

Diplomová práce



České
vysoké
učení technické
v Praze

F3

Fakulta elektrotechnická
Katedra mikroelektroniky

Analýza pohybu ruky u pacientů s poruchami chování v REM spánku pomocí kamerového systému

Bc. Petr Krýže

Školitel: Ing. Jan Rusz, Ph.D.
Obor: Elektronika a komunikace
Zaměření: Elektronika
Prosinec 2018

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Kryže** Jméno: **Petr** Osobní číslo: **406353**
Fakulta/ústav: **Fakulta elektrotechnická**
Zadávající katedra/ústav: **Katedra mikroelektroniky**
Studijní program: **Elektronika a komunikace**
Studijní obor: **Elektronika**

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

Analýza pohybu ruky u pacientů s poruchami chování v REM spánku pomocí kamerového systému

Název diplomové práce anglicky:

Analysis of Hand Motion in Patients with REM Sleep Behavior Disorder Using Camera System

Pokyny pro vypracování:

1. Seznamte se s možnostmi potenciálního využití a možnostmi objektivního měření pohybu ruky pro brzkou diagnostiku Parkinsonovy nemoci. Na základě literatury nastudujte vhodné metody pro vyhodnocení měření pohybů ruky s využitím kamerového systému.
2. Navrhněte algoritmy umožňující automatické vyhodnocení základních aspektů bradykineze ze záznamu pohybu ruky pořízeného pomocí kamerového systému.
3. Navržené metody otestujte na vybraném vzorku dat od jedinců s poruchou chování v REM spánku, pacientů s brzkou Parkinsonovou nemocí a zdravých jedinců kontrolní skupiny. Na dostupných datech proveďte vhodné statistické analýzy a vyhodnoťte nejvhodnější parametry pro odlišení zdravých jedinců od pacientů s Parkinsonovou chorobou.

Seznam doporučené literatury:

- [1] Ruzicka E, Krupicka R, Zarubova K, Ruzs J, Jech R, Zoltan S. Test of manual dexterity and speed in Parkinson's disease: Not all measure the same. *Park Relat Dis* 2016; 28:118-123.
[2] Krupicka R. Měření a analýza pohybu ruky u parkinsoniků. *Doktorská práce FBMI ČVUT, Kladno* 2013.

Jméno a pracoviště vedoucí(ho) diplomové práce:

Ing. Jan Ruzs, Ph.D., katedra teorie obvodů FEL

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

Datum zadání diplomové práce: **24.09.2018**

Termín odevzdání diplomové práce: **08.01.2018**

Platnost zadání diplomové práce: **19.02.2020**

Ing. Jan Ruzs, Ph.D.
podpis vedoucí(ho) práce

podpis vedoucí(ho) ústavu/katedry

prof. Ing. Pavel Ripka, CSc.
podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

Diplomant bere na vědomí, že je povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

Datum převzetí zadání

Podpis studenta

Poděkování

Rád bych velmi poděkoval panu vedoucímú práce Ing. Janu Ruzzovi, Ph.D., za velkou pomoc při formování práce a za dodání cenné motivace. Poděkování také patří Mgr. Radimu Krupičkovi, Ph.D., a prof. Evženu Růžičkovi za konzultace. Také chci poděkovat nezměrně mojí mámě a tátovi za podporu v celém životě i studiích. Díky také patří všem mým přátelům, kamarádům a kamarádkám, kteří mě těší svojí přítomností v mém životě.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne

.....
Podpis autora práce

Abstrakt

Náplní této práce je navázání na vyvinuté metody diagnostiky Parkinsonovy nemoci pomocí snímání pohybů prstů kamerovým systémem s cílem využít je pro včasné zachycení Parkinsonovy nemoci v prodromálním stádiu.

Byl sestaven datový soubor čítající zdravé kontroly, pacienty trpící Parkinsonovou nemocí a pacienty trpící spánkovou poruchou RBD (*Rapid Eye Movement Behavior Disorder*), která je jedním z raných příznaků objevujících se u pacientů s Parkinsonovou nemocí. Cílem práce bylo navrhnout sadu takových algoritmů, které by byly schopny statisticky odlišit skupinu zdravých kontrol a skupinu pacientů trpících RBD a odhalit jemné motorické změny při provádění úlohy "klepání prsty", zaznamenané pomocí kamerového systému.

Vytvořena byla programová struktura umožňující jednoduchou práci s měřeními datového souboru. Bylo zformováno 34 parametrů na čtyřech různých skupinách výpočtu z měření obou horních končetin - pomocí celkem 14 parametrů bylo možné nalézt statisticky významné rozdíly mezi rozděleními hodnot skupiny RBD a skupiny zdravých kontrol na hladině významnosti $p < 0.05$. Pomocí vybraných parametrů bylo také možné bez problému statisticky rozlišit skupinu pacientů trpících Parkinsonovou nemocí a skupinu zdravých kontrol.

Součástí práce byla také korelační analýza mezi jednotlivými hodnotami parametrů s cílem odhalit jejich závislost. V analýze laterality byl zkoumán vliv preference horní končetiny na výsledné hodnoty parametrů - žádný parametr nedosáhl hodnoty Spearmanova korelačního faktoru $\rho > 0.6$. Bylo seznáno, že hodnoty některých parametrů mají slabší tendenci vykazovat lepší výsledky na dominantní ruce, nicméně tento nálezný nebyl konzistentní přes všechny zkoumané parametry.

Z vybraných parametrů byl sestaven expertní systém klasifikátoru s využitím logistické regrese. Výkon klasifikátoru postaveného na kombinaci parametrů vykázal dobrou rozlišovací schopnost o hodnotě $AUC = 0.855$ při klasifikaci skupiny zdravých kontrol a skupiny pacientů trpících RBD, a výbornou schopnost oddělit skupiny zdravých kontrol a pacientů s Parkinsonovou nemocí o hodnotě $AUC = 0.91$.

Na výsledky této práce bude navazovat pokračující výzkum například ve formě longitudinální studie, mapující progresi nemoci u pacientů trpících RBD až k případné konverzi do Parkinsonovy nemoci. V budoucnu budou výsledky této práce použity pro včasnou diagnostiku Parkinsonovy nemoci již v prodromálním stádiu.

Klíčová slova: bradykineze, Parkinsonova nemoc, rapid eye movement sleep behavior disorder, analýza, včasný záchyt

Školitel: Ing. Jan Rusz, Ph.D.

Abstract

The goal of this thesis is to follow upon already developed diagnosis methods of the Parkinson's disease by utilising the motion capture of finger movements and use it for an early capture of the Parkinson's disease in its prodromal stage.

A group of healthy controls, a group of patients suffering from the Parkinson's disease and a group of patients suffering from the sleep disorder RBD (*Rapid Eye Movement Behavior Disorder*) were collected in a balanced dataset. RBD is one of the first early symptoms that appear in patients with the Parkinson's disease. The goal of this thesis was to suggest a set of parameters, that would be able to statistically differentiate the group of healthy controls and the group of patients suffering from RBD, and to reveal the subtle motoric changes during the finger-tapping task, as recorded by the motion capture system.

A program structure that enables easy measurement manipulation was established. A number of 34 parameters was formed on 4 different types of calculations from the measurements of both upper limbs - a total of 14 parameters showed statistical significant differences ($p < 0.05$) between the value distributions of the group of healthy controls and the group of patients suffering from RBD. Selected parameters could also differentiate the group of healthy controls from the group of patients with Parkinson's disease without any difficulty.

A correlation analysis between the distinct parameter values with the goal of finding any possible connections was also a part of this thesis. In the laterality analysis, an effect of upper limb preference on the calculated parameter values was observed - no parameter reached values of Spearman's correlation factor higher than $\rho > 0.6$. It was concluded, that the values of certain parameters have a weak ten-

dency towards having better performance on the patient's dominant hand, nevertheless this result wasn't consistent across all of the observed parameters.

An classification expert system on the basis of logistic regression was developed using a set of selected parameters. A classifier built on the combination of parameters performed well differentiating between the group of healthy controls and the group of patients suffering from RBD, having $AUC = 0.855$. Differentiation between the group of healthy controls and the group of patients with the Parkinson's disease performed very well, using a classifier of selected parameters, having $AUC = 0.91$.

A follow-up research, in a form of a longitudinal study that maps the progression of the disease in patients suffering from RBD towards the possible conversion to the Parkinson's disease, will connect on the results of this thesis. The results of this thesis will also be used in the future in the task of an early diagnosis of the Parkinson's disease in its prodromal stage.

Keywords: bradykinesia, Parkinson's disease, rapid eye movement sleep behavior disorder, analysis, early detection

Title translation: Analysis of Hand Motion in Patients with REM Sleep Behavior Disorder Using Camera System

Obsah

1 Úvod	1	6 Závěr	67
1.1 Parkinsonova nemoc	2	Bibliografie	69
1.1.1 Příznaky	2	A Použité zkratky	75
1.1.2 Příčiny	3	B Dodatečné grafy	77
1.1.3 Diagnóza a léčba	5	C Obsah příloženého CD	81
1.2 Kamerový systém a úloha klepání prsty	6		
1.2.1 Bradykineze	7		
1.3 RBD - REM Sleep Behavior Disorder	8		
1.3.1 Diagnóza	8		
1.3.2 Léčba	9		
1.3.3 Spojitost s PN	9		
1.4 Cíle práce	11		
2 Databáze	13		
2.1 Datový soubor	13		
2.2 Lateralita a rukost	16		
2.2.1 Edinburgh Handedness Inventory	18		
2.3 Struktura datového souboru	19		
2.3.1 Datová reprezentace měření	20		
2.4 Druhy FT měření	26		
2.5 Délka záznamu	26		
3 Metodika	29		
3.1 Parametry a třída <code>DataProperties</code>	29		
3.1.1 Volba parametrů	29		
3.1.2 Parametry	30		
3.1.3 Výběr ruky pro výpočet parametrů	37		
3.2 Statistická analýza	39		
3.3 Analýza laterality	41		
3.4 Korelační analýza	43		
3.5 <i>ROC</i> analýza	44		
4 Výsledky	47		
4.1 Statistická analýza	47		
4.1.1 Parametry typu <i>min</i> , <i>max</i>	47		
4.1.2 Parametry typu <i>diff</i>	50		
4.2 Analýza laterality	52		
4.3 Korelační analýza	55		
4.3.1 Korelace mezi parametry	55		
4.3.2 Korelace s klinickými daty	56		
4.4 <i>ROC</i> analýza	57		
5 Diskuze	63		
5.1 Další směřování	64		

Obrázky

1.1 Postup Lewyho tělísek, tvořených (zejména) proteinem alfa-synukleinem, skrze mozkovou tkáň (pohled z průřezu) v souvislosti s postupem PN.	4
2.1 Graf věkového rozdělení v jednotlivých skupinách datového souboru.	14
2.2 Rozdělení laterality mezi všemi subjekty z datového souboru.	16
2.3 Graf typu <i>beeswarm</i> zobrazující rozložení celkového skóre třetí sekce dotazníku MDS-UPDRS.	17
2.4 Hierarchie datové struktury vytvořené pro snadnou práci s daty měření.	21
2.5 Histogram časových délek měřených dat.	27
2.6 Graf zobrazující výsledky analýzy délky signálu u měření z datového souboru.	28
4.1 Grafy typu box-plot ilustrující distribuci hodnot mezi skupinami pro vybrané parametry typu <i>min/max</i>	49
4.2 Grafy typu box-plot ilustrující distribuci hodnot mezi skupinami pro vybrané parametry typu <i>diff</i>	52
4.3 Výsledky korelační analýzy mezi skóre z Edinburského dotazníku rukosti a výsledky vybraných parametrů.	53
4.4 ROC (<i>Receiver Operating Characteristic</i>) křivka pro logisticky regresní klasifikátor za použití dat parametrů typu <i>min</i> a <i>max</i>	58
4.5 ROC (<i>Receiver Operating Characteristic</i>) křivka pro logisticky regresní klasifikátor za použití dat parametrů typu <i>diff</i>	59
4.6 ROC (<i>Receiver Operating Characteristic</i>) křivka pro logisticky regresní klasifikátor za použití dat parametrů jak typu <i>min</i> a <i>max</i> , tak typu <i>diff</i>	60
4.7 Ilustrační zobrazení datových bodů pro různé kombinace použitých parametrů typu <i>min</i> a <i>max</i> v klasifikátoru na bázi logistické regrese.	61
4.8 Ilustrační zobrazení datových bodů pro různé kombinace použitých parametrů typu <i>diff</i> v klasifikátoru na bázi logistické regrese.	61
B.1 Skupinové srovnání a souhrnné zobrazení výsledných hodnot parametrů pro parametr <i>CloseVelNRMSD</i>	77
B.2 Skupinové srovnání a souhrnné zobrazení výsledných hodnot parametrů pro parametr <i>FreqSlopeMA</i>	78
B.3 Skupinové srovnání a souhrnné zobrazení výsledných hodnot parametrů pro parametr <i>MaxNRMSD</i>	79
B.4 Skupinové srovnání a souhrnné zobrazení výsledných hodnot parametrů pro parametr <i>MinDwellTimeNRMSD</i>	80

Tabulky

2.1 Vlastnosti jednotlivých skupin v datovém souboru využitém pro analýzu v této práci.	14
4.1 Výsledky statistické analýzy Mann-Whitneyho U-testu mezi výsledky parametrů skupiny kontrol a skupiny pacientů trpících RBD - <i>min/max</i> typ výpočtu	48
4.2 Výsledky statistické analýzy Mann-Whitneyho U-testu mezi výsledky parametrů skupiny kontrol a skupiny pacientů trpících RBD - <i>diff</i> typ výpočtu	51
4.3 Výsledky analýzy laterality horních končetin na základě vypočtených hodnot vybraných parametrů.	54
4.4 Tabulka vybraných výsledků korelací mezi dvojicemi parametrů.	56

Kapitola 1

Úvod

Dle výzkumu k roku 2010 se jen ve Spojených státech počet lidí trpících Parkinsonovou chorobou (Parkinsonovou nemocí, dále jen *PN*, viz 1.1) ve třicátých letech tohoto století přehoupne přes jeden milion. Zároveň je tento trend vzrůstající - jen mezi lety 2005 a 2030 se počítá se zdvojnásobením počtu výskytu této nemoci. Navzdory tomu, že v České republice, kde je *PN* druhá nejčastější neurogenerativní nemoc, v současné době neexistují ucelená data o počtu nemocných (narozdíl od jiných chorob), odhaduje se, že počet nemocných se pohybuje mezi 15 a 30 tisíci. Dá se ale předpokládat, že trend vývoje počtu nemocných bude následovat ostatní země vyspělého světa, a výskyt nemoci se do roku 2030 zdvojnásobí.¹ Zajímavostí také je, že data mapující prevalenci *PN* mezi obyvatelstvem Spojených států byla do vydání studie z roku 2010 čerpána (a extrapolována) z lokálního výzkumu v okrsku jménem *Copiah County* ve státě Mississippi, kde se tehdy v samotném průběhu testování odhalilo o 40 procent nových pacientů. - a to vše roku 1978.² Nová studie mapující počet nemocných tak jen dokazuje zvýšený zájem o výzkum a přehled o stavu *PN*.

Motivací pro intenzivní výzkum v oblasti *PN* jsou pak také faktory jako je vzrůstající průměrný věk obyvatel vyspělých zemí a s ním i ruku v ruce jdoucí širší výskyt chorob jako je *PN*, Alzheimerova nemoc či demence. To je nejen palčivou otázkou z hlediska kvality života lidí a často našich nejbližších, ale také díky vzrůstajícím materiálním nárokům států na péči o nemocné. Ve Spojených státech se kombinovaná cena za přímý a nepřímý vliv *PN* na společnost odhaluje na 25 miliard dolarů ročně. Všeobecná zdravotní pojišťovna, největší česká zdravotní pojišťovna³ s více než pěti miliony klientů, pak v článku z roku 2017 uvádí, že jí péče o klienty s *PN* je za uplynulý rok (2016) stála více než 455 milionů korun. Toto číslo je přitom více než narůstajícím počtem nemocných, kterých je podle VZP stále "stabilně nad hranicí 20 tisíc", ovlivněno novými (a nákladnými) léčivými přípravky. Studie z roku 2010, mapující celkové (přímé i nepřímé) náklady spojené s jedním pacientem trpícím *PN* pak uvádí, že pololetní výdaje vychází v průměru na 5510 euro na pacienta (s cenami pro rok 2008).^{4,5,6,7}

PN v současné době lidstvo není schopno vyléčit, pouze zmírnit její příznaky. To je také součástí motoru, který pohání výzkum této choroby vpřed. Výzkum léčby dokumentuje například snaha o zlepšení péče v České repub-

lice pomocí zavádění nových, specializovaných center umožňujících léčbu pomocí elektrod zaváděných do pacientova mozku, zaměření se na lepší propojení jednotlivých zdravotnických specialistů či investice do školení za účelem lepšího a včasného rozpoznání příznaků PN u pacientů.⁸

1.1 Parkinsonova nemoc

Parkinsonova nemoc (zkr. *PN*, anglicky *Parkinson's disease* a zkr. *PD*) je neurodegenerativní onemocnění postihující pacienty zejména v pozdějším věku, počínaje padesátým až šedesátým rokem života. Nemoc je progresivní a dlouhodobá, přičemž většinou začíná na jedné straně pacientova těla a postižení zůstává v průběhu nemoci asymetrické, ale není to zcela pravidlem.

1.1.1 Příznaky

Příznaky nemoci jsou zejména motorického charakteru, nicméně nemoc se také projevuje řadou příznaků nemotorických.

Motorické příznaky⁹:

- Klidový třes (tremor)
- Svalová ztuhlost (rigidita)
- Zpomalenost pohybů (bradykineze, viz 1.2.1)
- Celkové pohybové problémy (poruchy chůze)
- Poruchy řeči
- Postižení psaní rukou

Vedlejší příznaky¹⁰ (jež se mohou objevit daleko dříve než jsou viditelné motorické příznaky) jsou dvojího druhu:

- **Fyzické** nemotorické příznaky
 - *Bolesti* (například zad, ale také svalové křeče, buď přímo způsobené změnami v mozku (pálení, mravenčení) nebo nepřímo způsobené motorickými příznaky, ztuhlostí končetin, špatné držení těla apod.
 - *Senzorické poruchy*, zejména čichu (např. *hyposmie*, snížená schopnost vnímat zápachy a vůně), objevuje se až u 90 procent pacientů
 - *Snížení krevního tlaku*
 - *Zácpa* - společně s hyposmií může být jedním z prvních příznaků mozkové degradace
 - *Urogenitální dysfunkce* (poruchy močového a rozmnožovacího ústrojí)

■ Psychické (neuropsychiatrické) nemotorické příznaky

- *Deprese* spíše mírnějšího, nesebedestruktivního charakteru, ztráta zájmu, anhedonie (snížená úroveň prožitků emocí)
- *Úzkostné stavy*, a panické ataky
- *Kognitivní dysfunkce a demence*, zejména v pozdějších stádiích nemoci, společně s tím jak degenerace postupuje do vnějších částí mozku, projevuje se také v obtížném zvládnutí provádění několika úkolů současně (například hovor a chůze)
- *Psychózy* a halucinace (zejména vizuální), občas také zapříčiněny medikací, zejména u starších pacientů
- *Spánkové poruchy* - viz dále

Deprese, úzkosti a spánkové poruchy jsou významnými důsledky PN, první dvě zejména z hlediska kvality života pacientů a psychosomatických účinků na jejich zdravotní stav, spánkové poruchy pak z úhly pohledu včasného zachycení nemoci v jejím prodromálním stádiu. Vzhledem k tomu, že se tyto příznaky objevují v nejméně 40 procentech pacientů s PN, je nutné komplexní zaměření léčby PN, spojující schopnosti různých lékařských profesionálů, protože při pravidelných kontrolách je těžké psychické příznaky včas zachytit.¹¹

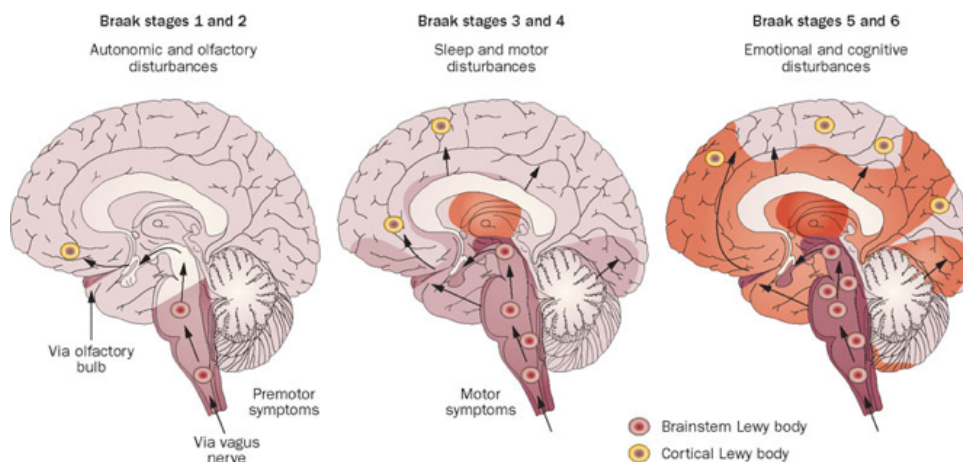
■ Spánkové poruchy

Spánkové poruchy spojené s PN jsou pak zajímavé zvláště se z hlediska této práce (více viz dále v části 1.3). Vyskytují se velmi často a zahrnují problémy s usínáním, noční křeče, časté buzení, syndrom neklidných nohou (SNN, anglicky *restless legs syndrome* - RLS), noční inkontinence, noční motorické problémy spojené s omezenou pohyblivostí na lůžku a zmatení či halucinace v noci či po probuzení. Degradace neuronů v mozku ovlivňuje i centra starající se o správný průběh lidského spánku, nacházející se zejména v částech mozku blíže míše (hypothalamus)¹² - dá se předpokládat, že poloha těchto center souvisí s brzkým nástupem spánkových problémů na časové ose PN vzhledem k tomu, že nemoc postupuje směrem od míchy přes mozkový kmen až do mozkové kůry (viz 1.1).¹³

■ 1.1.2 Příčiny

Příčiny PN nejsou v současné době zcela jasné, nicméně etiologie nemoci je předmětem intenzivního výzkumu. V mozku pacientů s PN se v tělech neuronů objevují tzv. *Lewyho tělíska* (LB, z ang. *Lewy bodies*) a *Lewyho neurity* (LN), což jsou abnormální agregace proteinu jménem alfa-synuklein, který zatím z nejasných příčin začne měnit svojí formu na jiné, patologické uspořádání. Protein v této nové struktuře pak má tendenci se agregovat s dalšími podobnými proteiny, tvořit tak LB/LN a postupovat skrze mozkovou tkáň. Neurony, které jsou zasaženy LB nebo LN, pak ztrácejí svojí normální funkci. Přítomnost LB/LN potažmo alfa-synukleinu charakterizuje všechny *synukleopatické* nemoci (kromě PN také DLB (*Dementia with Lewy bodies*, český

demence za přítomnosti Lewyho tělísek) nebo MSA (Multiple system atrophy, česky *Shyřiv-Dragerův syndrom*)).¹³ Nejviditelnější dopad poškození mozku je pak vidět v motorických příznacích nemoci. Ty jsou způsobeny poškození části středního mozku zvané *substantia nigra* (česky "černá substance"), nacházející se v bazálních gangliích. Tato struktura tvoří *dopamin*, chemický neurotransmitter zodpovědný za motorické funkce (kontrolu pohybů), emoční funkce (motivace, systém potěšení a odměn) a další. Při poškození struktur tvořících dopamin pak postupně jeho hladina v těle klesá až do bodu, kdy chybí přibližně 70% a začnou se viditelně projevovat motorické problémy.[14, str. 3] Do té doby je tělo schopné nedostatek dopaminu kompenzovat jeho zvýšenou produkcí nebo úpravou chování systémů v bazálních gangliích.¹⁵



Obrázek 1.1: Postup Lewyho tělísek, tvořených (zejména) proteinem alfa-synukleinem, skrze mozkovou tkáň (pohled z průřezu) v souvislosti s postupem PN. První obrázek (zleva) popisuje první dvě fáze PN (podle Braakovy teorie¹⁶), charakterizované změnami v oblasti čichu a funkce autonomního nervového systému. V počátku se (podle Braaka et al.) dostávají Lewyho tělíska do mozku směrem od míchy přes nervus vagus do mozkového kmene. Na druhém obrázku je vidět průběh třetí a čtvrté fáze, kdy začínají být na pacientovi pozorovatelné motorické a spánkové změny. Lewyho tělíska postupují hlouběji do mozku a vstupují do prostoru bazálních ganglií kde způsobují neurologické poškození *substantia nigra* (zodpovědné za produkci dopaminu). Na posledním obrázku je pak vyobrazen průběh páté a šesté fáze, kdy nemoc postupuje do prostoru mozkové kůry a zasahuje čelní, spánkový a temenní lalok. Postiženy jsou pak již i kognitivní a emoční centra mozku.

Zdroj: Doty et al. *Nature Reviews Neurology* 8(6) 329-339 (2012), doi:10.1038/nrneurol.2012.80 [<https://www.nature.com/articles/nrneurol.2012.80>] (19. 12. 2018)

■ 1.1.3 Diagnóza a léčba

Brzké rozpoznání PN je jedním z velkých výzev současného výzkumu. Důležitost brzkého zachycení PN je zejména v účinnějších možnostech léčby raných stádií PN, jak již pomocí standardních (a stále nepostradatelných) dopaminových suplementů (levodopa), dopaminových náhražek (agonistů) a MAO-B inhibitorů tak i pomocí moderních (a futuristických) metod, jako je například léčba implantací kmenových buněk do postižených center mozku, která je účinná zejména při velmi raných stádiích vývoje PN.^{17,18,19,20,21,22} Velkým problémem PN je v současnosti neexistence standardního diagnostického testu. Zjišťování PN u pacientů je založeno na klinických postupech, které nejen, že často mohou vést k chybným diagnózám (díky tomu, že symptomy onemocnění se často překrývají s jinými neurodegenerativními chorobami jako je DLB, Alzheimerova nemoc, MSA, ...), ale také nejsou schopny odhalit počáteční stádia nemoci, která jsou charakterizována změnami, které nejsou ve své podstatě motorické (pohybové) nebo jsou tak subtilní, že je lékař při klinickém vyšetření nemůže zaznamenat. Jinými slovy, ve chvíli kdy lékař s jistotou diagnostikuje pacientovi PN na základě motorických poruch, už je často nemoc v pokročilém stádiu (když je okolo 70% neuronů v postižených částech mozku již ztraceno[14, str. 3]), kdy s postupem nemoci v těle pacienta léčba standardními dopaminovými suplementy ztrácí na účinnosti, ale také už je (nebo v budoucnu bude) často pacient úplně vyřazen z možností léčby některými alternativními způsoby (kromě léčby kmenovými buňkami také například i z fyzioterapeutických cvičení, které pomáhají zmírnit motorické následky nemoci v pozdějších stádiích). Zároveň je také dobré poznamenat, že i pro účely výzkumu raných stádií nemoci a možností léčby v těchto stádiích je pro lékaře a vědce důležité mít dostatek pacientů, kteří jsou zachyceni již v počátcích PN. Bez dostatku vhodných pacientů v příslušných stádiích nemoci nelze vyvinout efektivní léčbu, jelikož neexistují pacienti, na kterých by se případná nová léčiva mohla otestovat. Díky tomu je nalezení brzkých biomarkerů zcela zásadní.¹⁴

Cílem výzkumů je tedy nalézt biomarkery - ukazatele (parametry), které by umožnily odhalit PN již v prodromálním (počátečním) stádiu. MDS (*Movement Disorder Society*, mezinárodní společnost pro pohybové vady založena roku 1985 v USA) například navrhla soubor kritérií pro diagnostiku PN v počátečních stádiích, jak pro účely výzkumu tak pro účely lékařské diagnostiky. Komplexní proces v jejich procesu využívá naivní Bayesovský klasifikátor kombinující mnoho klinických dat o pacientovi včetně výsledků testů pohybových i nepohybových symptomů, u kterých bylo potvrzeno, že mají schopnost predikovat nástup PN.^{23,24} Současně také MDS zrevidovala dotazník/škálu pro hodnocení stavu pacienta za účelem kvantifikování a také odhalení PN - tzv. **UPDRS: Unified Parkinson's Disease Rating Scale**²⁵. Ten hodnotí stav pacienta a sleduje progresi nemoci v čase. Hodnocení je pak rozděleno do několika sekcí, které hodnotí motorické i nemotorické příznaky nemoci včetně zpětné vazby z průběhu terapie. Některé sekce pacient vyplňuje sám, některé jsou vyhodnoceny lékařem při klinickém vyšetření. Výsledkem testu je pak skóre z jednotlivých částí současně s celkovým výsledkem testu.

Obdobným dalším používaným testem je také *Hoehn-Yahrova škála*²⁶, která také prošla revizí, nicméně je méně detailní než škála UPDRS a proto i dnes méně používaná.

Dalším přístupem je snaha najít způsob, jak již v prodromálním stádiu nemoci odhalit změny (motorických) parametrů, které při klinických vyšetřeních (vizuálně) sledují lékaři (viz například kritéria²⁷ definovaná společností *United Kingdom PD Society Brain Bank*: svalová rigidita/ztuhlost, klidový třes, problémy s rovnováhou, **bradykineze** (viz 1.2.1) - zpomalování pohybů) a pomocí kterých se v současné době odhaluje PN u postižených pacientů. Tyto změny jsou často tak malé, že není možné je při klinickém vyšetření pouhým okem odhalit. Proto se pro zlepšení (a kvantifikaci) evaluace motorických změn u pacientů například používají různé instrumentální testy (testy používající nástroje) - například tzv. *pegboard test*, při kterém se do speciální desky manuálně zapichují a vyndávají kolíčky (za současného měření času nebo počtu kolíčků) nebo *alternate tapping test*, při kterém se měří střídavě klepání prsty, buď analogově (pomocí speciálních počítadel) nebo pomocí počítačového rozhraní.^{28,29,30} Tyto testy také využívá studie dokumentující průběh motorických příznaků nemoci u pacientů s nemocí RBD, který je významným prekurzorem PN (viz sekce 1.3.3).

Jiným způsobem je přímé měření pohybu pacientů, nicméně použití nástrojů a technik, které se přímo dotýkají těla pacienta a měří jeho pohyb například pomocí akcelerometrů není ideální, protože se dá předpokládat, že jejich fyzický vliv ovlivní již tak drobné změny pohybu při počátečních stádiích nemoci.

1.2 Kamerový systém a úloha klepání prsty

Za účelem odpovědi na nedokonalosti používaných (kontaktních) technik přímého měření motorických změn u pacientů s PN byl vyvinut v rámci disertační práce Radima Krupičky^{31,32} speciální 3-D kamerový systém, který dokáže zaznamenávat změny v mobilitě pacienta bez toho, aby se ho dotýkal. Jediným zprostředkovatelem systému na těle pacienta jsou dvě reflexní značky umístěné na kloubech palce a ukazováčku. Jejich pohyb pak zaznamenají kamery a pomocí transformace vyvinutým softwarem je přeložen do dat obsahujících polohu obou značek v prostoru. Systém tak umožňuje neovlivněné měření vlivu PN na pohyb pacienta. Úlohu, kterou pacient provádí je tzv. *finger-tapping*, česky úloha "klepání prsty"- její náplní je stisk palce a ukazováku ve špičce střídaný jejich co největším oddálením. Tyto dvě pozice, respektive tento pohyb má pak pacient za úkol dělat s co největší četností a amplitudou. Tato úloha je také součástí dotazníku MDS-UPDRS v jeho sekci pro zhodnocení motoriky pacientů a byla vybrána v Krupičkově práci po konzultaci s experty jako úloha, na které se bude hodnotit míra postižení bradykinezí (1.2.1) u pacientů.

Průběh testu také přímo popisuje citace z disertační práce:

"Při testu „klepání prsty“ se měřená osoba pohodlně posadí a ohne horní končetinu v loketním kloubu tak, aby ohnutí nebylo pro ni omezující a bylo pohodlné. Na lékařův pokyn začne měřená osoba rychle a opakovaně klepat palcem o špičku ukazováku a to s největším následným oddálením palce a ukazováku a největší frekvencí. Frekvence klepání se pohybuje v rozmezí 2 - 3 Hz a s oddálením přibližně 8 - 13 cm při čemž pauza mezi flexí a extenzí prstu se pohybuje v rozmezí 20 až 50 ms." [31, str. 25]

Detailnější popis metodiky měření FT testu pomocí kamerového systému pak lze nalézt v práci R. Krupičky.[31, str. 55]

R. Krupička ve svojí práci pak definoval a vyhodnotil [31, str. 67] parametry na naměřených datech FT testu u 22 zdravých kontrol a 22 pacientů postižených PN. Výsledky jeho experimentů ukazují, že je možné s jistotou statisticky rozlišit tyto dvě skupiny pomocí parametrů maximálního oddálení prstů, maximální rychlosti pohybů prstů a maximální změny frekvence v čase. Hodnoty těchto parametrů pak byly ve shodě s hodnotou skóre dotazníku MDS-UPDRS. Pro výpočet parametrů pak ve svých experimentech použil více postiženou ruku, vybranou na základě subjektivního hodnocení dvou expertů ze zhlédnutého videozáznamu.

1.2.1 Bradykineze

Systém Radima Krupičky^{31,32} byl vyvinut zejména pro záznam *bradykineze* u pacientů s PN. **Bradykineze** (česky také *zpomalenost* pohybu) je jedním ze třech hlavních znaků PN, společně s třesem a ztuhlostí (rigiditou) a mezi nimi je nejvýraznější. Projevuje se, jak český název napovídá, zpomaleností pohybů pacienta, přičemž zpomalenost se vyazuje v průběhu vykonávané pohybové činnosti - *akineze* pak označuje problémy se samotným započítáním pohybu^{33,34}. To pak vyvolává následující:

- Obecné zpomalení každodenních pohybů a činností (čištění zubů, psaní, chůze, ...) a jejich zvýšená obtížnost
- Problémy se započítáním pohybů, například vstávání nebo započítání chůze
- Snížení rozsahu/amplitudy pohybů
- Omezení automatických pohybů, jako mrkání či pohyby rukou při chůzi
- Snížení mimiky a výrazovosti v obličeji

Dále je také vlastností bradykineze její nahodilost (stav a intenzita se mění rychle a často nahodile) a citlivost na *multi-tasking*, tedy děláni vícero činností najednou, například slovní komunikace a současné provádění pohybu. Tyto úkoly se pak stávají pro pacienty s bradykinezí obtížnými, a společně s důslednou sníženou mobilitou a nedobrovolnou změnou rychlosti životního stylu působí na pacienty negativně i po psychické stránce.

Kromě léčby příčiny bradykinesie v podobě PN pomocí levodopy a jiných druhů medikace pomáhá také pravidelná fyzioterapie a pohyb³⁵ (přes jeho obtížnost) stejně jako psychoterapie pro udržování dobrého psychického stavu pacientů. Zajímavostí také je efektivní využití hudební terapie pro zmírnění příznaků bradykinesie³⁶ a současné vylepšení kvality života a psychického stavu u pacientů s PN. Hudební terapie byla také využita pro zmírnění účinků PN na řeč pacientů.³⁷

Bradykineze intenzivně postihuje rychlé a repetitivní pohyby a je nejvýraznějším příznakem PN, proto je FT-test jedním z dobrých kandidátů na zaznamenání příznaků PN v počátečních stádiích.

1.3 RBD - REM Sleep Behavior Disorder

RBD, z anglického *REM Sleep Behavior Disorder*, česky pak *porucha chování v REM spánku* je, jak již název napovídá, spánková porucha (konkrétněji *parasomnie*), při které pacienti ztrácí v průběhu spánku, konkrétně při jeho REM fázích (*Rapid Eye Movement*, česky volně jako "rychlé pohyby očí"), přirozenou svalovou paralýzu a jejich tělo se může, zatímco sní, pohybovat, jinými slovy je jim nedobrovolně umožněno fyzicky vykonávat obsahy svých snů. Pokud se například pacientu postiženému RBD zdá, že s někým svádí pěstní souboj, spící bude v posteli provádět údery rukama do vzduchu, pokud bude mít sen jehož obsahem bude běh, bude v posteli hýbat nohama. Podobné chování je často viditelné například u psů, když "běhají ze spaní".

RBD se vyskytuje v dvou formách - idiopatické a symptomatické. V idiopatické formě je RBD pokud se vyskytuje samostatně, bez jiné, související a dříve diagnostikované nemoci, které by mohla být příznakem. V symptomatické formě se pak RBD nachází pokud je "pouze" příznakem jiné nemoci, například jedné ze synukeinopatických onemocnění.³⁸

1.3.1 Diagnóza

Diagnóza RBD je často komplikovaná, zejména u pacientů, kteří žijí sami, protože postižení RBD nemusí nutně narušovat kvalitu spánku samotného, a tak může trvat několik let, než pacient trpící RBD svojí nemoc nějakým způsobem zaznamená. Často dochází k záchytnu nemoci tak, že pacient na sobě po probuzení najde zranění, o kterých si není vědom, že by je utřil. Zranění spojená se spánkem (SRI, z anglického *sleep-related injuries*) mají pak u pacientů s PN nemocí velmi vysokou provázanost s diagnózou RBD.³⁹ Díky obtížnému určení počátku nemoci se pak pro výzkumné účely bohužel obtížně určuje i délka postižení (doba od chvíle, kdy se začaly objevovat první příznaky). Nejdůležitějším diagnostickým nástrojem je pak klasická *polysomnografie*, tedy záznam a analýza tělesných funkcí během spánku ve spánkové laboratoři, kde je možné přesně evidovat ztrátu svalové atonie během REM fáze spánku. Vedle polysomnografie je pak důležité konzultovat spolunocležníky pacientů a dokumentovat průběh spánkových incidentů v čase. Ke kvantifikaci problémů mohou posloužit také hodnotící dotazníky.^{38,40}

Diagnóza RBD pomocí zobrazovacích technik (MRI) má zatím omezené použití na záznam změn ve specifické oblasti mozku (kde nastává ztráta neuromelaninu). Předmětem výzkumu jsou také nálezy alfa-synukelinu u pacientů s RBD.⁴¹

1.3.2 Léčba

Přestože příčiny RBD se nedají úspěšně léčit (stejně jako se nedají zatím léčit neurodegenerativní onemocnění, která za RBD často stojí), efektivní symptomatická léčba existuje. Medikace podávaná pacientům s RBD se v současné době sestává buď z dávek *melatoninu* ("spánkového hormonu") nebo *klonazepamu* (v České republice prodáván pod názvem *Rivotril*, jednoho ze skupiny benzodiazepinů. Melatonin je pak zvláště prospěšný a preferovaný pro pacienty RBD, zejména díky své nenávykovosti a omezeným vedlejším účinkům. Kromě medikace je pak často nutné upravit prostředí ve kterém pacient tráví svůj spánek, například zabezpečit okolní nábytek tak, aby se minimalizovalo riziko poranění, přesunout postel či matraci blíže zemi aby se zabránilo pádům či odstranit z dosahu spícího nebezpečné předměty jako zbraně a podobně. Občas je pro pacienty prospěšné pořízení lokačního alarmu, který je může probudit při opuštění postele. Také je nutné pečlivě sledovat kontraindikaci jiných léčiv, která by mohla příznaky RBD rapidně zhoršit a dbát na zdravou spánkovou hygienu (vyhnout se spánkové deprivaci, mít pravidelný spánkový režim, omezit alkohol apod.).³⁸ Některé studie se také zabývají způsoby jak omezit neurodegenerativní změny v mozku nálezu RBD pomocí změn životního stylu a stravovacích návyků pacientů.⁴¹

1.3.3 Spojitost s PN

Důvodem, proč je RBD velmi významná ve spojitosti s PN je fakt, že velké procento pacientů postižených RBD je v budoucnu zasaženo nějakou ze synukleinopatických chorob jako je PN, DLB nebo MSA. Fenokonverzi do jedné z nich prodělá v rozmezí 2-5 let zhruba 15%-35% pacientů, v delším časovém odstupu 6-15 let pak dokonce mezi **41%-91,9% pacientů** původně postižených idiopatickou RBD. Kromě vysokého procenta konverze nemoci získaného z longitudinálních studií sledujících vývoj RBD je také nevyhnutelná spojitost s PN doložena klinickým mapováním výskytu neurodegenerativních biomarkerů. U pacientů s RBD se často vyskytují stejné počáteční/prodromální příznaky jako u lidí postižených PN, například nízký tlak, zácpy nebo problémy s čichem, a zároveň byly pomocí několika různých zobrazovacích technik zaznamenány u RBD pacientů neurodegenerativní změny v částech mozku (například mozkového kmene) vykazující náběh nemoci směrem k některé ze synukleinopatií. Také byly u zesnulých pacientů, kteří v době smrti trpěli idiopatickou RBD, nalezeny typické ukazatele v podobě Lewyho tělísek a neuritů a s nimi spojená tkáňová degenerace (a celkově byly příznaky synukleinopatií nalezeny při pitvě u 94% zesnulých pacientů, u kterých byla dříve polysomnograficky potvrzena diagnóza RBD). Obecně se tak diskutuje o možnosti, že RBD není samostatné onemocnění ale pouze jedním z příznaků

neurodegenerace způsobené například PN, nebo o kategorizaci PN s RBD jako samostatného a distingovaného podtypu PN, který se mimo jiné vykazuje intenzivnější formou nemoci, včetně motorických postižení.^{40,42}

Komplexní studii motorických změn u pacientů s RBD pak provedla a publikovala v roce 2012 skupina kanadských vědců.³⁰ Počínaje rokem 2004 sledovala 78 pacientů s idiopatickou formou RBD a každý rok byla provedena u každého z nich sada klinických testů pro zhodnocení jejich motorických schopností. Kromě standardního zhodnocení pomocí UPDRS dotazníku (viz 1.1.3) byl proveden test "střídavého klepání" (ang. *alternate-tap test* v jeho analogové formě, tedy s poklepáváním na kovové disky propojené s mechanickým počítadlem, test s využitím kolíčkové desky (ang. *Purdue Pegboard*) a test časovaného vstávání, chůze a zpětného posazení na židli (ang. *Timed up-and-go test*). Longitudinální forma studie umožnila přímo zaznamenat konverzi idiopatické RBD na neurodegenerativní onemocnění u 28 pacientů (velkou většinou na PN). Od bodu, kdy bylo u pacientů diagnostikováno neurodegenerativní onemocnění (proběhla konverze) pak vytvořili graf zpětně dokumentující vývoj zaznamenaného skóre jednotlivých motorických testů v čase. Progresi hodnot skóre jednotlivých testů pak podrobili regresi a proložili křivkou (v lineárním případě, který vývoj modeloval nejlépe, přímkou). Z dat pak vyhodnotili čas, kdy hodnoty motorických testů pacientů s RBD začaly vykazovat signifikantní statistickou deviaci od hodnot kontrol (zdravých pacientů). U všech testů se pak našel statisticky významný rozdíl zhruba 4 roky před diagnózou PN. Z jednotlivých součástí UPDRS dotazníku vyšlo zhodnocení bradykineze u pacientů s největším časovým intervalem mezi deviací parametru od kontrol a diagnózou PN, a to 4,2 let (bez započtení bradykineze v obličeji a hlase, které se manifestovaly velmi brzy). [30, str. 5]

Celkově v jejich studii shledali soubor jednoduchých motorických markerů jako efektivní a použitelný nástroj k zachycení neurodegenerativních nemocí v prodromálním stádiu. Je nutné dodat, jak autoři studie sami upozorňují, že diagnóza PN u pacientů byla provedena subjektivním klinickým vyšetřením, její čas není tedy naprosto precizní a při alternativním vyšetření jiným lékařem by se mohl pohybovat v rozmezí plus-mínus jednoho roku. Také upozorňují na možnost nelineárního průběhu motorických změn (a související nepřesnost využitého lineárního regresního modelu) s ohledem na průběh postižení před-a-po zasažení substantia nigra v mozku pacienta. Nakonec zmiňují existenci subtypu PN charakteristického přítomností RBD (tzv. *PN-RBD*), jehož příznaky a jejich progres v čase může být významně odlišná od průběhu nemoci u pacientů s PN bez příznaků RBD.

Výsledky této studie tvoří významnou motivaci pro obsah této práce, viz část 2.

1.4 Cíle práce

Cílem práce je ideologické a praktické propojení všech centrálních aspektů prezentovaných v úvodní části. Bradykineze jako příznak PN se u pacientů s RBD zatím nikdy objektivně neměřila. Pro účel včasného zachycení PN v prodromálním stádiu budu analyzovat výsledky FT testu, získané pomocí kamerového systému R. Krupičky, od pacientů s idiopatickou formou RBD. Dokázaná přítomnosti postupných motorických změn u pacientů s RBD (vedoucích na následné neurodegenerativní onemocnění) pomocí výsledků klinických testů z longitudinální studie je motivací pro úspěšné nalezení motorických biomarkerů pomocí dat z kamerového systému a FT testu. FT test je vhodným kandidátem pro evaluaci prodromálních motorických změn vzhledem k tomu, že akcentuje bradykineziální aspekt (prodromální) PN, který je jejím významným indikátorem před-i-po fenokonverzi z idiopatické RBD. Kamerový systém R. Krupičky úspěšně dokázal pomocí definovaných parametrů odhalit motorické změny u pacientů s PN.[31, str. 67] Při aplikaci ideologicky podobné analýzy na pacienty s iRBD bude cílem:

- Definovat na naměřených datech parametry, které by dokázaly statisticky rozlišit skupinu kontrol od pacientů s RBD a PD.
- Analyzovat vzájemný vliv laterality horních končetin a výsledných hodnot zvolených parametrů.
- Spojit vybrané efektivní parametry pomocí klasifikátoru a vyhodnotit jeho rozlišovací schopnosti pomocí AUC analýzy.

Výpočty a analýzy budou provedeny ve výpočetním prostředí programu *Matlab*⁴³ od vývojářské společnosti *MathWorks* ve verzi *R2017a*. Stejný program pak bude použit i na generování všech grafů. V následující části a části o metodice analýzy (implementaci analytických skriptů) se budu odkazovat na jednotlivé programové skripty napsané ve stejnojmenném programovacím jazyce, *Matlab*. Tyto skripty budou pak pod příslušnými jmény k nalezení na přiloženém disku. Všechny skripty a funkce jsou ve formátu *Matlabovských* souborů s příponou *.m*, tzv. *m-files* nebo počestně *m-fajly*. Datové balíky používané v mých skriptech jsou pak uloženy v *Matlabovském* datovém formátu s příponou *.mat*. Grafy jsou opět v *Matlabem* zpracovatelném a vygenerovaném formátu s příponou *.fig* (z ang. *figure*). Všechny grafy v této práci budou v anglickém jazyce, s překladem a českými vysvětlivkami v anotaci pod obrázkem. Rozhodl jsem se tak zejména z důvodu problematické podpory české diakritiky při procesu generování grafů.

Kapitola 2

Databáze

V úvodní části byla představena Parkinsonova nemoc (1.1), její vliv na společnost a základní charakteristiky. V části 1.1.3 byla shrnuta důležitost včasného zachytu nemoci v prodromálním stádiu pro efektivní léčbu příznaků pomocí současných i v budoucnu případně dostupných technik. Tato část také otevřela tematiku biomarkerů, které by se daly použít pro kvantifikaci a analýzu postupu PN a jejích symptomů u pacientů. V části 1.2 byl představen kamerový systém vyvinutý R. Krupičkou pro objektivní záznam jednoho z hlavních příznaků PN, bradykineze. Bylo zmíněno, že parametry definované na datech získaných ze zmíněného kamerového systému byly schopny úspěšně rozlišit pacienty s PN a kontrolní, zdravé jedince. V předposlední části úvodu byla představena RBD, spánková nemoc vystupující jako jeden z významných a velmi častých prekurzorů PN. Vztah s PN a existence motorických příznaků u pacientů trpících RBD ještě před fenokonverzí do PN byla popsána v sekci 1.3.3. Nakonec byly představeny cíle této práce.

2.1 Datový soubor

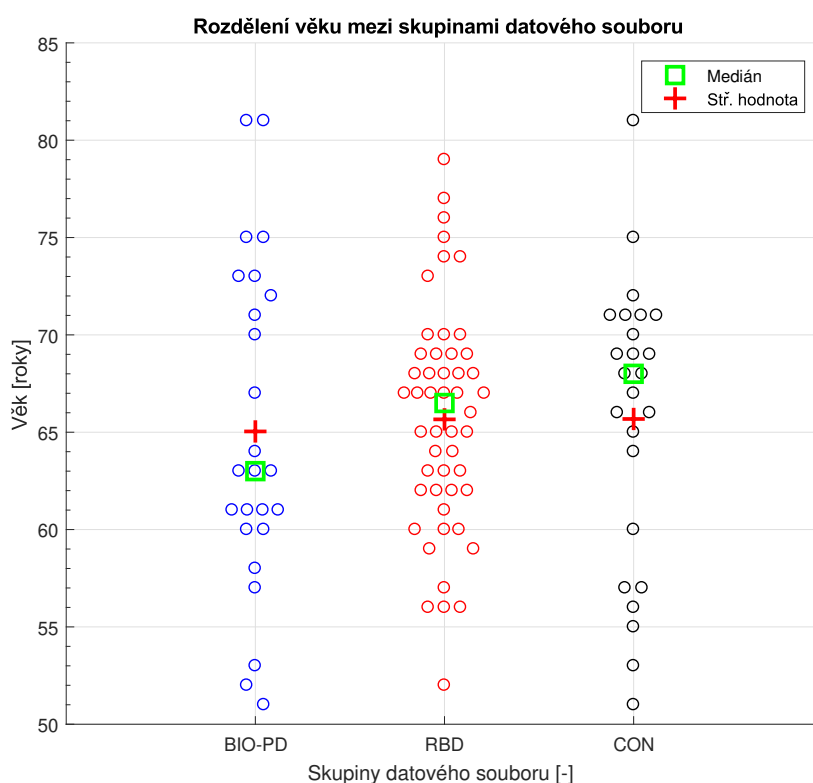
Z dostupných klinických záznamů byl sestaven vyvážený datový soubor dělicí se na tři skupiny měřených osob:

- **Pacienti s idiopatickou RBD** - dále také ozn. jako skupina *RBD*
- **Pacienti s nově diagnostikovanou PN** mírné až střední intenzity - dále také ozn. jako skupina *BIO-PD* či *BIO*
- **Zdravá, kontrolní skupina** - dále také ozn. jako skupina *CON*

Celková velikost datového souboru činí 100 jedinců, 50 z nich ve skupině RBD, 25 poté v každé ze dvou zbývajících skupin. Cílem bylo vytvořit vyvážený soubor z hlediska pohlaví, věku a dalších vlivů. Jedinci byli do skupin vybráni tak, aby byla obě pohlaví (s pochopitelným omezením se na zjednodušující binaritu pohlaví) zastoupena ve stejném poměru. Stejně tak byl kladen vliv na v průměru stejné věkové rozložení všech skupin. Výsledné parametry všech skupin ukazuje tabulka 2.1.

Skupina	<i>BIO-PD</i>	<i>RBD</i>	<i>CON</i>
Průměrný věk [let]	65	65,7	65,7
Max. věk [let]	81	79	81
Min. věk [let]	51	52	51
Počet mužů [-]	20	44	22
Počet žen [-]	5	6	3
Poměr ženy/celek [-]	0,2	0,12	0,12

Tabulka 2.1: Vlastnosti jednotlivých skupin v datovém souboru využitém pro analýzu v této práci. Skupinu s označením *BIO-PD* tvoří pacienti s diagnostikovanou PN. Skupinu *RBD* tvoří pacienti s idiopatickou RBD. Kontrolní zdraví jedinci pak tvoří skupinu označenou jako *CON*. Tabulka dokumentuje snahu o vyváženou skladbu jednotlivých skupin na základě věku a genderového rozložení. Maximální věk subjektů skrze všechny skupiny je 81 let, minimální 51 let.



Obrázek 2.1: Graf věkového rozdělení v jednotlivých skupinách datového souboru. Na vodorovné ose jsou popisky jednotlivých skupin, na svislé ose pak věk subjektů. Skupiny jsou od sebe barevně odlišeny, přičemž každý kruh reprezentuje jednoho člena příslušné skupiny. Zelený čtverec ukazuje medián stáří příslušné skupiny, červený kříž pak aritmetický průměr. Pro vygenerování tohoto grafu (typu *Beeswarm*) byl použit skript `plotSpread`.⁴⁴

U všech subjektů v datovém souboru byl dán k dispozici soubor (anonymizovaných) klinických dat, z nichž některá budou v rámci této práce použita.

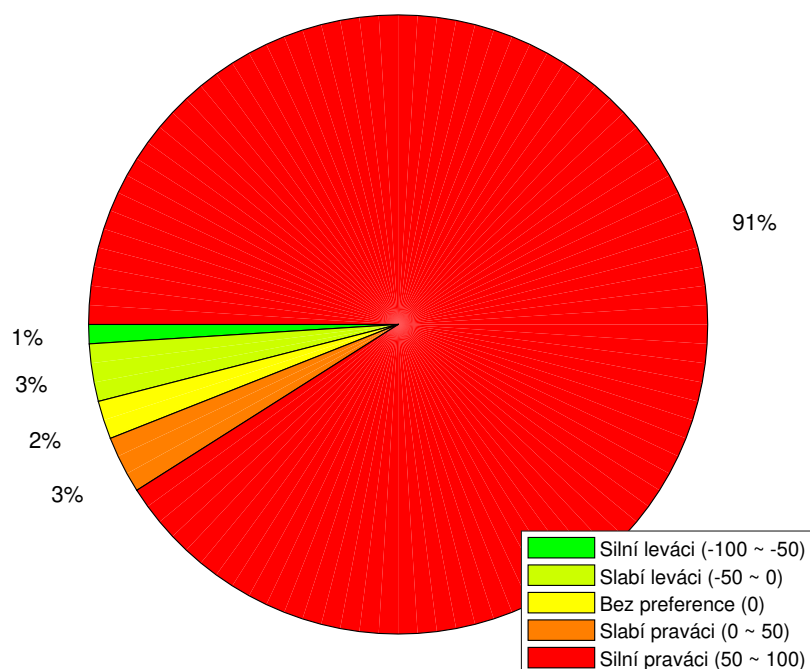
Jsou to:

- **Délka onemocnění u pacientů s PN** - časový interval (v letech) od diagnózy PN.
- **Délka onemocnění u pacientů s RBD** - časový interval (v letech) od diagnózy RBD. Problematický s ohledem na komplikovanou diagnózu RBD, viz sekce 1.3.1. Je možné, že subjekty trpěly příznaky již podstatně delší dobu, než byla RBD kvalifikovaně diagnostikována.
- **Skóre Edinburského dotazníku rukosti** - hodnota hrubě definující lateralitu subjektu. Více informací k Edinburskému dotazníku v sekci 2.2.1. Viz také 2.2.
- **Skóre dotazníku MDS-UPDRS**, konkrétně části MDS-UPDRS III, která se zabývá motorickými projevy pacientů. Zároveň jsou k dispozici i hodnocení jednotlivá hodnocení v této části dotazníku MDS-UPDRS, například (subjektivní) klinická hodnocení FT testu na levé a pravé ruce (hodnota 0-4, kde vyšší číslo indikuje větší postižení). Viz 2.3

Grafy 2.1, 2.2 a 2.3 byly vygenerovány pomocí skriptu `plotClinical.m`. Tento skript jednoduše nahraje klinická data předem uložená v Matlab formátu v souboru `clinical.mat`. V tomto souboru jsou přehledně uložena všechna klinická data vypsaná v seznamu výše. Nahraná data seřadí dle libovolně zvoleného pořadí skupin a pak jednotlivé zvolené datové sloupce společně s kategoričnými popisky předá v argumentu do funkce `plotSpread`, která se postará o komplexní vykreslení všech skupin v grafu typu `beeswarm`, včetně zobrazení mediánu a aritmetického průměru. V případě koláčového grafu rozdělení laterality se pak použije jednoduchá funkce `histcounts` pro výpočet četností v jednotlivých "binech" (česky volně "škatulkách", t.j. podskupinách celku s předem definovanými hranicemi). Hranice binů jsou definovány manuálně na základě rozpětí možných hodnot u EHI. Výsledek je vykreslen pomocí funkce `pie` pro tvorbu koláčových grafů a ručně obarven příhodnou barevnou škálou.

Všechny měření pacientů probíhaly v souladu s předem etablovaným postupem použitým při měření v rámci disertační práce R. Krupičky³¹, pod záštitou prof. Evžena Růžičky, vedoucího Neurologické kliniky 1. lékařské fakulty Univerzity Karlovy a Všeobecné fakultní nemocnice v Praze. Měření bylo schváleno etickou komisí Univerzity Karlovy a Všeobecné fakultní nemocnice v Praze a všichni pacienti podali informovaný souhlas. Vyšetření probíhala na půdě kliniky mezi roky 2015 a 2018. Datové záznamy byly vyhotoveny a měření realizována ještě před začátkem této diplomové práce. S odkazem na postup v disertační práci byly měření prováděna na pacientech, jimž byla náplň testu pečlivě vysvětlena. Měřené osoby pak prováděli FT test po dobu přibližně 20 sekund (více dále v sekci 2.5). Test byl proveden na obou rukou, pokud možno na každé ruce ve dvou iteracích s krátkým odpočinkem mezi nimi.

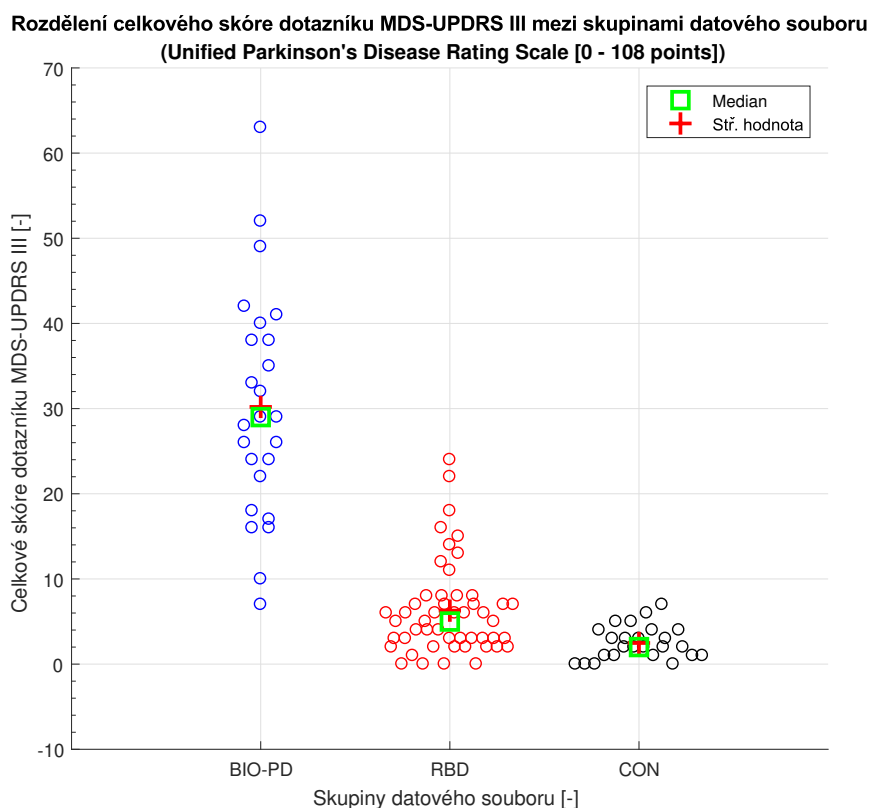
**Rozdělení rukosti všech osob z datového souboru (100 osob)
na základě Edinburského dotazníku rukosti**



Obrázek 2.2: Rozdělení lateralit mezi všemi subjekty z datového souboru. Kolačový graf zobrazuje rozdělení do následujících skupin: 1. Silní leváci (*Edinburgh Handedness Inventory* - dále jen EHI - skóre mezi -100 a -50), 2. Slabí leváci (EHI skóre mezi -50 a 0), 3. Bez preference (EHI skóre rovno 0), 4. Slabí praváci (EHI skóre mezi 0 a 50), 5. Silní praváci (EHI skóre mezi 50 a 100). Graf ukazuje masivní převahu praváků u subjektů v datovém souboru, 94%. Počet leváků je 4 %, což je pod soudobým odhadovaným počtem leváků v populaci (zhruba 10%⁴⁵). Nižší zastoupení leváků mezi subjekty datového souboru se dá vysvětlit věkovým rozložením, kde převažují starší lidé nad 50 let.⁴⁶ Více o Edinburském dotazníku rukosti v sekci 2.2.1.

2.2 Lateralita a rukost

Lateralita je lidská preference jedné strany lidského těla nad jinou. V kontextu horních končetin se pak podmnožina lateralit nazývá česky také *rukost* (specificky pak *levorukost* nebo *pravorukost*), anglicky *handedness* (analogicky *left-handedness*, *right-handedness*). Lateralita rukou popisuje lidskou vlastnost přirozené preference jedné ruky nad druhou při vykonávání běžných činností. Tato preference je zapříčiněna lepším "výkonem" jedné z rukou - výkonem v tomto kontextu a dále, v příslušných sekcích tohoto textu, je myšlena efektivita ruky při různých činnostech, například její rychlost, síla, obratnost a podobně. Pokud projevuje ruka při určité specifické činnosti/úkolu lepší výkon, tedy vykazuje na základě daného parametru objektivně lepší výsledky,



Obrázek 2.3: Graf typu *beeswarm* zobrazující rozložení celkového skóre třetí sekce dotazníku MDS-UPDRS. Svislá osa zobrazuje hodnotu tohoto skóre pro jednotlivé subjekty datového souboru, vyznačené barevnými kruhy. Horizontální osa obsahuje popisky jednotlivých skupin v datovém souboru. Skupiny jsou v grafu barevně rozlišeny. Zelený čtverec vyznačuje medián MDS-UPDRS III skóre vypočtený ze všech členů příslušné skupiny. Kříž červené barvy ukazuje aritmetický průměr přes všechny hodnoty MDS-UPDRS III skóre příslušné skupiny. Z grafu je evidentní distribuce MDS-UPDRS skóre ve skupině pacientů s PN v širokém spektru, s největším počtem hodnot okolo 30 bodů na škále a s maximy mezi hodnotami 40 až 70 bodů. Subjekty ze skupiny RBD mají pak největší koncentraci na škále pod hodnotou deseti bodů s očekávanou tendencí k vyšším hodnotám. Pro vygenerování tohoto grafu (typu *Beeswarm*) byl použit skript `plotSpread`.⁴⁴

můžeme o ní hovořit jako o *lepší ruce*, druhá ruka se pak obecně dá nazývat v rámci zvoleného hodnocení *horší rukou*. V souhrnném kontextu rukosti se pak člověkem preferovaná (a zpravidla lepší) ruka nazývá rukou **dominantní** (ang. *dominant*), opačná končetina je pak označována jako **nedominantní** (ang. *non-dominant*).

Rozložení leváků a praváků v populaci není rovnoměrné. Počet leváků je odhadován pouze na přibližně 10 procent s mírnou (23 procentní) početní převahou mužských leváků nad ženskými.^{45,47} Důvody pro asymetrické rozdělení laterality v populaci nejsou s jistotou zřejmé a jsou předmětem pokračujícího výzkumu. Převažující teorie však pracují s genetickými faktory

při vývoji plodu a komplexními dědičnými vzorci.⁴⁸ Pramenící z genetického vývoje jedince je i rozdělení mozkových funkcí mezi obě hemisféry. Mozkové hemisféry ovládají příslušné části těla zkříženě, levá hemisféra pravou část těla a naopak. Jedna z teorií naznačuje, že prevalence praváků v populaci je způsobena přítomností jazykového centra v lidském mozku většinou právě v levé hemisféře. Leváci potom mají buď zrcadlově obrácenou funkcionalitu mozku, nebo mají funkční centra rozloženy do obou hemisfér.⁴⁹ Vliv na nevyrovnaný poměr leváků a praváků v populaci může mít také dnes již naštěstí opuštěná praxe přeučování leváků na praváky (v naší zemi se od ní oficiálně opustilo vyhláškou ministerstva školství z roku 1967). Trend pak retrospektivně směrem do minulosti vykazuje nižší procentuální zastoupení leváků v populaci, než je tomu dnes, kdy je leváctví (alespoň v rozvinutých zemích) všeobecně přijímáno.⁴⁶

Shrnutím se dá říct, že role laterality v lidech je velmi komplexní a nedá se jednoduše a s jistotou pevně říci, že jedna strana těla nebo jedna ze dvou končetin bude mít při všech činnostech lepší výkon. Například autor tohoto textu je přirozený levák z hlediska psaní a většiny běžných činností (hraní na hudební nástroj, krájení nožem, držení hokejky, . . .), avšak u několika z nich používá pravou ruku, například u stříhání nůžkami. Zhodnocení laterality pomocí dotazníků, jakým je kupříkladu Edinburský dotazník rukosti blíže popsáný v části 2.2.1, dává obecný vhled do výkonnosti a preference končetiny u jedince, nedá se ale se skálopevnou jistotou říci, že člověk seznáný jako silný člen jedné ze skupin rukosti bude mít automaticky při každé činnosti příslušnou stranou těla lepší výsledky.

Analýza laterality s ohledem na výběr ruky pro výpočet parametrů na datech a další vlivy bude popsány v sekci 3.3 a její výsledky budou prezentovány v sekci 4.2.

■ 2.2.1 Edinburgh Handedness Inventory

Edinburský dotazník rukosti (ang. Edinburgh Handedness Inventory, zkr. EHI)⁵⁰ je škála a k ní příslušný hodnotící dotazník pro evaluaci rukosti (laterality), tedy vyhodnocení toho, která je dotazovaného preferovaná strana těla, respektive horní či dolní končetina. Jde o velice prostý dotazník čítající 10 otázek, které mohou být zodpovězeny přímo dotazovaným člověkem nebo za pomoci tazatele, který pozoruje hodnoceného člověka při činnostech, které jsou předmětem dotazníku.

V každé z deseti otázek se vybírá jedna z pěti možných odpovědí, na základě subjektivní a osobní preference dotazovaného u jednotlivých činnostech, které tvoří položky dotazníku. Za každou odpověď pro levou či pravou ruku se do příslušné skupiny připočte určitý počet bodů.

Možné odpovědi a příslušné body jsou:

1. *Silná preference pravé ruky* - 2 body pro pravou skupinu
2. *Slabá preference pravé ruky* - 1 bod pro pravou skupinu
3. *Bez preference* - 1 bod pro obě skupiny
4. *Slabá preference levé ruky* - 1 bod pro levou skupinu
5. *Silná preference levé ruky* - 2 body pro levou skupinu

Položky dotazníku, na které pak dotazovaný odpovídá pomocí možností výše, jsou následující: **a)** Psaní; **b)** Kreslení; **c)** Házení; **d)** Používání nůžek; **e)** Čištění zubů; **f)** Používání nože (bez použití vidličky); **g)** Používání lžice; **h)** Používání koštěte (dominantní ruka, t.j. ruka položená výše na koštěti); **i)** Škrtnání zápalkou, a **j)** Otevírání sklenic.

Vyhodnocení pak probíhá jednoduchým výpočtem pomocí vzorce 2.1.

$$E = 100 \cdot \frac{\sum R - \sum L}{\sum R + \sum L} \quad (2.1)$$

V rovnici 2.1 značí E výslednou hodnotu EHI pro dotazovaného, R tvoří soubor všech bodů pro pravou skupinu a L tvoří soubor všech výsledných hodnot pro levou skupinu. Suma platí přes všechny hodnoty v souboru.

Nutno dodat, že výpovědní hodnota EHI je spíše orientační a její síla tkví v jednoduchosti a rychlosti aplikace a vyhodnocení dotazníku. Sám autor seznává, že výsledky dotazníku nemusí být dostačující ve výzkumech kde manuální či cerebrální laterální tvoří podstatný prvek celku [50, str. 110].

2.3 Struktura datového souboru

Data z měření FT testu u tří skupin lidí byla v surovém formátu textového souboru. Každý soubor s příponou `.dat` obsahuje záznam jednoho z měření, jednoho pacienta na jedné ruce. Obsahem každého textového souboru jsou pak informace o pozicích reflexních značek kamerového systému v prostoru, přičemž každý vzorek v čase tvoří jeden samostatný řádek. Každý časový vzorek pak v první numerické hodnotě obsahuje časovou značku, následují tři číselné prostorové souřadnice x_1, y_1, z_1 první reflexní značky a tři číselné prostorové souřadnice druhé reflexní značky, x_2, y_2, z_2 . Všechny hodnoty jsou pak od sebe odděleny prostou mezerou.

Některá data byla předzpracována mírně odlišným způsobem, kdy je textový soubor uveden řetězcem `#BradykAn_Data_Set` a každá trojice souřadnic reflexních značek má před sebou ještě nulovou kolonku. Hodnoty jsou pak odděleny středníkem.

Všechny soubory obsahující data z jednotlivých měření byla shromážděna na základě zkompletovaného datového souboru v srpnu 2018. Celkem čítá datový soubor 1193 souborů s jednotlivými měřeními. Formát názvu datového souboru je následující:

G_H_N.dat

Kde:

- *G* je identifikátor (štítek, ang. *label*) subjektu: tvoří ho spojení identifikátoru skupiny (*RBD*, *BIO-PD* či *CON*) a následně za ním identifikační číslo osoby (celé číslo buď ve formátu prosté číslovky nebo prefixované nulami - 01, 02, ...)
- *H* je identifikátor měřené ruky a může nabírat následující hodnoty:
 - *L1*, *L2*, ... : standardní měření na levé ruce, číslo označuje pořadové číslo měření
 - *P1*, *P2*, ... : standardní měření na pravé ruce, číslo označuje pořadové číslo měření
 - *LC*, *PC* : měření FT na levé respektive pravé ruce při současném hlasitém počítání od jedničky do desítky, tzv. *double-task*, specificky *counting test* (česky volně "test s počítáním")
 - *LS*, *PS* : měření FT na levé respektive pravé ruce při současném třesení druhou rukou, také *double-task*, specificky *shaking test* (česky volně "test s třesením")
- *N* je pořadové číslo realizovaného měření

Například:

RBD3_L1_1.dat

Je datový soubor obsahující záznam prvního měření, prvního ze dvou v sezení, standardního testu na levé ruce, pacienta s identifikačním číslem 3 ze skupiny pacientů postižených RBD.

2.3.1 Datová reprezentace měření

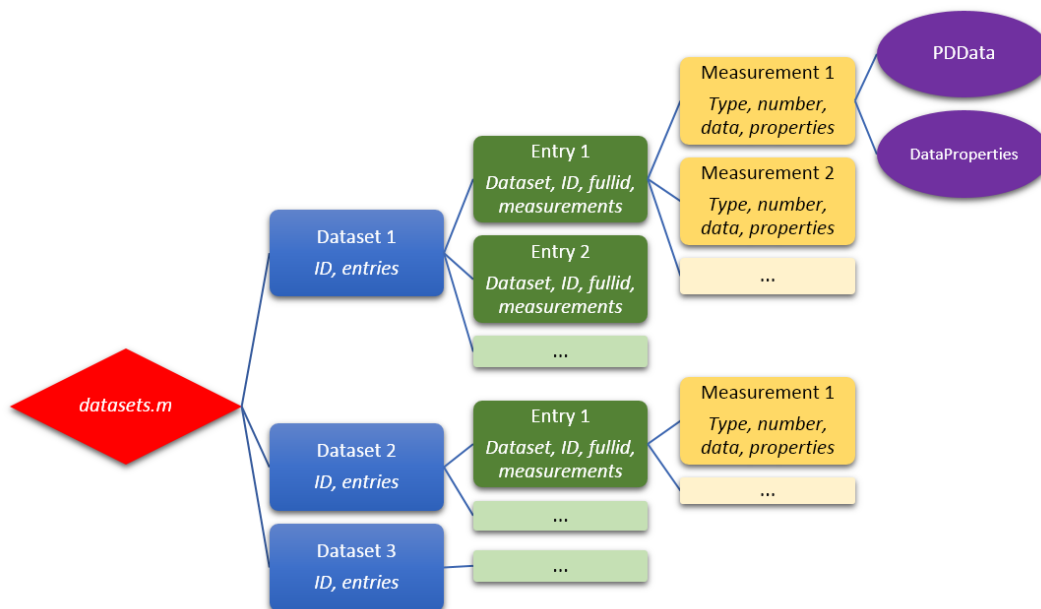
Pro potřeby analýzy v této práci je nutné mít ucelenou datovou reprezentaci jednotlivých měření, se kterou se bude dobře pracovat. Pro tyto účely jsem se rozhodl využít části již existujícího frameworku, který v rámci softwarové části svojí disertační práce³¹ navrhl R. Krupička. S jeho dovolením a pomocí jsem využil programovou třídu `PDData.m`, která slouží k uložení a programové reprezentaci jednoho měření z datového souboru. Jedná se o standardní matlabovskou třídu, obsahující sadu parametrů popisující jednotlivá měření. Pro účely analýzy byly pak používány jen některé z nich. Pro více informací se odkazují na disertační práci R. Krupičky.³¹

Centrální pro přehlednou práci s daty je pak funkční skript `getDataObjects.m`, jehož úkolem je následující:

1. Nahrát z příslušného adresáře na disku jednotlivé datové soubory s příponou `.dat`.
2. Získat z datového souboru data o pozicích reflexních značek.

3. Vytvořit instanci datové třídy `PDData` a vložit do ní data o pozici značek. Instance pak sama při inicializaci vypočte své parametry.
4. Vytvořit řetězec vnořených matlabovských struktur jako wrapper (obal) nad prostou třídou `PDData`, který strukturovaně rozdělí data do systému s logickou hierarchií.
5. Při vytváření vnořených struktur vyhodnotit sadu parametrů na datech měření.
6. Takto vytvořený soubor vnořených struktur uložit na disk pro rychlý přístup.

Přestože se může zdát výše uvedený postup složitý, je velmi praktický s ohledem na další práci s daty. Hierarchie struktur je možno názorně vidět v obrázku 2.4.



Obrázek 2.4: Hierarchie datové struktury vytvořené pro snadnou práci s daty měření. Obdélníky značí struktury, fialové kruhy jsou instance tříd. Červeně je datový soubor s příponou `.mat`. Modře jsou jednotlivé datové soubory, zeleně subjekty (pacienti, lidé) a žlutě jednotlivá měření.

- Červený kosočtverec značí soubor typu `.mat` s uloženými datovými sety.
- Druhá úroveň obsahuje jednotlivé datové sety, vyobrazené modrými obdélníky. Každý datový set reprezentuje jednu skupinu subjektů datového souboru (pacienti s RBD, pacienti s PN, zdravé kontroly). Každý datový soubor je struktura, všechny struktury jsou rozlišitelné podle obdélníkových okrajů. Struktura datového souboru obsahuje pole s identifikátory a jednotlivé vstupy (subjekty, pacienti) uložené v poli struktur `entries`.

- *Entries* obsahují jednotlivé subjekty jako struktury vyznačené zelenou barvou. Každý subjekt má pole s identifikátorem datového souboru, ke kterému patří (*dataset*), identifikační číslo subjektu *ID*, kombinaci předchozích dvou (*fullid*) a pole struktur s jednotlivými měřeními, které přísluší k subjektu (*measurements*).
- Nakonec, ve žlutých obdélnících jsou zobrazeny jednotlivé struktury s měřeními. Pole těchto struktur jsou *type*, obsahující typ měření (*L1*, *L2*, ...), *number* obsahující pořadové číslo měření, *data*, ve kterém je uložena instance třídy *PDData*, ve které jsou uloženy samotná data, viz sekce 2.3.1, a konečně *properties*, ve kterém je uložena instance třídy *DataProperties* s vypočtenými hodnotami parametrů (viz sekce 3.1).

Všechny subjekty v datových souborech jsou do struktur představených výše vkládány seřazené podle svých identifikačních čísel. To je důležité pro robustnost a přehlednost v dalších částech analýzy, kdy se na tuto seřazenost bude spoléhat. Do uloženého `mat`-souboru se také uloží aktuální datum a čas pro případnou referenci.

■ Třída *PDData*

Třída *PDData* má řadu vlastností (ang. *properties*, v kontextu třídy jsou to třídní proměnné).

Vytvoření instance třídy probíhá přes statickou metodu `Load`, která v argumentu bere absolutní cestu k datovému souboru s měřením (s příponou `.dat`). V těle této metody se vytvoří prázdná instance třídy, do které se následně uloží v podobě třídních proměnných plná cesta v souboru na disku. Poté se do další proměnné uloží kompletní data získaná ze souboru: buď se nahrají jednoduše pomocí funkce `load`, která je automaticky zformátuje a vytvoří z nich matici respektující uspořádání v datovém souboru (viz výše) nebo se pomocí standardní funkce `fgetl` postupně vyčtou všechny řádky v textovém souboru a zformátují se do řádků matice funkcí `sscanf`.

Důležitým poznatkem je také unifikování dat tak, aby časové značky byly v sekundách a pozice značek (a následně vypočtené vzdálenosti, viz dále) v centimetrech. V případě časových značek se ještě ode všech odečte hodnota první značky tak, aby začínaly od nuly. Do třídních proměnných `From` a `To` se uloží index první a poslední značky.

Jako poslední se v rámci metody `Load` zavolá další metoda, tentokrát `ComputeAll`. Ta už dostává v povinném argumentu předanou právě vytvořenou instanci třídy. Úkolem této metody je výpočet všech dalších třídních proměnných objektu *PDData*. Při nahrání objektu ze souboru a následným vytvořením datového objektu se tak celý objekt sám "naplní daty".

Výpočet vzdálenosti značek probíhá při zavolání getterové metody proměnné `DataDist` (z ang. *data distance*). Výpočet vzdálenosti je prostá vzdálenost dvou bodů v tří-dimenzionálním prostoru, viz rovnice 2.2.

$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (2.2)$$

V rovnici 2.2 je D vzdálenosti dvou bodů, x_1, y_1, z_1 jsou kartézské souřadnice prvního bodu a x_2, y_2, z_2 jsou souřadnice druhého bodu.

Třída si také vypočítává aproximaci derivace vzdálenostních dat (jejíž fyzikální reprezentace je rychlost pohybu obou prstů v prostoru). Výpočet je nastíněn v algoritmu 1.

Algoritmus 1: Výpočet aproximace první derivace vzdálenosti prstů podle času

vstup : $N \times 1$ Vektor `DateTime` s časy vzorků
 $N \times 1$ Vektor `DataDist` se vzdálenostmi v čase
výstup : $N \times 1$ Vektor `DDist` s aproximacemi derivace vzdálenosti prstů podle času

```

1 Function FirstDerivative(DateTime, DataDist)
2   DTime = zeros(N) ;           // Inicializace nulami
3   DTime = diff(DateTime) ;
4   DDist = zeros(N) ;
5   DDist = diff(DataDist) ;
6   DDist = DDist ./ DTime ;    // Dělení po prvcích
7 end

```

Zásadní je pak metoda třídy `PDData`, která vypočítává extrémy (minima a maxima) ze záznamu měření. Idealizovaná reprezentace měřených dat vzdálenosti prstů je sinusoida. V případě ideální sinusoidy by testovaný otevíral a zavíral prsty harmonicky, se stále stejnou frekvencí (rychlostí otevírání) a s konstantní amplitudou, tedy otevíral by prsty od sebe stále stejnou mírou.

Pokud si představíme takovou sinusoidu, tak její maximální amplitudy ("vrcholy") by odpovídaly momentu kdy má testovaný prsty co nejvíce od sebe, tedy v maximálním rozpětí. Naopak, nejnižší amplituda ("údolí") odpovídá momentu kdy má zkoušený prsty u sebe a ideálně se dotýkají. Nutno dodat, že díky fyzickému umístění značek na prstech subjektů nikdy nebude minimální amplituda nulová. Pro další parametrické analýzy je velmi důležité znát hodnoty a časy těchto extrémů.

Výpočet probíhá v metodě `ComputeExtremas`, jejíž obsah je stručně popsán v algoritmu 2. Celá metoda je empiricky nastavená tak, aby měla co nejlepší výkon: našla všechny významné extrémy a vyvarovala se označování chyb v signálu či nevýznamných zákmitů jako extrémů. "Šlechtění" metody probíhalo empirickou analýzou na manuálně vybrané sadě reprezentativních signálů s co nejširším spektrem možných tvarů záznamu. Manuálně jsem vybral čtyři reprezentanty datových záznamů odstupňované podle jejich subjektivní obtížnosti z hlediska hledání extrémů, od relativně pravidelného "hezkého" signálu až po velmi nepravidelný signál s mnoha neharmonickými změnami. Na vybrané sadě 4 reprezentantů jsem pak měnil koeficienty metody do té doby, než jsem byl s jejím chováním spokojen (ve smyslu vizuální kontroly výsledků metody, kdy jsem si nechal zobrazit v grafech záznamy měření a nalezené extrémy).

Koeficienty jsou následující:

- **Nrange** - koeficient upravující množství hodnot pro výpočet hraniční hodnoty prominence, viz dále.
- **widthThrCoef** - koeficient pro hraniční hodnotu šířky vrcholů.
- **promThrCoef** - koeficient pro hraniční hodnotu prominence vrcholu.

Prominencí vrcholu se v kontextu funkce `findpeaks`⁵¹ rozumí vzdálenost mezi vrcholem a bližším ze sousedních údolí. Šířka vrcholu je pak vzdálenost mezi body nalevo a napravo od vrcholu ve výšce odpovídající polovině prominence daného vrcholu. Principiálně je cílem mít vrcholy, které nemají příliš malou šířku a příliš malou prominenci. Hraniční hodnota pro minimální požadovanou prominenci je argumentem funkce `findpeaks`, a vypočítá se následovně. Nejdříve se vzorky záznamu seřadí. Vypočte se poměrný počet pozitivních a negativních vzorků na vycentrovaných datech. Tento počet upravuje koeficient `Nrange`. Ze seřazeného signálu se vezme škálovaného počet největších pozitivních hodnot a škálovaný počet největších negativních hodnot. Z obou souborů hodnot se udělají průměry a odečtou se - vyjde tak hrubá průměrná vzdálenost mezi horní a dolní obálkou signálu. Tato hodnota se ještě nakonec znásobí koeficientem `promThrCoef`.

Funkce `findpeaks` pak za pomoci dat, hraniční hodnoty prominence a hraniční hodnoty pro minimální vzdálenost vrcholů v počtu vzorků najde a vrátí indexy a hodnoty nalezených maxim. Společně s nimi vrátí i hodnoty šířky nalezených vrcholů. Za pomoci koeficientu `widthThrCoef` se z výsledku odstraní ty vrcholy, jejichž šířka je pod zvolenou hranicí.

Minima se naleznou snadno, prostým nalezením nejnižší hodnoty mezi dvěma nalezenými maximy. Pro vylepšení funkce jsem implementoval nástavbu - po nalezení minima se funkce podívá na obě strany, dokud nenajde vzorek, který se od kořenového minima vzdálí o více než 1% vzdálenosti mezi kořenovým minimem a maximem příslušné strany. Ve chvíli kdy se najdou takovéto vzorky na obou stranách od minima se vybere vzorek, který leží přesně mezi nimi. Pokud je takový vzorek původním kořenovým minimem, vezme se výsledné minimum to, jinak se vezme za výsledné minimum vybraný vzorek ve "středu údolí". Tato nástavba byla implementována potom, co jsem zjistil, že algoritmus při hledání extrémů má tendenci brát za minima vzorky na jednom z úbočí údolí, ve kterém minimum leží. To by mohlo působit chyby ve výpočtech pracujících s přesným časem nalezených minim.

Nakonec se ještě triviálně do vektoru extrémů přidá první a poslední extrém, které funkce `findpeaks` nehledá. To je zejména z důvodů robustnějšího chování při následném použití v dalších výpočtech.

Nalezené extrémy se pak uloží jako třídní proměnné vytvořeného objektu `PDDData` pro příslušné měření.

Algoritmus 2: Výpočet minim a maxim vzdálenosti prstů v čase**vstup** : $N \times 1$ Vektor DataDist se vzdálenostmi v čase**výstup** : $N \times 2$ Vektor mins, který má v prvním sloupci indexy minim
a v druhém sloupci jejich hodnoty $N \times 2$ Vektor maxs, který má v prvním sloupci indexy maxim

a v druhém sloupci jejich hodnoty

```

1 Function ComputeExtremas(DataDist)
   // Koeficienty
2   Nrange = 100 ;
3   widthThrCoef = 0.2 ;
4   promThrCoef = 0.15 ;
5   posN = floor(Nrange · poměrný počet poz. vzorků) ;
6   negN = floor(Nrange · poměrný počet neg. vzorků) ;
7   s = sort(DataDist, vzestupně) ;
8   promThr = mean (s[prvních posN-vzorků]) - mean (s[posledních
   negN-vzorků]) ;
9   promThr = promThr · promThrCoef ;
10  [maxv, maxix, width] = findpeaks(DataDist, promThr) ;
11  Zbav se příliš úzkých vrcholků
   (width < mean(width) · widthThrCoef) ;
   // Najdi minima mezi nalezenými maximy
12  for i = 1 → N - max - 1 do
13     frame = DataDist(max(i) až max(i + 1)) ;
14     [minv, minix] = min(frame) ;
   // Vylepši pozici minim nalezením okrajů údolí
15     j = 0, k = 0 ;
16     while abs(frame[minix - j] - minv) <= 1%min - max1vzd.
       do j ++ ;
17     while abs(frame[minix + k] - minv) <= 1%min - max2vzd.
       do k ++ ;
18     minix = (minix - j) + floor((j+k) / 2) ;
19  end
20  if První extrém je maximum then
21     | Přidej první vzorek jako minimum ;
22  else
23     | Přidej první vzorek jako maximum ;
24  end
25  if Poslední extrém je maximum then
26     | Přidej poslední vzorek jako minimum ;
27  else
28     | Přidej poslední vzorek jako maximum ;
29  end
30 end

```

2.4 Druhy FT měření

Jak bylo letmo zmíněno v sekci 2.3, kromě standardního FT testu byly u pacientů v rámci klinického vyšetření provedeny tzv. *dual-task* FT měření, při kterých měřená osoba kromě samotného FT testu klepání prsty vykonává ještě další činnost. Byly provedeny dva druhy *dual-task* testů, tzv. *counting* FT test, česky "test s počítáním", při kterém měřená osoba při provádění FT testu nahlas nebo potichu počítá (vzestupně nebo sestupně). Druhý typ testu byl *shaking* FT test, česky "test s třesením", při kterém měřená osoba rukou, kterou neklepe prsty, v průběhu FT testu určitým způsobem třese.

Dual-tasking, tedy česky provádění více (dvou) úloh naráz, se u pacientů s PN využívá pro akcentaci příznaků. Je známo, že motorický výkon se u pacientů s PN snižuje, pokud je pacient nucen vykonávat spolu s pohybem ještě další činnost. To se projevuje nejen u klepání prsty ale i při chůzi, kdy může pacient postižený PN tzv. "zamrzt".⁵² Motorika pacientů s PN se při vykonávání více úloh zároveň logicky zhoršuje, jelikož množství "zdrojů" v mozku, které jsou schopné kontroly a výpočtů úkolů je značně omezeno a mozek lidově řečeno "nestíhá".^{53,54}

Vzhledem ke komplexitě problematiky dual-taskingu se nebudu tomuto tématu v této práci dále hlouběji věnovat. Tématika dual-taskingu v souvislosti s analýzou FT testu u pacientů trpících RBD by vydala na celou další samostatnou práci. Nicméně data z naměřených testů jsem využil v analýze pro (pouze) explorativní náhled do možných souvislostí v kontextu RBD.

Pro potřeby analytických skriptů jsem si pak vytvořil enumerační třídu `TestType`, která v sobě všechny tři typy FT testu (`TestType.Normal`, `TestType.Counting` a `TestType.Shaking`) kategorizuje. Obsahuje také utilitární metodu pro získání správných štítků typu a ruky na základě předaných argumentů.

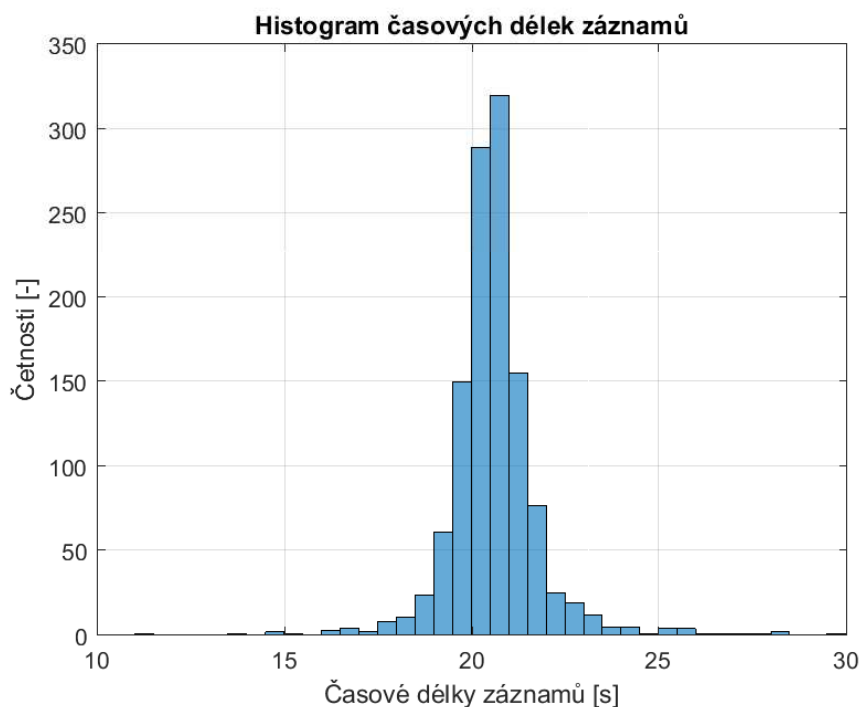
2.5 Délka záznamu

Vzhledem k tomu, že jednotlivá měření v datovém souboru mají relativně benevolentní délku, provedl jsem analýzu délek měření s cílem najít ideální hodnotu pro "ustříhnutí" (ang. *cutoff*) měření v čase. V odkazu na disertační práci R. Krupičky³¹, kde uvádí, že pro zpracování používali prvních 20 sekund signálu (z důvodu vlivu únavy měřených pacientů na měřený pohyb), jsem se chtěl ujistit, že dostatečný počet záznamů má délku alespoň 20 sekund.

Vytvořil jsem skript `analyzeDataLengths`, který si nahraje data pomocí funkce `getDataLengths` a analyzuje délky měření:

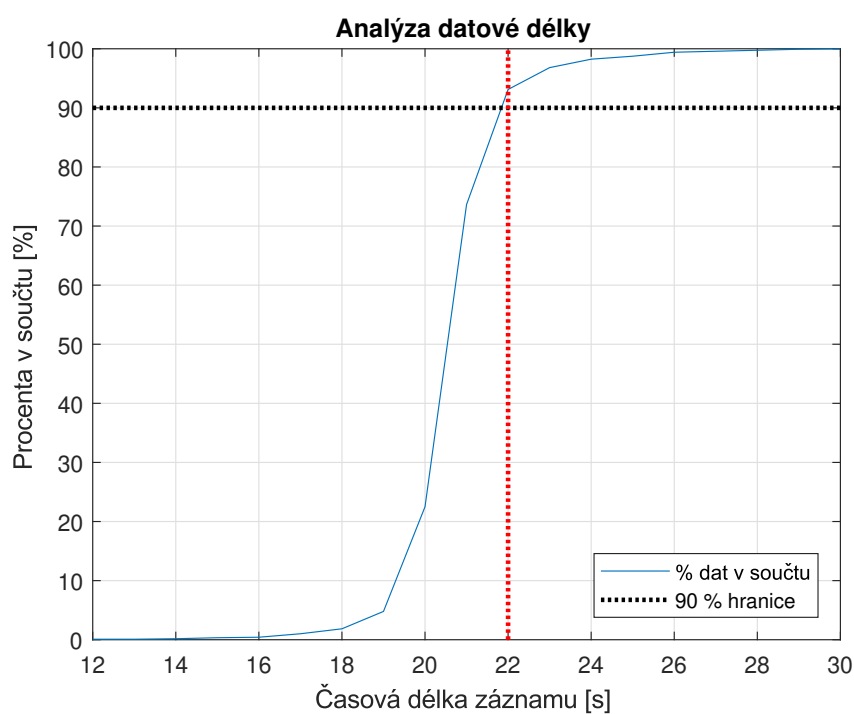
1. V cyklu projde přes všechny datové soubory, všechny jejich subjekty a všechny měření každého ze subjektů a uloží délky měření do vektoru.
2. Nalezne minimální (nejkratší) délku signálu ze všech měření.
3. Pomocí funkce `histcounts` rozdělí získané délky signálů do přihrádek o velikosti jedné sekundy a vypočte počet měření v každé přihrádce.

4. Iterativně projde přes všechny přihrádky od krátké délky záznamu po dlouhou a nasčítává počet všech měření v prošlých přihrádkách. Pokud napočítá dohromady více než 90% všech měření, uloží si číslo přihrádky a procento napočítaných měření.



Obrázek 2.5: Histogram časových délek měřených dat. Na vodorovné ose je vynesena čas v sekundách, na svislé pak počty měření v jednotlivých přihrádkách histogramu. Z grafu je patrné, že valná většina měření má délku mezi 20 a 22 sekundami.

Výsledky analýzy jsou vidět na grafech 2.5 a 2.6. Numerické výsledky pak ukazují, že nejkratší datový záznam má 11,41 sekund. Z histogramu v 2.5 je vidět největší koncentrace délek měření mezi 20 a 22 sekundami. Na grafu 2.6 je pak vidět průběh napočítaných měření přes časové přihrádky. Vodorovná černě tečkovaná čára vytyčuje hranici 90% všech vzorků napočítaných dohromady. Ji-protínající svislá červená čára pak ukazuje čas na vodorovné ose, pro který se výše popsaný algoritmus zastavil. Výsledná hodnota skriptu říká, že počet měření s délkou 22 sekund a méně je 93,12 procent. Jinými slovy, je možné oříznout signál přibližně v tomto čase, abychom zabránili vlivu únavy u zbylého procenta měření. Pro analýzu jsem nakonec zvolil hodnotu 23 sekund pro horní hranici délky signálu. Spodní hranici jsem po konzultaci a manuální kontrole tvaru signálu měření nestanovil, jelikož jsem neseznal, že by počáteční hodnoty měření trpěly jakoukoliv významnou vadou oproti zbytku.



Obrázek 2.6: Graf zobrazující výsledky analýzy délky signálu u měření z datového souboru. Na vodorovné ose je vynesena čas měření v sekundách. Na svislé ose je procentuální vyjádření počtu napočítaných měření z příslušných časových příhrádek. Z grafu je vidět, že zhruba po napočítání všech měření, které mají délku alespoň 22 sekund se překročí devadesátiprocentní hranice všech měření z datového souboru.

Kapitola 3

Metodika

V této kapitole se zaměřím na volbu a popsání parametrů, které budu získávat z naměřených dat. Na základě těchto parametrů pak provedu statistickou analýzu s cílem ověřit hypotézu, že je možné na hladině statistické významnosti $\alpha = 0.05$ rozlišit skupinu pacientů trpících RBD a zdravých kontrol. Zároveň provedu i následnou statistickou analýzu na stejné hladině významnosti, která vezme hodnoty parametrů jak pro zdravé kontroly a pacienty s RBD, tak pacienty trpící PN, tedy provedu ověření hypotézy, že se tyto tři skupiny dají statisticky odlišit, opět na základě hodnot zvoleného parametru. Tato druhá analýza bude spíše pro následnou grafickou ilustraci toho, tak se všechny tři skupiny, z úhlu pohledu zvoleného parametru, chovají vůči sobě.

Popis zvolených parametrů a způsob jejich výpočtu bude součástí této kapitoly. Zároveň v této kapitole popíšu postup, který jsem použil pro analýzu laterality, která úzce souvisí se způsobem výpočtu parametrů z naměřených dat - konkrétně se způsobem výběru toho, jakou ruku použít pro výpočet parametru. V kapitole 4 budu prezentovat výsledky těchto analýz v podobě grafů a tabulek. Nakonec, v kapitole 5 shrnu výsledky a pokusím se je dát do souvislostí.

3.1 Parametry a třída `DataProperties`

Volba parametrů, které budu vypočítávat z naměřených dat FT testu u všech tří skupin byla dlouhodobějším procesem. Byly diskutovány různé formy parametrů s ohledem na proběhlé práce v oblasti a také s ohledem na současně používané parametry, které používají lékaři v klinickém prostředí při vyšetřeních.

3.1.1 Volba parametrů

Výchozím bodem pro volbu parametrů byla předešlá práce R. Krupičky [31, str. 57], jejíž metodika je této práci příbuzná. Pro rozlišování zdravých kontrol a pacientů PN tam byly definovány parametry: **a)** Pokles vzdálenosti prstů; **b)** Průměrná maximální rychlost zavírání a otevírání prstů; **c)** Pokles maximální rychlosti zavírání a otevírání prstů; **d)** Průměrná frekvence; **e)** Pokles průměrné frekvence, a **f)** Hodnocení expertů z videozáznamu. Také

ve své práci používal různé parametry odvozené od délky/času záznamu, ze kterého byly vypočítávány (například posledních X sekund záznamu) - takovou analýzu jsem ve své práci nepoužil, hlavně s ohledem na neunifikované délky záznamů (nejkratší záznam měl délku okolo 11 sekund, průměrná délka byla okolo 20 sekund). Využití této formy analýzy bude možná součástí pokračujícího výzkumu na datech.

Pokles definovaný v práci R. Krupičky je vypočtený jiným způsobem, než v této práci. Pro úplný popis metodiky tohoto a dalších parametrů v jeho práci viz [31, str. 52, 4.3].

Pro volbu končetiny zvolil horší - tedy více postiženou - končetinu na základě subjektivního hodnocení dvou expertů z videozáznamu (s tím, že pokud je hodnocení stejné, je vybrána dominantní končetina).[31, str. 55]. Volba končetiny pro výpočet parametrů v této práci je diskutována v sekci 3.1.3.

Při výběru parametrů byly také zohledněny proběhlé studie a stávající parametry, které při vizuální kontrole subjektivně hodnotí experti provádějící klinické vyšetření pacientů.^{54,55,56} Tři parametry, které jsou nejčastěji používány jsou:

- Pokles amplitudy otevírání prstů (ang. *Amplitude decrement*)
- Maximální rychlost otevírání prstů (ang. *Maximum opening velocity*)
- Průměrná frekvence (ang. *Average frequency*)

■ 3.1.2 Parametry

Vycházejíc z výše uvedeného byl vytvořen soubor celkem 34 parametrů. U několika parametrů jsou vypočítávány podobné vlastnosti (sub-parametry), takže se dá říct, že počet hlavních, kořenových parametrů je 11.

V metodách, ze kterých získávám dříve vypočtená minima a maxima (viz 2.3.1), oříznutá podle zvolené hodnoty maximální délky záznamu, je provedeno *přepočtení vzdálenosti značek z centimetrů na metry* (prostým vydělením hodnot číslem 100).

U parametrů `Max`, `MinDwellTime`, `Freq`, `OpenVel`, `OpenVelPeak`, `CloseVel` a `CloseVelPeak` jsem u všech provedl vyhodnocení vlastností uvedených v seznamu níže. Celé jméno parametru pak tvoří název parametru spojený s názvem vlastnosti na něm zkoumané (například `MaxMed`, `FreqSlope`, `CloseVelStd`, ...).

- `Med`, zkratka pro *median*. Prostý medián všech hodnot příslušného parametru z měření. Pro výpočet byla použita funkce `median`.⁵⁷
- `Std`, zkratka z ang. *standard deviation*, tedy směrodatná odchylka.^a Byla použita funkce `std`.⁵⁸

^aVýpočet směrodatné odchylky podle vzorce $S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |x_i - \bar{x}|^2}$, kde N je počet pozorování, x_i je i -té pozorování, \bar{x} je aritmetický průměr náhodné veličiny x a S je výsledná směrodatná odchylka.

- **Slope**, česky významově směrnice. Principiálně uvádí velikost trendu parametru v čase, tedy informaci o tom, jestli se hodnoty trendu v čase zvyšují, snižují nebo jsou konstantní. Výpočet směrnice probíhá vytvořením lineárního regresního modelu na hodnotách naměřených parametrů v čase. Hodnoty se proloží přímkou, která je algoritmem optimalizována (je jí měněna směrnice a konstantní posun) tak, aby chyby mezi skutečnými hodnotami a hodnotami predikovanými modelem ve stejných časech byly co nejmenší.

$$\mathbf{M} = \min_j \left(\sum_{n=1}^N |\hat{p}_j[n] - p[n]| \right) \quad (3.1)$$

V rovnici 3.1 je M soubor parametrů modelu (pro lineární model je to směrnice a konstantní posun), j označuje j -té nastavení parametrů modelu M , N je celkový počet naměřených hodnot, ze kterých se model sestavuje, $\hat{p}_j[n]$ je hodnota modelu (bod na přímce) nastaveného parametry j na n -té pozici a $p[n]$ je n -tá skutečná hodnota parametru.

Výpočet prakticky probíhá pomocí funkce `polyfit`⁵⁹, jejímž argumenty jsou souřadnice x a y , ze kterých se má model vypočítat a řád polynomu, v případě lineární regrese prostá jednička. Funkce pak vrací ve vektoru vypočtené koeficienty polynomu, v případě přímky tedy dvě hodnoty. Z těchto dvou hodnot se poté jako výstup bere pouze sklon vypočtené lineární regresní přímky (pozitivní v případě, že má parametr, ze kterého byl model vypočítán stoupající tendenci a negativní v případě, že hodnoty parametru v čase klesají).

- **NRMSD**, zkratka z ang. *Normalized Root-Mean-Square Deviation*, česky volně jako "normalizovaná hodnota RMS odchylky (RMS - odmocnina z průměrů umocněných na druhou)". NRMSD je indikátorem "kvality" modelu - tedy míry toho, jak moc se hodnoty vypočteného (libovolného) modelu rozcházejí s hodnotami skutečnými. V případě lineárního modelu udává NRMSD hodnotu, která agreguje všechny chyby (vzdálenosti skutečných hodnot od modelové přímky) do jednoho pozitivního čísla. V případě, že by hodnota NRMSD byla nulová, byl by model perfektní a všechny skutečné hodnoty by ležely na modelové přímce. Nižší hodnoty NRMSD udávají větší věrnost modelu. Normalizace probíhá vydělením hodnoty RMSD (výpočet níže) rozsahem hodnot, tedy rozdílem mezi největší a nejmenší hodnotou dat, ze kterých se NRMSD vypočítává.

$$\text{RMSD} = \sqrt{\frac{\sum_{n=1}^N (\hat{p}[n] - p[n])^2}{N}} \quad (3.2)$$

$$\text{NRMSD} = \frac{\text{RMSD}}{\max(\mathbf{p}) - \min(\mathbf{p})}$$

V rovnici 3.2 RMSD označuje *root-mean-square-deviation* (RMS-odchylku), N je počet hodnot parametru, $\hat{p}[n]$ je n -tá hodnota (lineárního

regresního) modelu náhodné proměnné \mathbf{p} , $p[n]$ je n -tá skutečná hodnota a NRMSD je výsledná hodnota normalizované hodnoty RMS odchylky.

Výpočet modelu použitého pro vyhodnocení NRMSD je totožný s výpočtem lineárního regresního modelu u parametru *slope*. Využije se funkce `polyfit`⁵⁹ pro výpočet parametrů regresní přímky a následně se získají predikované n -té hodnoty modelu pomocí funkce `polyval`⁶⁰. Výpočet je dále totožný s uvedeným vzorcem výše.

Hodnoty parametru NRMSD mají fyzický význam v podobě nestálosti parametru. Předpokládejme například, že z měřených dat a vypočtených hodnot frekvence bude vycházet, že měřená osoba měla tendenci zrychlovat a zpomalovat kmitání prstů, náhodným způsobem, jinými slovy měla problémy s udržením pravidelného *rytmu* klepání prsty. Přestože směrnice lineární regresní přímky bude mít problémy takovou nestabilitu odhalit, parametr NRMSD dokáže agregovat odchylky od lineárního vývoje parametru v čase.

Parametr *DistAvg*

Zkratka z anglického *distance average*, česky průměrná vzdálenost. Tento parametr popisuje vzdálenost maximálního otevření prstů, zprůměrovanou přes celé měření. Výpočet probíhá vypočtením aritmetické střední hodnoty (**mean**) ze všech hodnot maxim a stejné střední hodnoty ze všech minim. Tyto hodnoty se od sebe poté odečtou.

$$P = \left(\frac{1}{N_{max}} \sum_{n=1}^{N_{max}} max_n \right) - \left(\frac{1}{N_{min}} \sum_{m=1}^{N_{min}} min_m \right) \quad (3.3)$$

V rovnici 3.3 N_{max} je počet maxim, N_{min} je počet minim, max_n je n -té maximum a min_m je m -té minimum.

Tento postup je analogický k tomu, kdyby se vzaly maxima a k nim příslušná minima, vypočetla se jejich vzdálenost a výsledné hodnoty by se zprůměrovaly počtem těchto dvojic. Problém toho parametru může být v porovnávání hodnot toho parametru mezi jednotlivými subjekty. Pokud bereme v potaz, že hodnoty jsou normalizované s ohledem na vzdálenost měřeného od kamery, měly by reflektovat skutečnou vzdálenost prstů při měření. Nedá se však předpokládat, že všechny subjekty budou mít stejně velké ruce a dlouhé prsty. Pokud bychom tedy chtěli pomocí tohoto parametru objektivně porovnávat intenzity otevření prstů u pacientů z měření, bylo by nutné hodnoty vzdálenosti rozevření prstů normalizovat například pomocí délky jejich prstů, či velikosti jejich rukou.

Parametry *Max* (MaxMed, MaxStd, MaxSlope, MaxNRMSD)

U parametrů tohoto typu se jako kořen používají maximální hodnoty amplitudy rozevření prstů (maxima). Ty indikují jak moc měřená osoba zvládá rozevírat při testu prsty od sebe. Medián a směrodatná odchylka nejsou normalizované. Slope udává směrnici lineární regresní přímky, kterou se hodnoty maximální amplitudy proloží v čase. Pokud osoba bude mít tendenci otevírat prsty postupem měření méně a méně, směrnice bude negativní. Pokud tomu

bude naopak, směrnice regresní přímky bude mít kladné hodnoty. Hodnota NRMSD pak přeneseně popisuje jak konzistentně (vůči trendu) osoba zvládá otevírat prsty. Pokud bude konstantně lineárně snižovat či zvyšovat amplitudu otevírání, hodnota NRMSD bude malá. Pokud bude chování nahodilé, a amplituda otevření prstů se bude v čase výrazně měnit, hodnota NRMSD bude vyšší.

Parametry *Min* (*MinStd*, *MinNRMSD*)

U parametrů, které berou jako kořenové hodnoty velikosti minim, byly vypočítány pouze směrodatná odchylka a NRMSD. Minimy se v kontextu FT testu rozumí nejnižší amplitudy vzdálenosti prstů, tedy momenty, kdy jsou při pohybu ruky prsty u sebe, nejlépe v dotyku. Ideálně by hodnoty minimální vzdálenosti měly dosahovat téměř nulových hodnot, v praxi ale díky nenulové vzdálenosti reflexních značek na prstech mají hodnoty o něco vyšší. Fyzicky potom měření na těchto hodnotách tohoto parametru zkoumají, jak dobře měřená osoba doklepává prsty k sobě. Vzhledem k tomu, že z vizuální analýzy měřených dat je evidentní, že osoby ve velké většině případů k sobě dovírají prsty konzistentně, omezil se výpočet parametrů pouze na směrodatnou odchylku a NRMSD, které mohou odhalit případné "falešné kmity" či opakující se nedovření prstů v průběhu měření.

Parametry *MinDwellTime* (*MinDwellTimeMed*, *MinDwellTimeStd*, *MinDwellTimeSlope*, *MinDwellTimeNRMSD*)

MinDwellTime, z ang. *Minima Dwelling Time*, označuje dobu, po kterou se měřená osoba "zdržuje" v minimech, jinými slovy, jak dlouhý je časový úsek od spojení prstů po jejich opětovné rozevření. Hodnoty tohoto parametru mají za úkol reflektovat motorické zpomalení či ztuhlost, která se u pacientů s PN vyskytuje. Na časových hodnotách *MinDwellTime* byly vypočítány všechny 4 druhy vlastností definovaných v úvodu této sekce.

Proces výpočtu délek údolí je ilustrován v algoritmu 3. Algoritmus se nejdříve dívá na levou stranu od výchozího minima do té doby, než najde vzorek jehož velikost je větší než jedno procento vzdálenosti mezi výchozím minimem a maximem na dané straně. Proces se analogicky opakuje pro pravou stranu od výchozího minima. Délka údolí se vypočte z časové vzdálenosti nalezených vzorků na koncích údolí po obou stranách od výchozího minima.

Parametr *Jitter*

Jitter je parametr označující frekvenční nestabilitu signálu. Jeho hodnota je v procentech a vypočítává se na základě rozdílu hodnot frekvencí všech dvojic sousedících period v signálu.

$$J = 100 \frac{\frac{1}{N_T-1} \sum_{n=1}^{N_T-1} |T[n+1] - T[n]|}{\bar{\mathbf{T}}} \quad (3.4)$$

V rovnici 3.4 je *J* výsledná hodnota jitteru, *N_T* je počet period v naměřených datech, **T** je vektor obsahující velikosti těchto period, $\bar{\mathbf{T}}$ je aritmetický průměr z hodnot period a *T*[*n*] je *n*-tá hodnota periody z vektoru.

Algoritmus 3: Výpočet času setrvávání v minimech (údolích)

vstup : $N \times 1$ Vektor DataDist se vzdálenostmi v čase
 $N \times 1$ Vektor DataTime s časy vzorků
percThr - konstanta udávající procentuální hraniční hodnotu
pro výšku stran hledaných údolí

výstup : $N \times 2$ Vektor valleys, který má v prvním sloupci indexy
minim v údolí a v druhém sloupci časové délky těchto údolí

```

1 Function getValleys(DataDist, DataTime, percThr)
2   dist = DataDist/100 ;           // Převod vzdálenosti na metry
   // Časově oříznuté hodnoty a časy extrémů
3   MaxV = getCroppedMax() ;
4   MinV = getCroppedMin() ;
5   i0 = 1 nebo 2 ;                 // Index prvního min. s údolím
6   iend = N nebo N - 1 ;       // Index posledního min. s údolím
7   for i = i0 → iend do
8     min = MinV(i) ;
9     max1 = MaxV(i - i0 + 1) ;           // Levé max.
10    max2 = MaxV(i - i0 + 2) ;         // Pravé max.
11    d1 = abs(max1 - min), d2 = abs(max2 - min) ;
12    j = 1 ;
13    rim1time = DataTime(minindex) ;    // Čas levého okraje
14    while true do
15      rim1value = dist(minindex - j) ;
16      if rim1value > (min + (percThr · d1)) then
17        rim1time = DataTime(minindex - j) ;
18        break ;
19      end
20      j ++ ;
21    end
22    j = 1 ;
23    rim2time = DataTime(minindex) ;    // Čas pravého okraje
24    while true do
25      rim2value = dist(minindex + j) ;
26      if rim2value > (min + (percThr · d2)) then
27        rim2time = DataTime(minindex + j) ;
28        break ;
29      end
30      j ++ ;
31    end
32    valleys(i - i0 + 1)index = minindex ;
33    valleys(i - i0 + 1)value = abs(rim2t - rim1t) ;
34  end
35 end

```

Parametr *Shimmer*

Shimmer je podobný parametru jitter ale označuje amplitudovou nestabilitu signálu opět se počítá na základě toho, jak moc se od sebe liší amplitudy dvojic sousedících vrcholků. Za velikost vrcholků se bere vzdálenost dvojic maxim a minim.

$$S = 100 \frac{\frac{1}{N_T-1} \sum_{n=1}^{N_T-1} |A[n+1] - A[n]|}{\bar{\mathbf{A}}} \quad (3.5)$$

V rovnici 3.5 je S výsledná hodnota shimmeru, N_T je počet period v naměřených datech, \mathbf{A} je vektor obsahující velikosti amplitud, $\bar{\mathbf{A}}$ je aritmetický průměr z hodnot amplitud a $A[n]$ je n -tá hodnota amplitudy vypočtená jako $A[n] = |max_n - min_n|$, kde max_n je hodnota n -tého maxima a min_n je hodnota n -tého minima.

Parametry *Freq* (*FreqMed*, *FreqStd*, *FreqSlope*, *FreqSlopeMA*, *FreqNRMSD*)

Tato sada parametrů pracuje s frekvencí signálu. Hodnoty frekvence signálu vypočítává jako reciproční hodnotu velikosti period. Hodnoty period jsou, stejně jako u předchozích parametrů, vypočítávány pomocí funkce `getPeriods`, která prochází všechny maxima v signálu a vrací jejich časové vzdálenosti mezi sebou.

$$f = \frac{1}{T} \quad (3.6)$$

V rovnici 3.6 označuje f frekvenci v hertzech a T je délka periody v sekundách.

Frekvence je jedním z významných parametrů při vyhodnocování FT testu a sleduje celkovou rychlost, kterou jsou schopni měřené subjekty klepání prsty provádět. Stejně jako u předchozích parametrů, i zde se provádí výpočet hodnoty mediánu ze všech hodnot frekvencí v signálu, výpočet směrodatné odchylky a také výpočet směrnice trendu frekvencí v čase (trend zpomalování nebo zrychlování) a NRMSD naměřených hodnot od lineární regresní přímky.

Součástí zkoumání vlastností frekvence je také výpočet směrnice lineární regresní přímky na hodnotách frekvence, které byly vyhlazené pomocí MA (*Moving Average*, česky klouzavé průměry) filtru. Výpočet tohoto vyhlazení provádí funkce `getSmoothPeriods`, jejíž argumenty jsou vektor nevyhlazených časů period a (poloviční) velikost vyhlazovacího okna. Vyhlazování pomocí MA filtru probíhá tak, že se okno "pohybuje" nad vyhlazovanými daty, vezme hodnoty aktuálního vzorku a *wlen*-vzorků před a po aktuálním a zprůměruje je aritmetickým průměrem. Poté se posune o vzorek dál a tento proces se opakuje až do konce signálu. Poloviční velikost okna jsem zvolil jako 3 vzorky, celková velikost průměrování je tedy 7 vzorků, 3 vzorky předcházejícímu, aktuální vzorek a 3 vzorky po aktuálním.

Parametry *OpenVel* (*OpenVelMed*, *OpenVelStd*, *OpenVelSlope*, *OpenVelNRMSD*)

Jak název napovídá, parametry *OpenVel* se budou týkat otevírací rychlosti prstů, z ang. *opening velocity*, tedy rychlost otevírání. V této práci byly použity dva různé přístupy k výpočtu rychlostí otevírání a zavírání: parametry *OpenVel* a *CloseVel* mají hodnoty průměrných rychlostí otevírání a zavírání pro každou periodu, parametry *OpenVelPeak* a *CloseVelPeak* mají hodnoty maximálních rychlostí otevírání a zavírání prstů.

Výpočet (průměrných) rychlostí otevírání prstů probíhá na základě triviálního fyzikálního vzorce $v = \frac{s}{t}$, kde v je průměrná rychlost v metrech za sekundu, s značí dráhu v metrech a t čas v sekundách, který trval pohyb o délce s . Dráhou v mém případě bude délka pohybu od minima k maximu, tedy od bodu dotyku prstů k jejich největšímu rozevření. Časem bude temporální vzdálenost těchto dvou extrémů.

Jako časy příslušné těmto rychlostem pak беру středové body na časové ose mezi okamžiky dotyku prstů (minima) a největšího rozevření (maxima).

$$\begin{aligned} dT &= |t_{max}[n] - t_{min}[n]| \\ dD &= \frac{1}{100} |A_{max}[n] - A_{min}[n]| \end{aligned} \quad (3.7)$$

$$V[n] = \frac{dD}{dT}$$

V rovnici 3.7 označuje dT časovou diferenci, dD je difference amplitud extrémů jednoho otevření, $t_{max}[n]$ a $t_{min}[n]$ jsou časy n -tých maxim respektive minim, $A_{max}[n]$ a $A_{min}[n]$ jsou velikosti n -té maximální respektive minimální amplitudy a $V[n]$ je n -tá hodnota velikosti průměrné otevírací rychlosti. Vydělení hodnotou 100 převede numerické hodnoty rychlosti na metry za sekundu (vzdálenosti jsou původně v centimetrech).

Na tomto parametru, stejně jako na všech dalších druhích vypočtených rychlostí jsou vyhodnoceny všechny 4 druhy vlastností uvedených na začátku této sekce (medián, směrodatná odchylka, směrnice lineární regresní přímky a NRMSD mezi skutečnými hodnotami a regresním modelem).

Parametry *CloseVel* (*CloseVelMed*, *CloseVelStd*, *CloseVelSlope*, *CloseVelNRMSD*)

Z ang. *closing velocity*, česky (průměrná) rychlost zavírání prstů. Tento parametr je principiálně identický k předchozímu s tím rozdílem, že se během výpočtu použije upravená indexace tak, aby se braly vzdálenosti a časy od maxim směrem k minimům (od chvíle největšího rozevření prstů k jejich stisku). Hodnoty jsou opět převedeny na metry za sekundu a časy, příslušející hodnotám rychlosti, jsou brány jako středobody na časové ose mezi maximem a minimem.

Parametry *OpenPeakVel* (*OpenPeakVelMed*, *OpenPeakVelStd*, *OpenPeakVelSlope*, *OpenPeakVelNRMSD*)

Podle ang. *peak opening velocity*, česky vrcholová (maximální) rychlost otevírání prstů. Na rozdíl od předchozích dvou typů rychlosti probíhá výpočet

maximální rychlosti otevírání (nebo dále zavírání) odlišně. Použijí se při něm hodnoty aproximace derivace (diference) vypočtené při nahrávání dat jak je popsáno v sekci 2.3.1. Hodnoty derivace/diference vzdálenosti prstů podle času mají fyzikální rozměr okamžité rychlosti. Pro nalezení maximální rychlosti otevírání se vždy vezme rámec mezi minimem a maximem, které ho následuje a v tomto úseku se najde maximální hodnota z již dříve vypočítaných diferencí.

$$v[n] = \left| \max_{j = \min_{index} \quad k = \max_{index}} \{d_j, d_{j+1}, \dots, d_{k-1}, d_k\} \right| \quad (3.8)$$

V rovnici 3.8 je $v[n]$ n-tá hodnota maximální rychlosti otevírání pro daný úsek, d_j je j-tá hodnota derivace/diference vzdálenosti prstů podle času, j značí index minima, které značí počátek úseku a k index maxima, který ukazuje konec úseku. Indexy se posouvají tak, aby byl předmětem hledání maxima vždy úsek od minima k maximu. Absolutní hodnota je přítomna proto, aby všechny rychlosti měly pozitivní hodnotu neohledně na jejich směr. Samozřejmě to je aplikovatelné (pouze) v případě rychlosti zavírání. Jako časová značka se k příslušným maximálním rychlostem přiřazuje vždy přesný čas ze záznamu, kdy maximum rychlosti nastává.

U tohoto druhu vypočtené rychlosti byly prověřeny všechny 4 druhy vlastností uvedené na začátku této sekce. Stejně tak tomu bylo i u následující analogické rychlosti, tedy maximální rychlosti zavírání prstů.

Parametry *ClosePeakVel* (*ClosePeakVelMed*, *ClosePeakVelStd*, *ClosePeakVelSlope*, *ClosePeakVelNRMSD*)

Z ang. *peak closing velocity*, česky vrcholová (maximální) rychlost zavírání prstů. Tento parametr je naprosto analogický k výše uvedenému (maximální rychlosti otevírání prstů). Hodnoty jsou zpracovány absolutní hodnotou. Na rozdíl od rychlosti otevírání prstů se zde jako výpočetní rámec pro hledání minima (minima z toho důvodu, protože směr pohybu je opačný, nejvyšší rychlost bude mít také opačné znaménko) bere úsek od maxima k minimu, které ho následuje. Jinak je výpočet totožný s výše uvedeným.

■ 3.1.3 Výběr ruky pro výpočet parametrů

Kromě návrhu parametrů bylo nutné navrhnout způsob, jak přistoupit ke kombinacím měření z levé a pravé ruky. Ke každému druhu měření jsou zaznamenána data FT z obou horních končetin, a to v několika vzorcích (z různých pokusů v sezení a jednotlivých sezení při klinickém vyšetření).

Můj zvolený přístup vzal všechny měření jednoho typu (například všechny měření normálního FT testu na levé ruce, *L1-1*, *L1-2*, ..., *L2-1*, *L2-2*, ...) a provedl na nich standardní aritmetický průměr (pomocí funkce `mean`. V rámci skriptu pro statistickou analýzu `statisticalAnalysis3` (s číslovkou 3 protože je třetí iterací verze tohoto skriptu) extrahuji tato data ze souboru všech měření pro daného pacienta z vybraného datového souboru. O kompletní cyklus přes všechna dostupná data se starají čtyři vnořené `for`-cykly:

1. Cyklus, který provede iteraci přes všechny datové soubory (pacienti s PN, pacienti s RBD a zdravé kontroly).
2. Cyklus, který v rámci vybraného datového souboru provede iteraci přes všechny subjekty (pacienty) v souboru. Tento a předchozí cyklus reflektují strukturu dat vytvořenou při jejich nahrávání a definovanou v sekci 2.3.
3. Cyklus, který u jednoho subjektu provede iteraci přes všechny parametry, které je potřeba vypočítat. Tyto parametry jsou definovány v sekci 3.1.2.
4. Nakonec cyklus, který provede iteraci přes všechny druhy výpočtu s ohledem na kombinaci končetin, tak jak bude popsáno v této sekci.

V rámci druhého cyklu se ze všech dat měření vyberou ta měření, která odpovídají zvolené končetině (levá, pravá). Z těchto měření se extrahuje hodnota zvoleného parametru (podle iterace třetího cyklu) a ze všech extrahovaných hodnot funkce `getPropertyMean` vypočte průměr. Výsledkem této operace jsou tedy dvě hodnoty:

- **Lmean** - aritmetický průměr hodnot zvoleného parametru, ze všech dat měření na **levé** končetině, pro zvolený druh FT měření u aktuálního subjektu z aktuálního datového souboru.
- **Rmean** - aritmetický průměr hodnot zvoleného parametru, ze všech dat měření na **pravé** končetině, pro zvolený druh FT měření u aktuálního subjektu z aktuálního datového souboru.

Z těchto dvou hodnot následně vytvářím 4 finální výstupy, na kterých nakonec provádím statistickou analýzu. Pro neexistenci lepšího termínu jsem tyto výstupy pojmenoval anglicky *hand calculation types*, přeloženo "typy/druhy výpočtu na základě použité ruky". Jsou to následující:

- **mean**: vypočítává se jako prostý aritmetický *průměr* mezi hodnotami pro levou a pravou končetinu ($\frac{Lmean+Rmean}{2}$).
- **diff**: *diference* mezi hodnotami pro levou a pravou ruku. Vypočítává se jako absolutní hodnota jejich rozdílu, tedy $|Lmean - Rmean|$.
- **max**: numericky *větší* hodnota z obou rukou. Vypočítává se prostým použitím funkce `max`. $\max\{Lmean, Rmean\}$
- **min**: numericky *menší* hodnota z obou rukou. Vypočítává se použitím funkce `min`. $\min\{Lmean, Rmean\}$

Cílem tohoto rozvětvení je možnost podívat se na laterální (stranovost) výkonu rukou měřeného člověka z hlediska parametrů. Pokud bude mít subjekt vysoké hodnoty diferenčního testu na parametru, pak to značí, že výkon obou jeho rukou je značně asymetrický (jedna je lepší než druhá). Nutno také dodat, že maximální a minimální hodnoty nemusí u všech parametrů znamenat vylepšení stejným směrem. Proto jsem se rozhodl kategorizovat parametry do dvou skupin, podle toho, jestli je výkon ruky člověka při provádění FT

testu lepší při větší, či menší hodnotě parametru. Nutno dodat, že tato rozhodnutí jsou subjektivní a nemusí reflektovat skutečné zlepšení. Tato kategorizace je založena na myšlence toho, že idealizovaná křivka průběhu vzdálenosti prstů při FT testu v čase by měla tvar harmonické sinusoidy s konstantní amplitudou, fází a frekvencí. V reálném světě by tedy idealizaci představoval takřikajíc robotický pohyb, kdy by měřená osoba pohybovala prsty stále stejnou rychlostí a se stejnou amplitudou po dobu měření. Také byla tato kategorizace vytvořena s ohledem na zadání testu, kdy je měřená osoba instruována aby prováděla pohyb prstů *co nejrychleji a s co největší amplitudou*.

Kategorizaci jsem (zejména z praktických důvodů kódy) charakterizoval pomocí čísel 1 a -1 . Kladným znaménkem jsem zařadil parametry, které vykazují lepší výkon (z hlediska výše uvedené idealizace) pokud mají vyšší hodnoty. Záporným znaménkem naopak ty, které nabývají lepších hodnot při menších hodnotách (i záporných).

Rozdělení je následující:

■ Pozitivní parametry:

- Parametr *DistAvg* (větší průměrná hodnota rozpětí prstů).
- Všechny parametry typu *Med* s výjimkou *MinDwellTimeMed*.
- Všechny parametry typu *Slope* s výjimkou *MinDwellTimeSlope*.

■ Negativní parametry:

- Všechny parametry typu *Std*.
- Všechny parametry typu *NRMSD*.
- Parametr *MinDwellTimeMed* (nižší medián doby setrvávání v mini-mech znamená větší rychlost pohybu).
- Parametr *MinDwellTimeSlope* (snižující se doba, po kterou subjekt setrvává v údolích může indikovat menší rigiditu a bradykinezi).
- Parametr *Jitter* (menší hodnota udává větší frekvenční stabilitu).
- Parametr *Shimmer* (menší hodnota udává větší amplitudovou stabilitu).

■ 3.2 Statistická analýza

Ve statistické analýze bude prvním cílem prokázat, že lze odlišit na základě hodnot parametrů skupinu pacientů postižených RBD a skupinu zdravých kontrol. Druhým cílem bude porovnání všech třech skupin ze statistického hlediska.

Formuji hypotézu, že na základě parametrů definovaných v sekci 3.1.2 lze rozlišit skupiny RBD a zdravých kontrol. Tuto hypotézu budu ověřovat pomocí dvoustranného, neparametrického *Mann-Whitneyho U testu*⁶¹, který nepředpokládá žádné pravděpodobnostní rozdělení vstupních vzorků, pouze

jejich nezávislost. Neparametrický test byl vybrán, protože ne-všechny vy-
počítávané veličiny mají normální rozdělení. Mann-Whitneyho U test (také
Wilcoxonův rank sum test) ověřuje nulovou hypotézu H_0 , že data ze vstup-
ních souborů jsou vzorky z rozdělení s identickými mediány. Alternativně
se dá také říci, že za předpokladu, že jsou všechny vzorky nezávislé a data
jsou ordinální (je možné libovolně porovnávat vzorky a určit větší a menší
z nich), pod nulovou hypotézou H_0 bude rozdělení obou datových souborů
identické. Alternativní hypotéza H_1 pak naopak říká, že rozdělení totožná
nejdou. Hypotézy budu ověřovat na hladině statistické významnosti $\alpha = 0.05$,
budu tedy hledat výsledné testové hodnoty $p - value < 0.05$. Poté mohu
zamítnout na příslušné hladině statistické významnosti nulovou hypotézu
 H_0 a na základě provedeného testu prohlásit, že vzorky z datových souborů
pocházejí z rozdílných rozdělení. Ve výpočetním prostředí Matlab použiji
funkci `ranksum`⁶², implementující tento test, která v argumentu bere datové
soubory na otestování a jako výstup vrací výsledné $p - value$ (p-hodnotu) a
výslednou hodnotu testovací statistiky U .

Druhá provedená statistická analýza bude za použití Kruskal-Wallisova
statistického testu, který v prostředí Matlab implementuje funkce `kruskal-`
`wallis`⁶³. Kruskal-Wallisův test (dále KW-test) je rozšířením výše uvede-
ného Mann-Whitneyho U testu na více, než dva datové soubory. KW-test
stejně jako předešlý test používá namísto skutečných datových hodnot jejich
hodnoty (hodnotami se rozumí jejich ordinální pořadí s ohledem na seřazený
kompletní soubor dat z vlastního a ostatních souborů). KW-test je nepara-
metrický a nepředpokládá žádné pravděpodobnostní rozdělení u testovaných
datových souborů vyjma nezávislosti všech pozorování. Výstupem funkce
`kruskalwallis` je $p - value$ (p-hodnota) pro nulovou hypotézu H_0 , že všechny
datové soubory vložené do argumentu funkce pocházejí ze stejného rozdělení,
přičemž alternativní hypotéza H_1 pak říká opak (tedy že ne-všechny vstupní
datové soubory pocházejí ze stejného rozdělení).⁶⁴ Funkce také navrácí hod-
noty testovacích statistik, které jsou užitečné pro následné zobrazení v grafu
funkcí `multcompare`⁶⁵. Výsledky tohoto testu jsou důležité zejména z hlediska
zobrazení, protože předešlý statistický test pomocí Mann-Whitneyho U testu
nám již dává námi hledanou informaci o tom, zda-li jde rozlišit skupiny
pacientů trpících RBD a zdravých kontrol na základě našich parametrů.

Porovnání a vizualizace výsledků KW-testu bude, jak je již zmíněno výše,
provedeno pomocí testu několikanásobných porovnání (ang. *multiple compari-*
son test), implementovaného ve funkci `multcompare`⁶⁵. Ověření proběhne na
hladině statistické významnosti $\alpha = 0.05$ pomocí Fisherovy metody nejméně
významných rozdílů (ang. *Fisher's least significant differences method*, zkr.
lsd metoda)⁶⁶. Výsledné grafy ilustrující statistické rozdíly mezi testovanými
skupinami z pohledu různých parametrů a druhů výpočtů z rukou prezentují
v sekci 4.1.

3.3 Analýza laterality

Součástí této práce je i obecná analýza vlivu laterální preference u horních končetin na výsledky FT testu. Pokusí se odpovědět na otázku zda-li u testových skupin odpovídají lepší hodnoty zkoumaných parametrů vypočtených z dat FT testu preferenci ruky z hlediska výsledků Edinburského testu rukosti, který byl prezentován v sekci 2.2.1. Výsledné hodnoty tohoto testu jsou dostupné pro každý subjekt, který byl zařazen do skupin pro testování - na grafu 2.2 je vidět 94% převaha praváků nad leváky. Dá se tedy s minimální chybou generalizovat a považovat všechny osoby za praváky. Takové zjednodušení bude výhodné zejména s ohledem na prováděné výpočty.

V první analýze (obsažené ve skriptu `handednessAnalysis`) provedu výpočet korelace mezi hodnotami vypočtených parametrů (viz sekce 3.1) a hodnotami Edinburského dotazníku rukosti. Výpočet proběhne následovně:

- Pro každý subjekt se načtou výsledky Edinburského testu rukosti. Jeho hodnoty se pohybují v rozmezí od -100 (silná levoruká preference) do 100 (silná pravoruká preference). Jinými slovy, čím větší hodnota, tím větší preference pravé ruky a naopak.
- Na základě typu parametru (viz sekce 3.1.3, +1 indikuje parametr jehož subjektivní výkon stoupá s vyšší pozitivní numerickou hodnotou, -1 indikuje parametr jehož subjektivní výkon stoupá s nižší až zápornou numerickou hodnotou) zvolím typ přístupu:
 - V případě pozitivního parametru (typ +1) vypočítám rozdíl mezi hodnotami parametru pravé a levé ruky, $d_i = h_{R,i} - h_{L,i}$, kde d_i je výsledný rozdíl pro i-tý parametr, $h_{R,i}$ je průměrná hodnota i-tého parametru ze všech měření standardního FT testu na pravé ruce, $h_{L,i}$ je průměrná hodnota i-tého parametru ze všech měření standardního FT testu na levé ruce. Pokud bude mít pravá ruka lepší výsledky, bude hodnota d_i pozitivní a bude se zvyšovat společně s dominancí pravé ruky z hlediska výkonu na parametru (za uvedeného zjednodušení výše, kdy považujeme všechny subjekty za praváky).
 - Pro negativní typ parametru (-1) se provede obrácený výpočet k předchozímu. Hodnota rozdílu se vypočítá jako $d_i = h_{L,i} - h_{R,i}$, kde mají uvedené proměnné stejný význam jako u parametru pozitivního typu. V tomto případě bude výsledek rozdílu tím větší, čím bude hodnota z pravé ruky menší oproti hodnotě z levé ruky. S narůstající výkonnostní dominancí pravé ruky se tak bude celková hodnota rozdílu d_i zvyšovat.
- Provede se výpočet korelace mezi hodnotami d_i pro i-tý parametr a hodnotami Edinburského dotazníku rukosti. Pro výpočet korelace byl zvolen Spearmanův koeficient pořadové korelace, ang. *Spearman's rank correlation coefficient*, také *Spearman's rho*⁶⁷. Spearmanův koeficient

navrací bezrozměrné číslo, které se dá reprezentovat jako míra toho, jako dobře se dá vztah dvou veličin popsat pomocí monotónní funkce. Základem je vzestupné seřazení hodnot náhodných proměnných x_i a y_i (zvláště pro oba soubory hodnot) přidělení pořadových hodnot $1, 2, 3, \dots$ od nejmenší po největší každé z nich. Výpočet koeficientu je uveden ve vzorci 3.9.

$$\rho = 1 - \frac{6 \sum_i (p_i - q_i)^2}{n(n^2 - 1)} \quad (3.9)$$

Ve vzorci 3.9 n označuje počet hodnot náhodných veličin, jejichž korelace je předmětem výpočtu, p_i je pořadové číslo (ang. *rank*) i -té hodnoty první veličiny, q_i je pořadové číslo i -té hodnoty druhé veličiny a ρ je výsledná hodnota korelačního koeficientu, která se pohybuje mezi hodnotami $+1$ a -1 pro dokonalé popsání monotónní funkcí vzrůstajícího respektive klesajícího typu.^b

Výpočet se provede odděleně pro všechny tři datové skupiny.

U vybraných parametrů na základě výsledků statistické analýzy poté bude provedeno vykreslení hodnot parametru na levé a pravé ruce v komplexním grafu, včetně dodaných procentuálních informací a rozložení výkonu na obou končetinách (jinými slovy, v kolika procentech případů/u kolika procent subjektů měla levá, respektive pravá ruka lepší výkon v hodnotách parametru). Součástí těchto grafů bude zobrazení mediánu hodnot zvláště pravé a zvláště levé ruky. Subjekty také budou seřazeny od nejkratší po nejdelší uvedenou klinickou dobu trvání nemoci (PN nebo RBD) a data proložena lineární regresní přímkou vypočtené z hodnot parametru pro levou a pravou rukou samostatně. Cílem je dodatečná vizuální reprezentace možné progresse nemoci v čase (simulace longitudinálních dat), s přihlédnutím na nepřesnost informace o délce trvání nemoci z důvodu obtížnosti diagnostiky obou nemocí. Tyto grafy lze najít v příloze B.

Nakonec bude opět u vybraných parametrů na základě výsledků statistické analýzy proveden výpočet procentuálního rozdělení lepších výsledků (opět s přihlédnutím ke kategorizaci ze sekce 3.1.3) na levé a pravé ruce u skupiny pacientů trpících RBD a u skupiny zdravých kontrol. Cílem také bude omezit vliv možné nepřesnosti, která vzniká pokud jsou si hodnoty parametru na levé a pravé ruce velmi podobné. Pak se nedá s jistotou určit, která ruka má z pohledu parametru lepší výkon, a subjekt by měla být z analýzy vyřazen nebo by měl být omezen jeho vliv na výsledek korelačního výpočtu. Nesnadné určení hranice pro vyloučení vedlo k pokusnému zavedení hraniční hodnoty v podobě aritmetického průměru z absolutních hodnot rozdílů výsledků na levé a pravé ruce (tedy na numerické rozdílnosti obou hodnot). Pokud byla vzdálenost hodnot mezi oběma rukama menší než průměr ze skupiny, byl subjekt vyřazen. Výsledky budou uvedeny jako pro uvedeným způsobem zredukovaná data tak pro kompletní datové soubory v sekci 4.2.

^bTento vzorec lze použít v případě, že všechny hodnoty v souboru dat každé náhodné veličiny mají přiděleny unikátní, neopakující se hodnoty pořadových čísel.

3.4 Korelační analýza

Provedl jsem několik druhů korelačních testů. První z nich je korelační test mezi hodnotami parametrů, u kterých by se dala předpokládat jejich závislost. U všech parametrů jsem vzal jejich vypočtené vlastnosti, tedy medián hodnot, směrodatnou odchylku, směrnici lineární regresní přímky a NRMSD skutečných hodnot od lineárního modelu. Těmito parametry jsou:

- `OpenVel` - průměrná rychlost otevírání prstů
- `CloseVel` - průměrná rychlost zavírání prstů
- `OpenVelPeak` - maximální rychlost otevírání prstů
- `CloseVelPeak` - maximální rychlost zavírání prstů
- `Freq` - hodnoty frekvence period signálu
- `MinDwellTime` - čas strávený v údolích/minimech vzdálenosti prstů

Korelační test jsem provedl na všech možných dvojicích parametrů. Použitá metodika je stejná jako v sekci 3.3, tedy s využitím Spearmanova korelačního koeficientu. Zobrazené výsledky v sekci 4 a tabulce 4.4 jsou pouze pro ty dvojice parametrů, které vykazují absolutní hodnotu korelačního koeficientu $|\rho| \geq 0.75$. Výsledné p-hodnoty testované hypotézy, že mezi dvojicemi dat neexistuje korelace (oproti alternativní hypotéze, že existuje nenulová korelace mezi nimi), navracené funkcí `corr` pro všechny dvojice s absolutní hodnotou korelačního koeficientu $|\rho| \geq 0.75$ měly ve všech případech hodnotu $p < 0.001$.

Kromě výše uvedené mezi-parametrové korelační analýzy jsem také provedl výpočet korelace mezi vypočtenými hodnotami parametrů a vybranými klinickými daty. Pokusil jsem se najít korelace mezi výsledky parametrů u diferenčního typu výpočtu z rukou (kdy je vypočtena absolutní hodnota rozdílu středních hodnot všech měření na levé a pravé ruce) a absolutní hodnoty rozdílu té podseky skóre z dotazníku MDS-UPDRS III, která klinicky hodnotí výkon při FT testu na levé ruce a té, která hodnotí totéž na pravé ruce (jinými slovy rozdíl ve výkonu obou rukou z klinického pohledu).

Dále jsem provedl obdobnou analýzu jako předešlou uvedenou, ale s vynecháním absolutní hodnoty z výpočtu a s přihlédnutím k typu parametru, jak byl definován v sekci 3.1.3. Vypočítal jsem tedy rozdíl klinických hodnocení FT testu z podseky MDS-UPDRS III dotazníku, viz vzorec 3.10.

$$d_{1,i} = FT_{L,i} - FT_{R,i} \quad (3.10)$$

Kde $FT_{L,i}$ ve vzorci 3.10 označuje hodnocení FT testu i-tého pacienta na levé ruce, $FT_{R,i}$ je hodnocení FT testu i-tého pacienta na pravé ruce a $d_{1,i}$ je jejich rozdíl pro i-tého pacienta.

Dále jsem v případě pozitivního parametru vzal průměrnou hodnotu vypočteného j-tého parametru ze všech měření na pravé ruce i-tého pacienta

a odečetl od něj průměrnou hodnotu parametrů z měření stejného pacienta na jeho levé ruce. U negativního typu parametru jsem odečetl naopak pravé hodnoty od levých.

$$\begin{aligned} d_{2,i,j} &= \bar{p}_{i,j,R} - \bar{p}_{i,j,L}; \text{ Pro } j\text{-tý parametr pozitivního typu} \\ &\text{nebo} \\ d_{2,i,j} &= \bar{p}_{i,j,L} - \bar{p}_{i,j,R}; \text{ Pro } j\text{-tý parametr negativního typu} \end{aligned} \quad (3.11)$$

$\bar{p}_{i,j,R}$ je ve vzorci 3.11 průměrná hodnota j -tého parametru, vypočtená ze všech měření standardního FT testu i -tého pacienta na pravé ruce, $\bar{p}_{i,j,L}$ je průměrná hodnota j -tého parametru, vypočtená ze všech měření standardního FT testu i -tého pacienta na levé ruce a $d_{2,i,j}$ jsou příslušné rozdíly mezi nimi pro i -tého pacienta a j -tý parametr.

Korelace se pak vypočítává mezi hodnotami $d_{1,i}$ a $d_{2,i,j}$. Princip je stejný, jako při výpočtu v sekci 3.3. Opět se uvažuje zjednodušení v podobě toho, že se všichni pacienti považují za praváky. Rozdíl jejich klinických hodnocení FT testu z MDS-UPDRS III dotazníku by tak v případě, kdy se odečítá skóre pravé ruky od levé, měl být pozitivní a tím větší, čím bude skóre na levé ruce větší (což indikuje větší postižení na levé ruce) a tím menší, čím je skóre na pravé ruce menší (což indikuje její menší postižení). Jinými slovy, výpočet tohoto rozdílu předpokládá vyšší hodnoty v případě těžšího postižení nedominantní ruky (a lepšího výkonu na dominantní ruce).

Hodnota vypočtená z rozdílu klinických hodnocení se podrobuje výpočtu korelace s výsledky parametrů, které jsou pomocí jejich klasifikace na pozitivní a negativní brány tak, že jejich rozdíl nabírá vždy vyšších hodnot při lepším výkonu pravé ruky z pohledu parametru.

Nutno dodat, že hodnoty klinického hodnocení FT testu z dotazníku MDS-UPDRS III se pohybují v rozmezí celých čísel od 0 do 4 včetně. Nedá se tedy předpokládat vysoká nuance výsledných hodnot korelace. Výsledky (ovlivněné tímto faktem) lze najít v sekci 4.3.2.

3.5 ROC analýza

ROC analýzou ve svojí práci nazývám analýzu výkonnosti možné klasifikace skupin na základě vybraných parametrů s využitím ROC (zkratka z ang. *Receiver operating characteristic*) křivky. ROC křivka⁶⁸ je grafická interpretace schopností binárního klasifikátoru, která dává do souvislosti senzitivitu (ang. také *true positive rate* nebo TPR - udává poměr počtu případů klasifikovaných jako pozitivní ku počtu skutečně pozitivních případů) a specificitu (ang. také *true negative rate* nebo TNR - udává poměr počtu případů klasifikovaných do negativní skupiny ku počtu skutečně negativních případů) klasifikátoru pro různá nastavení jeho hraničních hodnot. V kontextu této práce (a obecně medicínsky zaměřených klasifikací) označuje pozitivní skupina nemocné lidi a negativní skupina zdravé kontroly. TPR tedy udává přeneseně jak dobře klasifikátor hledá nemocné osoby a TNR ukazuje jak přesně dokáže odhalit zdravé osoby z datového souboru.

Užitečným údajem z ROC křivky je potom tzv. AUC, z ang. *area under curve*, tedy plocha pod ROC křivkou. Hodnota AUC se pohybuje mezi nulou a jedničkou, s tím, že vyšší hodnoty značí lepší klasifikaci. Pokud je hodnota AUC rovna 0.5, není klasifikátor rozhodování se podle hodu mincí.

Analýzu jsem rozdělil na dvě skupiny podle použitých parametrů. V první skupině jsem vybral 4 parametry s nejlepší schopností oddělit kontroly od pacientů s RBD, tedy s nejlepšími výsledky Mann-Whitneyho U testu (nejmenšími p-hodnotami, $p_{val} \leq 0.05$), které zároveň dávají logický smysl z hlediska posloupnosti dat (progrese hodnot parametru relativně vůči skupinám od kontrol, přes RBD až k pacientům s PN, viz výsledky v sekci 4.1) a jsou vypočteny z dat standardního FT testu lepší či horší ruky (skupiny *min* nebo *max*). Jsou to parametry:

- **MaxNRMSD - *min***: Hodnoty NRMSD mezi skutečnými hodnotami maxim a modelovými hodnotami lineární regresní přímky. Jsou vzaty menší hodnoty z obou rukou, tedy menší odchylka a větší pravidelnost.
- **FreqSlopeMA - *max***: Hodnoty směrnice lineární regresní přímky proložené skrz hodnoty frekvence vyhlazené pomocí MA filtru. Jsou brány vyšší hodnoty z obou rukou, tedy více stoupající/méně klesající trend frekvence klepání prstů.
- **CloseVelNRMSD - *min***: Hodnoty NRMSD mezi skutečnými hodnotami průměrné rychlosti zavírání prstů a modelovými hodnotami na lineární regresní přímce. Jsou brány menší hodnoty z obou rukou, tedy menší odchylka a větší pravidelnost.
- **CloseVelSlope - *min***: Hodnota směrnice lineární regresní přímky proložené skrz hodnoty průměrné rychlosti zavírání prstů. Jsou brány nižší hodnoty z obou rukou, tedy více stoupající/méně klesající trend prům. rychlosti zavírání prstů.

Druhá skupina je tvořená třemi parametry, které byly vybrány identickým způsobem jako parametry předchozí, ale tentokrát ze skupin typu *diff*, tedy vzdálenosti hodnot parametru mezi průměry levé a pravé ruky. Vybrané parametry druhé skupiny jsou:

- **MaxNRMSD - *diff***: Identicky k parametru z první skupiny, ale je vzata absolutní hodnota rozdílu průměru hodnot ze všech měření na levé ruce a průměru hodnot všech měření pravé ruce.
- **MinDwellTimeNRMSD - *diff***: Hodnoty NRMSD mezi skutečnými délkami setrvávání prstů v dotyku (místech s nejmenší amplitudou, údolích) a hodnotami na lineární regresní přímce proložené skrz skutečné hodnoty. Opět je vzata absolutní hodnota rozdílu průměru hodnot ze všech měření na levé ruce a průměru hodnot všech měření pravé ruce.
- **CloseVelNRMSD - *diff***: Identicky k parametru z první skupiny ale opět je použit rozdíl hodnot obou rukou.

Pro klasifikaci používám logistický regresní model vytvořený pomocí funkce `glmfit`⁶⁹ (která dokáže vytvořit různé druhy zobecněných lineárních modelů, ang. *generalized linear model*⁷⁰, dále GLM - logistický regresní model je druh zobecněného lineárního modelu), která dostává v parametru specifikované použití binomiálního modelu a logistické linkové funkce. GLM je takové zobecnění lineárního regresního modelu, které nepředpokládá normální rozdělení výstupní proměnné (ang. *response variable*) a umožňuje, aby se spolu se změnami vstupní proměnné místo výstupní proměnné samotné, měnila lineárně její předem definovaná *funkce*. Těto funkci se říká *linková* (propojovací) funkce, a v případě logistického regresního modelu je to logistická (resp. *logit*) funkce.

Na vstupu funkce `glmfit` jsou vyjma nastavení logistického regresního modelu skutečné kategorické popisky dat (ang. *labels*), které zaujímají hodnoty 1 pro zdravé kontroly a 0 pro pacienty s RBD, a soubory vypočtených parametrů (jejichž trojice jsou vyčteny výše). Funkce vypočítá takové nastavení (koeficienty) modelu, které odpovídá nadrovině nejlépe separující dvě vstupní datové třídy. Následně se pomocí funkce `glmval`⁷¹, která bere v argumentu vypočtený logistický model, navrácí hodnoty vypočtených klasifikačních skóre (v případě našeho modelu hodnoty mezi 0 a 1) pro jednotlivé datové vzorky.

Výsledné klasifikační skóre společně s reálnými popisky dat pak vstupují do funkce `perfcurve`⁷², která spočte hodnoty ROC křivky společně s hodnotou AUC. Dodatečné vyhodnocení se pak provede pomocí *leave-one-out* krosvalidace, ve které se nastaví model pomocí $N - 1$ vzorků (kde N je celkový počet vzorků) a n -tý, chybějící vzorek je tímto modelem klasifikován. Provede se N iterací přes všechny vzorky, vždy s n -tým vzorkem jako testovacím a zbytkem vzorků jako trénovacími daty. Výsledky cross-validace se promítnou do hodnot matice záměn (ang. *confusion matrix*), ze které se vypočtou hodnoty senzitivity a specificity.

$$\text{TPR} = \text{SENZ} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3.12)$$

$$\text{TNR} = \text{SPEC} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (3.13)$$

$$(3.14)$$

V rovnici 3.12 TPR značí senzitivitu, TP je počet korektně nalezených pozitivních vzorků, FN je počet negativních (nemocných) klasifikovaných do pozitivní (zdravé) skupiny, TNR je specificita klasifikátoru, TN je počet správně klasifikovaných negativních vzorků a FP je počet falešně pozitivních vzorků, tedy takových, kteří jsou negativní (zdraví) ale byli zařazeni jako pozitivní (nemocní).

Kapitola 4

Výsledky

Obsahem této kapitoly jsou výsledky analýz, které jsem navrhl a popsal v kapitole 3. Výsledky jsou ilustrovány grafy a tabulkami a dále okomentovány (zejména v případě ROC analýzy) v kapitole 5.

4.1 Statistická analýza

V této sekci budu prezentovat výsledky statistické analýzy, která byla popsána v kapitole 3, sekce 3.2. Popíšu výsledky analýzy zvláště pro parametry, které byly typu *min* nebo *max*, tedy jejichž hodnoty byly u všech subjektů brány z ruky s vyšší nebo nižší hodnotou (u daného subjektu), a pro parametry typu *diff*, jejichž hodnota se vypočítala z absolutní hodnoty rozdílu výsledků na levé a na pravé ruce (tedy z numerické vzdálenosti výsledků mezi rukami). Více o popisu výpočtu v sekci 3.1.3.

4.1.1 Parametry typu *min*, *max*

V tabulce 4.1 jsou vypsány výsledky Mann-Whitneyho U-testu (Wilcoxonova *rank sum* testu), jehož cílem bylo ověření hypotézy na hladině statistické významnosti $\alpha = 0.05$, zda-li pochází datové vzorky od skupiny zdravých kontrol a datové vzorky od skupiny osob trpících RBD ze stejného rozdělení. V tabulce jsou uvedeny pouze výsledky pro ty parametry, pro které byla navrácena *p-val* ≤ 0.05 - na základě výsledků testu jsou tedy tyto skupiny, z pohledu vypsáných parametrů, statisticky rozlišitelné. V tabulce jsou dále uvedeny mediány hodnot parametrů jednotlivých skupin a hodnoty testové statistiky Wilcoxonova *rank sum* testu, W .

Přepočítání hodnoty Wilcoxonovy testové statistiky W na Mann-Whitneyho U-testovou statistiku U probíhá pomocí vzorce 4.1.

$$U = W - \frac{n_x(n_x + 1)}{2} \quad (4.1)$$

Ve vzorci 4.1 označuje W hodnoty Wilcoxonovy testové statistiky, n_x je počet vzorků v souboru náhodné proměnné X a U je výsledná hodnota Mann-Whitneyho U-testové statistiky. Ve výpočtu pomocí funkce `ranksum` v

Název parametru	HCT	m_{CON}	m_{RBD}	m_{PD}	W	p -val
CloseVelNRMSD	min	1.343e-01	1.436e-01	1.593e-01	2076	0.0486
CloseVelSlope	min	-3.292e-05	-5.761e-05	-8.735e-05	1722	0.0461
FreqSlope	max	3.014e-04	-6.571e-03	-1.18e-02	1697	0.0229
FreqSlopeMA	max	1.913e-03	-6.779e-03	-1.06e-02	1681	0.0141
MaxNRMSD	min	1.639e-01	1.78e-01	1.802e-01	2119	0.0141
OpenVelNRMSD	max	1.99e-01	1.88e-01	1.91e-01	1647	0.00454
OpenVelPeakNRMSD	max	1.798e-01	1.712e-01	1.837e-01	1710	0.0332
OpenVelPeakSlope	max	-8.058e-01	-1.274	-8.119e-01	1694	0.0209

Tabulka 4.1: Výsledky statistické analýzy Mann-Whitneyho U-testu mezi výsledky parametrů skupiny kontrol a skupiny pacientů trpících RBD - výsledky pro typ výpočtu z rukou, kdy byla brána větší (max) nebo menší (min) hodnota vypočteného parametru z obou rukou. Sloupec HCT označuje vybraný druh dat. m_{CON} je medián hodnot parametru skupiny zdravých kontrol, m_{RBD} je medián z hodnot skupiny pacientů trpících RBD a m_{PD} je medián hodnot skupiny pacientů trpících PN. W je hodnota testové statistiky Wilcoxonova testu skupiny RBD (přepočten na Mann-Whitneyho U-testovou statistiku viz výše v této sekci). p -val je p -hodnota oboustranného Wilcoxonova *rank sum* testu nulové hypotézy na hladině významnosti $\alpha = 0.05$, že vypočtená data parametrů skupiny RBD a skupiny zdravých kontrol pocházejí ze stejného rozdělení. Počet vzorků (subjektů) ve skupině RBD byl roven $N_{RBD} = 50$, skupina zdravých kontrol pak obsahovala $N_{CON} = 25$ subjektů.

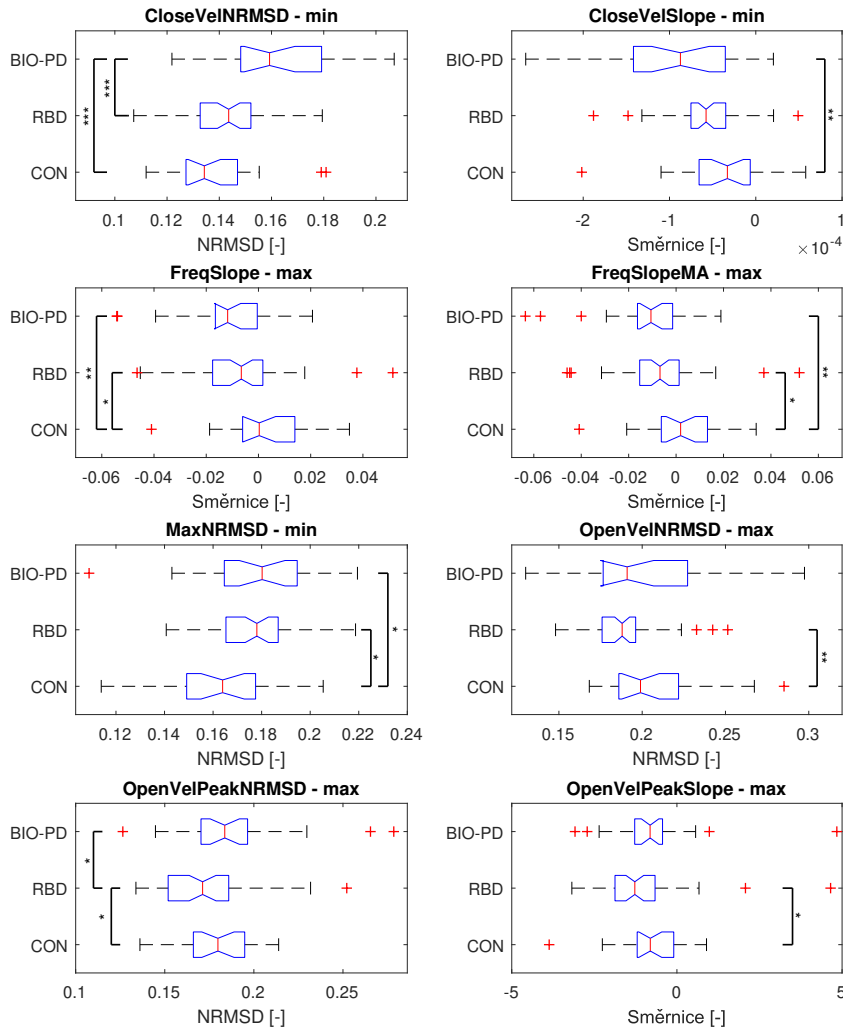
prostředí Matlab funkce vždy vrací hodnotu W pro první vstupní skupinu, v mém případě tedy pro data od skupiny RBD.

Podstatným údajem je také použitá hodnota výpočtu z rukou, v sloupci nadepsaném HCT (zkratka pro *hand calculation type*). Typ *min* znamená, že z hodnot vypočteného parametru na obou rukou (výpočet parametru proběhl na všech měření levé ruky, z jejich hodnot se spočetl aritmetický průměr - stejný postup se opakoval pro měření na pravé ruce) se vybrala *ta menší*. Mezi s sebou se tedy ve statistickém testu porovnávaly horší výsledky vždy v rámci dané osoby, nehledě na jeho dominantní ruce. Obdobně postup probíhal i u typu *max*, kdy se z obou hodnot z rukou vybrala vždy ta větší.

Na základě kategorizace popsané v sekci 3.1.3 je pak vidět, že u parametrů (z tabulky 4.1) *CloseVelNRMSD*, *FreqSlope*, *FreqSlopeMA*, *MaxNRMSD* a *OpenVelPeakSlope* se vybrala lepší, ze dvou hodnot na rukou. U parametrů *CloseVelSlope*, *OpenVelNRMSD* a *OpenVelPeakNRMSD* se vzala z hodnot na obou rukou u každé osoby ta horší.

Na obrázku 4.1 jsou vidět rozložení dat vypočtených parametrů včetně spojnic indikující nalezené statisticky významné rozdíly na základě dat z KW-testu tak, jak byl popsán v sekci 3.2. V grafu jsou pak hvězdičkami u spojnic uvedeny úrovně p -hodnot pro jednotlivá porovnání mezi skupinami, tak jak byly vypočteny pomocí testu několikanásobných porovnávaní (*multiple comparisons*) s Fisherovy procedury nejméně významných rozdílů. P -hodnoty podle počtu hvězdiček: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Ideální rozložení tří skupin v tom smyslu, že skupina pacientů trpících PN (skupina *BIO-PD*) bude zaujímat nejhorší hodnoty parametru (viz 3.1.3), skupina zdravých kontrol bude mít z hlediska parametru nejlepší výkon a



Obrázek 4.1: Grafy typu box-plot ilustrující distribuci hodnot mezi skupinami pro vybrané parametry typu *min/max*. Na obrázku jsou vidět skutečné rozdělení hodnot parametrů, které dokázaly statisticky oddělit skupiny RBD a CON na základě Wilcoxonova *rank sum* testu. Název parametru se nachází v hlavičce každého grafu, jednotlivé skupiny jsou popsány na svislé ose a uvádí příslušné řádky hodnot. Hranice modrých boxů přísluší hodnotám prvního a třetí kvartilu, červená středová linie označuje medián hodnot skupiny. Přerušované čáry po obou stranách boxu sahají do přibližné vzdálenosti $\pm 2.7\sigma$ a vzdálenější hodnoty jsou značeny červenými kříži.⁷³ Vedle řádků jsou spojnici vyznačeny výsledky post-hoc testu (na výstupech KW-testu) vzájemných porovnání s použitím Fisherova procesu nejméně významných rozdílů (zkr. lsd, *Fisher's Least Significant Difference*). Hvězdičky označují p-hodnoty porovnání mezi vyznačenými skupinami se statisticky významnými rozdíly. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

skupina RBD zaujme pozici mezi nimi, by odpovídalo představě o progresi nemoci od zdravé osoby, přes diagnózu RBD (jako možný prodromální příznak PN) až k diagnóze PN. Takové rozdělení dat zaujímají hodnoty parametrů:

- **CloseVelNRMSD** - menší hodnoty z obou rukou (typ *min*), z pohledu tohoto parametru tedy lepší hodnota mezi rukama. Hodnoty skupiny PN mají nejvyšší odchylky, zdravé kontroly nejmenší.
- **CloseVelSlope** - menší hodnoty z obou rukou (typ *min*), z pohledu tohoto parametru tedy horší hodnota mezi rukama. Hodnoty skupiny pacientů s PN zaujímají nejvíce záporné hodnoty, což odpovídá poklesu průměrné rychlosti zavírání v čase. Zdravé kontroly mají pak směrnici regresní přímkou blíže nulové hranici, což indikuje minimální změnu v čase.
- **FreqSlope** a **FreqSlopeMA** - větší hodnota z obou rukou, z pohledu tohoto parametru tedy lepší hodnota mezi rukama. U obou druhů směrnic zaujímá skupina osob s PN nejvíce záporné hodnoty, což odpovídá poklesu frekvence klepání prsty v čase. Medián hodnot zdravých kontrol je téměř přesně na nulové hodnotě, což odpovídá konstantní frekvenci.
- **MaxNRMSD** - menší hodnota z obou rukou, z pohledu tohoto parametru tedy lepší hodnota mezi rukama. Zde je zajímavé, že hodnoty skupiny pacientů s PN a pacientů trpících RBD jsou si velmi blízko. Skupina zdravých kontrol má pak nejmenší hodnoty odchylek od regresní přímkou proložené skrze maxima.

Ostatní tři parametry týkající se rychlosti otevírání prstů nerespektují výše popsané rozložení a proto je dále nebudu uvažovat.

■ 4.1.2 Parametry typu *diff*

Výsledky statistické analýzy z hlediska parametrů typu *diff* je možné vidět v tabulce 4.2. Šest parametrů prokázalo statistickou schopnost rozlišit skupinu osob trpících RBD a skupinu zdravých kontrol pomocí Wilcoxonova *rank sum* testu na hladině statistické významnosti $\alpha = 0.05$. V tabulce jsou uvedeny p-hodnoty příslušící k jednotlivým parametrům včetně hodnot mediánů a hodnot testovací statistiky *W*.

Hodnoty parametrů typu *diff*, jak bylo popsáno v sekci 3.1.3, se vypočítávají jako numerická vzdálenost průměrovaných hodnot vypočteného parametru ze všech měření na levé a na pravé ruce. Pokud je označíme průměr vypočtených parametrů ze všech měření jedné osoby na levé ruce písmenem *L*, a identicky průměr vypočtených parametrů ze všech měření jedné osoby na pravé ruce písmenem *P*, hodnota použitá v parametru typu *diff* se vypočítá jako $|L - P|$. Takový to údaj značí to, jak moc jsou si (z hlediska zvoleného parametru) obě ruce vzdáleny. Pokud jedna ruka bude mít výborný výsledek a druhá mizerný, výsledkem bude vysoká hodnota parametru tohoto typu. Pokud obě ruce budou mít stejně dobré výsledky, parametr typu *diff* bude mít malou hodnotu.

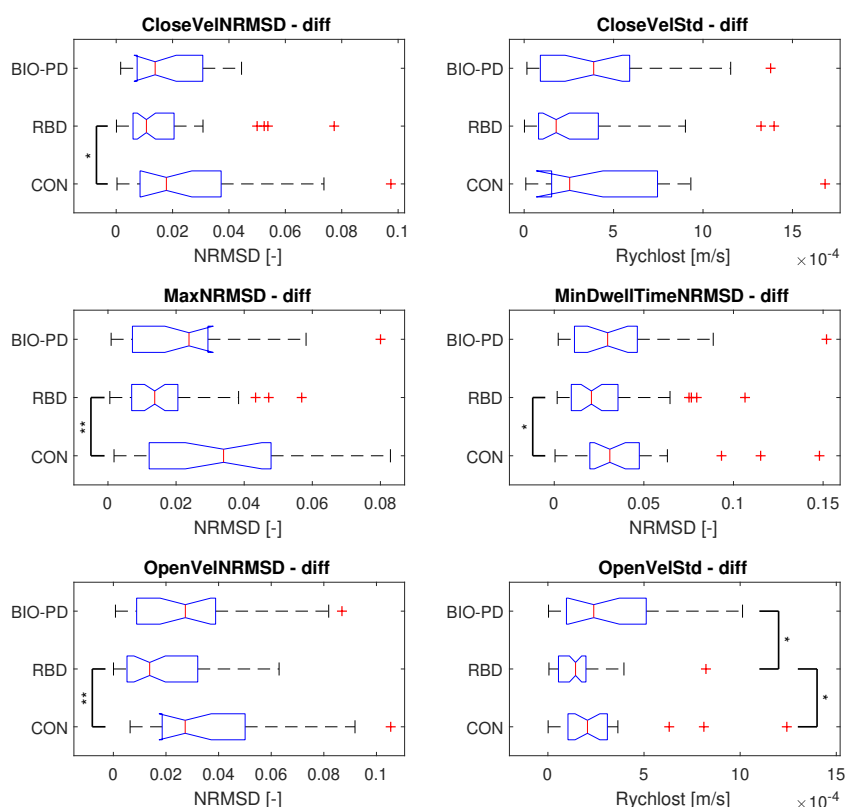
Z tabulky vidíme, že extrémně dobrou rozlišovací schopnost prokázal parametr *OpenVelNRMSD*, tedy rozdíl výkonu obou rukou z hlediska normalizované odchylky skutečné průměrné rychlosti otevírání prstů od lineární regresní

Název parametru	HCT	m_{CON}	m_{RBD}	m_{PD}	W	p -val
CloseVelNRMSD	diff	1.78e-02	1.072e-02	1.382e-02	1695	0.0215
CloseVelStd	diff	2.546e-04	1.791e-04	3.888e-04	1722	0.0461
MaxNRMSD	diff	3.395e-02	1.373e-02	2.38e-02	1636	0.00306
MinDwellTimeNRMSD	diff	3.102e-02	2.074e-02	2.979e-02	1717	0.0403
OpenVelNRMSD	diff	2.73e-02	1.38e-02	2.736e-02	1603	0.000861
OpenVelStd	diff	2.059e-04	1.439e-04	2.386e-04	1705	0.0288

Tabulka 4.2: Výsledky statistické analýzy Mann-Whitneyho U-testu mezi výsledky parametrů skupiny kontrol a skupiny pacientů trpících RBD - výsledky pro typ výpočtu z rukou, kdy byla brána absolutní hodnota rozdílu vypočtených parametrů z obou rukou. Sloupec HCT označuje vybraný typ dat - diff. m_{CON} je medián hodnot parametru skupiny zdravých kontrol, m_{RBD} je medián z hodnot skupiny pacientů trpících RBD a m_{PD} je medián hodnot skupiny pacientů trpících PN. W je hodnota testové statistiky Wilcoxonova testu skupiny RBD (přepočten na Mann-Whitneyho U-testovou statistiku viz výše v této sekci). p -val je p -hodnota oboustranného Wilcoxonova *rank sum* testu nulové hypotézy na hladině významnosti $\alpha = 0.05$, že vypočtená data parametrů skupiny RBD a skupiny zdravých kontrol pocházejí ze stejného rozdělení.

přímky skrz tyto hodnoty rychlosti proložené. Dobrý rozlišovací výkon ještě ukazuje parametr *MaxNRMSD*, uvedený již v předchozí sekci.

Na obrázku 4.2 pak můžeme vidět skutečné rozdělení hodnot parametrů typu *diff* včetně spojnic indikující nalezené statisticky významné rozdíly na základě dat z KW-testu tak, jak byl popsán v sekci 3.2. V grafu jsou pak hvězdičkami u spojnic uvedeny úrovně p -hodnot pro jednotlivá porovnání mezi skupinami, tak jak byly vypočteny pomocí testu několikanásobných porovnání (*multiple comparisons*) s Fisherovy procedury nejméně významných rozdílů. P -hodnoty podle počtu hvězdiček: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Díky postupu výpočtu je logické, že všechny parametry "začínají" (téměř) v nule, neboť to značí takové osoby, jejichž hodnoty parametru na obou rukou si byly velmi blízké. Zajímavé také je, že zástupci osob s vyrovnaným výkonem na obou rukou se nachází v téměř každé skupině. Dalším poznatkem z grafů na obrázku 4.2 je velké množství extrémních hodnot (*outlierů*) u téměř všech skupin a parametrů - tyto hodnoty ukazují osoby s velmi rozdílným výkonem obou rukou. U všech parametrů pak můžeme pozorovat, že skupina RBD má konzistentně nejmenší rozdíly mezi hodnotami na obou rukou. U parametrů typu *NRMSD* má skupina osob trpících PN spíše menší rozdíly mezi rukama, zejména u parametru *MaxNRMSD*, kde je zároveň i velmi velký rozptyl hodnot u skupiny zdravých kontrol. Grafy hezky ilustrují, jak parametry typu *diff*, nebo-li jednoduše řečeno rozdíly mezi rukami, dobře rozlišují skupinu osob trpících RBD a skupinu zdravých kontrol, naopak skupina pacientů trpících PN je od zdravých kontrol z hlediska těchto parametrů od zdravých kontrol obtížně rozlišitelná.



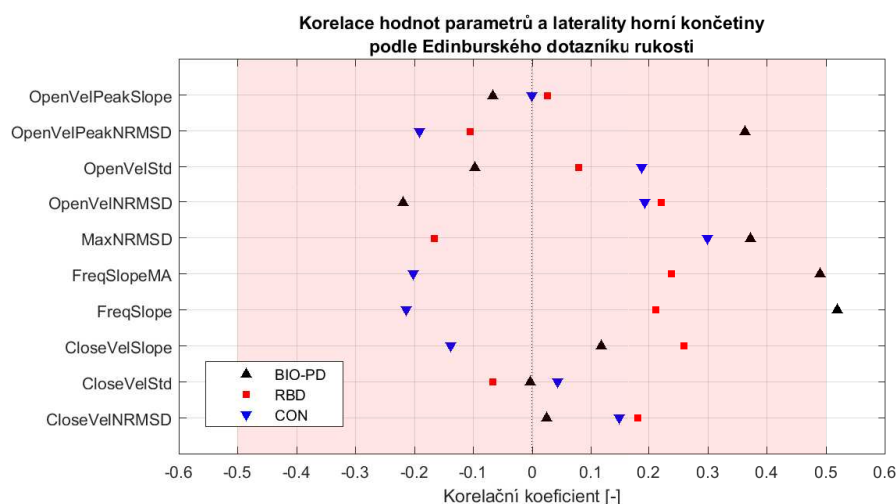
Obrázek 4.2: Grafy typu box-plot ilustrující distribuci hodnot mezi skupinami pro vybrané parametry typu *diff*. Na obrázku jsou vidět skutečné rozdělení hodnot parametrů, které dokázaly statisticky oddělit skupiny RBD a CON na základě Wilcoxonova *rank sum* testu. Název parametru se nachází v hlavičce každého grafu, jednotlivé skupiny jsou popsány na svislé ose a uvádí příslušné řádky hodnot. Hranice modrých boxů přísluší hodnotám prvního a třetího kvartilu, červená středová linka označuje medián hodnot skupiny. Přerušované čáry po obou stranách boxu sahají do přibližné vzdálenosti $\pm 2.7\sigma$ a vzdálenější hodnoty jsou značeny červenými kříži.⁷³ Vedle řádků jsou spojnici vyznačeny výsledky post-hoc testu (na výstupech KW-testu) vzájemných porovnání s použitím Fisherova procesu nejméně významných rozdílů (zkr. *lsc*, *Fisher's Least Significant Difference*). Hvězdičky označují p-hodnoty porovnání mezi vyznačenými skupinami se statisticky významnými rozdíly. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

4.2 Analýza laterality

V této sekci představím výsledky analýz vlivu laterality na hodnoty parametrů pomocí tabulek a grafů. Vliv laterality je zajímavý z hlediska výběru ruky pro výpočet parametru. Pro účely klasifikace by bylo samozřejmě neefektivnější použít hodnoty z ruky, která je nejvíce postižena (možnou) postupující neurodegenerativní chorobou. Zjištění postižené strany těla, či potažmo konkrétně horní končetiny není, zejména v prodromálních stádiích PN snadné, tím spíše u pacientů trpících RBD, kde jsou motorické projevy možné budoucí konverze

do PN velmi subtilní. Jednou z možných myšlenek by byla teorie, že postižení by se mohlo více projevovat na dominantní nebo nedominantní ze dvou horních končetin. I když moje práce se na toto téma specificky nezaměřuje, mohl jsem zkusit tuto myšlenku ověřit díky již vypočteným datům. Zkombinoval jsem kategorizaci provedenou v sekci 3.1.3, vypočtené hodnoty parametrů na obou horních končetinách pacientů a data o preferenci horní končetiny z Edinburského dotazníku rukosti.

Ve skupině pacientů postižených PN jsou na základě dat z Edinburského dotazníku rukosti všichni praváci (slabí i silní, viz 3.3). Ve skupině pacientů s RBD je praváků 94 procent, ve skupině zdravých kontrol je praváků nejméně, a to 88 procent. Osoby bez preference (se skóre rovným nule z Edinburského dotazníku) byly přiřazeny do skupiny praváků, jedná se ale pouze o jednotky lidí.



Obrázek 4.3: Výsledky korelační analýzy mezi skóre z Edinburského dotazníku rukosti a výsledky vybraných parametrů. V tomto grafu jsem vybral (z praktických důvodů) jen výsledky analýzy pro ty parametry, které dokázaly statisticky rozlišit skupinu RBD od zdravých kontrol. Postup analýzy je popsán v sekci 3.3. Na horizontální ose jsou vyneseny hodnoty korelačního koeficientu (*Spearmanova* ρ , řádky, popsané na svislé ose, pak odpovídají vybraným parametrům. Symboly popsané v legendě určují hodnoty korelace pro data z jednotlivých skupin. Z grafu je patrné, že žádný parametr nedosahuje vyšší hodnoty ρ než 0.6, a že chování napříč skupinami je více méně náhodné.

Z výsledků v grafu 4.3 je vidět, že vliv rukosti na výsledky parametrů je více méně zanedbatelný. Žádný z vybraných parametrů (které dokázaly oddělit statisticky skupiny RBD a zdravých kontrol (viz sekce 4.1) nemá hodnotu Spearmanova $\rho > 0.6$, přítomnost lepších výsledků na dominantní nebo nedominantní končetině tedy má analýza neukázala. U výsledků ostatních parametrů (které nejsou graficky vyobrazeny) je chování stejné. Není přítomna ani viditelná konzistence výsledků mezi jednotlivými skupinami - u valné většiny parametrů některé skupiny vykazují kladnou a jiné zas zápornou korelaci. Nepomohlo ani omezení výpočtu korelace na ty subjekty, které mají

větší než (na skupinu) průměrný rozdíl hodnot parametru z obou končetin.

Tabulka 4.3 ukazuje výsledky analýzy laterality pro ty parametry, které byly vybrány do klasifikační úlohy, jejíž výsledky jsou v sekci 4.4. Data v tabulce byla vypočtena tak, že se nahrály hodnoty parametru jednotlivých subjektů pro levou a pravou ruku a bylo spočítán počet subjektů, které měly vyšší, respektive nižší hodnotu na levé, respektive pravé ruce. Z toho i vyplývá viditelná zrcadlovost hodnot v tabulce (např. pokud měl subjekt vyšší hodnotu na levé ruce, musel mít nižší hodnotu na pravé atd.). Černou barvou byl označeny hodnoty, které jsou pro daný parametr, určený řádkem v tabulce, subjektivně lepší, na základě kategorizace parametrů ze sekce 3.1.3.

Název parametru	Skup.	R_{max} [%]	L_{max} [%]	R_{min} [%]	L_{min} [%]
MaxNRMSD	PD	28	72	72	28
MaxNRMSD	RBD	38	62	62	38
MaxNRMSD	CON	32	68	68	32
MinDwellTimeNRMSD	PD	60	40	40	60
MinDwellTimeNRMSD	RBD	48	52	52	48
MinDwellTimeNRMSD	CON	36	64	64	36
FreqSlopeMA	PD	64	36	36	64
FreqSlopeMA	RBD	76	24	24	76
FreqSlopeMA	CON	64	36	36	64
CloseVelNRMSD	PD	48	52	52	48
CloseVelNRMSD	RBD	46	54	54	46
CloseVelNRMSD	CON	24	76	76	24

Tabulka 4.3: Výsledky analýzy laterality horních končetin na základě vypočtených hodnot vybraných parametrů. Tato tabulka zobrazuje procentuální rozdělení vyšších a nižších hodnot parametrů, které byly vybrány do klasifikační úlohy (jak diferenčního typu tak i typu vyšší/nižší hodnoty). Sloupce uvádějí procentuální počet subjektů ze skupiny na příslušném řádku, který měl vyšší, respektive nižší výsledek na pravé, respektive levé ruce. Sytě černou barvou jsou zaznamenány hodnoty, které pro příslušný parametr dávají subjektivně lepší výsledky (nižší hodnoty odchylky u parametrů typu *NRMSD* a vyšší kladné hodnoty směrnice u parametru typu *Slope*).

Obecná interpretace hodnot z tabulky může být taková, že vzhledem k tomu, že ve všech skupinách je vysoká převaha praváků nad leváky (viz výše v této sekci) a za teoretického předpokladu, že by dominance ruky měla být určujícím prvkem, který ovlivňuje (subjektivní) kvalitu výsledků mezi horními končetinami, tak by měly být černě vyznačené hodnoty ve sloupcích, které přísluší pravé ruce, výrazně vyšší než počty minim nebo maxim na levé ruce.

Kupříkladu pokud vezmeme směrnici lineární regresní přímky vyhlazených hodnot frekvence (tedy parametr *FreqSlopeMA*), můžeme tvrdit, že lepšímu výkonu z hlediska zadání FT testu přísluší kladné pozitivní hodnoty této směrnice, které odpovídají tomu, že měřená osoba při kmitání prsty nezpomaluje. Dále můžeme vyslovit předpoklad, že dominantní ruka měřené osoby bude mít lepší motorické vlastnosti než nedominantní. Na základě (aproximovaného) procentuálního zastoupení pravorukých osob (dle výsledků Edinburského dotazníku rukosti) bychom měli ideálně vidět vyšší hodnoty (*max*) na pravých rukou měřených osob, nejlépe v takovém procentuálním

počtu, který se blíží počtu praváků ve skupinách.

Výsledkem zjednodušené analýzy je, že se nedá s jistotou tvrdit, že na dominantní ruce se objevují, z hlediska definovaných parametrů, automaticky lepší výsledky. Z vybraných parametrů se nejbližší rozdělení praváků blíží výsledky parametru *CloseVelNRMSD* u skupiny kontrol a *FreqSlopeMA* u skupiny pacientů, oba s hodnotou, která odpovídá tomu, že 76 procent subjektů (z 94 procent u skupiny RBD a 88 procent u skupiny kontrol) má lepší výkon parametru na dominantní ruce. Výsledky ale nejsou konzistentní mezi skupinami a sic ve většině případů vykazují velmi lehkou tendenci tíhnout k vyšším zastoupením na pravé ruce, u žádné skupiny a u žádného parametru (z celého souboru, včetně té většiny, která není v tabulce 4.3 uvedena) nedosáhlo procentuální zastoupení lepších výsledků na pravé ruce více než 80 procent. Parametry *MaxNRMSD*, *MinDwellTimeSlope* a *FreqSlope* (*FreqSlopeMA*) jako jediné vykazovaly lepší výsledky (více než 60 procent subjektů) na dominantní ruce přes všechny tři skupiny.

4.3 Korelační analýza

V této sekci stručně prezentuji výsledky analýzy korelace, tak jak byla popsána v části 3.4.

4.3.1 Korelace mezi parametry

Tabulka 4.4 ukazuje vybrané dvojice parametrů (z celkového počtu 828 vypočtených), které ukázaly vyšší míru korelace (či antikorelace) na základě výpočtu hodnoty Spearmanova ρ . Očekávaně se prokázaly logické vztahy mezi hodnotami průměrné otevírací a zavírací rychlosti prstů a to na obou druhých výpočtů z rukou (při braní vyšší i nižší hodnoty průměru vypočtených parametru ze všech měření na obou rukou jedné osoby). Principiálně identický vztah ukázala korelaci i mezi hodnotami maximální rychlosti otevírání a zavírání prstů. Dále se prokázal vztah mezi hodnotami směrodatných odchylek průměrné rychlosti otevírání a zavírání prstů. Tento vztah se neukázal u maximálních rychlostí, kde z důvodů, které nedokážu vysvětlit, korelační koeficient mezi směrodatnou odchylkou maximální rychlosti otevírání a směrodatnou odchylkou maximální rychlosti zavírání prstů dosahuje pouze hodnoty $\rho = 0.425$. Logická je také negativní korelace mezi mediánem frekvence FT testu a mediánem času, který osoby "tráví" v minimech (místech dotyku prstů) - se stoupajícím časem zdržení se v minimech musí klesat i celková frekvence kmitů.

Na posledních dvou řádcích tabulky je pak vidět i vztah mediánových hodnot průměrné a maximální rychlosti otevírání prstů. Dodatečně ještě uvádím případné hodnoty, které by čtenáři mohly v tabulce logicky chybět. Kombinace parametrů *CloseVelPeakMed* a *CloseVelMed*, tedy mediány průměrné a maximální rychlosti zavírání prstů vykazovaly na nižší i vyšší hodnotě s rukou korelační koeficient vyšší než 0.7. Stejně tak korelace parametru *FreqMed* a *MinDwellTimeMed* na nižší hodnotě z rukou (*min*) měla zápornou hodnotu

Kombinace parametrů	HCT	$\rho \cdot 100$
CloseVelMed - OpenVelMed	max	79.992
CloseVelMed - OpenVelMed	min	79.094
CloseVelPeakMed - OpenVelPeakMed	max	89.411
CloseVelPeakMed - OpenVelPeakMed	min	83.602
CloseVelStd - OpenVelStd	max	75.692
CloseVelStd - OpenVelStd	min	82.001
FreqMed - MinDwellTimeMed	max	-75.923
OpenVelPeakMed - OpenVelMed	max	76.829
OpenVelPeakMed - OpenVelMed	min	82.925

Tabulka 4.4: Tabulka vybraných výsledků korelací mezi dvojicemi parametrů. Do tabulky byly zahrnuty jen ty dvojice parametrů, u nichž vyšla absolutní hodnota korelačního koeficientu Spearmanova ρ vyšší, než $|\rho| \geq 0.75$ (při hodnotě $p < 0.001$), což značí nadprůměrnou míru (pozitivní či negativní) korelace mezi hodnotami parametrů. V prvním sloupci jsou zapsány dvojice parametrů, pro které byla korelace vypočítána, sloupec HCT pak označuje druh výpočtu z rukou, určující konkrétní typ obou parametrů k výpočtu korelace.

korelačního koeficientu $\rho = -0.712$, značící (očekávanou) negativní korelaci těchto parametrů.

4.3.2 Korelace s klinickými daty

Korelační analýza mezi hodnotami parametrů a klinickými údaji nepřinesla dobré výsledky. V první analýze byl hledán vztah mezi izolovanými údaji klinického zhodnocení FT testu v rámci třetí sekce dotazníku MDS-UPDRS III. Byly přitom brány v potaz pouze hodnoty pro pacienty postižené PN, jelikož ostatní skupiny měly v převážné většině případů u výsledků klinického zhodnocení FT testu nuly. Následující parametry měly p-hodnoty korelačního testu menší, než $p < 0.05$:

- *MinStd - diff*: rozdíl hodnot parametru směrodatné odchylky minim mezi oběma končetinami vykázal korelaci s rozdílem výsledků klinického zhodnocení FT testu o hodnotě $\rho = -0.586$ (hodnota $p = 0.0026$).
- *FreqStd - diff*: rozdíl hodnot parametru směrodatné odchylky frekvencí mezi oběma končetinami vykázal korelaci s rozdílem výsledků klinického zhodnocení FT testu o hodnotě $\rho = -0.516$ (hodnota $p = 0.00982$).
- *MinNRMSD - diff*: rozdíl hodnot NRMSD, vypočtených z minim, mezi oběma rukama vykázal korelaci s rozdílem výsledků klinického zhodnocení FT testu o hodnotě $\rho = -0.406$ (hodnota $p = 0.0488$).
- *OpenVelPeakSlope - diff*: rozdíl hodnot směrnice regresní přímky, proložené skrz hodnoty maximální otevírací rychlosti, mezi oběma rukama vykázal korelaci s rozdílem výsledků klinického zhodnocení FT testu o hodnotě $\rho = 0.414$ (hodnota $p = 0.0442$).

Vzhledem k nízkým hodnotám korelace a celkové absenci korelace u jiných, očekávatelnějších parametrů, seznávám tyto výsledky jako neprůkazné.

U druhého typu testu korelace hodnot parametrů s klinickými daty, kde jsou brány v potaz výsledky na konkrétních rukou (viz sekce 3.4), pouze jeden parametr dosáhl vyšší hodnoty korelačního koeficientu než 0.5, a to sice parametr *CloseVelPeakMed*, udávající medián maximální rychlosti zavírání rukou. Stejně jako u předchozí analýzy považuji tento výsledek za neprůkazný. Výčet parametrů, které při korelační analýze beroucí v potaz výsledky jednotlivých rukou vykázaly hodnotu $p < 0.05$, je následující:

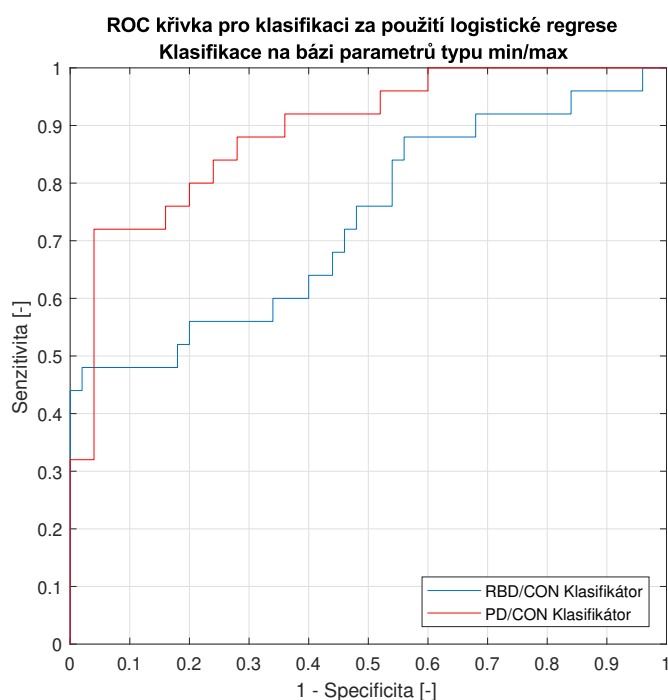
- *FreqMed* (L-R): rozdíl hodnot mediánu frekvencí period FT testu levé a pravé ruky - $\rho = 0.431$ (hodnota $p = 0.0353$).
- *OpenVelMed* (L-R): rozdíl hodnot průměrné rychlosti otevírání prstů levé a pravé ruky - $\rho = 0.447$ (hodnota $p = 0.0284$).
- *OpenVelPeakMed* (L-R): rozdíl hodnot maximální rychlosti otevírání prstů levé a pravé ruky - $\rho = 0.469$ (hodnota $p = 0.0207$).
- *CloseVelPeakMed* (L-R): rozdíl hodnot maximální rychlosti zavírání prstů levé a pravé ruky - $\rho = 0.569$ (hodnota $p = 0.00371$).

Nakonec, v testu korelace mezi hodnotami celkového MDS-UPDRS III skóre pacientů trpících PN a horšími ze dvou hodnot vypočtených na ruku (dle klasifikace v sekci 3.1.3) ani jeden z parametrů nekoreloval s hodnotou korelačního koeficientu vyšší než 0.5 (ani s hodnotou $p < 0.05$).

4.4 ROC analýza

Postup ROC analýzy za použití klasifikace logistickou regresí jsem popsal v sekci 3.5. V této části uvádím ROC křivky pro tři druhy vypočtených klasifikátorů, klasifikátor za použití parametrů typu *min* a *max*, klasifikátor za použití parametrů typu *diff* a klasifikátor kombinující oba druhy parametrů. Pro každý typ jsem provedl výpočet klasifikátoru a vykreslení ROC křivky pro klasifikaci skupiny RBD od skupiny zdravých kontrol a pro klasifikaci skupiny PD od skupiny zdravých kontrol.

První byl klasifikátor nastavený hodnotami vypočtených parametrů typu *min* a *max*. Výslednou vypočtenou ROC křivku, jak pro klasifikaci RBD od kontrol tak klasifikaci PD od kontrol, je možné vidět na obrázku 4.4. Jak je vidět, navržený klasifikátor nemá ideální schopnost rozdělení skupin RBD a zdravých kontrol. Přes relativně uspokojující hodnotu senzitivity (tedy klasifikace skutečně nemocných jako nemocných) $SENZ = 0.84$ je velmi nízká hodnota specificity (tedy klasifikace skutečně zdravých jako zdravých), $SPEC = 0.48$. To by praxi znamenalo, že tento klasifikátor je při zařazování skutečně zdravých lidí do zdravé skupiny zhruba stejně efektivní, jako je házení mincí. Klasifikátor tak označuje příliš mnoho osob jako nemocných. Dodatečně, hodnota AUC (plochy pod křivkou) tohoto klasifikátoru je $AUC = 0.73$. Nedostatky tohoto klasifikátoru diskutuji v kapitole 5. Oddělení skupiny

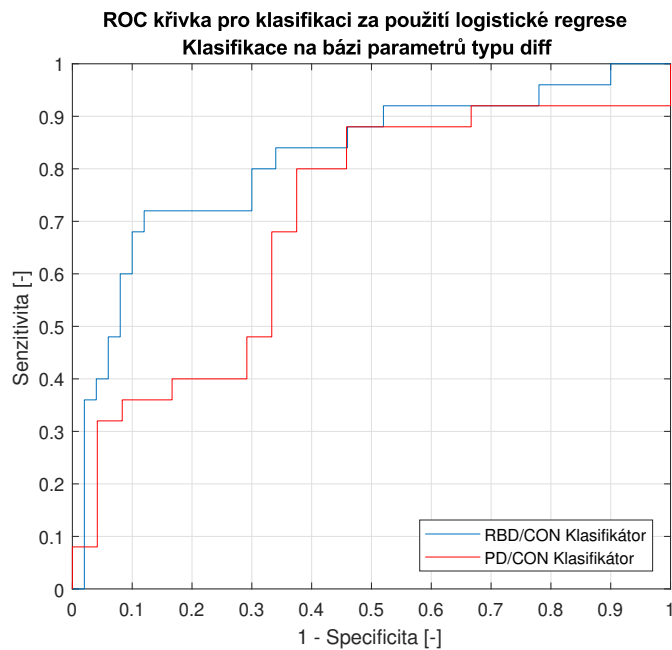


Obrázek 4.4: ROC (*Receiver Operating Characteristic*) křivka pro logisticky regresní klasifikátor za použití dat parametrů typu *min* a *max*. Výkon klasifikátoru mezi skupinami RBD a CON: SENZ = 0.84, SPEC = 0.48, AUC = 0.734. Výkon klasifikátoru mezi skupinami PD a CON: SENZ = 0.76, SPEC = 0.76, AUC = 0.89.

pacientů s PN a zdravých kontrol má pak celkem výrazně lepší výkon. To souhlasí s rozložením vypočtených hodnot parametrů, na kterých je tento typ klasifikátoru postaven, kdy jsou všechny skupiny v logické posloupnosti tak, že skupina pacientů s PN má hodnoty více vzdálené od kontrol, než skupina RBD (viz obrázek 4.1).

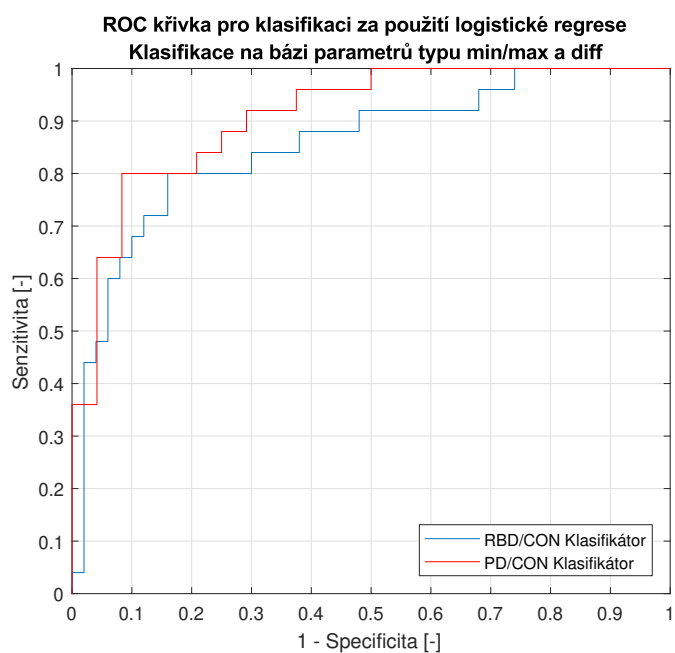
Obdobná situace nastává i u druhého klasifikátoru, který byl nastaven za použití hodnot parametrů typu *diff*. Jejich výčet je opět možné najít v sekci 3.5. Díky mírně odlišné struktuře hodnot parametrů typu *diff* (viz diskuze v následující kapitole) má tento klasifikátor o něco lepší výkon při klasifikaci skupiny RBD od skupiny zdravých kontrol, jak je vidět z příslušné ROC křivky na obrázku 4.5. Výsledná hodnota senzitivity SENZ je rovna 0.9, což je více, než u předchozího klasifikátoru. Hodnota specificity je pak totožná, tedy SPEC = 0.48. Hodnota AUC pro tento klasifikátor je pak logicky také o něco větší, AUC = 0.82.

Při klasifikaci skupiny pacientů s PN od zdravých kontrol dochází k významnému poklesu výkonu. Tento pokles je způsoben tím, že parametry typu *diff* nejsou pro klasifikaci skupiny PN postaveny, jak je ostatně vidět z příslušných rozdělení hodnot v obrázku 4.2. Více tento efekt rozvádím v části 5.



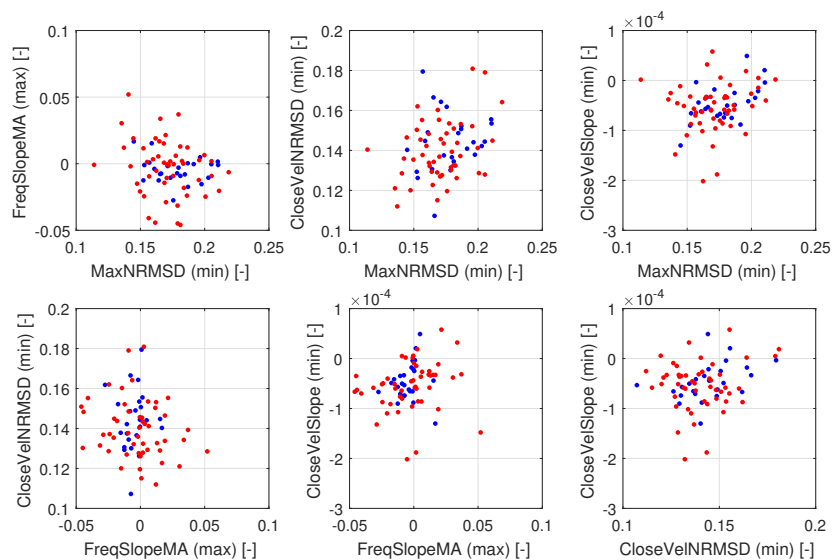
Obrázek 4.5: ROC (*Receiver Operating Characteristic*) křivka pro logistický regresní klasifikátor za použití dat parametrů typu *diff*. Výkon klasifikátoru mezi skupinami RBD a CON: SENZ = 0.9, SPEC = 0.48, AUC = 0.82. Výkon klasifikátoru mezi skupinami PD a CON: SENZ = 0.64, SPEC = 0.56, AUC = 0.702.

Na posledním grafu této části je vidět provedená klasifikace za použití kombinace parametrů obou předchozích typů. Výsledky klasifikátoru, který odděluje skupiny RBD a zdravých kontrol jsou lepší z hlediska hodnoty AUC, byť jen o relativně malou hodnotu. Klasifikace skupiny pacientů s PN od skupiny zdravých kontrol je pak prakticky totožná od klasifikace jen za pomocí parametrů typu *min* a *max*, což odpovídá tomu, že hodnoty klasifikace těchto dvou skupin pomocí parametrů typu *diff* nemohly prakticky přinést do klasifikátoru žádnou novou informaci.

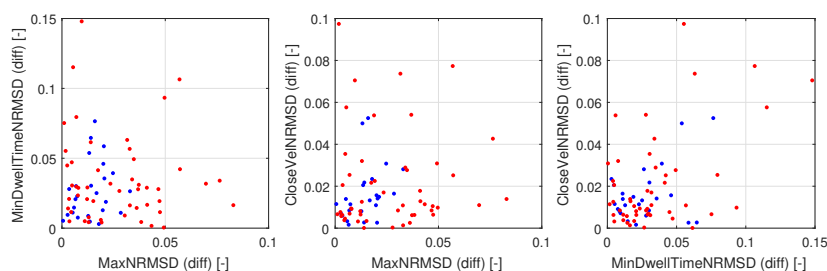


Obrázek 4.6: ROC (*Receiver Operating Characteristic*) křivka pro logisticky regresní klasifikátor za použití dat parametrů jak typu *min* a *max*, tak typu *diff*. Výkon klasifikátoru mezi skupinami RBD a CON: SENZ = 0.88, SPEC = 0.48, AUC = 0.855. Výkon klasifikátoru mezi skupinami PD a CON: SENZ = 0.76, SPEC = 0.76, AUC = 0.91.

Navíc uvádím sérii grafů ilustrující rozdělení hodnot obou skupin (pacientů s RBD a zdravých kontrol) podle dvojic parametrů použitých pro konstrukci klasifikátoru. Na obrázku 4.7 a 4.8 jsou modrou barvou označeny datové body patřící zdravým kontrolám a červeně datové body patřící pacientům s RBD. Je možné si povšimnout, že rozložení hodnot mezi oběma skupinami se liší zejména s ohledem na jejich variabilitu. U skupin parametrů typu *diff* je pak rozložení ještě význačně tím, že hodnoty typu *diff* nabývají vždy pozitivních hodnot od nuly výše.



Obrázek 4.7: Ilustrační zobrazení datových bodů pro různé kombinace použitých parametrů typu *min* a *max* v klasifikátoru na bázi logistické regrese. Datové body skupiny pacientů trpících RBD jsou vyznačeny červeně a datové body zdravých kontrol jsou vyznačeny modře.



Obrázek 4.8: Ilustrační zobrazení datových bodů pro různé kombinace použitých parametrů typu *diff* v klasifikátoru na bázi logistické regrese. Datové body skupiny pacientů trpících RBD jsou vyznačeny červeně a datové body zdravých kontrol jsou vyznačeny modře.

Kapitola 5

Diskuze

V této části budu diskutovat výsledky provedených analýz, navrhopvat možná vylepšení a zamýšlet se nad přesahy této práce.

Z výsledků statistické analýzy byly vybrány parametry s nejlepší schopností oddělit skupinu pacientů trpících RBD a zdravých kontrol. Ve skupině parametrů typu *min* a *max* bylo nalezeno 8 parametrů, pomocí kterých se na hladině statistické významnosti $\alpha = 0.05$ podařilo najít pomocí Mann-Whitneyho U-testu významné rozdíly mezi rozděleními hodnot skupiny RBD a skupiny zdravých kontrol. Ze skupiny parametrů typu *diff* bylo stejným způsobem nalezeno 6 parametrů schopných statisticky rozlišit tyto skupiny.

Na základě parametrů z obou skupin byli vybráni nejvhodnější kandidáti, na kterých byl navržen logisticky regresní klasifikátor, který s kombinací obou dvou druhů parametrů dokázal úspěšně klasifikovat obě dvě skupiny. AUC tohoto klasifikátoru dosáhlo slušných hodnot, AUC = 0.855 pro klasifikátor s kombinací parametrů. První typ parametrů dokázal dle očekávání oddělit výborně skupinu pacientů s PN a zdravých kontrol při hodnotě $AUC = 0.89$.

Vzhledem k tomu, že rozdělení hodnot parametrů jak typu *min* a *max*, tak typu *diff* se mezi skupinami pacientů s RBD a skupinou zdravých kontrol liší zejména ve své variabilitě (rozptylu), použité klasifikátory trpí nízkými hodnotami specificity. Jak je možné vidět na uvedených grafech 4.7 a 4.8, obě skupiny nejsou triviálně oddělitelné s vysokou efektivitou pomocí lineární dělicí (nad)plochy. Pro efektivnější způsob klasifikace by bylo v budoucnu lepší použít sofistikovanější metodu, která dokáže vytvořit nelineární (například kruhovou) dělicí plochu mezi hodnotami obou skupin, například v podobě SVM (*Support Vector Machine*) nebo jiného klasifikátoru (např. neuronové sítě).

Výsledky korelační analýzy mezi parametry v sekci 4.3 prokázali předpokládaný vztah mezi parametry popisujícími rychlost otevírání a zavírání prstů, a parametry frekvence a času stráveného v minimech jejich vzdálenosti. Na druhou stranu, korelační analýza mezi klinickými hodnotami a hodnotami parametrů nepřinesla dobré výsledky. U korelací se skóre klinického zhodnocení FT testu by se daly špatné výsledky (nenalezeného vztahu s parametry) přisoudit malé jemnosti hodnotící škály, která se sestává z celých čísel mezi 0 a 4. U korelace hodnot parametrů s celkovým skóre dotazníku MDS-UPDRS je pak jedním z možných vysvětlení nenalezeného vztahu u pacientů trpících

PN vliv všech ostatních příznaků PN, které dotazník MDS-UPDRS hodnotí. Jejich rozličná intenzita pak může posouvat hodnoty celkového skóre na obě strany dál, než je schopen úroveň postižení FT test odrazit v naměřených hodnotách.

Analýza laterality a její výsledky v sekci 4.2 byly překvapivé. Hypotéza, že dominance (preference) končetiny bude mít vliv na výsledky parametrů, nebyla potvrzena. Z výsledků práce vyplývá, že z hlediska navržených parametrů, nemá dominance ruky žádný vliv na jejich hodnoty. Očekávání, že dominantní ruka bude mít (například alespoň u skupiny zdravých kontrol) téměř vždy lepší výkon, se nepotvrdilo, přesto, že existuje malá tendence k tomuto trendu. Jedna ze studií⁷⁴ ukazuje, že dominantní ruka (resp. dominantní strana těla) je zasažena jako první ve většině případů, nicméně strana, na které nemoc začne, nemá vliv na celkové MDS-UPDRS skóre (a celkovou intenzitu PN) pacientů. Jiná studie⁷⁵ z roku 2014, že procentuální zastoupení pacientů, jejichž symptomy začaly na dominantní straně těla, nejsou zas tak vysoká. Pokud by symptomy počaly na dominantní straně těla, a ta byla postupem PN zasažena více, tak by se dalo očekávat, že se pacientů trpících PN ukáže významná antikorelace mezi hodnotami Edinburského dotazníku rukosti a hodnotami parametrů. Pokud by byla nemoc symetrická, dalo by se očekávat, že dominantní ruka bude mít lepší výsledky. Nicméně lepší výsledky na dominantní ruce se z hlediska mých parametrů nepotvrdily ani u zdravých osob. Problematika laterality (vzhledem k tomu, že lateralita sama u lidí není triviální vlastností, která by se dala lehce vysvětlit) má ve vztahu s PN (potažmo jejím počátkem) mnoho proměnných, a je zapotřebí hlubší analýzy, na kterou v této práci již nebylo místo. S určitou pravděpodobností se ale dá říct, že pokud by v mých statistických analýzách byly souhrnně u všech osob brány jen výsledky dominantní (či nedominantní) ruky, nebylo by možné skupiny zdravých osob a osob trpících RBD statisticky rozlišit. Toto a další vlivy laterality budou předmětem pokračujícího výzkumu, například za pomoci klinických dat měření úbytku dopaminu v levých a pravých postižených částech mozku.

5.1 Další směřování

Posledním tématem k zamyšlení bude krátká diskuze nad výsledky statistických testů za použití hodnot parametrů typu *diff*. Při hloubkové kontrole výsledků všech parametrů jsem si povšiml zajímavého trendu, a to sice že v nezanedbatelném počtu výsledků se opakuje následující trend. Při pohledu na výsledky od všech tří skupin, měla skupina RBD velmi často nejmenší rozdíl mezi výsledky obou končetin. Při bližším ohledání výsledků lepší i horší ruky jsem zjistil, že skupina RBD má tendenci na své horší ruce (nehledě na laterality) setrvávat s výkonem podobně nebo trochu hůře než skupina zdravých kontrol, a na své lepší ruce má pak vůči zdravým kontrolám významný propad směrem k hodnotám, které zaujímá skupina pacientů s PN. Tento propad sám o sobě není statisticky významný, nicméně pokud se na něj podíváme z pohledu rozdílu obou rukou, je situace odlišná. Jinými slovy, u zdravých lidí

by se dalo uvažovat, že budou mít přirozený rozdíl v motorických schopnostech horních končetin. Potom je má myšlenka taková, že u pacientů s RBD (a s možnou budoucí konverzí do PN) je lepší z těchto končetin (nehledě na dominanci) zasažena jako první a její výkon je postižen tak, že se vyrovnává výkonu končetiny, která byla původně (za zdravého stavu) přirozeně horší. Při další progresi nemoci a možné konverzi do PN výkon prvně postižené končetiny klesá ještě dále, a rozdíl mezi výkonem končetin, sic v opačném "směru", se znovu objeví.

K potvrzení této myšlenky v praxi by bylo nutné vypracování longitudinální studie, která by pečlivě sledovala vývoj bradykineze (z pohledu parametrů definovaných v této práci) v čase. Pokud by byla potvrzena výše uvedená hypotéza progresu postižení u pacientů trpících RBD, zcela jistě by to otevřelo dveře k dalším, efektivním možnostem brzké diagnostiky PN.

Zároveň bude možné zlepšit výkon klasifikace za použitých parametrů použitím sofistikovanějšího klasifikátoru a pečlivého naladění klasifikace s cílem dosáhnout nejlepšího klasifikačního výkonu. Tyto cíle budou součástí pokračujícího výzkumu navazujícího na tuto práci.

Jedním z dalších možných vylepšení v analýzách, které jsem použil, by mohlo být použití mezi-kvartilové vzdálenosti pro normalizaci dat ve výpočtu NRMSD, čímž by se omezil vliv extrémních hodnot. Také bych mohl vyzkoušet vliv časového oříznutí z počátku signálu na výsledné hodnoty parametrů a následné statistické analýzy. Ze strany výběru datového souboru by bylo možné omezit pacienty PN na tu skupinu, která má formu nemoci s příznaky RBD. Skupinu RBD pak šlo teoreticky, například pomocí diagnostiky za pomoci magnetické rezonance, omezit na ty pacienty, jejichž forma nemoci předvídá neurodegenerativní progresi.

Kapitola 6

Závěr

Tato práce jako první na světě objektivně analyzovala vliv bradykineze u pacientů s RBD a ukázala, že je možné tento vliv použít pro detekci PN v její prodromální fázi. Vycházejíc ze souboru parametrů, které byly použity pro detekci bradykineze u pacientů s PN³¹, byl navržen nový soubor rozšířených algoritmů a parametrů vytvořených s cílem zachytit subtilní změny motorických schopností v raných stádiích PN. Všechny parametry byly prověřeny pomocí statistické analýzy na velkém vzorku pacientů trpících PN, pacientů trpících RBD a zdravých kontrol. Bylo vybráno celkem 14 parametrů, pomocí kterých byly nalezeny statisticky významné odlišnosti v rozděleních skupin zdravých kontrol a pacientů s RBD. S využitím těchto parametrů byl navržen expertní systém, který dokázal úspěšně klasifikovat skupinu zdravých kontrol a skupinu pacientů trpících RBD s relativně vysokou hodnotou $AUC = 0.855$. Na výsledcích této práce se nadále bude stavět v podobě pokračujícího výzkumu, jehož cílem bude vylepšení klasifikace a výkonu parametrů, stejně jako případná longitudinální studie mapující progresi onemocnění u pacientů s RBD směrem k PN. V budoucnu se pak výsledky této práce budou používat pro prodromální detekci PN, která je zásadní pro její včasnou a efektivní léčbu.

Bibliografie

- [1] Anonym. *Parkinsonova nemoc v ČR - aktuální průzkum*. Edukafarm. 12. dub. 2015. URL: <http://www.edukafarm.cz/c1421-parkinsonova-nemoc-v-cr-aktualni-pruzkum> (cit. 12. 12. 2018).
- [2] B. S. Schoenberg, D. W. Anderson a A. F. Haerer. „Prevalence of Parkinson’s disease in the biracial population of Copiah County, Mississippi“. In: *Neurology* 35.6 (1. červ. 1985). DOI: 10.1212/WNL.35.6.841.
- [3] T. Cikrt a Š. Křeček. *Ekonomická analýza zdravotních pojišťoven v České republice. Jak zdravé jsou zdravotní pojišťovny?* Analýza. Národohospodářská fakulta VŠE, 2018.
- [4] C. Marras, J. C. Beck, J. H. Bower et al. „Prevalence of Parkinson’s disease across North America“. In: *npj Parkinson’s Disease* 4.21 (1 2018). DOI: 10.1038/s41531-018-0058-0.
- [5] P. foundation. *Parkinson’s foundation. Understanding Parkinson’s - Statistics*. URL: <https://parkinson.org/Understanding-Parkinsons/Statistics> (cit. 12. 12. 2018).
- [6] O. Tichý. *Parkinson stál loni VZP 455 milionů. Náklady na nejmodernější léčbu stouply za tři roky o 41 procent*. Všeobecná zdravotní pojišťovna - VZP. 6. dub. 2017. URL: <https://www.vzp.cz/o-nas/aktuality/parkinson-stal-loni-vzp-455-milionu-naklady-na-nejmodernejsi-lecbu-stouply-za-tri-roky-o-41> (cit. 12. 12. 2018).
- [7] Y. Winter, S. von Campenhausen, H. Brožová et al. „Costs of Parkinson’s disease in eastern Europe: a Czech cohort study.“ In: *Parkinsonism & related disorders* 16 1 (2010), s. 51–6.
- [8] Anonym. *Nepřítel jménem Parkinson. Česká republika chce výrazně zlepšit péči o nemocné*. Česká televize, ČT24. 10. dub. 2018. URL: <https://ct24.ceskatelevize.cz/veda/2445511-nepritel-jmenem-parkinson-ceska-republika-chce-vyrazne-zlepsit-peci-o-nemocne>.
- [9] A. A. Moustafa, S. Chakravarthy, J. R. Phillips et al. „Motor symptoms in Parkinson’s disease: A unified framework“. In: *Neuroscience & Biobehavioral Reviews* 68 (zář. 2016), s. 727–740. DOI: 10.1016/j.neubiorev.2016.07.010.
- [10] W. Poewe. „Non-motor symptoms in Parkinson’s disease“. In: *European Journal of Neurology* 15 (s1 17. břez. 2008), s. 14–20. DOI: 10.1111/j.1468-1331.2008.02056.x.
- [11] L. M. Shulman, R. L. Taback, A. A. Rabinstein et al. „Non-recognition of depression and other non-motor symptoms in Parkinson’s disease“. In: *Parkinsonism & Related Disorders* 8 (3 led. 2002), s. 193–197. DOI: 10.1016/s1353-8020(01)00015-3.
- [12] Anonym. *How is Sleep Regulated? Under the Brain’s Control*. Healthy Sleep - Division of Sleep Medicine at Harvard Medical School. 18. pros. 2007. URL: <http://healthysleep.med.harvard.edu/healthy/science/how/neurophysiology> (cit. 19. 12. 2018).

- [28] J. Desrosiers, R. Hébert, G. Bravo et al. „The Purdue Pegboard Test: Normative data for people aged 60 and over“. In: *Disability and Rehabilitation* 17 (5 srp. 1995), s. 217–224. DOI: 10.3109/09638289509166638.
- [29] M. Memedi, T. Khan, P. Grenholm et al. „Automatic and Objective Assessment of Alternating Tapping Performance in Parkinson’s Disease“. In: *Sensors* 13 (12 9. pros. 2013), s. 16965–16984. DOI: 10.3390/s131216965.
- [30] R. B. Postuma, A. E. Lang, J. F. Gagnon et al. „How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behaviour disorder“. In: *Brain* 135 (6 2012), s. 1860–1870. DOI: 10.1093/brain/aww093.
- [31] R. Krupička. „Měření a analýza pohybu ruky u parkinsoniků“. Disertační práce. Kladno: České vysoké učení technické v Praze, Fakulta biomedicínského inženýrství - Katedra biomedicínské informatiky, 2013. URL: <http://hdl.handle.net/10467/37372> (cit. 18. 12. 2018).
- [32] R. Krupička, Z. Szabo, S. Viteckova et al. „Motion Capture System for Finger Movement Measurement in Parkinson Disease“. In: *Radioengineering* 23 (červ. 2014), s. 659–664. URL: <http://hdl.handle.net/11012/36464> (cit. 18. 12. 2018).
- [33] Anonym. *Parkinsonova nemoc - Hlavní pohybové příznaky*. 2017. URL: <https://parkinson-help.cz/parkinsonova-nemoc-pn/priznaky-motoricke/zakladni-priznaky/> (cit. 18. 12. 2018).
- [34] Anonym. *About Parkinson’s - Bradykinesia*. Led. 2018. URL: <https://www.epda.eu.com/about-parkinsons/symptoms/motor-symptoms/bradykinesia/> (cit. 18. 12. 2018).
- [35] B. E. Fisher, A. D. Wu, G. J. Salem et al. „The Effect of Exercise Training in Improving Motor Performance and Corticomotor Excitability in People With Early Parkinson’s Disease“. In: *Archives of Physical Medicine and Rehabilitation* 89 (7 čvc 2008), s. 1221–1229. DOI: 10.1016/j.apmr.2008.01.013.
- [36] C. Pacchetti, F. Mancini, R. Aglieri et al. „Active Music Therapy in Parkinson’s Disease: An Integrative Method for Motor and Emotional Rehabilitation“. In: *Psychosomatic Medicine* 62 (3 květ. 2000), s. 386–393. DOI: 10.1097/00006842-200005000-00012.
- [37] E. Haneishi. „Effects of a Music Therapy Voice Protocol on Speech Intelligibility, Vocal Acoustic Measures, and Mood of Individuals with Parkinson’s Disease“. In: *Journal of Music Therapy* 38 (4 1. pros. 2001), s. 273–290. DOI: 10.1093/jmt/38.4.273.
- [38] E. K. St Louis a Boeve. „REM Sleep Behavior Disorder: Diagnosis, Clinical Implications, and Future Directions“. In: *Mayo Clinic Proceedings* 92 (11 lis. 2017), s. 1723–1736. DOI: 10.1016/j.mayocp.2017.09.007.
- [39] C. L. Comella, T. M. Nardine, N. J. Diederich et al. „Sleep-related violence, injury, and REM sleep behavior disorder in Parkinson’s disease“. In: *Neurology* 51 (2 1. srp. 1998), s. 526–529. DOI: 10.1212/WNL.51.2.526.
- [40] E. K. St Louis, A. R. Boeve a B. F. Boeve. „REM Sleep Behavior Disorder in Parkinson’s Disease and Other Synucleinopathies“. In: *Movement Disorders* 32 (5 17. květ. 2017), s. 645–658. DOI: 10.1002/mds.27018.
- [41] D. Arnaldi, E. Antelmi, E. K. St. Louis et al. „Idiopathic REM sleep behavior disorder and neurodegenerative risk: To tell or not to tell to the patient? How to minimize the risk?“ In: *Sleep Medicine Reviews* 36 (pros. 2017), s. 82–95. DOI: 10.1016/j.smrv.2016.11.002.
- [42] B. F. Boeve, M. H. Silber, T. J. Ferman et al. „Clinicopathologic correlations in 172 cases of rapid eye movement sleep behavior disorder with or without a coexisting neurologic disorder“. In: *Sleep Medicine* 14 (8 srp. 2013), s. 754–762. DOI: 10.1016/j.sleep.2012.10.015.
- [43] I. MathWorks. *MATLAB - MathWorks. MATLAB & Simulink*. The MathWorks, Inc. URL: <https://www.mathworks.com/products/matlab.html> (cit. 26. 12. 2018).

- [61] H. B. Mann a D. R. Whitney. „On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other“. In: *Annals of Mathematical Statistics* 18 (1 1947), s. 50–60. DOI: 10.1214/aoms/1177730491.
- [62] Mathworks. *Wilcoxon rank sum test. MATLAB ranksum*. The Mathworks, Inc. 2018. URL: <https://www.mathworks.com/help/stats/ranksum.html> (cit. 27. 12. 2018).
- [63] Mathworks. *Kruskal-Wallis test. MATLAB kruskalwallis*. The Mathworks, Inc. 2018. URL: <https://www.mathworks.com/help/stats/kruskalwallis.html> (cit. 27. 12. 2018).
- [64] W. H. Kruskal a W. A. Wallis. „Use of Ranks in One-Criterion Variance Analysis“. In: *Journal of the American Statistical Association* 47 (260 1952), s. 583–621. DOI: 10.1214/aoms/1177730491.
- [65] Mathworks. *Multiple comparison test. MATLAB multcompare*. The Mathworks, Inc. 2018. URL: <https://www.mathworks.com/help/stats/multcompare.html> (cit. 02. 01. 2019).
- [66] Mathworks. *Multiple comparisons. Least Significant Difference*. The Mathworks, Inc. 2018. URL: <https://www.mathworks.com/help/stats/multiple-comparisons.html#bum7uib-1> (cit. 02. 01. 2019).
- [67] J. L. Myers a A. D. Well. *Research Design and Statistical Analysis*. 2. vydání. Lawrence Erlbaum, 2003. ISBN: 978-0-8058-4037-7.
- [68] T. Fawcett. „An introduction to ROC analysis“. In: *Pattern Recognition Letters* 27 (8 2006), s. 861–874. DOI: 10.1016/j.patrec.2005.10.010.
- [69] Mathworks. *Generalized linear model regression. MATLAB glmfit*. The Mathworks, Inc. 2018. URL: <https://www.mathworks.com/help/stats/glmfit.html> (cit. 02. 01. 2019).
- [70] P. McCullagh a J. A. Nelder. *Generalized Linear Models*. 2. vydání. Chapman a Hall, 1989. ISBN: 0-412-31760-5.
- [71] Mathworks. *Generalized linear model values. MATLAB glmval*. The Mathworks, Inc. 2018. URL: <https://www.mathworks.com/help/stats/glmval.html> (cit. 02. 01. 2019).
- [72] Mathworks. *Receiver operating characteristic (ROC) curve or other performance curve for classifier output. MATLAB perfcurve*. The Mathworks, Inc. 2018. URL: <https://www.mathworks.com/help/stats/perfcurve.html> (cit. 02. 01. 2019).
- [73] Mathworks. *Box plot. MATLAB boxplot*. The Mathworks, Inc. 2018. URL: <https://www.mathworks.com/help/stats/boxplot.html> (cit. 03. 01. 2019).
- [74] M. J. Barrett, S. A. Wylie, M. B. Harrison et al. „Handedness and motor symptom asymmetry in Parkinson’s disease“. In: *Journal of Neurology, Neurosurgery & Psychiatry* 82 (10 2011), s. 1122–1124. DOI: 10.1136/jnnp.2010.209783.
- [75] J. Shi, J. Liu a Q. Q. „Handedness and dominant side of symptoms in Parkinson’s disease“. In: *Medicina Clínica* 142 (4 2014), s. 141–144. DOI: 10.1016/j.medcli.2012.11.028.

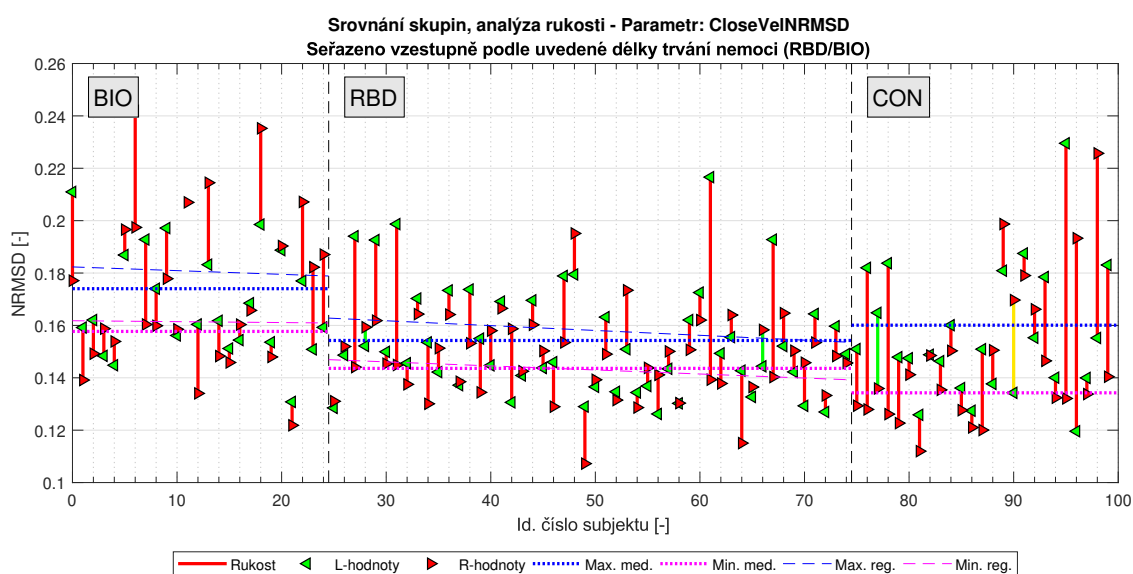
Příloha A

Použité zkratky

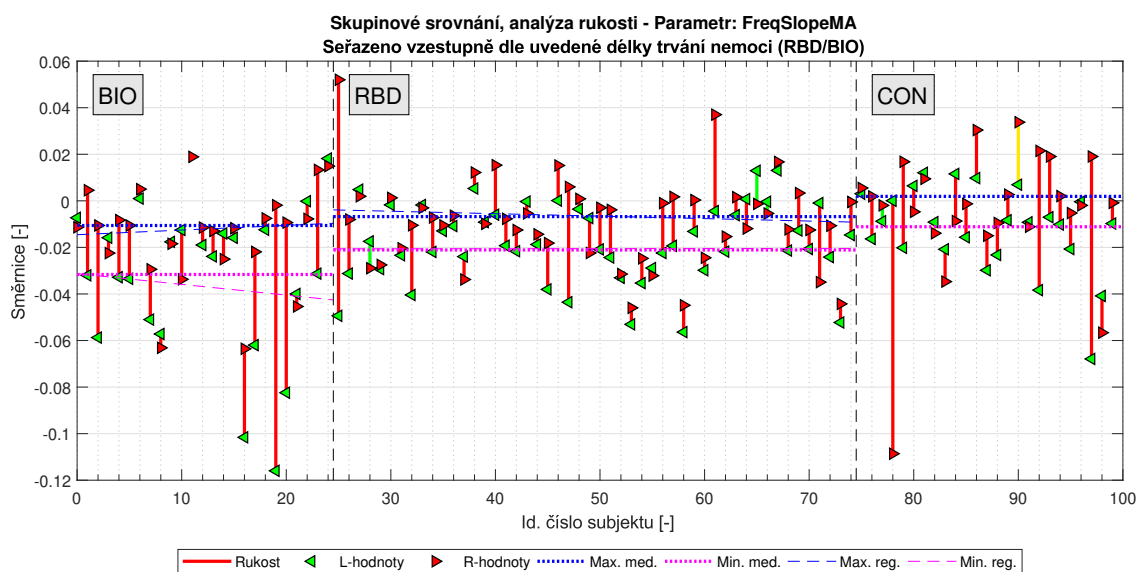
- AD - Alzheimer's disease (Alzheimerova nemoc)
- AN - Alzheimerova nemoc
- AUC - Area Under Curve (Oblast pod křivkou)
- DLB - Dementia with Lewy bodies (Demence s Lewyho tělísky)
- EHI - Edinburgh Handedness Inventory (Edinburský dotazník rukosti/laterality)
- FT - Finger-tapping ("klepání prstem")
- GLM - Generalized linear model (Zobecněný lineární model)
- iRBD - Idiopathic RBD (Idiopatická forma RBD)
- L-dopa - Levodopa
- LB - Lewy bodies (Lewyho tělíška)
- LN - Lewy neurites (Lewyho neurity)
- MA - Moving Average (Klouzavý průměr)
- MDS-UPDRS - Movement Disorder Society UPDRS
- MRI - Magnetic resonance imaging (Magnetická rezonance)
- MSA - Multiple system atrophy (Multisystémová atrofie, také *Shyův-Dragerův syndrom*)
- NRMSD - Normalized Root Mean Square Deviation (Normalizovaná RMS odchylka)
- PD - Parkinson's disease (Parkinsonova nemoc)
- PN - Parkinsonova nemoc
- RBD - Rapid eye movement sleep behavior disorder (Porucha chování v REM spánku)
- RMS - Root-mean-square (Odmocnina z průměrů umocněných na druhou)
- ROC - Receiver operating characteristic (Operační charakteristika přijímače)
- SVM - Support Vector Machine
- UPDRS - Unified Parkinson's Disease Rating Scale (Sjednocená škála pro hodnocení Parkinsonovy nemoci)

Příloha B

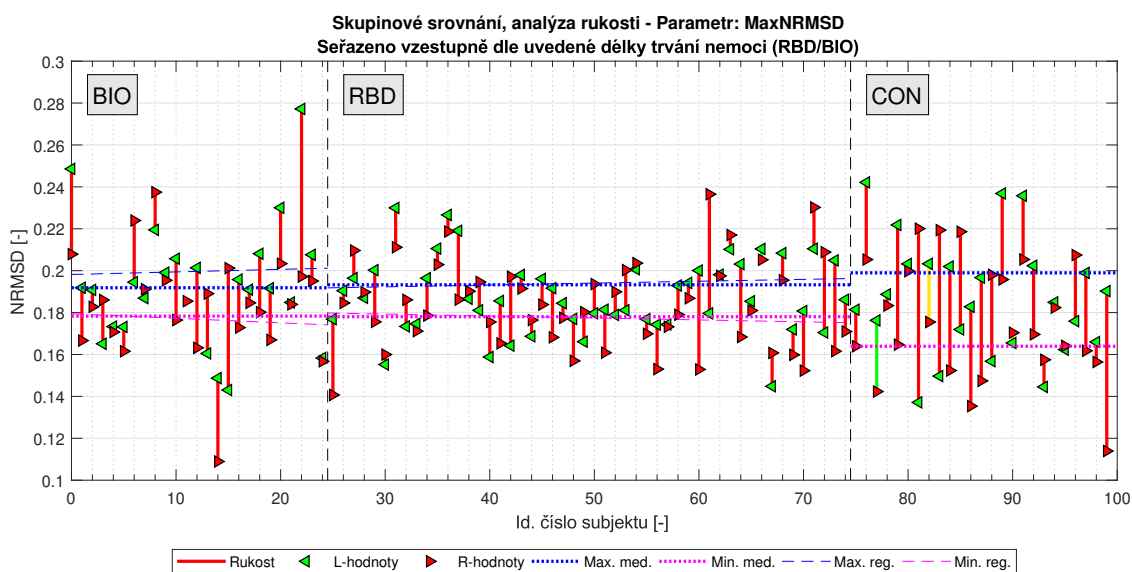
Dodatečné grafy



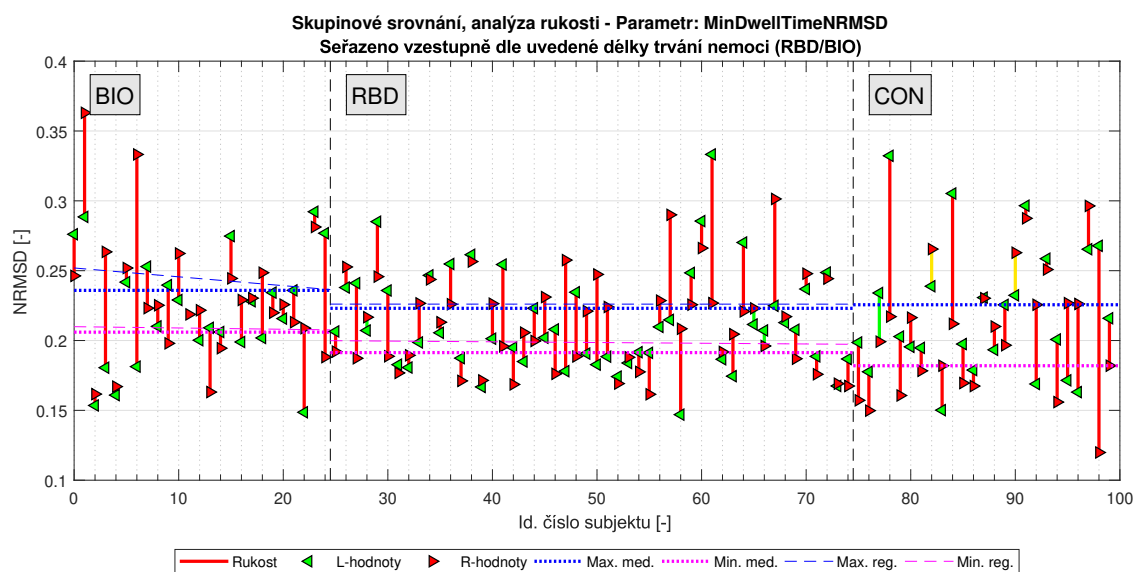
Obrázek B.1: Skupinové srovnání a souhrnné zobrazení výsledných hodnot parametrů pro parametr *CloseVelNRMSD*. Jednotlivé sloupce v grafu přísluší jednotlivým měřeným osobám. Zelené trojúhelníčky jsou hodnoty naměřené na levé ruce, červené trojúhelníčky jsou hodnoty naměřené na pravé ruce. Tečkované čáry označují mediány maxim a minim, viz legendu. Čárkované čáry jsou lineární regresní přímky proložené skrz hodnoty maxim a minim. Data skupiny BIO a skupiny RBD byla seřazena zleva doprava na základě vzrůstající uvedené hodnoty délky trvání nemoci. Barva spojovníků datových bodů určuje rukost měřené osoby, červená určuje praváka, zelená leváka, žlutá pak osobu bez preference.



Obrázek B.2: Skupinové srovnání a souhrnné zobrazení výsledných hodnot parametrů pro parametr *FreqSlopeMA*. Jednotlivé sloupce v grafu přísluší jednotlivým měřeným osobám. Zelené trojúhelníčky jsou hodnoty naměřené na levé ruce, červené trojúhelníčky jsou hodnoty naměřené na pravé ruce. Tečkované čáry označují mediány maxim a minim, viz legendu. Čárkované čáry jsou lineární regresní přímky proložené skrz hodnoty maxim a minim. Data skupiny BIO a skupiny RBD byla seřazena zleva doprava na základě vzrůstající uvedené hodnoty délky trvání nemoci. Barva spojovníků datových bodů určuje rukost měřené osoby, červená určuje praváka, zelená leváka, žlutá pak osobu bez preference.



Obrázek B.3: Skupinové srovnání a souhrnné zobrazení výsledných hodnot parametrů pro parametr *MaxNRMSD*. Jednotlivé sloupce v grafu přísluší jednotlivým měřeným osobám. Zelené trojúhelníčky jsou hodnoty naměřené na levé ruce, červené trojúhelníčky jsou hodnoty naměřené na pravé ruce. Tečkované čáry označují mediány maxim a minim, viz legendu. Čárkované čáry jsou lineární regresní přímky proložené skrz hodnoty maxim a minim. Data skupiny BIO a skupiny RBD byla seřazena zleva doprava na základě vzrůstající uvedené hodnoty délky trvání nemoci. Barva spojovníků datových bodů určuje rukost měřené osoby, červená určuje praváka, zelená leváka, žlutá pak osobu bez preference.



Obrázek B.4: Skupinové srovnání a souhrnné zobrazení výsledných hodnot parametrů pro parametr *MinDwellTimeNRMSD*. Jednotlivé sloupce v grafu přísluší jednotlivým měřeným osobám. Zelené trojúhelníčky jsou hodnoty naměřené na levé ruce, červené trojúhelníčky jsou hodnoty naměřené na pravé ruce. Tečkované čáry označují mediány maxim a minim, viz legendu. Čárkované čáry jsou lineární regresní přímky proložené skrz hodnoty maxim a minim. Data skupiny BIO a skupiny RBD byla seřazena zleva doprava na základě vzrůstající uvedené hodnoty délky trvání nemoci. Barva spojovníků datových bodů určuje rukost měřené osoby, červená určuje praváka, zelená leváka, žlutá pak osobu bez preference.

Příloha C

Obsah přiloženého CD

- Zdrojové skripty, použité pro výpočty v této práci

(Formát: Matlab m-file)

- **Analýza Laterality**

(Složka se skripty pro analýzu rukosti)

- **AUC Analýza**

(Složka se skripty pro analýzu AUC z ROC křivky)

- **Korelační Analýza**

(Složka se skripty pro korelační analýzy)

- **Ostatní**

(Ostatní pomocné skripty pro práci s datovým souborem a výpočet parametrů)

- **Statistická analýza**

(Složka se skripty pro statistickou analýzu)

- **Vykreslení grafů**

(Složka se skripty určenými pro vykreslování výsledků analýz)

- DP_Kryze_2018.pdf

(Tato práce v elektronické podobě)