise plots, the feature distributions of the two groups of patients are more overlapping than the previous distributions of large effect size features, making clustering more arduous. In fact, as it can be seen from the following tables, performance deteriorates, especially in the number of misclassified patients, except for the Gaussian mixture model with principal components.

Significant features

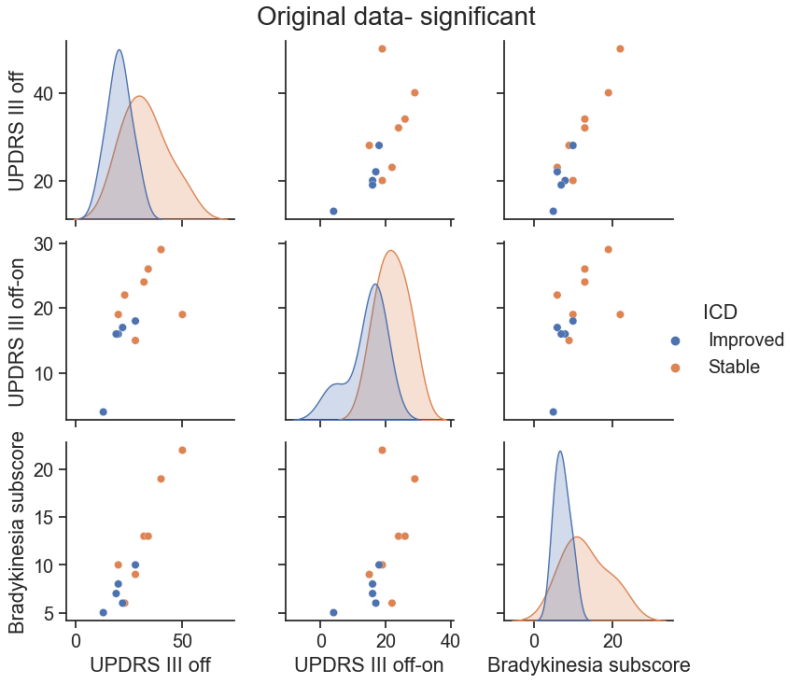


Figure 3.7: Pairwise plot with significant features.

Significant features - Spectral clustering

Hypermarameters					
Affinity	Eigen_solver	Assign_labels	n_clusters	random_state	n_neighbors
nearest-neighbors	Arpack	k-means	2	0	3
Unsupervised and supervised metrics					
Silhouette score		Homogeneity	Completeness	V_measure	
0.518		0.351	0.374	0.362	

Table 3.9: Hyperparameters and performance metrics of spectral clustering with significant features.

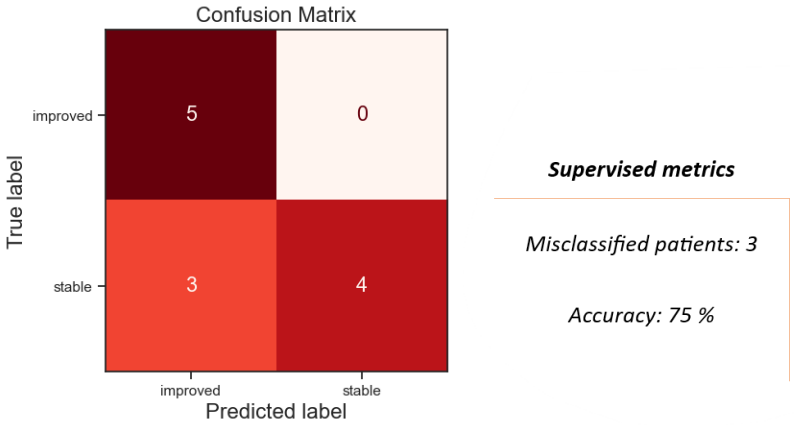


Figure 3.8: Significant features - Spectral clustering. a) Confusion matrix; b) supervised metrics.

Significant features - Gaussian mixture model

Hypermarameters			
Covariance_type	init_params	n_components	random_state
full	random	2	0
Unsupervised and supervised metrics			
Silhouette score	Homogeneity	Completeness	V_measure
0.369	0.004	0.006	0.005

Table 3.10: Hyperparameters and performance metrics of Gaussian mixture model with significant features.

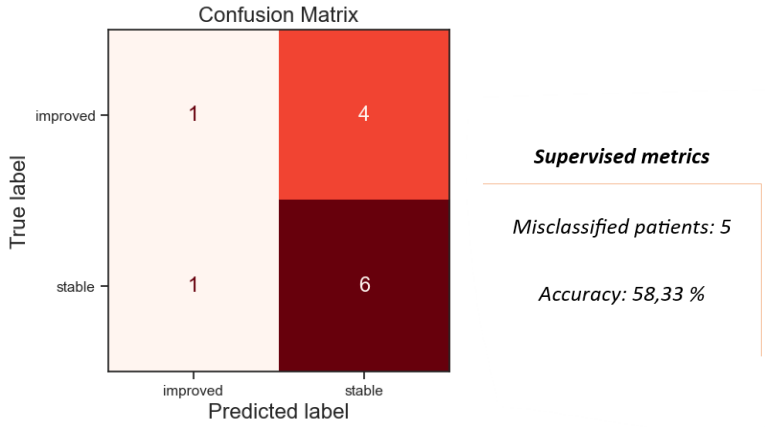


Figure 3.9: Significant features - Gaussian mixture model. a) Confusion matrix; b) supervised metrics.

Principal components

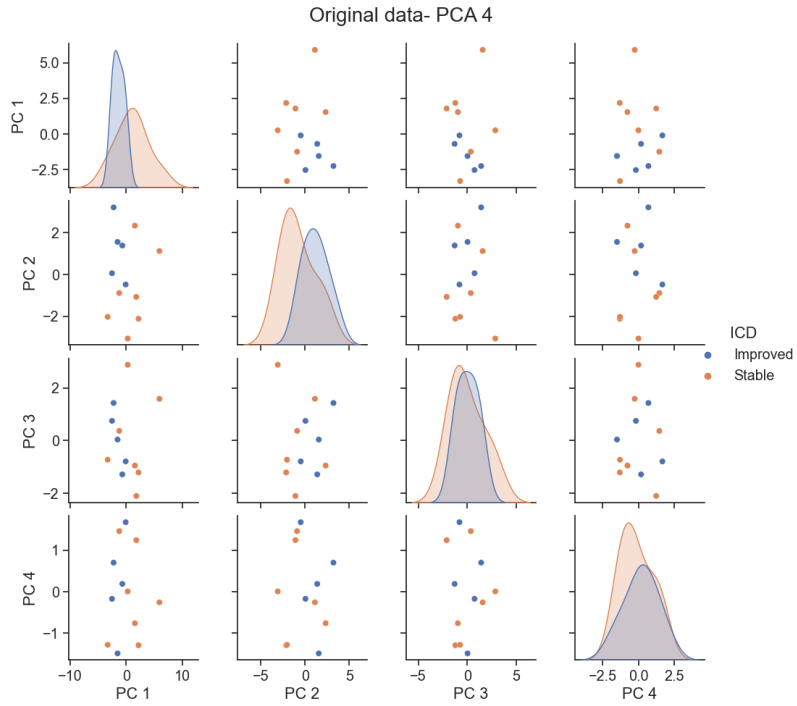


Figure 3.10: Pairwise plot with the first four PCs.

Principal components - Spectral clustering

Hypermarameters					
Affinity	Eigen_solver	Assign_labels	n_clusters	random_state	n_neighbors
nearest-neighbors	Arpack	discretize	2	0	3
Unsupervised and supervised metrics					
Silhouette score	Homogeneity	Completeness	V_measure		
0.200748	0.106493	0.106493	0.106493		

Table 3.11: Hyperparameters and performance metrics of spectral clustering with 4 PCs.

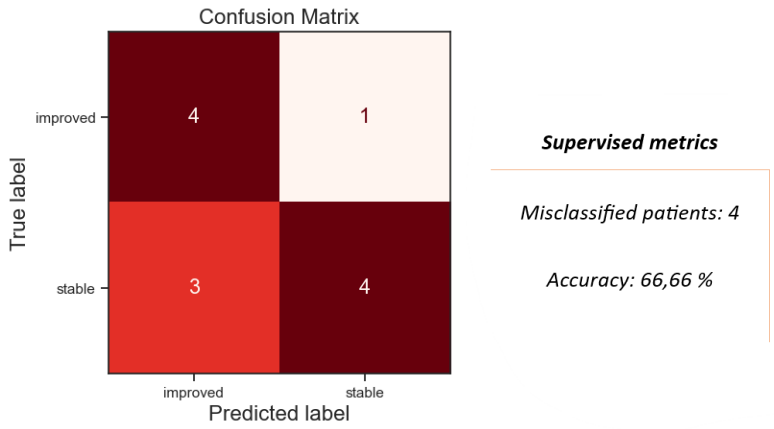


Figure 3.11: Principal components - Spectral clustering. a) Confusion matrix; b) supervised metrics.

Principal components - Gaussian mixture model

Hypermarameters			
Covariance_type	init_params	n_components	random_state
diag	k-means	2	0
Unsupervised and supervised metrics			
Silhouette score	Homogeneity	Completeness	V_measure
0.230	0.486	0.486	0.486

Table 3.12: Hyperparameters and performance metrics of Gaussian mixture model with 4 PCs.

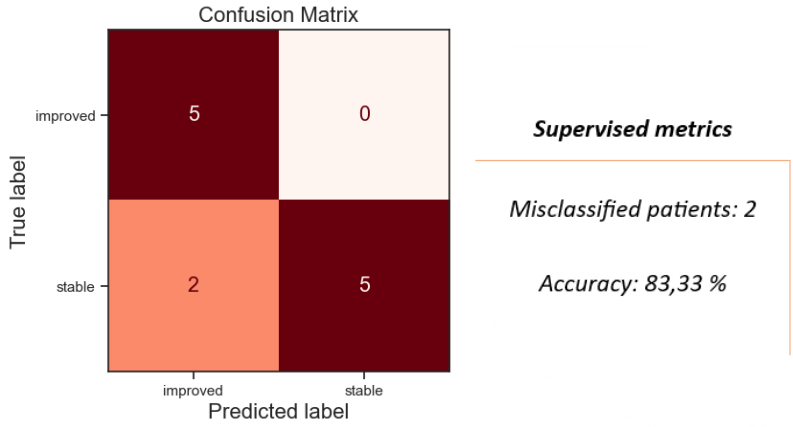


Figure 3.12: Principal components - Gaussian mixture model. a) Confusion matrix; b) supervised metrics.

3.2 Predicting outcome from surgical evaluation

3.2.1 Statistical differences on neural features within ICD+

Table 3.13 shows the comparison between the stable and improved groups, performed with the Mann-Whitney U test for continuous variables and the Fisher Exact test for categorical variables. The compared features are neural data recorded during surgery, and the line below the IBI separates continuous and binary variables. Significant differences are highlighted in orange.

		ICD stable vs ICD improved		
	Feature	p-value	Cohen's d	Cohen's d interpretation
Mann-Whitney U test with Holm-Bonferroni Correction	CV	0.911	0.011	small
	Firing rate	0.103	0.161	small
	Firing regularity	0.728	0.034	small
	Beta oscillation frequency	0.774	0.028	small
	Beta oscillation amplitude	***	0.350	small
	Ccbeta	0.323	0.079	small
	IBF (Hz)	0.075	0.176	small
	Theta oscillation frequency	0.04	0.197	small
	Theta oscillation amplitude	0.208	0.123	small
	Cctheta	0.246	0.084	small
	Burst duration (sec)	0.069	0.180	small
	IBI (sec)	0.003	0.291	small
Fisher Exact Test	Is tonic	0.523		
	Is bursting	0.659		
	Is irregular	0.842		
	Beta oscillatory	0.332		
	Theta oscillatory	0.293		

Table 3.13: Significantly different neural features between the improved and stable groups.

*** $p<0.001$

Beta oscillation amplitude, theta oscillation frequency and interburst interval (IBI) are significantly different between the two groups and have higher effect sizes than the other features, although they are small. The boxplots below show the distributions of these three features, comparing them between the improved group (in orange) and the stable group (in dark red). The p-value of the Mann-Whitney U test with Holm-Bonferroni correction is shown to highlight significant differences. Diamond-shaped dots are used to indicate the presence of outliers.

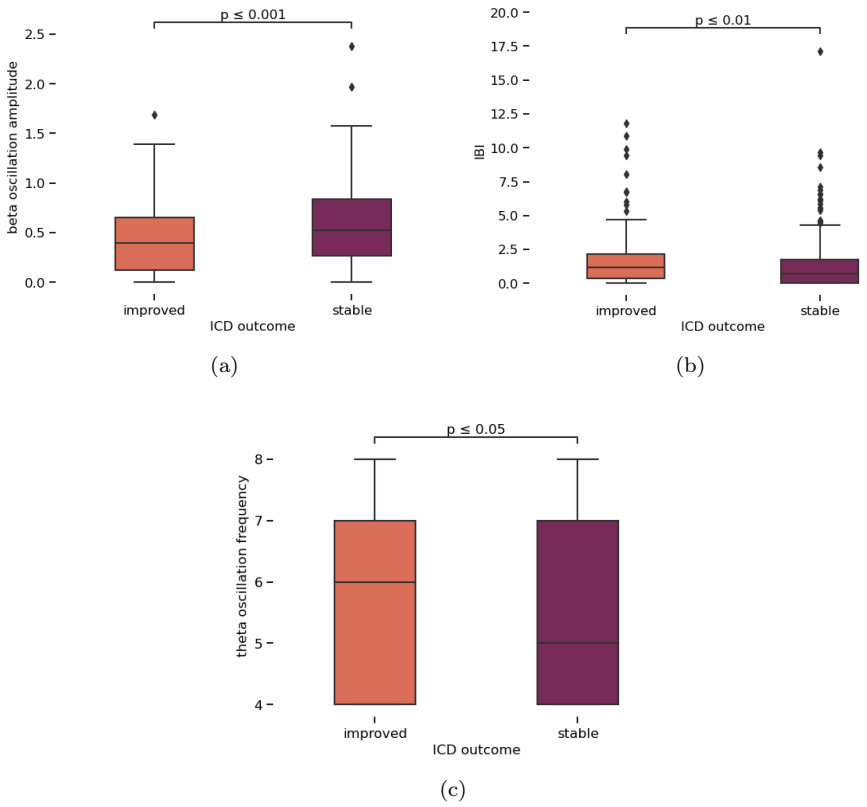


Figure 3.13: Boxplots of significantly different neural features between the improved and the stable groups. a) beta oscillation amplitude; b) interburst interval (IBI); c) theta oscillation frequency.

Boxplots reveal that the amplitude of beta oscillations is higher on average in the stable group, whereas the theta-band oscillation frequency is lower and intervals between bursts are shorter than in the group that recovered from ICD.

3.2.2 Prediction of the class of individual neurons based on MER features

The purpose of classifying the improved and stable groups at the time of intervention was pursued using the neural features of all neurons (412 in total). Three types of features selection were performed:

- Significant features with a p-value lower than 0.01: beta oscillation amplitude and IBI.
- Significant features with a p-value lower than 0.05: beta oscillation amplitude, IBI and theta oscillation frequency.
- Principal component analysis (PCA) was performed with all continuous neural features and a scree plot was displayed, as shown in Figure 3.14. Kaiser's rule was followed, as in PCA before clustering, and the first five principal components were retained:

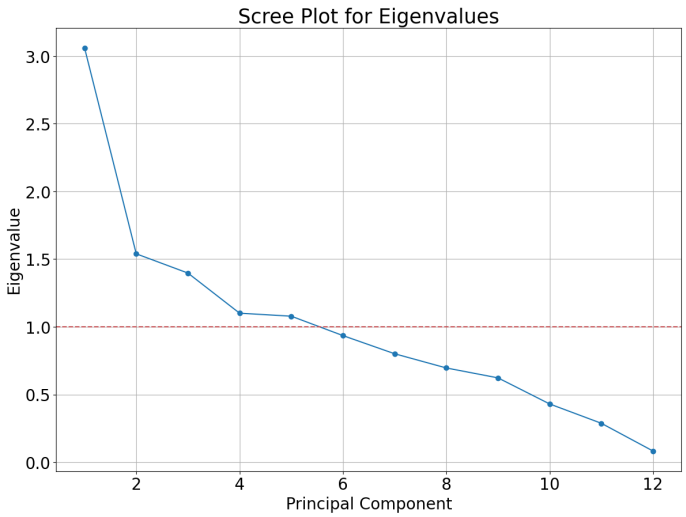


Figure 3.14: Scree plot for eigenvalues. The first five PCs have an eigenvalue greater than 1.

Figure 3.15 shows correlations between the original variables and the principal components, calculated as in the PCA performed before clustering.

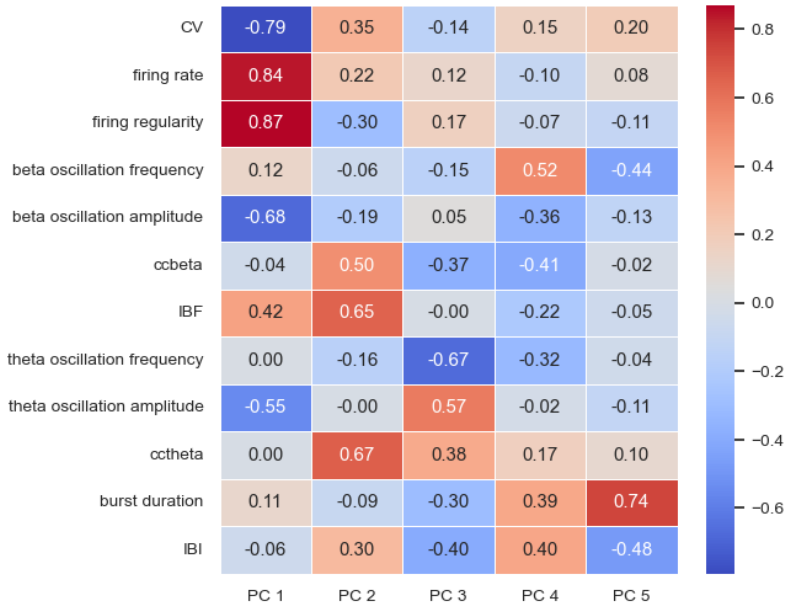


Figure 3.15: Heatmap representing correlations between neural features and the first five PCs.

It can be seen that among the significant features only the amplitude of beta oscillations is strongly correlated with PC 1. The coefficient of variation (CV) and the amplitude of theta oscillations are positively correlated with PC 1 whereas firing rate and firing regularity are negatively correlated.

Figure 3.16 shows the classification performances in terms of balanced accuracy and weighted AUC of each classifier tested and of the two voting classifiers. The voting classifiers were both composed of random forest and Gaussian processes classifier: one was constructed using significant features with a p-value lower than 0.01 (defined as Significant 1), whereas the other with all significant features (defined as Significant 2). Among all, the performance of random forest with Significant 1 stands out from others. Table 3.14 shows the selected hyperparameters of random forest and its performance, expressed in terms of accuracy, AUC, and F1 scores (described in 2.5.4). Figure 3.17 shows a confusion matrix normalized over the true condition (the sum of each row is 1).

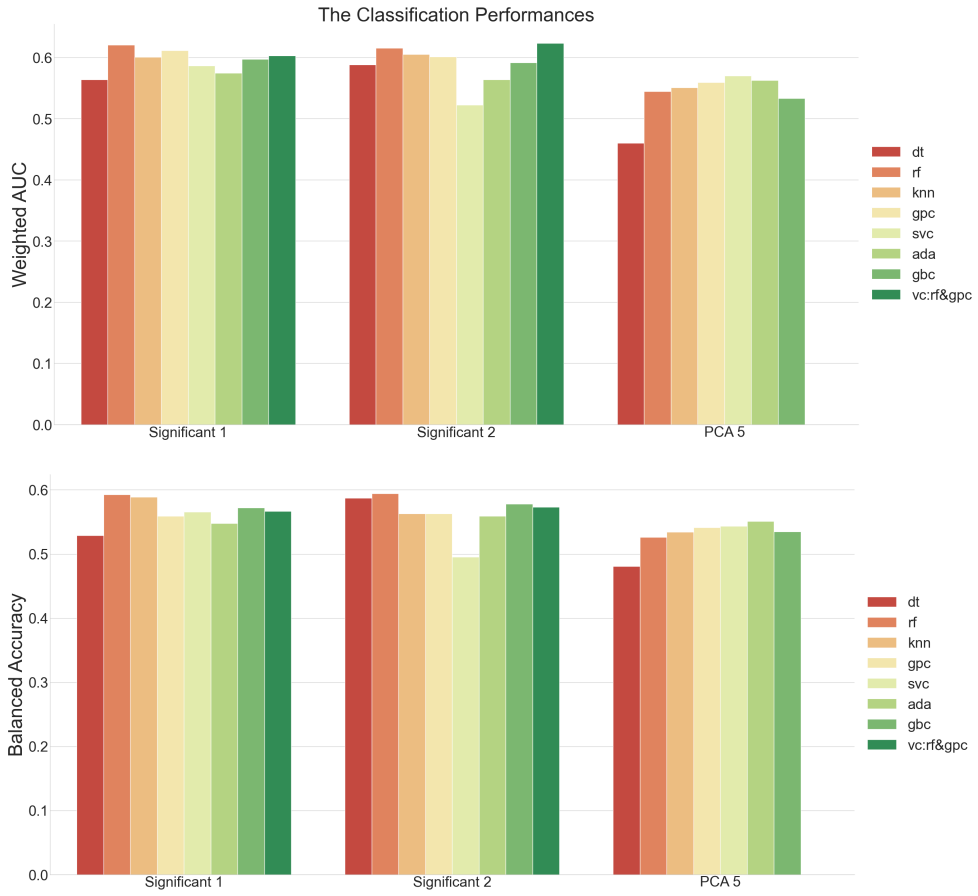


Figure 3.16: Weighted AUC and balanced accuracy of all classifiers. dt=decision tree; rf=random forest; knn=k-nearest neighbors; gpc=gaussian process classifier; ada=adaboost; gbc=gradient boosting classifier; vc=voting classifier.

Hypermarameters						
Criterion	Max depth	Max leaf nodes	Min samples leaf	Min samples split	N estimators	Random state
entropy	1	6	36	2	50	0
Metrics						
	Mean accuracy		Mean AUC		Mean F1	
	0.593		0.620		0.590	

Table 3.14: Random forest hyperparameters and performances.

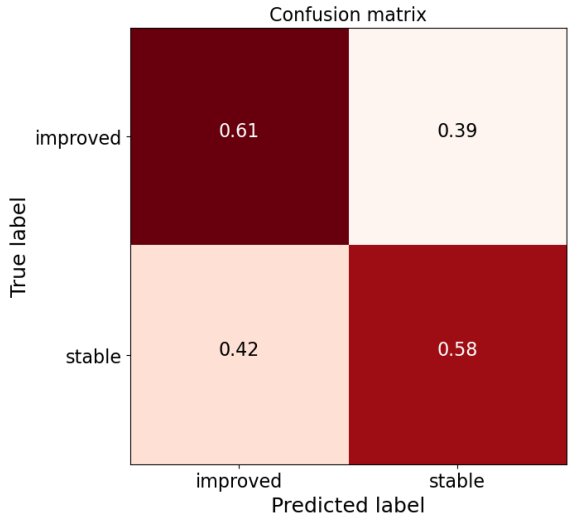
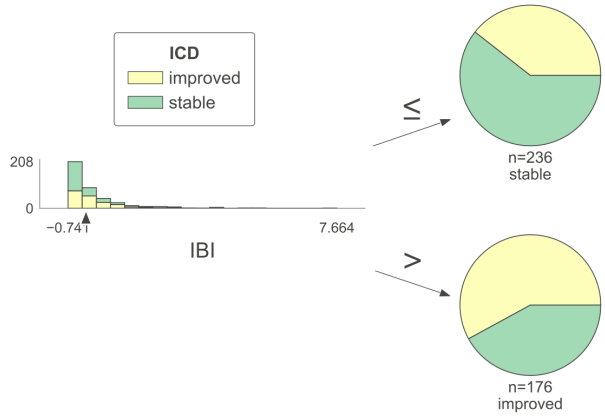


Figure 3.17: Normalized confusion matrix of all neurons' outcomes.

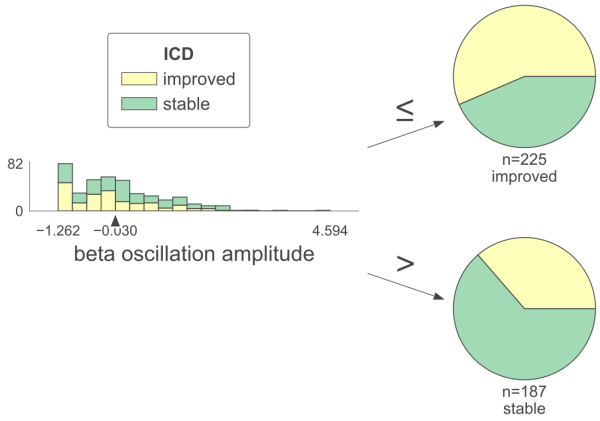
As shown in the normalized confusion matrix above, the random forest selected with features of beta oscillation amplitude and IBI is able to classify correctly 61% of the improved neurons and 58% of the stable ones.

3.2.3 Individual patient class prediction and post-classification analysis

To better understand the role of features in the random forest classifier, the Dtreeviz library was applied, which allows detailed visualization of individual decision trees. For each, the features selected as input and along the way are shown, the thresholds used in each partition and the results of the leaf nodes. Below there are two examples of decision tree estimators, one with IBI as the selected feature and the other with the amplitude of oscillations in the beta band.



(a) Decision tree n°60.

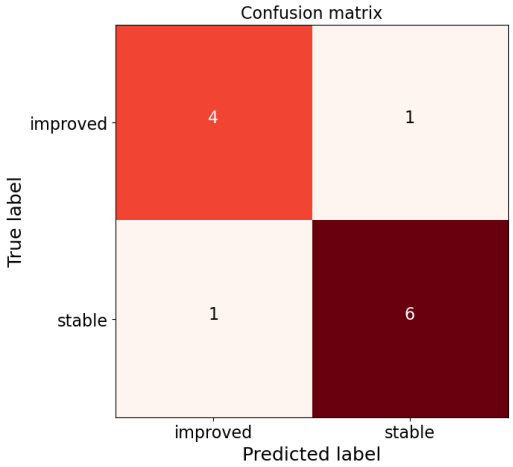


(b) Decision tree n°96.

Figure 3.18: Random forest estimators.

In both examples, the subset of features is composed of only one. Usually, individual trees can be more complicated with more branches and nodes, but in this case, the random forest has only two features overall, thus each decision tree is composed by selecting the first or second feature as input. In the first Figure 3.18(a), IBI is the selected feature and the threshold is set to -0.741 sec: if the neuron has longer intervals between bursts it is more likely to be an improved neuron, otherwise, it will be classified as a stable neuron (in this given decision tree). In the second case (Figure 3.18(b)), instead, if the neuron has an amplitude greater than -0.03 mV, it will be classified as stable. Figure 3.18 also shows the total number of neurons classified as improved or stable, whereas the pie charts indicate the actual ratio between the two groups.

Since the main interest is in classifying patients, the individual patient's prediction class is obtained by majority voting, using the classification of all neurons belonging to the patient. The results are reported in a confusion matrix, and the percentage of winning neurons for each patient is also shown next to it.



Y pred	Y true	Perc (%)
1	1	68.92
1	0	72.73
0	0	56.25
0	1	56.00
1	1	53.85
1	1	52.63
1	1	66.67
1	1	57.14
0	0	67.24
1	1	52.63
0	0	72.00

Figure 3.19: Prediction of patient class. On the left is the confusion matrix; on the right is the prediction class compared to the true class and the majority vote of neurons is expressed as a percentage of the total. Y pred=classifier prediction; Y true = true class; 0=improved; 1=stable.

From the confusion matrix, it can be seen that 10/12 patients are correctly classified and that the misclassified patients are balanced between the two classes. Looking at the percentage of winning neurons, it can be noticed that in some cases this is just over 50%, either in the case of the correct prediction or in one of the two cases of incorrect prediction.

The above results represent the classification obtained using all neurons. The same steps were applied using dorsal and ventral neurons (see subsection 2.5.1) but the results are not reported since they are similar to those obtained with all neurons.

3.3 Post-surgical analysis

To study the evolution of the patients after the surgery, a statistical analysis was conducted to compare the clinical features of the different groups at the 1-month follow-up, when the improved group recovered from ICD, and at the first year of follow-up.

3.3.1 1-month follow-up

At the 1-month follow-up, the Mann-Whitney U test with Holm-Bonferroni correction was performed to compare the differences between the stable and improved groups, in terms of drug doses (levodopa, dopamine agonists, and their total amount). On the right side of Table 3.15 a comparison of dose difference from the presurgical baseline was also performed.

	ICD stable vs ICD improved			ICD stable vs ICD improved (The difference from T0)		
Feature	p-value	Cohen's d	Cohen's d interpretation	p-value	Cohen's d	Cohen's d interpretation
Levodopa Dose	0.530	0.568	medium	0.343	0.530	medium
LEDD	0.530	0.185	small	0.755	0.412	small
DA LEDD	0.076	1.312	large	1	0.108	small

Table 3.15: Mann-Whitney U test with Holm-Bonferroni correction to compare the improved and stable groups at the 1-month follow-up.

Figure 3.20 shows the monthly doses of LEDD and DA from the time of surgery until the 1-year follow-up: absolute dose values are shown on the left, and relative doses at the presurgical baseline (T0) are shown on the right. Each line represents the dose trend of a specific patient, red lines for improved patients and blue lines for stable patients.

Table 3.15 shows that between the improved and stable groups, drug doses were not statistically different, both at the 1-month follow-up and in their difference from T0. As it can be seen from Figure 3.20 even in the first 12 months after surgery, the trend of LEDD and DA doses does not seem to be different between the improved and stable patients: in most cases, for both groups, there was a noticeable decrease in the first month, whereas in the following months, the doses increased until they stabilized, being on average lower than the preoperative doses, as is evident in the graphs to the right.

3.3.2 1-year follow-up

At the 1-year follow-up, the Mann-Whitney U test with Holm-Bonferroni correction was performed (unpaired test) to compare the differences in clinical features between the stable and improved groups. On the right side of Table 3.16 the differences between the two groups in the differences in characteristics between follow-up and baseline condition was also performed. Table 3.17 shows the same type of comparison between the ICD+ and ICD- groups. Table 3.18 and Table 3.19 show paired tests performed with Wilcoxon signed-rank test within each group (stable, improved, ICD+ and ICD-) to compare clinical features between the 1-year follow-up and the baseline condition.

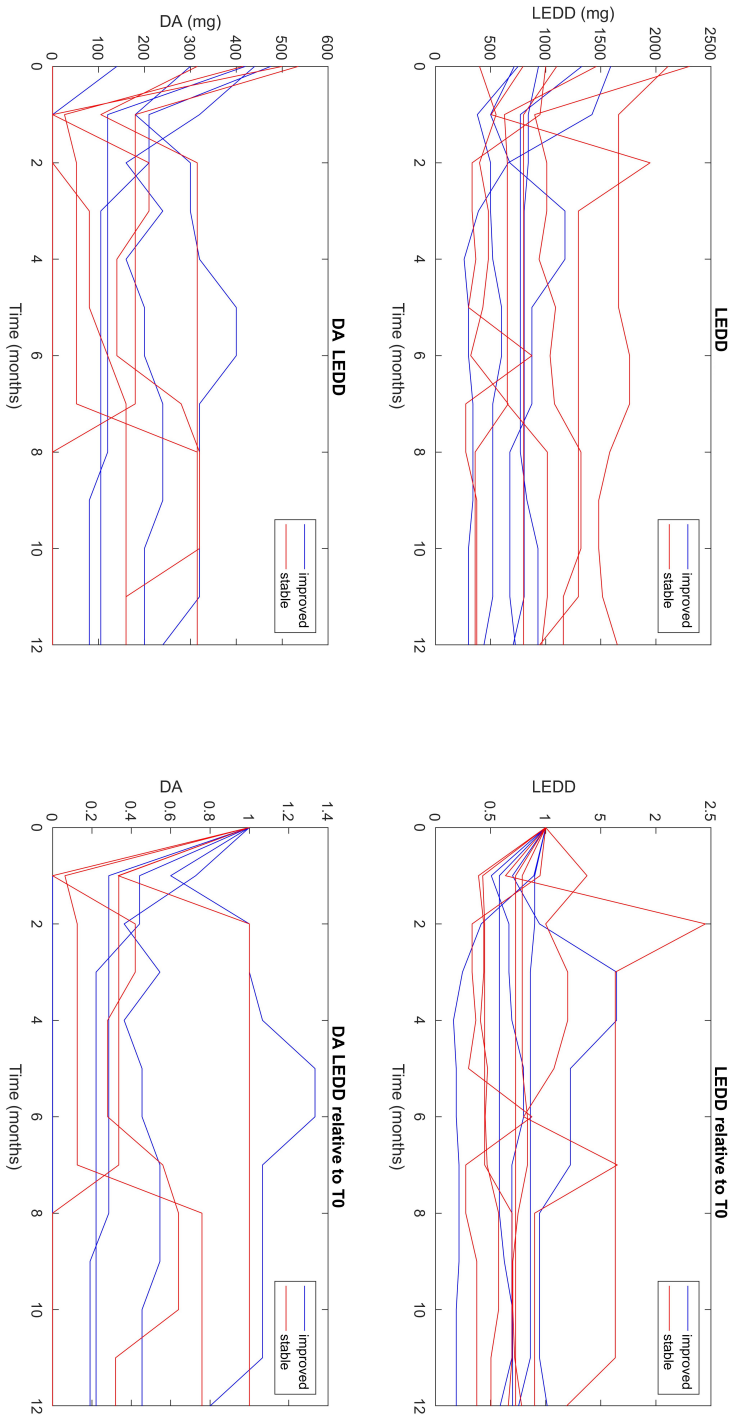


Figure 3.20: Monthly LEDD and DA doses.

Unpaired tests

Mann-Whitney U test with Holm-Bonferroni Correction	ICD stable vs ICD improved			ICD stable vs ICD improved (The difference from T0)		
Feature	p-value	Cohen's d	Cohen's d interpretation	p-value	Cohen's d	Cohen's d interpretation
Hoehn-Yahr Off condition	0.618	0.612	medium	0.559	0.369	small
Hoehn-Yahr On Condition	0.390	0.594	medium	0.805	0.156	small
Hoehn-Yahr On-Off Difference	0.734	0.267	small	1.000	0.148	small
UPDRS III Off	0.050	0.887	large	1.000	0.157	small
UPDRS III On	0.083	1.086	large	0.807	0.357	small
UPDRS III On-Off Difference	0.140	0.513	medium	0.684	0.466	small
Levodopa Dose	0.371	0.624	medium	0.569	0.209	small
LEDD	0.202	0.715	medium	0.876	0.060	small
DA LEDD	1.000	0.085	small	0.219	0.560	medium

Table 3.16: Unpaired tests between the improved and stable groups on clinical features at the 1-year follow-up and on their differences from T0.

Mann-Whitney U test with Holm-Bonferroni Correction	ICD+ vs ICD-			ICD+ vs ICD- (The difference from T0)		
Feature	p-value	Cohen's d	Cohen's d interpretation	p-value	Cohen's d	Cohen's d interpretation
Hoehn-Yahr Off condition	0.716	0.158	small	0.069	0.770	medium
Hoehn-Yahr On Condition	0.806	0.143	small	0.818	0.100	small
Hoehn-Yahr On-Off Difference	0.295	0.526	medium	0.129	0.679	medium
UPDRS III Off	0.043	0.924	large	0.406	0.141	small
UPDRS III On	0.288	0.355	small	0.455	0.326	small
UPDRS III On-Off Difference	0.055	0.909	large	0.433	0.343	small
Levodopa Dose	0.586	0.237	small	0.478	0.310	small
LEDD	0.135	0.703	medium	0.621	0.215	small
DA LEDD	0.937	0.034	small	0.555	0.257	small

Table 3.17: Unpaired tests between the ICD+ and ICD- groups on clinical features at the 1-year follow-up and on their difference from T0.

In both comparisons, only UPDRS III off is significantly different between the groups at the 1-year follow-up, and being different even at the baseline condition, there is no difference between groups in the difference between features collected before and after the intervention.

Paired tests

Wilcoxon signed-rank test
Holm-Bonferroni Correction

	ICD+			ICD-		
Feature	p-value	Cohen's d	Cohen's d interpretation	p-value	Cohen's d	Cohen's d interpretation
Hoehn-Yahr Off condition	0.158	0.764	medium	0.047	1.647	large
Hoehn-Yahr On Condition	1	0.121	small	1	0.329	small
Hoehn-Yahr On-Off Difference	0.209	0.841	large	0.047	1.841	large
UPDRS III Off	0.158	0.793	medium	0.446	0.615	medium
UPDRS III On	1	0.171	small	1	0.113	small
UPDRS III On-Off Difference	0.030	1.460	large	0.502	0.628	medium
Levodopa Dose	0.209	0.596	medium	0.068	1.371	large
LEDD	0.030	0.869	large	0.047	1.593	large
DA LEDD	0.158	0.970	large	0.048	1.404	large

Table 3.18: Paired tests within the ICD+ and ICD- groups on clinical features between baseline condition and the 1-year follow-up.

Mann-Whitney U test with
Holm-Bonferroni Correction

	ICD improved			ICD stable		
Feature	p-value	Cohen's d	Cohen's d interpretation	p-value	Cohen's d	Cohen's d interpretation
Hoehn-Yahr Off condition	0.820	0.894	large	0.529	0.826	large
Hoehn-Yahr On Condition	1	0.0	small	0.938	0.172	small
Hoehn-Yahr On-Off Difference	1	0.894	large	0.625	0.798	medium
UPDRS III Off	1	0.88	large	0.375	0.885	large
UPDRS III On	1	0	small	0.938	0.242	small
UPDRS III On-Off Difference	1	1.023	large	0.141	2.133	large
LEDD	0.875	1.375	large	0.375	0.712	medium
DA LEDD	0.563	1.944	large	0.921	0.580	medium
Levodopa Dose	1	0.670	medium	0.469	0.606	medium

Table 3.19: Paired tests within improved and stable groups on clinical features between baseline condition and the 1-year follow-up.

Paired tests in the improved and stable groups reveal that there are no significant differences between clinical features presurgery and those at the 1-year follow-up. However, there are significant differences within the ICD+ group in LEDD and UPDRS III off-on values, and within the ICD- group in the Hoehn-Yahr scale (off and off-on condition) in LEDD and DA LEDD values.

4 | Discussion

4.1 Conclusions

The purpose of this thesis was to understand the factors that led to recovery from ICD in some patients after DBS implant. We were indeed able to delineate the profile of the patients improving after DBS by clinical features acquired before the implant. Moreover, we found the relevant neural features in this process, shedding light on the subthalamic nucleus's role in decision-making. We finally investigated the role of the therapies at the 1-month and the 1-year follow-ups.

The improved group, indeed, showed significantly lower values than the stable group in UPDRS III off, UPDRS III off-on, and bradykinesia subscore, but the rigidity subscore also showed a low p-value (0.06) and a high effect size. Moreover, since the whole ICD+ group has significantly lower values in UPDRS III off than the ICD- group, this means that the improved patients were those who showed the least impairment of motor symptoms before the intervention. This condition was also maintained at the 1-year follow-up. Before the surgery, the difference between UPDRS III off (without drug activity), and UPDRS III on (with drug activity), is significantly higher in the improved patients, which could indicate greater drug sensitivity in the improved group.

Regarding drug doses, no significant differences were found between the two groups either before the intervention or at follow-ups. Although there was a decrease in doses in the first month after the intervention, it was not significantly different between the two groups. The results from clustering patients' profiles confirmed the relevance of UPDRS III off score and its subscales (bradykinesia and rigidity), UPDRS III off-on and Hoehn Yahr off scores: with these five features, Gaussian mixture model and spectral clustering achieved both an accuracy of 83,3% in separating the improved from the stable group.

Statistical analysis of neural features recorded during surgery, revealed a lower oscillation amplitude in the beta band, a higher frequency in the theta band, and a longer interval between bursts (IBI) in the improved group than in the stable group. The features just mentioned are generally associated with the presence of milder PD symptoms. Beta oscillation amplitude and IBI were used to train different classifiers to recognize individual improved or stable neurons. Among them, random forest shows the best results, being able to classify correctly 61% of the improved neurons and 58% of the stable ones. Neurons from each individual patient were used to assign the class to the patient by majority voting, reaching an accuracy value of 83,3%, as did clustering performance.

This section focuses on comparing the results found with other studies on ICDs in Parkinson's disease and the role of DBS in ICD development. First of all, it is important to consider the incidence of ICD in PD patients. In a huge cross-sectional study [148] was found that 13,6% of PD patients have at least 1 type of ICD. In our study, the percentage is higher (50%), but it depends on the inclusion and exclusion criteria adopted for patients undergoing DBS [149]. The higher prevalence of ICDs in the population of PD patients undergoing DBS [150] may be explained by higher dopamine treatment and long disease duration as inclusive criteria, which are risk factors for developing ICDs [151]. Post-intervention prevalence in our study is comparable to the incidence of 25.8% in a retrospective STN-DBS study: 29,2% at the 1-month follow-up and 33,3% at the three-year follow-up [151]. Numerous studies suggested that patients with ICD may improve their impulsive condition following deep brain stimulation (DBS) of the subthalamic nucleus (STN). This may be a consequence of reduced dopaminergic treatment or the direct effects of stimulation [152, 153]. In some cases, ICD may worsen or appear following STN-DBS (de novo ICD) [154, 155]. One of 24 patients developed de novo ICD 3 years after surgery. A hypothesis [156] suggested that one of the reasons may be a micro traumatic effect of the surgery, due to STN compact structure. It is indeed difficult to affect only the motor part without touching nearby structures involved in motivational behaviors. Another reason may be a wrong electrode location: stimulation of the dorsal STN is worse than stimulation of the ventral STN in terms of its effect on decision-making [157].

In a prospective study [158], the recovery from ICD is related to dopaminergic drug reduction in patients with milder PD, even though these patients are more likely to become apathetic after surgery. Similarly, in our case, the patients who improved had milder motor symptoms (UPDRS III off) at baseline condition and this difference persisted after the intervention, although less pronounced. The bradykinesia subscore reported this significant difference and also the rigidity subscore had a high effect size and a low p-value (0.06), whereas the tremor subscore did not differ between the improved and stable groups. Some studies have shown that PD patients with tremors are usually less cognitively impaired than PD patients with akinesia or gait dysfunction [159, 160]. Another study found that the rigid-akinetic form of PD is positively associated with compulsive shopping development [161].

On the other hand, dopaminergic treatment did not differ significantly between the improved and stable groups either at baseline condition or after DBS surgery. Paired tests revealed a decreasing trend in drug doses at the 1-year follow-up in both groups, although it was not significant. This contrasts with the hypothesis that the recovery from impulsive behavior following STN-DBS may be a consequence of the reduction in dopaminergic treatment [154]. Comparing dopaminergic treatment in the ICD+ and ICD- groups, instead, significantly lower postoperative LEDD was observed compared with preoperative LEDD in both groups. The same was also found regarding DA doses for the ICD+ group, whereas there was no evidence to support dose differences between the two groups. A recent study found the opposite [162], observing a difference between the groups and not an association with the pre or postoperative condition. However, it can be argued that the mismatch may be due to the way the study [162] was conducted, that is, trying to vary the dopaminergic treatment as little as possible over time.

Among neural features retrieved through microelectrode recordings (MERs) during surgery, beta oscillation amplitude, theta oscillation frequency and interburst interval (IBI) differ significantly between the improved and stable groups. Beta oscillation amplitude is lower in the improved group,

and it may be related to their milder symptoms. In fact, LFP registered in PD patients at rest without any treatment is characterized by an excess of beta activity, and their suppression means an improvement in motor symptoms [49]. IBI and theta oscillation frequency show significantly higher values in the improved group, which is in line with studies on local field potential in PD patients. Pathological beta activity, indeed, is characterized not only by greater amplitude but also by longer beta bursts, which are associated with disease severity in terms of motor symptoms [54, 55].

4.2 Limitations

There are some limitations in this study. First of all, the limited number of patients creates some difficulties in being able to generalize the results found, even if non-parametric tests were used. The small sample size adds to the way the data were collected: this is a retrospective study, which means that not all data were available. For example, the BIS was used to measure the degree of impulsivity and not the more common Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease-Rating Scale (QUIP-RS), which limits comparison with other studies. Moreover, the BIS was not available after surgery: it would have been interesting to evaluate the BIS at the 1-month and the 1-year follow-ups. It could have been determined in which domains of impulsivity the improved group healed and whether the stable group improved even if only slightly.

Another limitation regards the low resolution in the STN localization and the difficulty in distinguishing between the ventral and dorsal STN. Indeed, a better resolution would have provided insight into which part of the STN is most involved in recovery from ICDs. Finally, the chosen classification and clustering methods allow the identification of which features are most involved and in which direction (e.g., the improved group showed lower values on the UPDRS III scale) but do not allow for the definition of quantitative measures, which would be useful for clinicians to predict postoperative outcomes of impulsivity disorders.

4.3 Further studies

Some analyses could be added to overcome the above limitations and to complete the study. First of all, one could study the correlations between clinical and neural features, since, for example, a correlation between beta oscillations and motor symptoms has been found in literature [48]. Second, different classification methods could be tried to obtain mathematical expressions that can predict postoperative outcomes. Third, an analysis of impulsivity ratings with clinical scales at the 1-month and the 1-year follow-ups could be added. Moreover, one could consider the relationship between ICD and apathy, found in literature [158], to shed light on the complex hyper- and hypo-dopaminergic mechanisms. Finally, trained algorithms could be tried with new patients undergoing DBS to increase their robustness.

Widening the horizon of possibilities, it would be interesting to conduct a prospective study with anatomical localization of the STN performed using T2, SWI, and T1 sequences of the preoperative MRI with a higher resolution than 1.5T. A higher resolution would allow us to know the STN region where the DBS contacts are located and to extract neural features with their exact locations. A prospective study may follow several methodologies, for example, one might try not to vary dopaminergic treatment much in the postoperative period, thus having only one variable [162]. To

study the incidence of de novo ICD, it would be appropriate to conduct a study of more patients with worse impulsive conditions after DBS. In this way, it would be possible to achieve more comprehensive inclusion or exclusion criteria for undergoing DBS. Another possibility would be to consider each patient's entire PD history, not only preoperative profiles, to better distinguish between patients' progress over the years and the extent of DBS effects.

4.4 Clinical implications

The observation that patients who recover from ICD are those with milder PD conditions could lead to considering patients with ICD and moderate motor symptoms as inclusion criteria, despite their possible young age. The relevance of neural features in recognizing the improved group leads to a confirmation of the validity of closed-loop DBS with local field potential as a biomarker. The prediction results of clustering methods, having achieved the same accuracy as the classification method, may have clinical relevance. Indeed, they make it possible to analyze the patient's profile and predict future outcomes before surgery, thus better evaluating the procedure to be followed.

Bibliography

[1] Paul Cisek and John F Kalaska. "Neural mechanisms for interacting with a world full of action choices". In: *Annual review of neuroscience* 33 (2010), pp. 269–298.

[2] Peter Redgrave, Tony J Prescott, and Kevin Gurney. "The basal ganglia: a vertebrate solution to the selection problem?" In: *Neuroscience* 89.4 (1999), pp. 1009–1023.

[3] Hagai Bregman. *The hidden life of the basal Ganglia: at the base of brain and mind*. MIT Press, 2021.

[4] Jonathan W Mink. "The basal ganglia and involuntary movements: impaired inhibition of competing motor patterns". In: *Archives of neurology* 60.10 (2003), pp. 1365–1368.

[5] EAC Pereira and TZ Aziz. "Parkinson's disease and primate research: past, present, and future". In: *Postgraduate medical journal* 82.967 (2006), pp. 293–299.

[6] James Knierim,Ph.D., Department of Neuroscience, The Johns Hopkins University. *Neuroscience Online*. URL: https://nba.uth.tmc.edu/neuroscience/.

[7] John D Salamone and Mercè Correa. "The mysterious motivational functions of mesolimbic dopamine". In: *Neuron* 76.3 (2012), pp. 470–485.

[8] Tsuyoshi Yamaguchi et al. "Mesocorticolimbic glutamatergic pathway". In: *Journal of Neuroscience* 31.23 (2011), pp. 8476–8490.

[9] Bernard W Balleine, Mauricio R Delgado, and Okihide Hikosaka. "The role of the dorsal striatum in reward and decision-making". In: *Journal of Neuroscience* 27.31 (2007), pp. 8161–8165.

[10] Charles R Gerfen and D James Surmeier. "Modulation of striatal projection systems by dopamine". In: *Annual review of neuroscience* 34 (2011), pp. 441–466.

[11] Daniel Kahneman. *Thinking, fast and slow*. macmillan, 2011.

[12] Roger L Albin, Anne B Young, and John B Penney. "The functional anatomy of basal ganglia disorders". In: *Trends in neurosciences* 12.10 (1989), pp. 366–375.

[13] Weixing Shen et al. "Dichotomous dopaminergic control of striatal synaptic plasticity". In: *Science* 321.5890 (2008), pp. 848–851.

[14] J Parkinson. "An essay on the shaky palsy". In: *London: Sherwood, Neely and Jones* (1817), pp. 1–6.

[15] Douglas J Gelb, Eugene Oliver, and Sid Gilman. "Diagnostic criteria for Parkinson disease". In: *Archives of neurology* 56.1 (1999), pp. 33–39.

[16] Roberta Balestrino and AHV Schapira. "Parkinson disease". In: *European journal of neurology* 27.1 (2020), pp. 27–42.

[17] Seyed-Mohammad Fereshtehnejad and Ronald B Postuma. "Subtypes of Parkinson's disease: what do they tell us about disease progression?" In: *Current neurology and neuroscience reports* 17 (2017), pp. 1–10.

[18] Janice M Beitz. "Parkinson's disease: a review". In: *Frontiers in Bioscience-Scholar* 6.1 (2014), pp. 65–74.

[19] Lorraine V Kalia and Anthony E Lang. "Parkinson's disease". In: *The Lancet* 386.9996 (2015), pp. 896–912.

[20] Andrea Lee and Rebecca M Gilbert. "Epidemiology of Parkinson disease". In: *Neurologic clinics* 34.4 (2016), pp. 955–965.

[21] Thomas Gasser. "Genetics of Parkinson's disease". In: *Clinical genetics* 54.4 (1998), pp. 259–265.

[22] Anna Delamarre and Wassilios G Meissner. "Epidemiology, environmental risk factors and genetics of Parkinson's disease". In: *La Presse Medicale* 46.2 (2017), pp. 175–181.

[23] Charles B Breckenridge et al. "Association between Parkinson's disease and cigarette smoking, rural living, well-water consumption, farming and pesticide use: systematic review and meta-analysis". In: *PloS one* 11.4 (2016), e0151841.

[24] David K Simon, Caroline M Tanner, and Patrik Brundin. "Parkinson disease epidemiology, pathology, genetics, and pathophysiology". In: *Clinics in geriatric medicine* 36.1 (2020), pp. 1–12.

[25] Koichi Wakabayashi et al. "The Lewy body in Parkinson's disease and related neurodegenerative disorders". In: *Molecular neurobiology* 47 (2013), pp. 495–508.

[26] Maria Grazia Spillantini et al. "alpha-Synuclein in Lewy bodies". In: *Nature* 388.6645 (1997), pp. 839–840.

[27] PK Morrish et al. "Measuring the rate of progression and estimating the preclinical period of Parkinson's disease with [18F] dopa PET". In: *Journal of Neurology, Neurosurgery Psychiatry* 64.3 (1998), pp. 314–319.

[28] AH Schapira. "Jenner P". In: *Etiology and pathogenesis of Parkinson's disease. Movement Disorders* 26 (2011), pp. 1049–54.

[29] Hsiuchen Chen and David C Chan. "Mitochondrial dynamics–fusion, fission, movement, and mitophagy–in neurodegenerative diseases". In: *Human molecular genetics* 18.R2 (2009), R169–R176.

[30] Alicia M Pickrell and Richard J Youle. "The roles of PINK1, parkin, and mitochondrial fidelity in Parkinson's disease". In: *Neuron* 85.2 (2015), pp. 257–273.

[31] Daniel Erny et al. "Host microbiota constantly control maturation and function of microglia in the CNS". In: *Nature neuroscience* 18.7 (2015), pp. 965–977.

[32] JA Zimmer et al. *Lanabecestat: Neuroimaging results in early symptomatic Alzheimer's disease. Alzheimer's Dement Transl. Res. Clin. Inter. 7, 1–11.* 2021.

[33] Tianfeng Hao and Patricia Rockwell. "Signaling through the vascular endothelial growth factor receptor VEGFR-2 protects hippocampal neurons from mitochondrial dysfunction and oxidative stress". In: *Free Radical Biology and Medicine* 63 (2013), pp. 421–431.

[34] DIMITRI P. AGAMANOLIS, M. D. *Neuropathology-Chapter 9 DEGENERATIVE DISEASES*. URL: https://neuropathology-web.org/chapter9/chapter9dPD.html#pd.

[35] Joseph Jankovic. "Parkinson's disease: clinical features and diagnosis". In: *Journal of neurology, neurosurgery & psychiatry* 79.4 (2008), pp. 368–376.

[36] AJ Jagadeesan et al. "Current trends in etiology, prognosis and therapeutic aspects of Parkinson's disease: a review". In: *Acta Bio Medica: Atenei Parmensis* 88.3 (2017), p. 249.

[37] Francois JG Vingerhoets et al. "Which clinical sign of Parkinson's disease best reflects the nigrostriatal lesion?" In: *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* 41.1 (1997), pp. 58–64.

[38] C Lozza, RM Marie, and Jean-Claude Baron. "The metabolic substrates of bradykinesia and tremor in uncomplicated Parkinson's disease". In: *Neuroimage* 17.2 (2002), pp. 688–699.

[39] James A Cooper et al. "Slowed central processing in simple and go/no-go reaction time tasks in Parkinson's disease". In: *Brain* 117.3 (1994), pp. 517–529.

[40] Joseph Jankovic, Kenneth S Schwartz, and William Ondo. "Re-emergent tremor of Parkinson's disease". In: *Journal of Neurology, Neurosurgery & Psychiatry* 67.5 (1999), pp. 646–650.

[41] Emmanuel Broussolle et al. "Contribution of Jules Froment to the study of parkinsonian rigidity". In: *Movement disorders* 22.7 (2007), pp. 909–914.

[42] David R Williams, Hilary C Watt, and Andrew J Lees. "Predictors of falls and fractures in bradykinetic rigid syndromes: a retrospective study". In: *Journal of Neurology, Neurosurgery & Psychiatry* 77.4 (2006), pp. 468–473.

[43] JD Schaafsma et al. "Characterization of freezing of gait subtypes and the response of each to levodopa in Parkinson's disease". In: *European journal of neurology* 10.4 (2003), pp. 391–398.

[44] Magdalena Sabate et al. "Obstructive and restrictive pulmonary dysfunctions in Parkinson's disease". In: *Journal of the neurological sciences* 138.1-2 (1996), pp. 114–119.

[45] Simon-Shlomo Poil et al. "Critical-state dynamics of avalanches and oscillations jointly emerge from balanced excitation/inhibition in neuronal networks". In: *Journal of Neuroscience* 32.29 (2012), pp. 9817–9823.

[46] Andrea A Kühn et al. "High-frequency stimulation of the subthalamic nucleus suppresses oscillatory β activity in patients with Parkinson's disease in parallel with improvement in motor performance". In: *Journal of Neuroscience* 28.24 (2008), pp. 6165–6173.

[47] A Priori et al. "Rhythm-specific pharmacological modulation of subthalamic activity in Parkinson's disease". In: *Experimental neurology* 189.2 (2004), pp. 369–379.

[48] John-Stuart Brittain and Peter Brown. "Oscillations and the basal ganglia: motor control and beyond". In: *Neuroimage* 85 (2014), pp. 637–647.

[49] Ashwini Oswal et al. "Deep brain stimulation modulates synchrony within spatially and spectrally distinct resting state networks in Parkinson's disease". In: *Brain* 139.5 (2016), pp. 1482–1496.

[50] Christos Tsiokos et al. "Pallidal low β-low γ phase-amplitude coupling inversely correlates with Parkinson disease symptoms". In: *Clinical Neurophysiology* 128.11 (2017), pp. 2165–2178.

[51] Jon López-Azcárate et al. "Coupling between beta and high-frequency activity in the human subthalamic nucleus may be a pathophysiological mechanism in Parkinson's disease". In: *Journal of Neuroscience* 30.19 (2010), pp. 6667–6677.

[52] Neil J Vickers. "Animal communication: when i'm calling you, will you answer too?" In: *Current biology* 27.14 (2017), R713–R715.

[53] Megan H Trager et al. "Subthalamic beta oscillations are attenuated after withdrawal of chronic high frequency neurostimulation in Parkinson's disease". In: *Neurobiology of Disease* 96 (2016), pp. 22–30.

[54] Gerd Tinkhauser et al. "Beta burst dynamics in Parkinson's disease OFF and ON dopaminergic medication". In: *Brain* 140.11 (2017), pp. 2968–2981.

[55] Gerd Tinkhauser et al. "The modulatory effect of adaptive deep brain stimulation on beta bursts in Parkinson's disease". In: *Brain* 140.4 (2017), pp. 1053–1067.

[56] Roxanne Lofredi et al. "Beta bursts during continuous movements accompany the velocity decrement in Parkinson's disease patients". In: *Neurobiology of disease* 127 (2019), pp. 462–471.

[57] M Alegre et al. "Changes in subthalamic activity during movement observation in Parkinson's disease: is the mirror system mirrored in the basal ganglia?" In: *Clinical Neurophysiology* 121.3 (2010), pp. 414–425.

[58] DAVID Devos et al. "Predominance of the contralateral movement-related activity in the subthalamo-cortical loop". In: *Clinical neurophysiology* 117.10 (2006), pp. 2315–2327.

[59] Andrea A Kühn et al. "Modulation of beta oscillations in the subthalamic area during motor imagery in Parkinson's disease". In: *Brain* 129.3 (2006), pp. 695–706.

[60] Huiling Tan et al. "Frequency specific activity in subthalamic nucleus correlates with hand bradykinesia in Parkinson's disease". In: *Experimental neurology* 240 (2013), pp. 122–129.

[61] Gaia Giannicola et al. "The effects of levodopa and deep brain stimulation on subthalamic local field low-frequency oscillations in Parkinson's disease". In: *Neurosignals* 21.1-2 (2013), pp. 89–98.

[62] Fortes Alonso-Frech et al. "Slow oscillatory activity and levodopa-induced dyskinesias in Parkinson's disease". In: *Brain* 129.7 (2006), pp. 1748–1757.

[63] Maria C Rodriguez-Oroz et al. "Involvement of the subthalamic nucleus in impulse control disorders associated with Parkinson's disease". In: *Brain* 134.1 (2011), pp. 36–49.

[64] WRG Gibb and AJ Lees. "Gibb 1988 UK Brain Bank Criteria". In: *J. Neurol. Neurosurg. Psychiatry* 51 (1988), pp. 745–752.

[65] Margaret M Hoehn, Melvin D Yahr, et al. "Parkinsonism: onset, progression, and mortality". In: *Neurology* 50.2 (1998), pp. 318–318.

[66] Józef A Opara et al. "Motor assessment in Parkinsons disease". In: *Annals of Agricultural and Environmental Medicine* 24.3 (2017).

[67] Christopher G Goetz et al. "Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations the Movement Disorder Society Task Force on rating scales for Parkinson's disease". In: *Movement disorders* 19.9 (2004), pp. 1020–1028.

[68] SRLE Fahn. "Unified Parkinson's disease rating scale". In: *Recent developments in Parkinson's disease* (1987), pp. 153–163.

[69] Robert S Schwab. "Projection technique for evaluating surgery in Parkinson's disease". In: *Third symposium on Parkinson's disease*. E&S Livingstone. 1969, pp. 152–157.

[70] Christopher G Goetz et al. "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results". In: *Movement disorders: official journal of the Movement Disorder Society* 23.15 (2008), pp. 2129–2170.

[71] Roongroj Bhidayasiri and Pablo Martinez-Martin. "Clinical assessments in Parkinson's disease: scales and monitoring". In: *International review of neurobiology* 132 (2017), pp. 129–182.

[72] Arvid Carlsson, Margit Lindqvist, and TOR Magnusson. "3, 4-Dihydroxyphenylalanine and 5-hydroxytryptophan as reserpine antagonists". In: *Nature* 180 (1957), pp. 1200–1200.

[73] Hornykiewicz Ehringer and Oleh Hornykiewicz. "Verteilung von Noradrenalin und Dopamin (3-Hydroxytyramin) im Gehirn des Menschen und ihr Verhalten bei Erkrankungen des extrapyramidalen Systems". In: *Klinische Wochenschrift* 38 (1960), pp. 1236–1239.

[74] George C Cotzias, Melvin H Van Woert, and Lewis M Schiffer. "Aromatic amino acids and modification of parkinsonism". In: *New England Journal of Medicine* 276.7 (1967), pp. 374–379.

[75] Marwan Hariz and Patric Blomstedt. "Deep brain stimulation for Parkinson's disease". In: *Journal of internal medicine* 292.5 (2022), pp. 764–778.

[76] Joachim K Krauss et al. "Technology of deep brain stimulation: current status and future directions". In: *Nature Reviews Neurology* 17.2 (2021), pp. 75–87.

[77] J Lawrence Pool. "Psychosurgery in older people". In: *Journal of the American Geriatrics Society* 2.7 (1954), pp. 456–466.

[78] CW Sem-Jacobsen. "Depth-electrographic observations related to Parkinson's disease. Recording and electrical stimulation in the area around the third ventricle." In: *Journal of neurosurgery* 24.1 (1966), Suppl–388.

[79] AL Benabid et al. "Chronic VIM thalamic stimulation in Parkinson's disease, essential tremor and extra-pyramidal dyskinesias". In: *Advances in Stereotactic and Functional Neurosurgery 10: Proceedings of the 10 th Meeting of the European Society for Stereotactic and Functional Neurosurgery Stockholm 1992*. Springer. 1993, pp. 39–44.

[80] Jessica Frey et al. "Past, present, and future of deep brain stimulation: hardware, software, imaging, physiology and novel approaches". In: *Frontiers in Neurology* 13 (2022).

[81] Jean Siegfried and Bodo Lippitz. "Bilateral chronic electrostimulation of ventroposterolateral pallidum: a new therapeutic approach for alleviating all parkinsonian symptoms." In: *Neurosurgery* 35.6 (1994), pp. 1126–9.

[82] Patricia Limousin et al. "Effect on parkinsonian signs and symptoms of bilateral subthalamic nucleus stimulation". In: *The Lancet* 345.8942 (1995), pp. 91–95.

[83] U.S. Food and Drug Administration. *Premarket Approval (PMA)*. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?ID=P960009.

[84] U.S. Food and Drug Administration. *Premarket Approval (PMA)*. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?ID=P960009S007.

[85] Aristide Merola et al. "New frontiers for deep brain stimulation: directionality, sensing technologies, remote programming, robotic stereotactic assistance, asleep procedures, and connectomics". In: *Frontiers in Neurology* 12 (2021), p. 694747.

[86] Cameron C McIntyre et al. "Cellular effects of deep brain stimulation: model-based analysis of activation and inhibition". In: *Journal of neurophysiology* 91.4 (2004), pp. 1457–1469.

[87] Günther Deuschl et al. "Stimulation of the subthalamic nucleus at an earlier disease stage of Parkinson's disease: concept and standards of the EARLYSTIM-study". In: *Parkinsonism & related disorders* 19.1 (2013), pp. 56–61.

[88] Zixiao Yin et al. "Local field potentials in Parkinson's disease: A frequency-based review". In: *Neurobiology of Disease* 155 (2021), p. 105372.

[89] Dustin A Heldman et al. "Telehealth management of Parkinson's disease using wearable sensors: an exploratory study". In: *Digital biomarkers* 1.1 (2017), pp. 43–51.

[90] David C Wong et al. "Supervised classification of bradykinesia for Parkinson's disease diagnosis from smartphone videos". In: *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE. 2019, pp. 32–37.

[91] Aristide Merola et al. "New frontiers for deep brain stimulation: directionality, sensing technologies, remote programming, robotic stereotactic assistance, asleep procedures, and connectomics". In: *Frontiers in Neurology* 12 (2021), p. 694747.

[92] Aparna Wagle Shukla et al. "DBS programming: an evolving approach for patients with Parkinson's disease". In: *Parkinson's Disease* 2017 (2017).

[93] Lucia K Feldmann et al. "Subthalamic beta band suppression reflects effective neuromodulation in chronic recordings". In: *European Journal of Neurology* 28.7 (2021), pp. 2372–2377.

[94] Gaia Giannicola et al. "Subthalamic local field potentials after seven-year deep brain stimulation in Parkinson's disease". In: *Experimental neurology* 237.2 (2012), pp. 312–317.

[95] Joohi Jimenez-Shahed. "Device profile of the percept PC deep brain stimulation system for the treatment of Parkinson's disease and related disorders". In: *Expert Review of Medical Devices* 18.4 (2021), pp. 319–332.

[96] Suneil K Kalia, Tejas Sankar, and Andres M Lozano. "Deep brain stimulation for Parkinson's disease and other movement disorders". In: *Current opinion in neurology* 26.4 (2013), pp. 374–380.

[97] Marie Abbes et al. "Subthalamic stimulation and neuropsychiatric symptoms in Parkinson's disease: results from a long-term follow-up cohort study". In: *Journal of Neurology, Neurosurgery & Psychiatry* 89.8 (2018), pp. 836–843.

[98] Matteo Vissani, Ioannis U Isaias, and Alberto Mazzoni. "Deep brain stimulation: a review of the open neural engineering challenges". In: *Journal of neural engineering* 17.5 (2020), p. 051002.

[99] Samuel B Guze. "Diagnostic and statistical manual of mental disorders, (DSM-IV)". In: *American Journal of Psychiatry* 152.8 (1995), pp. 1228–1228.

[100] Liana Schreiber, Brian L Odlaug, and Jon E Grant. "Impulse control disorders: updated review of clinical characteristics and pharmacological management". In: *Frontiers in psychiatry* 2 (2011), p. 1.

[101] Jon E Grant and Marc N Potenza. "Escitalopram treatment of pathological gambling with co-occurring anxiety: an open-label pilot study with double-blind discontinuation". In: *International clinical psychopharmacology* 21.4 (2006), pp. 203–209.

[102] DSM-5 Task Force American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: DSM-5™ (5th ed.)* Am Psychiatric Assoc, 2013.

[103] Valerie Voon and Susan H Fox. "Medication-related impulse control and repetitive behaviors in Parkinson disease". In: *Archives of neurology* 64.8 (2007), pp. 1089–1096.

[104] G Giovannoni et al. "Hedonistic homeostatic dysregulation in patients with Parkinson's disease on dopamine replacement therapies". In: *Journal of Neurology, Neurosurgery & Psychiatry* 68.4 (2000), pp. 423–428.

[105] Daniel Weintraub et al. "Clinical spectrum of impulse control disorders in Parkinson's disease". In: *Movement Disorders* 30.2 (2015), pp. 121–127.

[106] Santiago Perez-Lloret et al. "Prevalence and pharmacological factors associated with impulse-control disorder symptoms in patients with Parkinson disease". In: *Clinical neuropharmacology* 35.6 (2012), pp. 261–265.

[107] AW Willis et al. "Epidemiology and neuropsychiatric manifestations of Young Onset Parkinson's Disease in the United States". In: *Parkinsonism & related disorders* 19.2 (2013), pp. 202–206.

[108] Ingrid de Chazeron et al. "Hypersexuality and pathological gambling in Parkinson's disease: A cross-sectional case–control study". In: *Movement disorders* 26.11 (2011), pp. 2127–2130.

[109] Daniel Weintraub et al. "Validation of the questionnaire for impulsive-compulsive disorders in Parkinson's disease". In: *Movement disorders: official journal of the Movement Disorder Society* 24.10 (2009), pp. 1461–1467.

[110] Patricia de la Riva et al. "Course of psychiatric symptoms and global cognition in early Parkinson disease". In: *Neurology* 83.12 (2014), pp. 1096–1103.

[111] Pieter L du Toit et al. "Comparison of obsessive-compulsive disorder patients with and without comorbid putative obsessive-compulsive spectrum disorders using a structured clinical interview". In: *Comprehensive psychiatry* 42.4 (2001), pp. 291–300.

[112] Claire Ardouin et al. "Assessment of hyper-and hypodopaminergic behaviors in Parkinson's disease". In: *Revue neurologique* 165.11 (2009), pp. 845–856.

[113] Jim H Patton, Matthew S Stanford, and Ernest S Barratt. "Factor structure of the Barratt impulsiveness scale". In: *Journal of clinical psychology* 51.6 (1995), pp. 768–774.

[114] Ernest S Barratt. "Factor analysis of some psychometric measures of impulsiveness and anxiety". In: *Psychological reports* 16.2 (1965), pp. 547–554.

[115] Ernest S Barratt. "Impulsiveness subtraits: Arousal and information processing". In: *Motivation, emotion and personality* 99 (1985), pp. 137–146.

[116] Steven P Reise et al. "The Barratt Impulsiveness Scale–11: Reassessment of its structure in a community sample." In: *Psychological assessment* 25.2 (2013), p. 631.

[117] Lynne Steinberg et al. "New tricks for an old measure: The development of the Barratt Impulsiveness Scale–Brief (BIS-Brief)." In: *Psychological assessment* 25.1 (2013), p. 216.

[118] G Giovannoni et al. "Hedonistic homeostatic dysregulation in patients with Parkinson's disease on dopamine replacement therapies". In: *Journal of Neurology, Neurosurgery & Psychiatry* 68.4 (2000), pp. 423–428.

[119] Stéphane Thobois et al. "Non-motor dopamine withdrawal syndrome after surgery for Parkinson's disease: predictors and underlying mesolimbic denervation". In: *Brain* 133.4 (2010), pp. 1111–1127.

[120] Paolo Amami et al. "Impulse control behaviours in patients with Parkinson's disease after subthalamic deep brain stimulation: de novo cases and 3-year follow-up". In: *Journal of Neurology, Neurosurgery & Psychiatry* 86.5 (2015), pp. 562–564.

[121] Chayut Kasemsuk, Genko Oyama, and Nobutaka Hattori. "Management of impulse control disorders with deep brain stimulation: a double-edged sword". In: *Journal of the neurological sciences* 374 (2017), pp. 63–68.

[122] Federico Micheli et al. "Impulsivity markers in parkinsonian subthalamic single-unit activity". In: *Movement Disorders* 36.6 (2021), pp. 1435–1440.

[123] Matteo Vissani et al. "Impulsivity is associated with firing regularity in parkinsonian ventral subthalamic nucleus". In: *Annals of Clinical and Translational Neurology* 9.4 (2022), pp. 552–557.

[124] Ahmet Kaymak et al. "Patterns of neural activity and clinical outcomes in a juvenile Huntington's Disease patient undergoing Deep Brain Stimulation of the subthalamic nucleus". In: *Deep Brain Stimulation* 1 (2023), pp. 15–19.

[125] Samuel Sanford Shapiro and Martin B Wilk. "An analysis of variance test for normality (complete samples)". In: *Biometrika* 52.3/4 (1965), pp. 591–611.

[126] Frank Wilcoxon. *Individual comparisons by ranking methods*. Springer, 1992.

[127] Henry B Mann and Donald R Whitney. "On a test of whether one of two random variables is stochastically larger than the other". In: *The annals of mathematical statistics* (1947), pp. 50–60.

[128] Ronald A Fisher. "The logic of inductive inference". In: *Journal of the royal statistical society* 98.1 (1935), pp. 39–82.

[129] Sture Holm. "A simple sequentially rejective multiple test procedure". In: *Scandinavian journal of statistics* (1979), pp. 65–70.

[130] Jacob Cohen. *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

[131] Jiawei Han, Jian Pei, and Hanghang Tong. *Data mining: concepts and techniques*. Morgan kaufmann, 2022.

[132] Henry F Kaiser. "The application of electronic computers to factor analysis". In: *Educational and psychological measurement* 20.1 (1960), pp. 141–151.

[133] Andrew Ng, Michael Jordan, and Yair Weiss. "On spectral clustering: Analysis and an algorithm". In: *Advances in neural information processing systems* 14 (2001).

[134] Ulrike Von Luxburg. "A tutorial on spectral clustering". In: *Statistics and computing* 17 (2007), pp. 395–416.

[135] X Yu Stella and Jianbo Shi. "Multiclass spectral clustering. In null". In: *IEEE* 5 (2003), p. 313.

[136] Anil Damle, Victor Minden, and Lexing Ying. "Simple, direct and efficient multi-way spectral clustering". In: *Information and Inference: A Journal of the IMA* 8.1 (2019), pp. 181–203.

[137] Sebastian Raschka. *Python machine learning*. Packt publishing ltd, 2015.

[138] Nitesh V Chawla et al. "SMOTE: synthetic minority over-sampling technique". In: *Journal of artificial intelligence research* 16 (2002), pp. 321–357.

[139] imbalanced-learn.org. *Compare over-sampling samplers*. URL: https://imbalanced-learn.org/dev/auto_examples/over-sampling/plot_comparison_over_sampling.html.

[140] Leo Breiman. "Random forests". In: *Machine learning* 45 (2001), pp. 5–32.

[141] Friedman Breiman. "Olshen, and Stone". In: *Classification and Regression trees* (1984).

[142] Jon Louis Bentley. "Multidimensional binary search trees used for associative searching". In: *Communications of the ACM* 18.9 (1975), pp. 509–517.

[143] Stephen M Omohundro. *Five balltree construction algorithms*. International Computer Science Institute Berkeley, 1989.

[144] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. "A training algorithm for optimal margin classifiers". In: *Proceedings of the fifth annual workshop on Computational learning theory*. 1992, pp. 144–152.

[145] Corinna Cortes and Vladimir Vapnik. "Support-vector networks". In: *Machine learning* 20 (1995), pp. 273–297.

[146] Carl Edward Rasmussen and Christopher K Williams. *Gaussian processes for machine learning, vol. 1*. 2006.

[147] Jerome H Friedman. "Greedy function approximation: a gradient boosting machine". In: *Annals of statistics* (2001), pp. 1189–1232.

[148] Daniel Weintraub et al. "Impulse control disorders in Parkinson disease: a cross-sectional study of 3090 patients". In: *Archives of neurology* 67.5 (2010), pp. 589–595.

[149] Robert S Eisinger et al. "Medications, deep brain stimulation, and other factors influencing impulse control disorders in Parkinson's disease". In: *Frontiers in Neurology* 10 (2019), p. 86.

[150] Lucy Gee et al. "The influence of bilateral subthalamic nucleus deep brain stimulation on impulsivity and prepulse inhibition in Parkinson's disease patients". In: *Stereotactic and Functional Neurosurgery* 93.4 (2015), pp. 265–270.

[151] Young Eun Kim et al. "Impulse control and related behaviors after bilateral subthalamic stimulation in patients with Parkinson's disease". In: *Journal of Clinical Neuroscience* 20.7 (2013), pp. 964–969.

[152] Eugénie Lhommée et al. "Subthalamic stimulation in Parkinson's disease: restoring the balance of motivated behaviours". In: *Brain* 135.5 (2012), pp. 1463–1477.

[153] Eugénie Lhommée et al. "Behavioural outcomes of subthalamic stimulation and medical therapy versus medical therapy alone for Parkinson's disease with early motor complications (EARLYSTIM trial): secondary analysis of an open-label randomised trial". In: *The Lancet Neurology* 17.3 (2018), pp. 223–231.

[154] Paolo Amami et al. "Impulse control behaviours in patients with Parkinson's disease after subthalamic deep brain stimulation: de novo cases and 3-year follow-up". In: *Journal of Neurology, Neurosurgery & Psychiatry* 86.5 (2015), pp. 562–564.

[155] Chayut Kasemsuk, Genko Oyama, and Nobutaka Hattori. "Management of impulse control disorders with deep brain stimulation: a double-edged sword". In: *Journal of the neurological sciences* 374 (2017), pp. 63–68.

[156] M Sensi et al. "Explosive-aggressive behavior related to bilateral subthalamic stimulation". In: *Parkinsonism & related disorders* 10.4 (2004), pp. 247–251.

[157] HMM Smeding et al. "Pathological gambling after bilateral subthalamic nucleus stimulation in Parkinson disease". In: *Journal of Neurology, Neurosurgery & Psychiatry* 78.5 (2007), pp. 517–519.

[158] Marie des Neiges Santin et al. "Impact of subthalamic deep brain stimulation on impulse control disorders in Parkinson's disease: a prospective study". In: *Movement Disorders* 36.3 (2021), pp. 750–757.

[159] Magdalena Eriksson Domellöf, Eva Elgh, and Lars Forsgren. "The relation between cognition and motor dysfunction in drug-naive newly diagnosed patients with Parkinson's disease". In: *Movement Disorders* 26.12 (2011), pp. 2183–2189.

[160] Joo Young Oh et al. "Relationship between clinical phenotypes and cognitive impairment in Parkinson's disease (PD)". In: *Archives of gerontology and geriatrics* 49.3 (2009), pp. 351–354.

[161] Mauricio Iván Garcıa-Rubio et al. "Analysis of impulse control disorders (ICDs) and factors associated with their development in a parkinson's disease population". In: *Healthcare*. Vol. 9. 10. MDPI. 2021, p. 1263.

[162] P Justin Rossi et al. "Measures of impulsivity in Parkinson's disease decrease after DBS in the setting of stable dopamine therapy". In: *Parkinsonism & related disorders* 44 (2017), pp. 13–17.