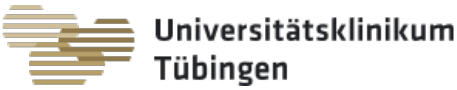
Post-operative DBS Implantation Recording Analysis & data-driven Assessment of Biomarkers

A laboratory report submitted by
Fabian Klopfer, M.Sc.

Supervised by
Prof. Dr. Alireza Gharabaghi



Institute for Neuromodulation & -technology
University Hospital Tübingen
and
Neural & Behavioral Science
Graduate Training Centre of Neuroscience
Faculty of Medicine
University of Tübingen

Tübingen, 2022

Contents

1 Parkinson's Disease

Parkinson's disease (PD) (ICD-10-CM G20 [30] is one of the most prevalent neurodegenerative diseases with a prevalence increasing monotonically with age, with an average overall prevalence of 315 in 100000 [31] between 1985 and 2010 worldwide. The incidence in the USA is 108 per 100,000 [43], while it is approximately 84.1 per 100000 in Germany [18] (2018). Overall there are approximately 6 milion patients suffering from PD in 2016 which is more than double than in 1990 (2.5 milion) [10].

Symptoms can be categorized in motor and non-motor symptoms. Motor symptoms include tremor, muscular rigidity, slowness of movement, mask-like facial expressions, slurrish speech, a lack of postural reflexes among others. Non-motor symptoms are for example sleeping problems, comorbidities with depression and anxiety. The disease trajectory can be classified into stages depending on the presence and severity of symptoms e.g. using the Hoehn and Yahr scale [20].

PD can be assessed by different means. One of them is the Movement Disorder Society's version of the Unified PD Rating Scale (MDS-UPDRS) [15]. It consists of 4 parts dealing with

I. non-motor aspects of daily living like mood, depression, pain, sleeping problems,

II. motor aspects of daily living like Speech, eating, dressing, hygiene, handwriting,

III. examination of motor symptoms by a clinican with and without the administration of medication like rigidity, handmovements and speech,

IV. treatment-related motor complications like dyskinesia.

Another possibility is to do single photon emission computed tomography (SPECT) with a contrast agent marking dopamine (DA) reuptake transporters. In a patient with PD the DA transporter density is reduced in the substantia nigra pars compacta (SNc) [41].

The latter option is based on the cellular and molecular mechanisms of PD. Gradually the patients dopamine-producing/melanin containing neurons in the SNc die causing a lack of DA in the nigrostriatal pathway projecting to the dorsal part of the striatum.

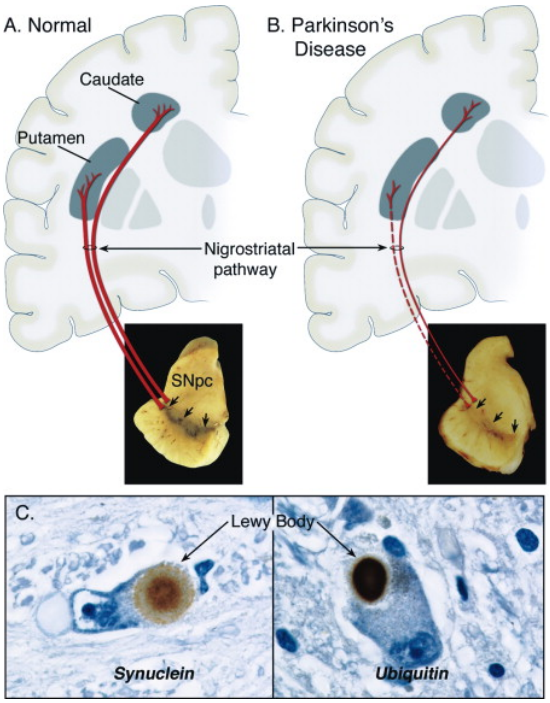


Figure 1 ***A. & B.*** *SNc and projections to the dorsal striatum in healthy subjects and patients with PD. In the healthy subject, the SNc is still highly pigmented due to the melanin-containing dopamine-producing cells being intact. The SNc projects to the striatum and delivers normal amounts of DA into the basal ganglia (BG) circuit. If the dopamine-producing neurons undergo apoptosis, the pigmentation decreases and so does the amount of dopamine adimnistered to the striatum.* ***C.*** *Photomicrographs of Lewy bodies.*

A broad schematic of this is illustrated in 1 A and B. Additionally Lewy bodies (LB) accumulate in the SNc but also in other areas (e.g. locus coeruleus, amygdala, basal nc. of Maynert and cortex), spreading increasingly [7, 3]. A photomicrograph of LB is shown in 1 C. Lately, two sub-types have been hypothesized: the PNS-first and the CNS-first subtypes. The authors describe the gut-first variant only where LB spread through retrograde transport via the vagus nerve first to the dorsal motor nucleus in stage I associated with REM sleep behavior disorder (RBD) and only spread to

the SNc in stage III. The CNS-first PD variant does not show RBD in the early/prodormal stage of the disease and the spread of LB originates in the CNS and spreads anterogradely to the PNS [2, 23].

The implications of the degradation of the dopamine-producing neurons in the SNc become apparent when considering the nigrostriatal pathway and the basal ganglia and associated motor cortex circuits, depicted in 2. We can see that the SNc projects to the basal ganglia, more precisely to the striatum (among others) [21].

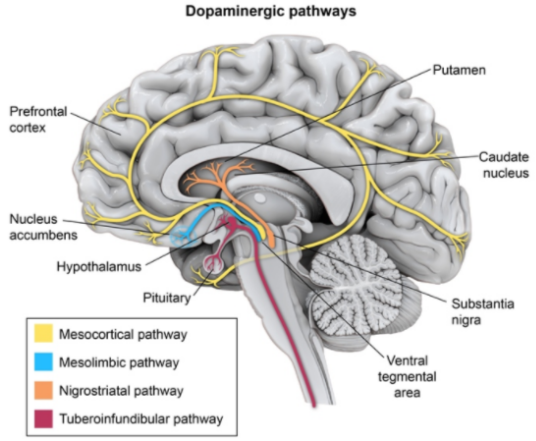


Figure 2 *The dopaminergic pathways in the CNS. The mesocortical pathway projects from the ventral tegmental area (VTA) to the cortex (yellow). The mesolimbic patheay projects from the VTA to the ventral striatum including the nucleus accumbens and is the major pleasure and value system (along with the orbitofrontal cortex, the ventromedial cortex and the dorsal anterior cingulate cortex) (blue). The nigrostriatal system projects from the SNc to the dorsal striatum (orange). Note that the nigrostriatal pathway is also involved in action selection and reward [21]. Finally, the tuberoinfundibular pathway projects from the hypothalamus to the pituary gland and the Hyopthalamocortical pathway, which connects the Hypothalamus to the ciliospinal center, which controls e.g. pupil dialation in the ipsilateral eye.*

We can also see, that mainly the basal ganglia are affected, in comparison to other limbic and cortical structures. For an overview of the precise effects on taget region receptor densities, gene expression and firing rate abnormalities, see table 13 in the appendix.

When we take a closer look at the basal ganglia circuits shown in 3, we can see that the implications of less dopaminergic innervation affects the excitation-inhibition balance in the circuit via three pathways:

- the direct pathway originating from the cortex inhibitory via the putamen inhibitory ending in the globus pallidus pars interna (GPi) and the substantia nigra pars reticulata which in turn excite the ventrolateral nucleus of the thalamus (VL),
- the indirect pathway again starting inhibitorry from the cotex to the putamen again inhibitory, via the globus pallidus pars externa (GPe) and the STN (sequentially) both inhibitory to the GPi and the SNr excitatory, finally GPi and SNr again inhibit the VL, and
- the hyperdirect pathway from the cortex to the STN which excites the SNr and GPi again inhibiting the VL.

All three pathways ultimately output to the VL inhibitory via the GPi and the SNr, which finally feeds back excitatory to the cortex [27]. In a healthy patient, these three pathways interact to establish a homeostasis with respect to excitation, and inhibition, mediated by DA [21]. One model which tries to explain this homeostasis is called the brake accelerator model [16]. It suggests that the direct pathway causes less inhibition of the VL via D1 receptors exciting the striatum, inhibiting the GPi & SNr i.e. the initiation of movement and acts as an accelerator. The indirect pathway leads to more inhibition of the VL mediated by D2 receptors in the striatum inhibiting GPi, disinhibiting the STN, exciting GPi & SNr, thus acting as a break. Together the "break" and the "accelerator" balance each other out as neccessary depending on the cortical (and other) signals.

If the DA supply decreases with the degradation of DA-ergic SNc neurons, two things happen:

1. in the direct pathway the putamen is stimulated by DA which in turn inhibits the GPi and SNr. When DA decreases, the putamens inhibitory effect on the GPi and

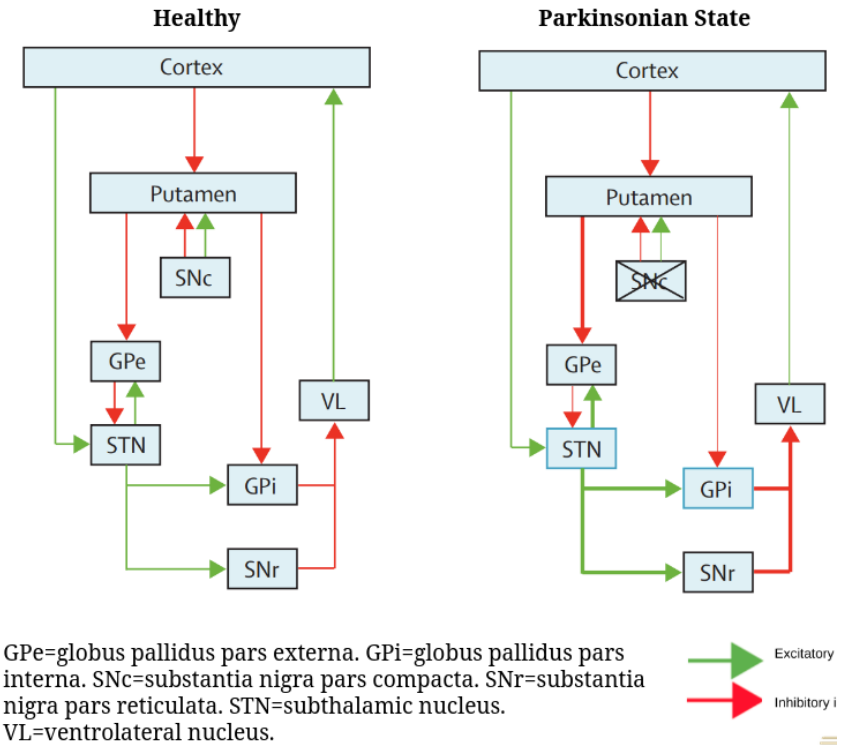


Figure 3 *The basal ganglia circuitry in a healthy state with intact homeostatic control of the VL (left) and in patients with disturbed homeostasis due to disturbed DA signalling in PD [32].*

SNr decreases too, leading to higher activity in these two structures, and

2. in the indirect pathway, the putamen is inhibited by DA. Thus the inhibitory effect of the putamen on the GPe decreases and in turn the inhibitory effect of the GPe on the STN is decreased and the STN in turn excites the GPi and SNr a lot stronger.

Overall the lacking inhibition of GPi & SNr along with the additional excitation causes GPi and SNr to inhibit the VL more than in healthy subjects along with decreased cortical activity (due to VLs decreased excitatory effect) [32].

As the VL is a central input and output site for motor circuits — as shown for the motor output circuit in figure 14 — this explains many of the motor symptoms of PD quite well. Other symptoms might be explained transiently, like bradykinesia: in PD patients off medication (see later in the text), the hyperactive STN exhibits strong activity in the beta band around 20Hz, which propagates through basal ganglia, thalamic and cortical structure causing increased neural synchronity. Hypokinesia might be explained by the overinhibition of brain stem and downstream central pattern generators [32]. This leads to lack of time-locked inhibition of the GPi and in turn to a lack of motor neurons which are recruited for intended movements [32].

However, when considering especially the non-motor symptoms, we have to take a closer look at the sub-networks of the basal ganglia, shown in figure 4 [32]. First of all, the motor-part of the BG circuitry is organized somatotopically as the motor cortex [32], illustrated in the appendix in figure 15. Besides the gradient induced by somatotopy, there is another organizational scheme which is governed by the cortical targets: the limbic associative and motor sub-networks.

For example, when stimulating the premotor cortex magnetically increases excitability in both healthy controls and PD patiens with medication on. In contrast, patients with PD without medication do not exhibit this excitability, suggesting that the associative circuit shows is altered behavior, too. The precise mechanisms remain unclear [32].

Executive dysfunctions have been shown to be associated with the loss of the nigrostriatal

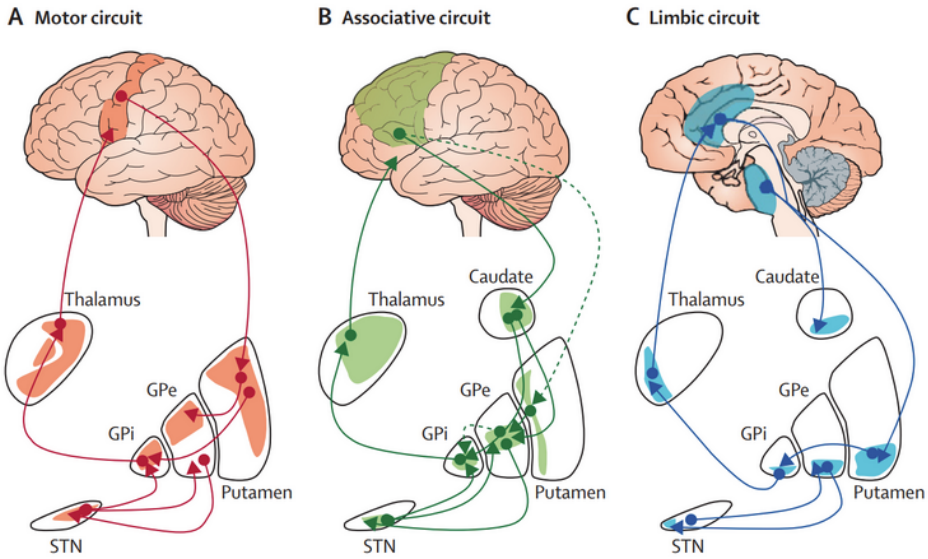


Figure 4 *Sub-networks of the basal ganglia by projection target.*

pathways [33], indicating, that the associative circuit is affected.

Other neurotransmitters than DA also play a role. Acetylcholine (ACh) is in balance with dopamine in healthy subjects. In PDs the reduction of DA is accompanied by an increase of ACh [29]. For example giant, aspiny cholinergic interneurons are lost in late PD [12]. These neurons among other functions modulate DA secretion [36]. As ACh modulates wakefulness & sleep, arrousal among others [21], its deregulation is one possible mechamism for some of the cognitive functions [24].

An extensive review of the anatomy, dynamics, pathology and lesions of the basal ganglia circuitry is laid out by J.W. Mink [27] and by Rodriguez-Oroz et al. [32].

A mixture of genetic and environmental factors is hypothesized as the cause of PD but so far the exact etiology of PD remains unknown [1]. Genetic factors are for examle Pink1 and Parkin [38, 14, 8], mitochondrial proteins inducing autophagy to clear out depolarized or damaged mitochondria. If these damaged mitochondria are not cleared out by auto-/mitophagy, they release Cytochrome C (CytC) and AIF (apoptotic protease) through the mitochondrial permeability transition pore facilitated by the voltage dependent anion channel. CytC and AIF trigger apoptosis by activating Caspases [4]. Since LB induce oxidative stress and promote mitochondrial dysfunction and the neurons in the SNc already have a higher oxidative stress level due to the breakdown of DA using monoamine oxidase (MAO), this leads to the decreased survival rate of these dopamine producing neurons [6].

Another example for a potentially PD inducing mutation is the SNCA gene which encodes α-synuclein. Especially mutations with locus at the membrane bound part causes changes in protein conformation which leads to aggregations and ultimately to the formation of LBs [21]. Finally, lysosome mutations are associated with PD. Lysosomes are crucially involved in degenerating proteins marked for disassembly by ubiquitin-regulated autophagy. As α-synuclein degeneration is mainly done in lysosomes, a disfunction leads to aggregation of α-synucleins.

Environmental factors could be e.g. overexposure to pesticides like rotenone, some heavy metals & metalloids, and traumatic brain injury [1].

General (i.e. not sub-group specific) gene therapy has four different means to mitigate the effects causing PD [25]:

- Protection: neurotrophic factors are ensuring cell survival. AAVs loaded with e.g. GDNF encodng the glial-derived neurotrophic factor or NTRN encoding neuturin are injected into the Putamen, where

they are transported antero- and retrogradely to other connected structures — like the SNc.

- Modulation: the GAD_{65} gene encodes glutamic acid decarboxylase which catalyzes the production of GABA and CO_2 from glutamic acid (Glu). AAVs containing the GAD_{65} gene are injected in the subthalamic nucleus (STN) where it decreases the amount of Glu and increases the amount of GABA. Thus it reduces the amount of excitation within and adjacent areas, while it increases the amount of inhibtion within the STN. Why this is effective is explained later when we come to deep brain stimulatin (DBS).

- Restoration of DA synthesis rate: the AADC gene encodes aromatic L-amino acid decarboxylase, the rate limiting enzyme that catalyzes the synthesis from L-DOPA to DA. Further, the GTPCH gene encodes GTP cyclohydroxylase which is the first rate-limiting enzyme for the synthesis of tetrahydrobiopterin, an essential cofactor required by the enzyme aromatic amino acid hydroxylase for the synthesis of monoamine neurotransmitters including catecholamines, thus dopamine and the reaction of phenylalanine to tyrosine which is the first step in DA/catecholamine production. Additionally tetrahydrobiopterin binds free radicals and thus has an antioxidative function and prevents ferroptosis. The later is hypothesized to be involved in the cell death of the DA-producing neurons of the SNc [17].

- Enhancement of lysosomal function: as α-synuclein are degenerated in lysosomes, a reduced functioning of lysosomes leads to the aggregation of α-synucleins. Thus enhancing the lysosomal functioning increases the clearance of α-synuclein preventing or at least reducing aggregations and ultimately the formation of LB.

Pharmacological agents are listed in figure 16 in the appendix. Most pharmacological treatments rely on the DA system. L-DOPA is one of the most important medications. It is the direct precursor to DA and is thus metabolized to DA in one step catalyzed by an aromatic amino acid decarboxylase (AADC). DA agonists can be used instead.

Another type of medication is monoamine oxidase ingibitors (MAO-I), or catechol-O-methyltransferase inhibitors (COMT-I) which are the two molecules involved in the two step breakdown of catecholamines like DA.

Anti-cholinerigc drugs help to reestablish the DA-ACh ratio; amantadine's mechanisms are not fully understood as it affects many sites in the brain but include an increase of DA, DA reuptake reduction and anti-cholinergic effects. Clozapine is used to treat possible psychotic symptoms when occuring due to L-DOPA medication, i.e. too high DA levels in the mesolimbic & mesocortical pathways when the disease hasn't progressed enough to seriously affect the VTA with Lewy bodies.

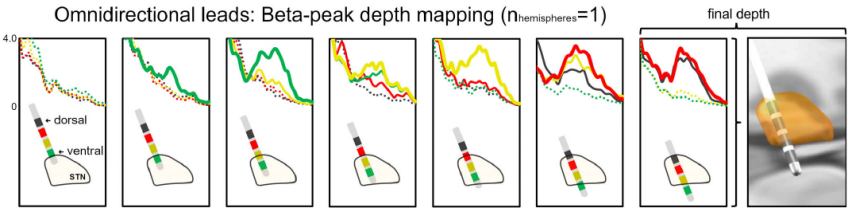
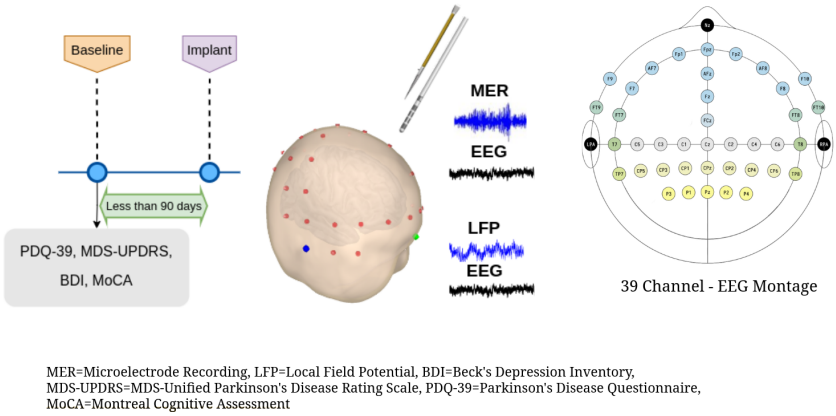


Figure 5 *Stepwise insertion of the deep brain stimulation electrode. PSDs recorded per depth (left to right) and contact (colors). When a contact enters the STN, there is a peak in the higher part of the β band [26].*



MER=Microelectrode Recording, LFP=Local Field Potential, BDI=Beck's Depression Inventory,
MDS-UPDRS=MDS-Unified Parkinson's Disease Rating Scale, PDQ-39=Parkinson's Disease Questionnaire,
MoCA=Montreal Cognitive Assessment

Figure 6 *An overview of all data that is acquired per patient for DBS electrode implantation and post-operative evaluation.*

Deep brain stimulation (DBS) is an effective treatment option for patients who experience motor symptoms that are not adequately controllable with medication. It involves the implantation of an electrode, which is connected to a battery-powered device that delivers electrical stimulation. DBS has been shown to significantly improve motor symptoms, quality of life, and reduce medication-related side effects in individuals with PD [34].

The implantation targets vary based on the profile of the patient. Usually the electrode is placed in the hyperactive STN, which is then inhibited by high-frequency stimulation [34]. The surpression of the STN decreases its excitatory effect on the SNr & GPi and thus the amount of inhibition acted upon the VL, as visible in figure 7 . Gait issues can be decreased by implanting the first contact into the SNr [42].

The patients are assessed 90 days before the surgery using two different PD scales including the MDS-UPDRS, with the Beck depression index and the Montreal cognitive assessment. Before the surgery, the clincal team plans the implantation using frame-based stereotactic surgery [22]. With CT and T_1-MRI the position and the angle for insertion along with the target depth are infered. During the surgery, with all surgical equippment in place, the DBS electrode is inserted at the position and angle. Depthwise, the electrode is positioned approximately 10mm before the target region. Stepwise, the signals picked up by the contacts is meassured and a PSD is calculated. Increased β-oscillations are a hallmark of PD and thus a good selective biomarker [35]. However, β-oscillations are not PD-specific [40]. Never the less, about 80% of the patients show this elevated β component, thus it is a reasonable biomarker of the STN. If the PSD shows a rise in beta activity, this indicates, that the electrode reached the STN. If this rise is not present, the electrode is probably outside of the STN and the electrode is lowered by a milimeter. An illustration of this is shown in figure 5 Additional measurements like micro-electrode measurements can be done to observe spiking behavior to delineate the structural borders

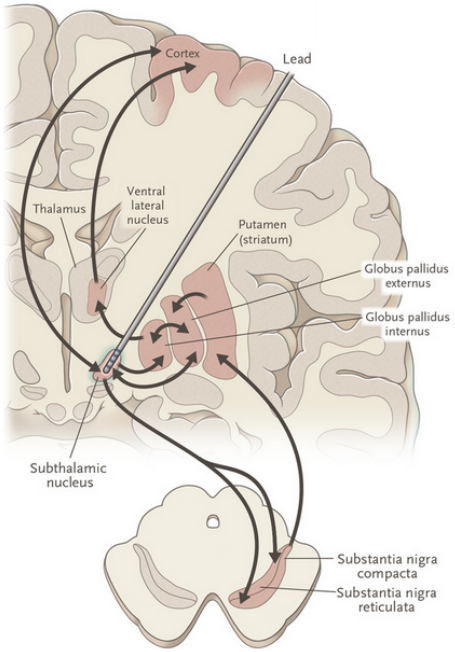


Figure 7 *A coronal section is shown with a DBS electrode implanted in the STN, the basal ganglia circuitry highlighted and connections indicated by arrows. The effects take place via the two efferents of the STN to the GPi on the one handside and to the SNr on the other.*

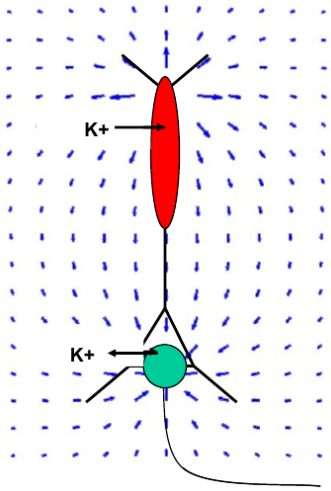


Figure 8 *A visualization of how an electrical field is created by a single neuron. When an action potential is fired, potassium channels open and potassium flows out of the cell. Na^+/K^+-ATPases pump K^+ ions back in. These two processes together with the temporal delay between channel opening in different compartments create the necessary soruce and sink for the electrical field.*

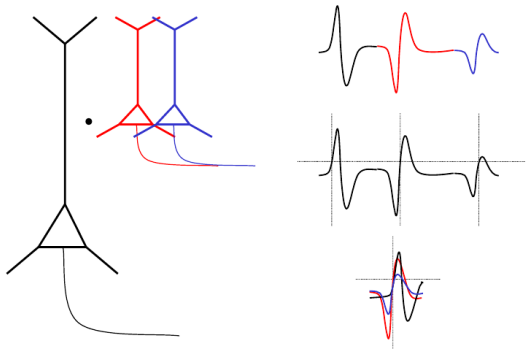
and EEG recordings for post-operative evaluation. An overview of the data collected is shown in figure 6 and the guidance procedure is described in more detail in [26].

2 Local Field Potentials at Scale: EEG, Micro-Electrodes, Deep Brain Stimulation Electrodes

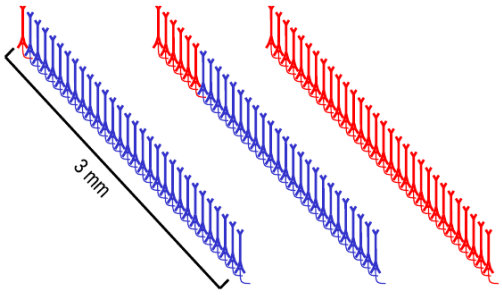
Neural activity induces changes in the concentration of ions across the cell membrane and gradients among different compartment of the neuron and between neurons, as shown in figure 8.

These changes induce an electrical field which can be measured using electrodes. The high frequency components above 500 Hz of this signal is closely related to the spiking activity of neurons in close vincity of the electrode. Lower frequency parts of the signal are called local field potentials (LFPs), which do not allow for an as straight forward interpretation when it comes to the underlying neural activity [11, 19]. LFPs can be meassured on different scales ranging from single unit recordings over multi unit recordings with different setups (e.g. micro-electrode, multi-electrode arrays, deep brain electrodes) up to extra-cranial recordings, placing the electrode on the skin head.

The different scales require each a different approach for analysis and interpretation: While single unit recordings (SUR) allow to track action potentials (APs) and subthreshold potentials, multi-unit recordings (MUR) capture aggregated neural activity of populations. Aggregation means, that the meassured voltage differences are caused by many different neurons, as visualized and explained in figures 9a and 9b. The size of the population that is recorded depends on the recording setup, e.g. the distance between two contacts/electrode tips, the electrode position and the size of the electrode itself. The broadest scale signals like electroencephalographies (EEGs) only allow the analysis of large groups of populations. For that, different frequency bands have been identified already by the inventor of EEG Hans



(a) *When the meassuring electrode is placed not very close to one neuron but to multiple neurons, the recorded signal contains contributions from many neurons with their magnitude depending on the distance from the electrode and on the frequency (see below).*



(b) *Depending on the distance of the neuron, different frequency components of the neurons contribute: When the neuron is close, also high frequency components are contributing, while whith increasing distance only lower components do. This is due to the size of the generating source: The field generated by single APs is rather small, while synchronous firing of a population causes the electrical field to extend further.*

Berger, whom described them first from an intracranial EEG in a dog and also found these in humans [37]. So the commonly used frequency bands are solely based on EEGs and do not characterize any properties of the recorded substrate. Rather they act as an odd binning. For a comprehensive review on myths in LFP processing, see [19]. Another example is the differenct between LFPs recorded from microelectrodes and DBS electrodes. The former has a very thin and sharp tip and is capable of recording single to a few cells. The latter's tip thickness is on the milimeter scale and records populations rather than single or a few units.

To summarize: With decreasing spatial resolution, the aggregation of the currents emitted by the neurons increases, i.e. the signal is more ambiguous to interpret in terms of the sources [11, 19]. This yields information on a more global scale or that is more network-oriented.

On the other hand with an increase in spatial resolution, the neccessary temporal resolution or the frequency at which the signal has to be sampled to caputre the information has to increase to capture all phenomena, i.e. $f_s^{\text{EEG}} < f_s^{\text{DBS}} < f_s^{\text{ME}}$.

Thus for single unit recordings it makes sense to detect action potentials by using an absolute threshold, while for multi-unit recordings, a threshold can be used to detect peaks or waveforms in the signals but they correspond by no means to action potentials but rather to a high synchronity in the recorded neural population [19]. Population recordings capture mesoscopic local activation patterns and may explain possible network effects in certain setups, e.g. when co-recording with another region or when presented with a stimulus. Global electrical recordings, like EEG report — mostly cortical — region based activity and can identify current gradients between regions.

3 Data Analysis Pipeline

3.1 Existing Analysis Framework

As we span a wide range of spatial scale, and per patient there are at least 300s recorded, the amount of data is quite large. This causes issues when it comes to computability. Without the application of reduction methods, certain algorithms, that will be briefly described later, will not finish in a reasonable amount of time ($<$ 1 week).

Thus reduction techniques have to be applied. One is down-sampling, which reduces the amount of samples in exchange for temporal resolution. The temporal resolution is half of the sampling frequency, i.e. the highest measurable frequency is half of the sampling frequency. Depending on the feature of interest, this can reduce the problem size coniderably. LFPs are usually evaluated in a range from 1 Hz to 500 Hz for multi-unit recordings [11], which is an order of magnitude lower than the sampling frequency without the loss of relevant information. When investigating spiking behavior using micro-electrode recordings however, applying downsampling would discard the

most relevant information. To date we use EEG and DBS electrode recordings, which are both in the lower frequency (<500 Hz) realm.

Another method is epoching. Data is segmented into equal-duration chunks, providing multiple advantages:

- When there are temporary impedance deviations, these epochs can be dropped without affecting the rest of the signal,
- all subsequent computations can be done in parallel/vectorized to speed up the computation, and
- bad epochs can be interpolated.

To clean the data, high pass, low pass and notch filters can be applied, e.g. to cut out 50 Hz line noise. Bad channels EEG or DBS channels can be detected, and then interpolated/"repaired" or dropped. Detrending can be applied to get rid of voltage drifts. Eye components can be removed from the EEG using the pairwise differences between horizontal and vertical frontal electrodes. Finally, rereferencing can be applied to normalize the differences between electrodes. Without this, differences could appear small because the potential difference between DBS and ground is high [39].

Firstly, due to the facts laid out in [26], the PSD is computed. Atop of that there is a new method called FOOOF which aims at separating periodic from aperiodic components in the signal [9]. The algorithm provided by the authors is applied to recieve the aperiodic component and subtract it from the original PSD to gain the periodic-only PSD. Additionally, the parameters of the function fit to the aperiodic component is stored. These two quantities are computed for both EEG and MEG. The median absolute deviation is a measure for activity in general, including bursting behavior. Functional connectivity — also called coherence — is computed between DBS and EEG channels. The current source density can be computed, which tries to reconstruct the currents between channels. Phase-amplitude couplings are one of the most complex quantities in terms of computability due to its high runtime complexity. Finally different measures of information content — entropies — are computed per channel.

These features are then exported to a data frame and persisted to disk for further analysis. Each row in the data frame corresponds to one contact at a certain depth, along with all features that were computed. Additionally, the data is safed not only on a whole-group but also on a per patient basis.

The software design and file organisation before refactoring is shown in figures 17 18 in the appendix.

3.2 Novel Pipeline Design

From a data analytical perspective, changing the code organisation of the dataset generation should not alter the result of the analysis, as data should remain unchanged. This goal was successfully reached. With approximately 12000 lines of code added with 6000 lines removed, the code base grew by 50%. Almost every single line of code was moved, restructured or rewritten. Even though no difference was the goal, some improvements have accompanied the reorganization:

- a GUI was implemented for 2D and 3D visualization using Qt,
- more quantities was added like the computation of the median absolute deviation, more entropy types and more functional connectivity normalization methods,
- by using lead dbs [**reco**] for DBS electrode trajectory reconstruction, the final position of each contact can be determined and the rows are annotated respectively,
- the database of patient recordings was refactored,
- data was cleaned, when there were mismatches in the recording protocol
- target implantation depths were reinitialized from track sheets as far as possible,
- an iterator was implemented to simplify and deduplicate the code as well as to ease maintenance and possible usage of additional multi-processing capabilities,
- code comments were added for most functions,

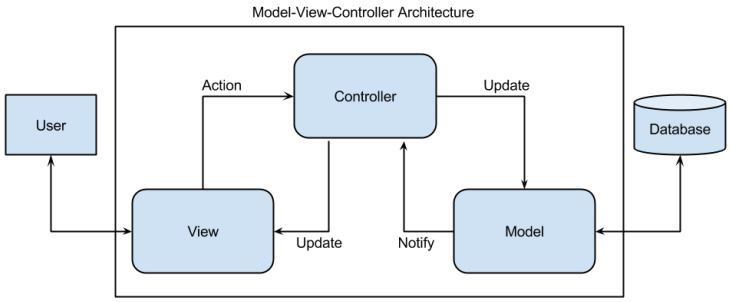


Figure 10 *Schematic of the model-view-controller architecture.*

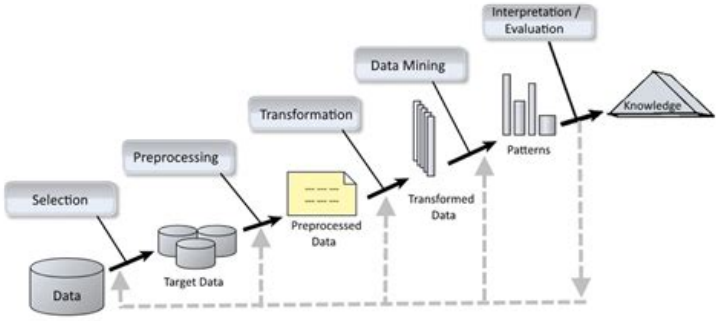


Figure 11 *Knowledge discovery in databases process diagram as proposed in the seminal paper by Fayyad et al. [13].*

- the size of the dataset file is reduced by a factor of 8 while more information is stored in the file,
- the patients raw signal is preserved,
- configuration of the pipeline is now done via configuration files
- code for questionaire analysis has been collected and begun to merge into the framework.
- software design that easily supports extension to other data types like wearables (ROAM, Adroid) and sleep EEGs (Pasitea),
- synchronisation of raw & patient data and the full preprocessed and aggregated dataset to the server,
- many scripts conducting data analysis have been written, producing overall approximately 200 different plots between December and January, with additional 60 in February and March.

The specific contributions of the author include: Implementation of an iterator, restructuring all file organizations, rewriting large parts to incoorporate the iterator, documentation, integration of new quantities and anatomical annotations, reinitiation of target implantation depths from track sheets, database file reduction, and some smaller and larger other fixes. Conceptually, the PSD is now computed for a larger frequency range to extend to scope of biomarker exploration.

The new code design follows the notion of a model-view-controller (MVC), shown in figure 10. Here data is the model, what changes data is the controller and the view provides visualizations of the model. This aligns with the data analysis pipeline proposed by the knowledge discovery in databases (KDD) process [13] shown in figure 11 To combine to previous two designs, each step of the KDD process has been assigned to one of the MVC components as shown in figure 12. Overall this then yields the structure after refactoring that is shown in figures 19 20 in the appendix.

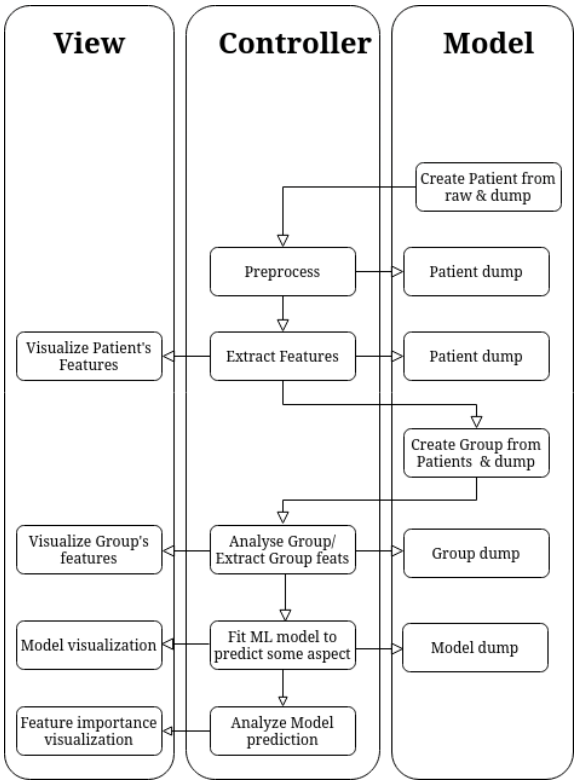


Figure 12 *Functionwise decomposition of the KDD process into components of a MVC.*

4 Discussion

Several limitations have been encountered:

- So far micro-electrode recordings are not analyzed (this is work in progress).
- To compare different sub-regions of the STN physiologically, there is a need for highly reliable labels. Otherwise assumptions are made initiating circular reasoning.
- The feature evaluation framework is very basic as the author decided to focus on understanding and restructuring the data generation pipeline.
- Feature importance is broken down to a mean value per feature but interactions between features could play an important role.

For example when assuming that the motor part has more β and the limbic part more θ, we already assign labels based on that and all data-driven methods will only be capable of reproducing that assumptions. Simply clustering with all available information will likely not yield the desired separation and characterizations. So a mixture is necessaty: exploration of potential features and their ranking (i.e. well-labelled data) as well as exploration for segregating features and exploitation of high feature importance in the hope that these features yield plausible clusters. For the data-driven analysis that was presented in the seminar, a simplistic clasifier was used. More sophisticated classifiers like neural networks which are capable of extracting features for themselves (not straightforward to interpret though) or transforming and combinging given features in a certain way. Methods for investigating the latent space of such classifiers exist [5, 28].

Acknowledgements

I'd like to thank Farzin Neghabani, Dr. Enrico Ferea and Prof. Dr. Alireza Gharabaghi for teaching me and giving me this opportunity.

References

[1] Giorgia Adani et al. “Selenium and other trace elements in the etiology of Parkinsons disease: a systematic review and meta-analysis of case-control studies”. In: *Neuroepidemiology* 54.1 (2020), pp. 1–23.

[2] Per Borghammer and Nathalie Van Den Berge. “Brain-first versus gut-first Parkinsons disease: a hypothesis”. In: *Journal of Parkinson's disease* 9.s2 (2019), S281–S295.

[3] Heiko Braak et al. “Staging of brain pathology related to sporadic Parkinsons disease”. In: *Neurobiology of Aging* 24.2 (2003), pp. 197–211. ISSN: 0197-4580. DOI: https://doi.org/10.1016/S0197-4580(02)00065-9. URL: https://www.sciencedirect.com/science/article/pii/S0197458002000659.

[4] Jiyang Cai, Jie Yang, and DeanP Jones. “Mitochondrial control of apoptosis: the role of cytochrome c”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1366.1-2 (1998), pp. 139–149.

[5] Jonathan Crabbé et al. “Explaining latent representations with a corpus of examples”. In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 12154–12166.

[6] Esther Dalfó et al. “Evidence of Oxidative Stress in the Neocortex in Incidental Lewy Body Disease”. In: *Journal of Neuropathology & Experimental Neurology* 64.9 (Sept. 2005), pp. 816–830. ISSN: 0022-3069. DOI: 10.1097/01.jnen.0000179050.54522.5a. eprint: https://academic.oup.com/jnen/article-pdf/64/9/816/9556740/64-9-816.pdf. URL: https://doi.org/10.1097/01.jnen.0000179050.54522.5a.

[7] William Dauer and Serge Przedborski. “Parkinson's disease: mechanisms and models”. In: *Neuron* 39.6 (2003), pp. 889–909.

[8] Ted M Dawson and Valina L Dawson. “The role of parkin in familial and sporadic Parkinson's disease”. In: *Movement disorders* 25.S1 (2010), S32–S39.

[9] Thomas Donoghue et al. “Parameterizing neural power spectra into periodic and aperiodic components”. In: *Nature neuroscience* 23.12 (2020), pp. 1655–1665.

[10] E Ray Dorsey et al. “Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016”. In: *The Lancet Neurology* 17.11 (2018), pp. 939–953.

[11] Gaute T Einevoll et al. “Modelling and analysis of local field potentials for studying the function of cortical circuits”. In: *Nature Reviews Neuroscience* 14.11 (2013), pp. 770–785.

[12] Fatemeh N Emamzadeh and Andrei Surguchov. “Parkinsons disease: biomarkers, treatment, and risk factors”. In: *Frontiers in neuroscience* 12 (2018), p. 612.

[13] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. “From data mining to knowledge discovery in databases”. In: *AI magazine* 17.3 (1996), pp. 37–37.

[14] Sven Geisler et al. “PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1”. In: *Nature cell biology* 12.2 (2010), pp. 119–131.

[15] Christopher G Goetz et al. “Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): process, format, and clinimetric testing plan”. In: *Movement disorders* 22.1 (2007), pp. 41–47.

[16] Ann M Graybiel. “The basal ganglia”. In: *Current biology* 10.14 (2000), R509–R511.

[17] Stephanie J Guiney et al. “Ferroptosis and cell death mechanisms in Parkinson's disease”. In: *Neurochemistry international* 104 (2017), pp. 34–48.

[18] Sebastian Heinzel et al. “Do we need to rethink the epidemiology and healthcare utilization of Parkinson's disease in Germany?” In: *Frontiers in neurology* 9 (2018), p. 500.

[19] Oscar Herreras. “Local field potentials: myths and misunderstandings”. In: *Frontiers in neural circuits* 10 (2016), p. 101.

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

[21] Eric R Kandel et al. *Principles of neural science*. Vol. 6. McGraw-Hill New York, 2021.

[22] Kathrin Machetanz et al. "Time efficiency in stereotactic robot-assisted surgery: an appraisal of the surgical procedure and surgeons learning curve". In: *Stereotactic and Functional Neurosurgery* 99.1 (2021), pp. 25–33.

[23] Connie Marras and K Ray Chaudhuri. "Nonmotor features of Parkinson's disease subtypes". In: *Movement Disorders* 31.8 (2016), pp. 1095–1102.

[24] Matthew M McGregor and Alexandra B Nelson. "Circuit mechanisms of Parkinsons disease". In: *Neuron* 101.6 (2019), pp. 1042–1056.

[25] Aristide Merola et al. "Gene therapy for Parkinsons disease: Contemporary practice and emerging concepts". In: *Expert review of neurotherapeutics* 20.6 (2020), pp. 577–590.

[26] Luka Milosevic et al. "Online mapping with the deep brain stimulation lead: a novel targeting tool in parkinson's disease". In: *Movement disorders* 35.9 (2020), pp. 1574–1586.

[27] Jonathan W Mink. "The basal ganglia: focused selection and inhibition of competing motor programs". In: *Progress in neurobiology* 50.4 (1996), pp. 381–425.

[28] Christoph Molnar. *Interpretable machine learning*. Lulu. com, 2020.

[29] Martin Niethammer, Andrew Feigin, and David Eidelberg. "Functional neuroimaging in Parkinsons disease". In: *Cold Spring Harbor perspectives in medicine* 2.5 (2012), a009274.

[30] World Health Organization. *International Statistical Classification of Diseases and related health problems: Alphabetical index*. Vol. 3. World Health Organization, 2004.

[31] Tamara Pringsheim et al. "The prevalence of Parkinson's disease: a systematic review and meta-analysis". In: *Movement disorders* 29.13 (2014), pp. 1583–1590.

[32] Maria C Rodriguez-Oroz et al. "Initial clinical manifestations of Parkinson's disease: features and pathophysiological mechanisms". In: *The Lancet Neurology* 8.12 (2009), pp. 1128–1139.

[33] Nobukatsu Sawamoto et al. "Cognitive deficits and striato-frontal dopamine release in Parkinson's disease". In: *Brain* 131.5 (2008), pp. 1294–1302.

[34] WMM Schuepbach et al. "Neurostimulation for Parkinson's disease with early motor complications". In: *New England Journal of Medicine* 368.7 (2013), pp. 610–622.

[35] Alessandro Stefani et al. "Mechanisms of action underlying the efficacy of deep brain stimulation of the subthalamic nucleus in Parkinson's disease: central role of disease severity". In: *European Journal of Neuroscience* 49.6 (2019), pp. 805–816.

[36] Asami Tanimura et al. "Striatal cholinergic interneurons and Parkinson's disease". In: *European Journal of Neuroscience* 47.10 (2018), pp. 1148–1158.

[37] Mario Tudor, Lorainne Tudor, and Katarina Ivana Tudor. "Hans Berger (1873-1941)–the history of electroencephalography". In: *Acta medica Croatica: casopis Hravatske akademije medicinskih znanosti* 59.4 (2005), pp. 307–313.

[38] Enza Maria Valente et al. "Hereditary early-onset Parkinson's disease caused by mutations in PINK1". In: *Science* 304.5674 (2004), pp. 1158–1160.

[39] Wim Van Drongelen. *Signal processing for neuroscientists*. Academic press, 2018.

[40] Doris D Wang et al. "Subthalamic local field potentials in Parkinson's disease and isolated dystonia: an evaluation of potential biomarkers". In: *Neurobiology of disease* 89 (2016), pp. 213–222.

[41] Ling Wang et al. “SPECT molecular imaging in Parkinson’s disease”. In: *Journal of Biomedicine and Biotechnology* 2012 (2012).

[42] Daniel Weiss et al. “Nigral stimulation for resistant axial motor impairment in Parkinsons disease? A randomized controlled trial”. In: *Brain* 136.7 (2013), pp. 2098–2108.

[43] AW Willis et al. “Incidence of Parkinson disease in North America”. In: *npj Parkinson’s Disease* 8.1 (2022), p. 170.

	Putamen	GPe	GPi	STN	SNr
Parkinsonian state	• Increased expression of D2 receptors, enkephalin, preproenkephalin (rat, monkey) • Decreased expression of D1 receptors, dynorphin, preprotachykinin and substance P (rat, monkey)	Decreased rate of neuronal firing (rat, monkey)	• Increased rate of neuronal firing (monkey, patients with PD) • Increased expression of CO-I and GAD (rat, monkey) • STN lesion reduces CO-I and GAD (monkey) and neuronal firing rate (monkey, rat)	• Increased rate of neuronal firing (rat, monkey, patients with PD) • Increased expression of CO-I (rat, monkey) • Decreased expression of 2-DG (monkey) • STN lesion improves parkinsonism (monkey, rat, patients with PD)	• Increased expression of CO-I and GAD (monkey, rat) • STN lesion reduces expression of CO-I and GAD (monkey)
Treatment with L-dopa or apomorphine	• Normalisation of expression of D1 and D2 receptor (rat, monkey) • Normalisation of expression of preprotachykinin and substance P (rat, monkey)	Normalisation of rate of neuronal firing (monkey)	• Decreased rate of neuronal firing (rat, monkey, patients with PD) • Decreased expression of CO-I and GAD (rat, monkey)	• Decreased rate of neuronal firing (rat, monkey, patients with PD) • Decreased expression of CO-I (monkey)	• Decreased expression of CO-I and GAD (monkey)

2-DG=2-deoxyglucose measured by autoradiography. 6-OHDA=6-hydroxydopamine. CO-I=cytochrome oxidase-I measured by in situ hybridisation. GAD=glutamic acid decarboxylase measured by in situ hybridisation. GPe=globus pallidus pars externa. GPi=globus pallidus pars interna. MPTP=1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine. PD=Parkinson's disease. SNr=substantia nigra pars reticulata. STN=subthalamic nucleus.

Figure 13 *Summary of cellular and molecular effects of PD on animal models (6-OHDA rat and MPTP modkey).*

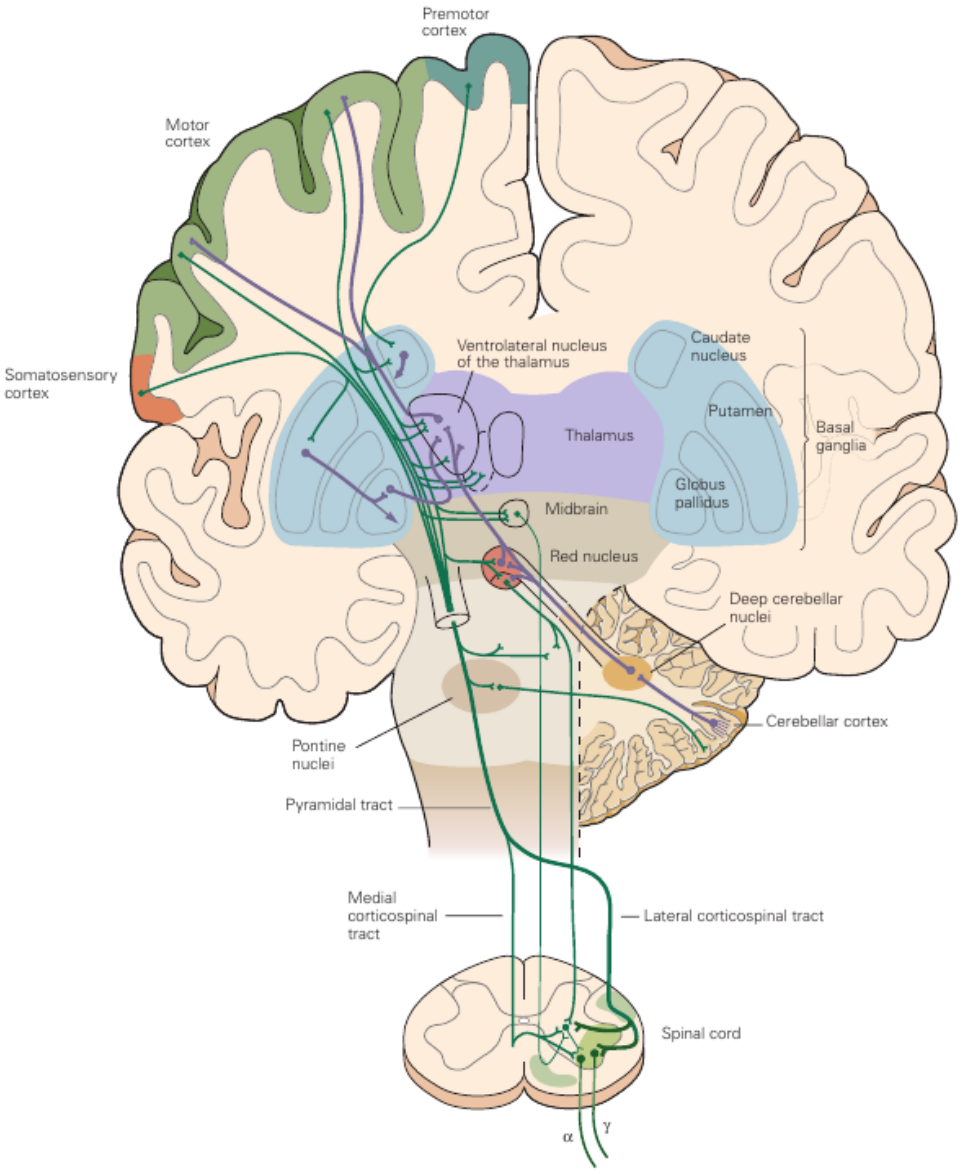


Figure 14 *The motor circuitry of the brain involves the coordinated functioning of principal components: the motor cortex in green, the somatosensory cortex in darker orange, the premotor cortex in turquoise, the basal ganglia in blue, the thalamus in purple, the midbrain in brown, the cerebellar nuclei and cortex (light orange) and finally the spinal cord at the bottom. The efferents originate mainly in the motor cortex, relayed via the basal ganglia and the thalamus to the brain stem & spinal cord to the muscles. Green projections indicate principal descending projections, purple connections indicate local circuitry and feedback.*

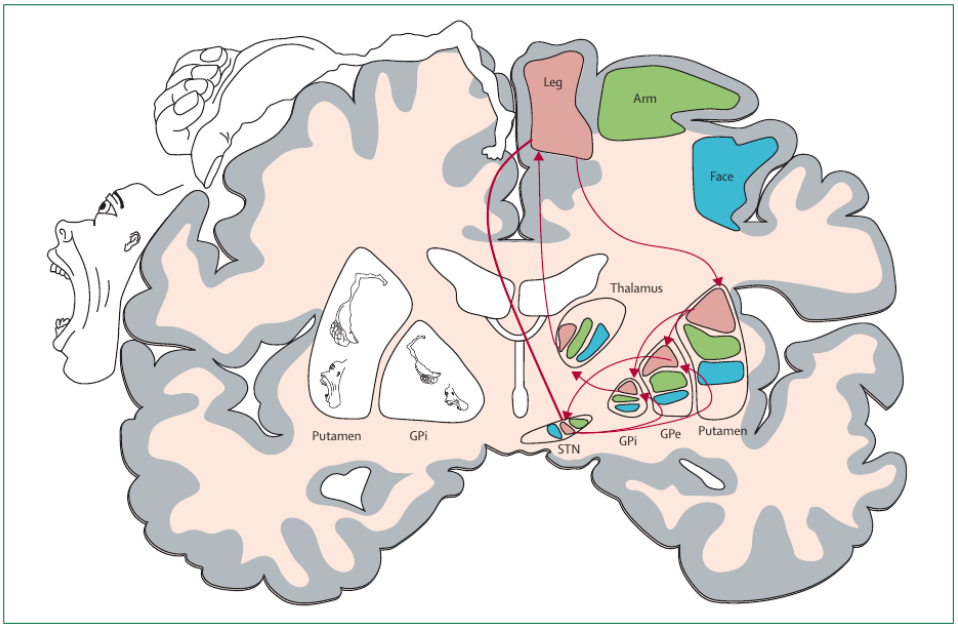


Figure 15 *The basal ganglia viewed from an efferent perspective: regions are shown by what they innervate. That is not only the motor and somatosensory cortices are organized somatotopically but also the basal ganglia follow this oganization principle.*

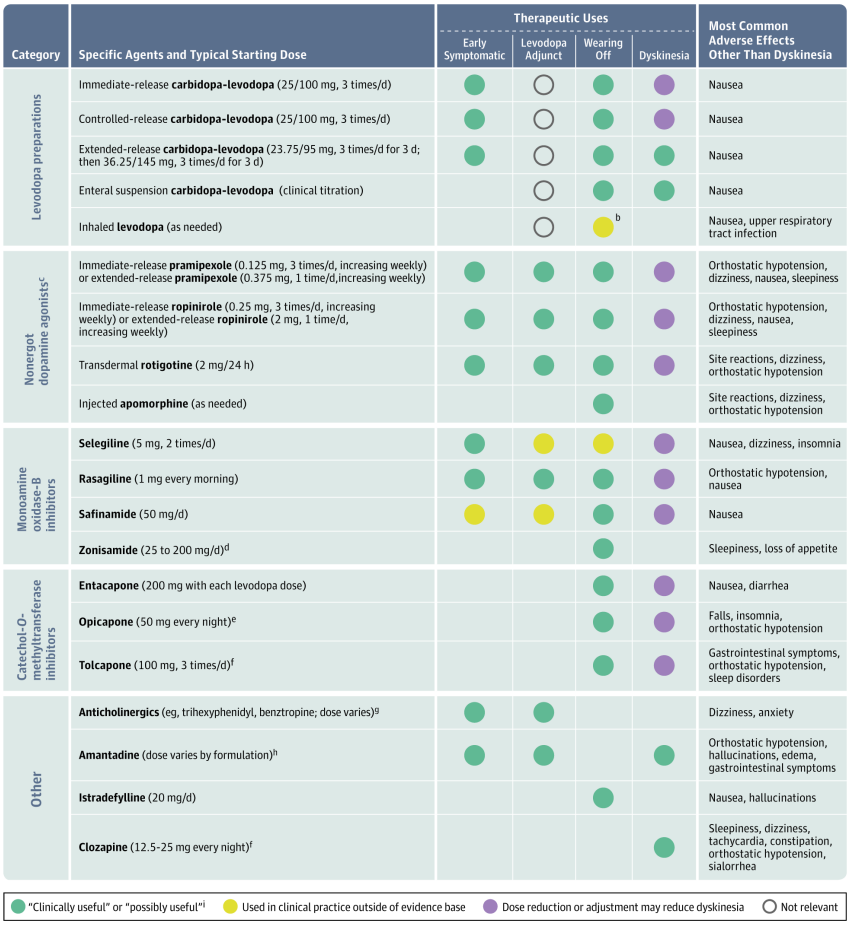


Figure 16 *a — Inclusion in the table does not imply US Food and Drug Administration (FDA) approval for any specific indication.*
b — Not included in International Parkinson and Movement Disorders Society review; approved by the FDA for motor fluctuations (off time).
c — Conversely, ergot dopamine agonists include cabergoline, pergolide, and bromocriptine are typically not used given adverse event risks including cardiac valvulopathy. d — Mechanism of action not completely certain; inhibition of monoamine oxidase-B is thought to be one contributing mechanism. Zonisamide is approved for use in Parkinson disease in Japan, but it is not commonly used for this purpose in the United States, where it is approved for use as an antiepileptic medication. e — Under review by the FDA at time of publication. f — Requires specialized monitoring (liver function for tolcapone; complete blood count for clozapine). g — Anticholinergic agents should be used sparingly in clinical practice given common adverse effects such as cognitive slowing. h — Amantadine is more commonly used for treatment of dyskinesias rather than as early symptomatic or adjunctive treatment. i — Indicates usefulness as determined by the International Parkinson and Movement Disorder Society Evidence-Based Medicine Review

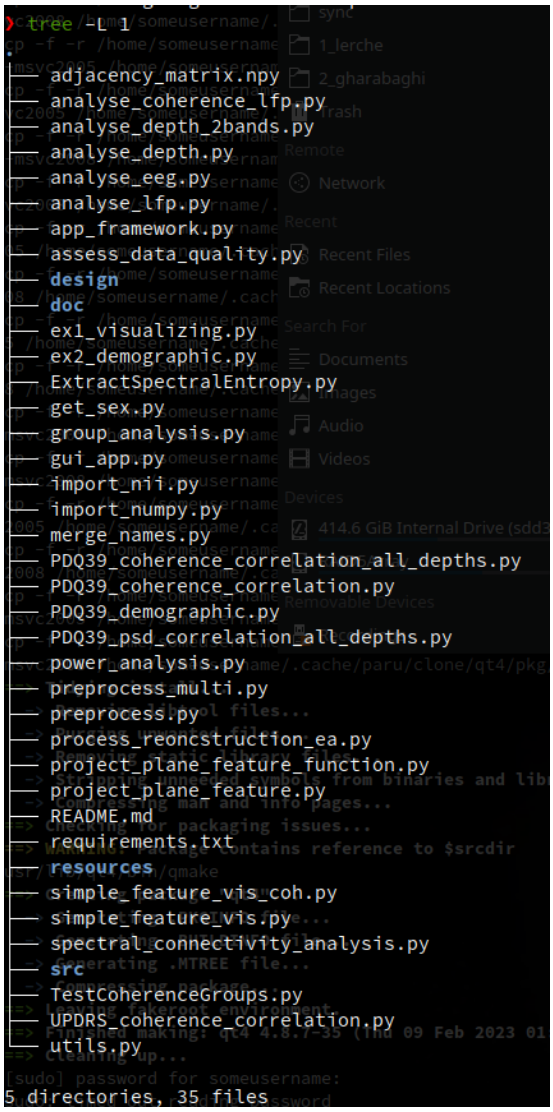


Figure 17 *Top level directory of the data analysis pipeline before refactoring.*

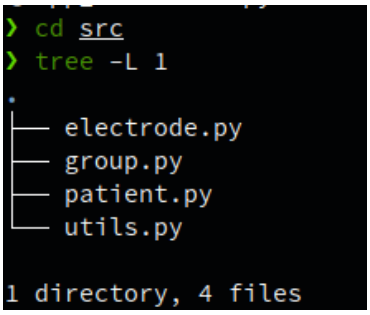


Figure 18 *Files included in the src folder before refactoring.*

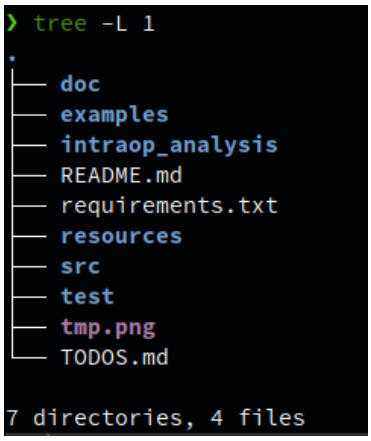


Figure 19 *Top level directory of the data analysis pipeline after refactoring.*

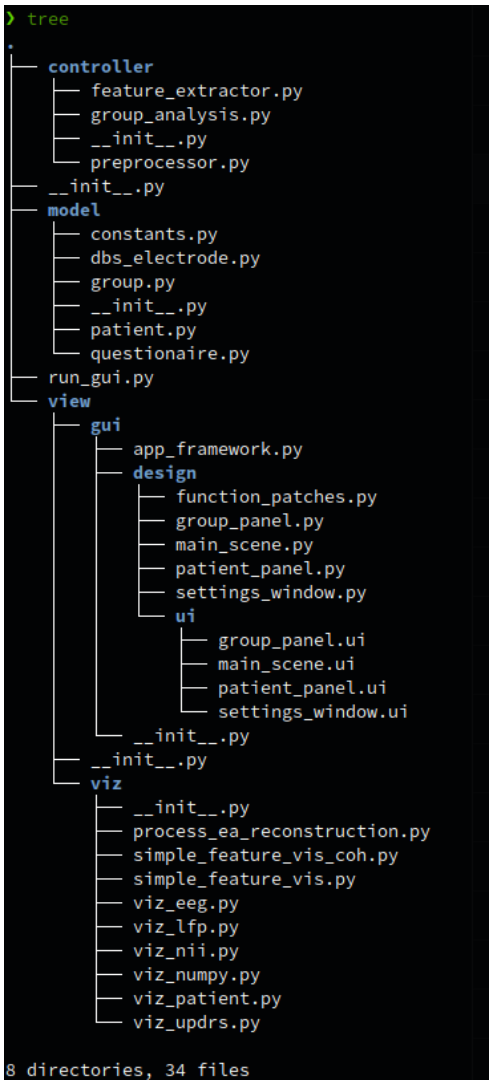


Figure 20 *Files included in the src folder after refactoring.*