EARLY DIAGNOSIS OF PARKINSON'S DISEASE THROUGH MACHINE LEARNING ANALYSIS OF HANDWRITING BASED ON THE PAHAW DATASET.

QUANTIFYING THE ECONOMIC BURDEN OF PARKINSON'S DISEASE IN CZECHIA.

By

Sunčica Rosić

Submitted to Central European University – Private University

Undergraduate Studies

In partial fulfilment of the requirement for the Bachelor of Science/Bachelor of Arts in Data Science and Society

Supervisor: Mark Wittek

COPYRIGHT NOTICE

ABSTRACT

This thesis investigates the predictive potential of handwriting in the early detection of Parkinson's Disease (PD) (Study 1). Currently, PD diagnostics mostly involves clinical methods like imaging or Cerebrospinal fluid and serum tests. In recent years, we have observed a shift towards computer-aided diagnosis (CAD) based on behavioural data such as text, speech and handwriting. Despite contributions to the handwriting-based detection of PD, most of the literature has neglected the implications of disease severity. I use a PaHaW data set from Czechia, containing handwriting data from 75 participants, to identify features used to detect PD individuals in the early stages based on the UPDRS scale. I utilise the machine learning (ML) models Logistic Regression, Random Forest, SVC and XGBoost as binary classifiers to discriminate between PD and HC based on handwriting features. Results show that velocity parameters, pressure and number of changes in pressure constitute a diagnostic criterion for early detection of PD. In a longitudinal analysis of the economic implications of PD (Study 2), I estimate its economic burden in Czechia between 1996 and 2018, by utilising data from a variety of sources, including the IHME GBD database, Our World in Data and the Czech Statistical Office. Results show that PD prevalence among the working-age population represents a considerable economic burden, as productivity loss due to PD accounted for more than 0.02% of GDP in 2015. Those findings illustrate that handwriting-based CAD, statistics and ML are important tools for complementing clinical PD diagnostics to detect the disease earlier and reduce the economic burden.

AUTHOR'S DECLARATION

I, the undersigned, Sunčica Rosić, candidate for the BA/BSc degree in Data Science and Society, declare herewith that the present thesis titled "Early Diagnosis of Parkinson's Disease through Machine Learning Analysis of Handwriting based on the PaHaW Dataset. Quantifying the economic burden of Parkinson's Disease in Czechia." is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright. I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 26th May 2025

Sunčica Rosić

CONTENTS

List of Figures

List of Tables

Introduction

This thesis examines the use of handwriting in the early detection of Parkinson's Disease (PD) and assesses the economic burden of the illness in Czechia. The first study analyses handwriting samples from the PaHaW dataset, focusing on participants in the early stages of PD as defined by the Unified Parkinson's Disease Rating Scale (UPDRS) (Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, 2003). Through feature extraction and classification using several machine learning (ML) models, results show the differential potential of velocity and pressure. The second study uses publicly available sources—including IHME, Our World in Data, and the Czech Statistical Office—to estimate the productivity loss due to PD among the working-age population between 1996 and 2018. Findings indicate that this loss ranges from approximately 0.02 % to above 0.04%, depending on the time period. Together, these studies demonstrate how statistical analysis and behavioural data can support early diagnosis and underline the broader socioeconomic implications of PD in Czechia. The proposed methodology makes a novel contribution to the field of CAD detection of PD and economic burden quantification by focusing on the early-stage PD identification and productivity loss in Czechia, respectively.

Neurological perspective

PD is the second most recurrent neurodegenerative disorder, which remains incurable (Santos et al., 2022). It is characterised by motor symptoms such as akinesia (loss of spontaneous, voluntary muscle movement), bradykinesia (slowness of movement), rigidity and tremors (Drotár et al., 2016). These motor disorders are reflected in handwriting, which has motivated the research in PD dysgraphia – PD motor deficits specific to handwriting (Letanneux et al., 2014). Frequent characteristics of people with PD are impaired hand function caused by

deterioration of hand dexterity (Wong-Yu et al., 2021). Poorer manual skill, reduced grip strength and lower self-perceived functional hand ability compared to controls are examples where PD showcased motor impairments due to brain dopamine deficiency (Błaszczyk, 1998). Handwriting represents an important and valuable activity, conditional on complex sensorimotor, perceptual and cognitive skills (Bonney, 2010). If one of these skills diminishes, handwriting may degrade, which is common in neurological conditions like PD (Drotár et al., 2016) and multiple sclerosis (Bisio et al., 2017).

Since PD consists of such a combination of symptoms and cannot be termed as "a single nosological entity", diagnosing PD through a single metric remains difficult (Calne, Snow, and Lee, 1992). There are no laboratory tests that represent absolute indicators of PD (Calne, Snow, and Lee 1992), which leaves neurologists with physical and neurological exams to identify the disease. Although PD is progressive, the progression is rather slow, which might result in a failure to diagnose it early. Neurologists differentiate between different severities of PD based on the UPDRS, used to guide researchers in establishing novel PD detection methods at its prodromal stage. PD persists to be a major public health challenge and no novel diagnosis ruling out existing challenges has been implemented. This emphasises the urge for novel approaches to an early diagnosis, as it can improve the odds of intervention and diagnose the disease at the prodromal stages.

In recent years, the diagnostic approaches of PD have undergone a notable shift—from an early reliance on neuroimaging techniques to the growing adoption of ML that analyses behavioural signals. Santos et al. (2022) studied how certain natural peptides, like LL-37, reduce the harmful effects of protein aggregates linked to PD. Numerous authors explored imaging approaches to PD diagnosis, such as structural and functional imaging techniques that distinguish PD from atypical variants by detecting dopamine deficiency (Brooks, 2010).

Although existing techniques can mitigate early motor system symptoms, they are ineffective as the disease progresses, which stresses the need for "disease modifying agents" that interfere with PD progression (Santos et al., 2022). A growing body of research has been shifting towards ML methods of identifying behavioural indicators of PD, manifested through text, speech and handwriting (Islam et al., 2024). Drotár et al. (2016) showcase the motor symptoms associated with PD, including their manifestation in handwriting, which reflects the deterioration of motor abilities. Letanneux et al. (2014) and Wong-Yu et al. (2021) both discuss the connection between motor deficits in PD and the observable changes in handwriting, such as micrographia, a decrease in the letter size of handwritten text (Larner, 2016).

Although handwriting impairments are not officially part of the diagnostic criteria for PD (Chahine & Stern, 2011), they might indicate that the individual requires further medical attention. There has been growing interest in new objective biomarkers aimed at enabling earlier diagnosis and better monitoring of PD progression (Wu et al., 2011). Handwriting analysis has emerged as one such promising tool, offering an efficient, non-invasive and attenable way to signify early signs of PD (Phillips et al., 1991). As a complex fine motor skill demanding precision and coordination, handwriting is especially susceptible to disruption in individuals with PD, as PD micrographia affects letter size (Margolin & Wing, 1983).

Despite these advancements in understanding PD's neurological symptoms, early detection remains a challenge. Clinical features may correspond to those of other neurodegenerative diseases, and tests or biomarkers do not provide a deterministic diagnosis from the earliest stages (Tolosa et al., 2021). There are no definitive laboratory tests for PD, making it reliant on clinical exams with suboptimal results. Neurodegenerative processes normally start decades before the occurrence of clinical symptoms. As a result, diagnosis becomes feasible only after substantial dopaminergic neuronal loss has occurred,

making current treatments primarily symptomatic rather than disease-modifying (Leggio et al., 2017). This has created a need for alternative methods to detect PD early, aiming to close the gap between the beginning of neurodegenerative processes and symptomatic treatment. Recent efforts have explored dysgraphia-based methods, including handwriting analysis to showcase sensorimotor impairments pertinent to early detection (Bisio et al., 2017).

The present study aims to investigate whether handwriting features can support the classification of PD at early stages of disease progression based on UPDRS. I utilise the PaHaW dataset (Drotár et al., 2016) with handwriting samples from PD and healthy controls (HC), each group including 37 and 38 participants respectively and containing metadata with differing disease severity. Participants were asked to complete eight handwriting tasks, including drawing a simple Archimedean spiral, through a digital tablet. The tablet acquired data for each task across seven variables: x- and y-coordinates, time (sampled at 150 Hz in milliseconds), stylus azimuth and altitude, pressure, and contact with the tablet surface. The proposed methodology pertains to the extraction of features from handwriting tasks and statistical feature selection. Several ML models are used: Logistic Regression, Random Forest, Support Vector Classification (SVC) and XG Boost, followed by an evaluation of their discriminative capacity. The results show that the proposed methodology provides a feasible technique to differentiate between HC and early-stage PD, given the accuracy of 68% and recall of 73%.

Whilst existing literature has focused on diagnosing PD through motor symptoms and handwriting analysis without accounting for the disease stage, this thesis contributes by focusing on handwriting-based PD differentiation through ML for patients at the prodromal stage of PD. By using data from Czechia, this thesis provides a more holistic understanding of the limitations of current clinical methods, emphasising the need for early diagnosis and pre-screening. I argue that the proposed method contributes to an effective tool that aims to mitigate

the risk of delayed diagnosis and complement existing clinical methods.

Economic perspective

PD represents a public health challenge complemented with socio-economic consequences (K. Ray Chaudhuri et al., 2024), including medical costs, affected quality of life (Europe PMC, 2016) and caregiver burden (Rosqvist et al., 2022). The progressive nature of the disease results in a gradual decline in patients' life quality, leading to an increase in medical and societal demands. Aside from direct medical costs, such as hospitalisations, medications and outpatient care, a significant economic impact stems from indirect costs (Yang et al., 2020), including early retirement, reduced workforce participation, lost productivity and informal caregiving. These losses become especially apparent in working-age patients forced to exit the labour market prematurely.

Several authors quantified the economic and indirect PD burden on individual and social levels in the US (Yang et al., 2020) and Western and Northern Europe (WNE) (Andlin-Sobocki et al., 2005). Yang et al. (2017) estimated that the total PD economic burden was $51.9 billion, and it is projected that by 2037, PD will exceed the total economic burden of $79 billion in the US. This projection equates the entire 2023 budget of the U.S. Department of Housing and Urban Development to 71.9 billion (U.S. DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT, 2023). The total annual cost of brain disorders in Europe was €798 billion in 2010, with PD accounting for €13.9 billion of this amount, almost 2% of the annual cost of all brain disorders (Olesen et al., 2011), exceeding the amount dedicated to the European Social Fund for Europe's human capital. This shows that PD is not only a growing public health challenge but also an economic burden straining national budgets in both the US and Europe.

Despite efforts to estimate the economic cost of PD in the US and WNE, few efforts quantified the PD burden in Central and Eastern Europe (CEE). Andlin-Sobocki et al. (2005) observed that available PD cost data failed to provide cost information for CEE. Although a few studies (Winter et al., 2010, Sonja von Campenhausen et al., 2011) identified cost patterns associated with PD in CEE, this region's contributions are below those from the US and WNE. It was estimated that, although per-patient costs due to PD are lower in Czechia compared to WNE, "the proportion of costs that fall on patients is higher because of lower incomes" (Winter et al., 2010). Despite notable contributions to identifying the PD cost burden in Czechia, little work has been done to establish total productivity and salary loss.

I utilise prevalence and labour productivity combined aggregate data to estimate TPL. I calculate lost productivity per person by subtracting the retirement age of PD patients from the average national retirement age and multiplying the difference by annual working hours and output per hour worked, where PD retirement age and working hours vary to account for differences in labour market exit and hours spent at work. This number is multiplied by the annual estimate of PD working-age patients in Czechia, which represents the total annual loss. Salary loss is obtained by multiplying the average monthly age for full-time employees by the number of PD working-age patients. I compare the obtained estimates to other public spending sectors as a percentage of GDP.

This thesis contributes to a more comprehensive quantification of the PD burden on society and the imposed national budget constraints, emphasising the need for effective early diagnosis and pre-screening to delay premature retirement. Whilst existing economic burden literature has focused on quantifying medical, outpatient and caregiver burden in the US and WNE, this thesis contributes by complementing handwriting-based diagnostic tools with a quantitative estimation of productivity and economic loss due to PD in the CEE. By

using aggregate data from Czechia, this research sheds light on the broader economic consequences, bridging the gap between clinical and economic aspects of PD. This thesis aims to explore the multifaceted nature of PD, including the diagnostic and economic dimensions of PD.

Background Literature

Computational Neurology Literature

Neurodegenerative disorders progressively alter the structure and function of the brain, leading to declines in cognitive, behavioural, and functional capacities. Among these conditions, PD is one of the most prevalent and disabling, characterised by motor symptoms such as bradykinesia, akinesia, tremor and rigidity (Kalia & Lang, 2015), as well as non-motor features including depression, sleep disturbances and cognitive impairment (Li et al., 2021). At present, there is no definitive cure for PD, and accurate diagnosis remains difficult, with recent studies identifying diagnostic accuracy to be only 76% (Rajput et al., 1991). This emphasises a need for an early diagnosis in order to enable timely medical intervention and prolong workforce survival.

In this context, there is growing interest in the development of CAD systems within the broader field of e-health (Hirschauer et al., 2015). Such systems have the potential to support clinical decision-making by providing non-invasive, cost-effective and complementary diagnostic tools. Given that handwriting impairments in PD have been well-documented, handwriting analysis emerged as a promising digital biomarker (Impedovo et al., 2018). The act of handwriting involves complex motor coordination, visuospatial processing, and cognitive engagement, making it a sensitive indicator of neuromotor decline.

Several studies have demonstrated the utility of ML methods in distinguishing PD patients from HC based on kinematic features extracted from handwriting tasks. These features, derived from time-series data captured via digital tablets, include pen trajectory, pressure, tilt and in-air movements. A model based on Bidirectional Gated Recurrent Units (BiGRU) was employed to evaluate the predictive potential of handwriting-based sequential properties in the detection of Parkinsonian symptoms (Momina Moetesum et al., 2020). Some authors have leveraged the hybrid potential of static and dynamic handwriting analysis by creating a dynamically enhanced static handwriting method to retain temporal features (Diaz et al., 2019).

With the exception of one paper, which relied on the generation of images based on the PaHaW data set (Impedovo et al., 2018b), most studies have focused solely on binary classification of HC versus PD subjects, without accounting for heterogeneity in disease severity. This step is crucial to test and evaluate whether the proposed CAD method is also effective in the prodromal stages of the disease, to slow down the progression of the disease and prevent patients from premature labour market exit.

The development of graphic tablet technology has advanced the study of handwriting kinematics, referring to the measurement of the physical movements involved in handwriting, in individuals with PD. Unlike traditional pen-and-paper assessments (Alty et al., 2017), digital tools enable the precise measurement of temporal and spatial handwriting dynamics, including parameters such as pressure, velocity and stroke duration. This technological property has allowed for the extraction of quantitative features that were previously inaccessible, making an advancement in sample acquisition and methodological analysis within the field.

In light of these advancements, numerous graphomotor tasks have been proposed to capture motor impairments associated with PD, such as tremors, dysgraphia (Letanneux et al.,

2014) and bradykinesia. One of the most frequent diagnostic tasks is the drawing of an Archimedean spiral (Infocommunications Journal, 2024), a clinically recognised assessment that has captured significant attention in CAD research. Other geometric patterns, including circles and meanders (Pereira et al., 2016), have similarly been used to distinguish between the handwriting of HC and PD.

Beyond drawing-based tasks, researchers have explored handwriting exercises involving cursive letters, repeated character sequences (e.g., 'lll' or loops) and the writing of participants' names and addresses (Rosenblum et al., 2013). These tasks have been effective in identifying hallmark PD symptoms such as tremor and reduced letter size, as well as broader manifestations like dysgraphia, referring to "loss of previously intact writing function" (Gubbay & de Klerk, 1995) and agraphia, a combination of neurological classifications of writing impairments (Sitek et al., 2012). In some studies (Drotár et al., 2016), pauses or in-air times between words have also been analysed as indicators of cognitive decline associated with neurodegenerative conditions, including PD. These handwriting-based tasks provide a low-cost, non-invasive foundation for the data acquisition and differential feature selection that leverage PD dysgraphia to detect fine motor impairments in PD patients.

The adoption of digital handwriting analysis has further enabled the extraction of novel features such as pen pressure, entropy and stroke sequencing (Senatore & Marcelli, 2019), enriching the diagnostic potential of handwriting-based assessments. Over the past decade, these developments have positioned online handwriting analysis as a valuable and non-invasive tool for both the diagnosis and monitoring of PD.

To date, few authors have addressed the importance of analysing the handwriting of patients at the early stages of PD, as in the later disease stages, the proposed CAD handwriting method is likely to be replaced by clinical methods. Although a vast body of literature has

addressed the application of statistical methods for identifying relevant handwriting features, few studies have leveraged the potential of ML in testing the predictive potential of these features. As an exception, Impedovo et al. (2018b) incorporated both early handwriting-based PD detection, defined by the UPDRS scale, and ML models such as CNN. However, the usage of CNN was conditional on generating images from tabular handwriting data. The method proposed in this thesis, on the other hand, removes this extra step of image generation by feeding the models with the handwriting features extracted from coordinates from the PaHaW data set.

Economic Literature

PD is one of the most rapidly growing neurodegenerative disorders globally, accounting for the biggest increase in disability-adjusted life year (DALY) among neurological disorders, equal to a 61% rise between 1990 and 2017 (Deuschl et al., 2020). Between 1990 and 2016, the number of individuals with PD exceeded 6 million cases and resulted in over 200,000 PD-related deaths worldwide (Feigin et al., 2019). During this period, the global age-standardised prevalence increased by 22%, and projections suggest that by 2040, the number of people affected could reach between 12 and 17 million (Dorsey et al., 2018). This rise in Parkinson's prevalence has been attributed to several factors, including an ageing population, increased disease duration, potential shifts in environmental or social risk factors, and improved recognition and diagnosis within the routine medical practice (Aarsland et al., 2003).

Alongside motor symptoms, PD is associated with non-motor manifestations. These involve dysfunction of the gastrointestinal and autonomic nervous systems, as well as disturbances in sleep, mood, and cognitive function, all of which significantly reduce quality of life and contribute to disability (Schapira et al., 2017).

As PD advances, individuals often become increasingly dependent on care, much of which is provided by informal caregivers. The burden placed on caregivers has been linked to several disease-related factors, including the duration of illness and caregiving and the presence of non-motor symptoms such as anxiety, depression and sleep disturbances. This burden may lead to personal and occupational challenges, including reduced participation in family, social and leisure activities, sleep disruption and mental health difficulties (Zhong et al., 2015).

One study employing a human capital approach to estimate the future economic impact of PD in the US found that the most indirect costs stemmed from lost income due to premature morbidity and mortality, as well as reduced employment, absenteeism and early retirement (Yang et al., 2020). Research involving Medicare beneficiaries with PD has shown that individuals with advanced disease stages face a markedly higher economic burden compared to those with milder forms of the condition (Dahodwala et al., 2020). These findings stress the importance of understanding the incremental economic burden of PD across varying levels of disease severity and within different healthcare systems. Despite these insights, the economic burden of PD in CEE remains inadequately defined, particularly concerning the working years lost due to PD as the disease progresses.

The Economic Impact of Early Diagnosis in Parkinson's Disease

The burden imposed by PD can be observed through the perspective of national, regional and global figures, both in terms of DALY and death rates and economic impact. PD burden should not be observed in a vacuum, as delayed clinical diagnosis may hinder the odds of slowing down disease progression, thus affecting an individual's workforce survival. Besides this burden stemming from restricted labour market participation, PD is also associated with medical costs and caregiver burden (Rosqvist et al., 2022). Additionally, the global burden of PD has more than doubled compared to 1990, not only due to an increasing ageing population, emphasising the potential moderating effect of other factors (Dorsey, Elbaz, et al., 2018). This stresses the need for suitable medical, economic and scientific responses, indicating that cross-sector collaboration, like the one that extends traditional clinical methods through computational approaches, may be required. Although PD is incurable, data-driven early detection could assist clinicians in patient pre-screening and reduce the odds of late self-reported symptoms to neurologists, possibly increasing the likelihood of longer workforce survival and saved productive hours as result from treatment and medication. Thus, early PD diagnosis has the potential to reduce long-term economic costs through enabling earlier intervention, delaying disease progression and reducing complications and caregiver burden.

Data, Methods and Results

Handwriting Perspective

PaHaW Data Set

The main dataset used in this study is the Parkinson's Disease Handwriting Database (PaHaW), developed as part of the work by Drotár et al. (2014). It contains handwriting data from 37 individuals diagnosed with PD (19 men and 18 women) and 38 age- and gender-matched HC (20 men and 18 women). Each participant completed eight handwriting tasks through a digital Wacom tablet. As shown in Figure 1, these tasks range in complexity, from drawing a simple Archimedean spiral to writing a full sentence in Czech Latin cursive. The tablet recorded data for each task across seven variables: x- and y-coordinates, time (sampled at 150 Hz in milliseconds), stylus azimuth and altitude, pressure, and button state (contact with the tablet surface).

The data is stored in SVC files, supplemented by separate metadata tables for each participant. These metadata files include nationality, age, gender and PD progression, measured using Part 5 of UPDRS V. For this study, both the handwriting data and metadata were preprocessed and combined into a single data frame. Whilst it is also of clinical interest to be able to identify patients at all disease stages, the focus of this research is on leveraging statistical methods to identify patients at prodromal stages of PD, as this might help slow down the progression. My analytical sample considers 33 patients whose score on the UPDRS scale is lower than 3 and 36 HC.

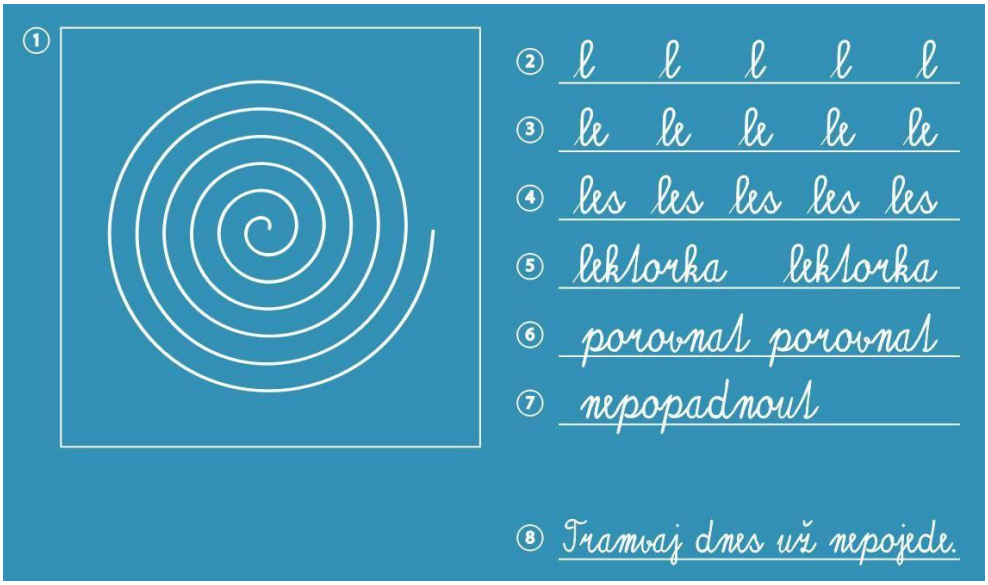

Figure 1 PaHaW writing task template - Figure reproduced from (Drotár et al., 2016), with permission from Artificial Intelligence in Medicine.

Statistical and Machine Learning Analysis of Handwriting

Feature extraction

The handwriting features were derived from on-surface movements (in the form of Cartesian coordinates) and pressure. While features like X and Y coordinates, time stamp, button state, azimuth, altitude and pressure were originally available in the data, other features were quantitatively derived. Note that pressure was adjusted dynamically depending on the threshold obtained based on a 10th percentile, such that pressure changes per stroke are based on its own pressure distribution. Only points where pressure is above the dynamic threshold are kept for velocity computation and feature analysis. Out of these 7 features, I decided to utilise x and y coordinates and pressure for further analysis, as the available literature suggested that pressure and kinematic features successfully discriminated between HC and PD (Drotár et al., 2016). The kinematic features used in this study are listed in Table 1.

Table 1Kinematic features

Maximum velocity	Highest velocity during handwriting.
Minimum velocity	Lowest velocity during handwriting.
Maximum pressure	Highest pressure applied during handwriting.

Minimum pressure	Lowest pressure applied during handwriting.
Stroke duration	Time duration of each stroke during handwriting.
NCV	Number of changes in velocity.
NCP	The number of changes in pressure.
Mean velocity	Mean velocity during handwriting.
Velocity standard deviation	Standard deviation of velocity during handwriting.
Mean pressure	Mean pressure applied during handwriting.
Pressure standard deviation	Standard deviation of pressure applied during handwriting.

Statistical analysis

Table 2 Left: Significant features before dropping UPDRS > 3 rows, Middle: Significant features after dropping UPDRS > 3 rows, Right: Significant features after Bonferroni correction

Whole Data Set			**Filtered Data Set based on UPDRS <3**			**After Bonferroni Correction**		
feature	**task**	**p-value**	**feature**	**task**	**p-value**	**feature**	**task**	**p-value**
maximum pressure	1	0.052	NCV	1	0.070	maximum velocity	2	0.000
NCV	1	0.068	NCP	1	0.099	velocity standard deviation	3	0.008
maximum velocity	2	0.013	NCP	2	0.094	velocity standard deviation	8	0.002
NCP	2	0.02	maximum velocity	2	0.000	maximum velocity	8	0.006
maximum pressure	3	0.051	velocity standard deviation	2	0.033	mean velocity	8	0.006
NCP	3	0.029	NCP	3	0.042			
NCP	4	0.096	maximum velocity	3	0.038			
pressure standard deviation	4	0.077	velocity standard deviation	3	0.008			
maximum pressure	5	0.035	maximum velocity	4	0.058			
pressure standard deviation	5	0.042	velocity standard deviation	4	0.063			
NCP	6	0.07	maximum velocity	5	0.060			

maximum pressure	6	0.044	velocity standard deviation	5	0.083			
velocity standard deviation	7	0.032	velocity standard deviation	7	0.041			
mean pressure	7	0.044	maximum velocity	8	0.006			
maximum velocity	8	0.03	mean velocity	8	0.006			
NCP	8	0.081	velocity standard deviation	8	0.002			
mean velocity	8	0.01						
velocity standard deviation	8	0.003						

Statistical hypothesis testing was implemented to find significant features which discriminate between HC and PD. I employ Mann Whitney U-Test, a non-parametric statistical test to compare two groups, allowing me to find whether the groups are drawn from populations with different levels of a variable of interest. I look for statistical differences between features associated with PD and HC, respectively. I analyse 9 features, as minimums were excluded for each task for all 8 tasks.

I obtain a list of significant features for data before and after selecting the data set based on UPDRS criteria, to grasp which features are meaningful in differentiating between PD and HC in early disease stages, and which ones are only meaningful once the disease has progressed significantly. In Table 2 in the middle, descriptive statistics features extracted from velocity populate the list, with a small exception of NCP for tasks 2 and 3. On the left-hand side, before filtering out rows where UPDRS is at most 3, I observe the following: 1) more features become

significant – the number of significant features increases from 14 to 18 – 2) fewer velocity parameters are statistically significant, and pressure becomes more prominent. Additionally, I include Bonferroni correction to compensate for the increase in the odds of incorrectly rejecting a null hypothesis and thus increasing Type I Error. I use the Bonferroni correction at an alpha level of 0.1 and 9 hypothesis tests, given 9 features, as I compare HC and PD groups while fixing the feature and a task. This step has narrowed down the list of final selection of features, which are represented in column 3 of Table 2. Although these features represent attributes based on which one can discriminate between healthy individuals and with Parkinson's, results from this table provide no information on the direction or intensity of the group-specific feature. Thus, I obtain box plots to gain a more informative and contextual image of the significant features (Figure 2).

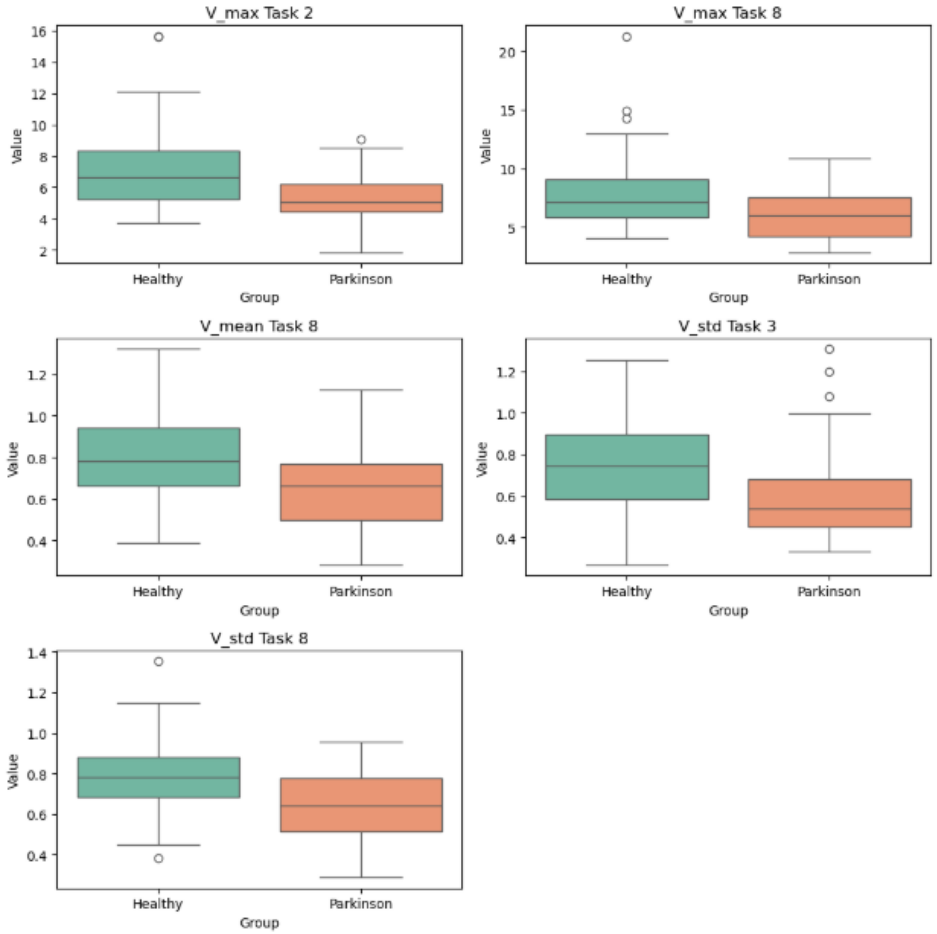

Figure 2 Box plots for features filtered after Bonferroni Correction

The individual with PD requires more time to complete a task, as shown by a smaller V_max for tasks 2 and 8. This is consistent with V_mean being lower for PD in task 8. Velocity standard deviation also appears as a significant feature in tasks 3 and 8. This leads to an assumption that within the PaHaW dataset, after applying Bonferroni correction, task 3 (cursive writing of 'le') and 8 (sentence) where velocity is contained, seem to be the most significant in distinguishing between PD and HC.

Predictive Model Design

I employ four models with implemented cross-validation to compare different models' performance and avoid potential overfitting associated with a train-test split. The models used are the following: Logistic Regression, Random Forest, SVC and XGBoost. While the first three models are trained on selected features after Bonferroni correction, XGBoost takes all features to leverage its tree boosting system.

All models are optimised thanks to `GridSearchCV`, which allows finding the optimal hyperparameters for models by exhaustively searching through a predefined set of hyperparameter values. After parameter fine-tuning best model is automatically chosen through a built-in function and then applied to the input data to generate predictions.

Given the model's objective to pre-screen participants, I specifically optimise it to maximise the recall of PD patients when using XGBoost, to ensure that all actual PD participants are classified as such, at the expense of falsely flagging HC. Recall is maximised by lowering the threshold, which increases the odds of the classifier assigning a predicted label to PD class. This is reasonable because such a model's use case is not its use as a diagnostic tool as such. Rather, it would serve as a scalable tool to pre-screen population-size handwriting data and flag those who might be susceptible to suffering from PD.

Predictive model results

The tree models are evaluated based on accuracy, precision, recall, F1 score and ROC AUC. The evaluation metrics provide rather mediocre results, with all three models scoring below 70 % of accuracy. On the contrary, recall exceeds 70 % for Logistic Regression, meaning that the model have over 70% chance of correctly identifying positive instances (PD). This can be also seen from confusion matrices, where Logistic Regression and SVC, respectively, identify 5 and 4 more TPs compared to Random Forest.

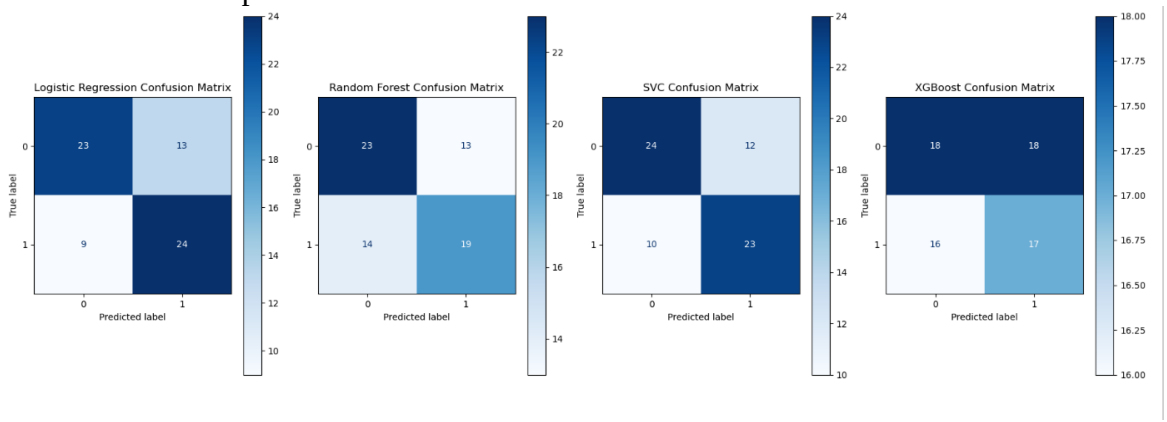

Figure 3 Confusion matrices

The aforementioned three models perform better than XGBoost, whose accuracy is 58% and 57% for default threshold and threshold of 0.1, respectively. A threshold is a value used to convert probability predictions into binary class decisions. If the predicted probability is greater than or equal to the threshold, the sample is classified as the positive class; otherwise, it is classified as the negative class. The default threshold is 0.5 (50%). Adjusting the threshold allows a trade-off between precision and recall. Lowering the threshold to 0.1 makes the model more likely to predict the positive class, increasing recall but decreasing precision, which is reflected in higher recall in Table 5 (88%) but lower precision (53%) for the target class. Thus, the classifier achieves maximising the recall of PD participants, when limited so that the PD recall is the highest at a threshold of 0.1. As visible in Tables 4 and 5, the accuracy is just 57% with threshold of 0.1 and a slightly higher 58% otherwise. This means that just 50-60 percent

of participants flagged as PD sufferers are actual PD sufferers.

Given the fact that XGBoost, as a tree-based model, does not undergo statistical feature selection, I show feature importance as part of this model's implicit feature selection process. In Figure 4, not only is NCV for tasks 6 and 4 the best predictor, but also the most frequent, since NCV combined with other tasks takes 5 out of 10 most important features. At the same time, the difference compared to other features is not as stark, meaning that there is useful and extractable information in derived handwriting features alone.

Table 3 Evaluation metrics for Logistic Regression, Random Forest and SVC

Model	**Logistic Regression**	**Random Forest**	**SVC**
Accuracy	0.681	0.609	0.681
Precision	0.649	0.594	0.657
Recall	0.727	0.576	0.697
F1 Score	0.686	0.585	0.677
ROC AUC	0.757	0.658	0.708

Table 4 Classification report of the classifier, threshold = 0.5

Class	**Precision**	**Recall**	**F1-Score**
Healthy	0.61	0.56	0.58
Diseased	0.56	0.61	0.58
Accuracy			**0.58**
Macro Avg	0.58	0.58	0.58
Weighted Avg	0.58	0.58	0.58

Table 5 Classification report of the classifier, threshold = 0.1

Class	**Precision**	**Recall**	**F1-Score**
Healthy	0.71	0.28	0.40
Diseased	0.53	0.88	0.66
Accuracy			0.57
Macro Avg	0.62	0.58	0.53
Weighted Avg	0.62	0.57	0.52

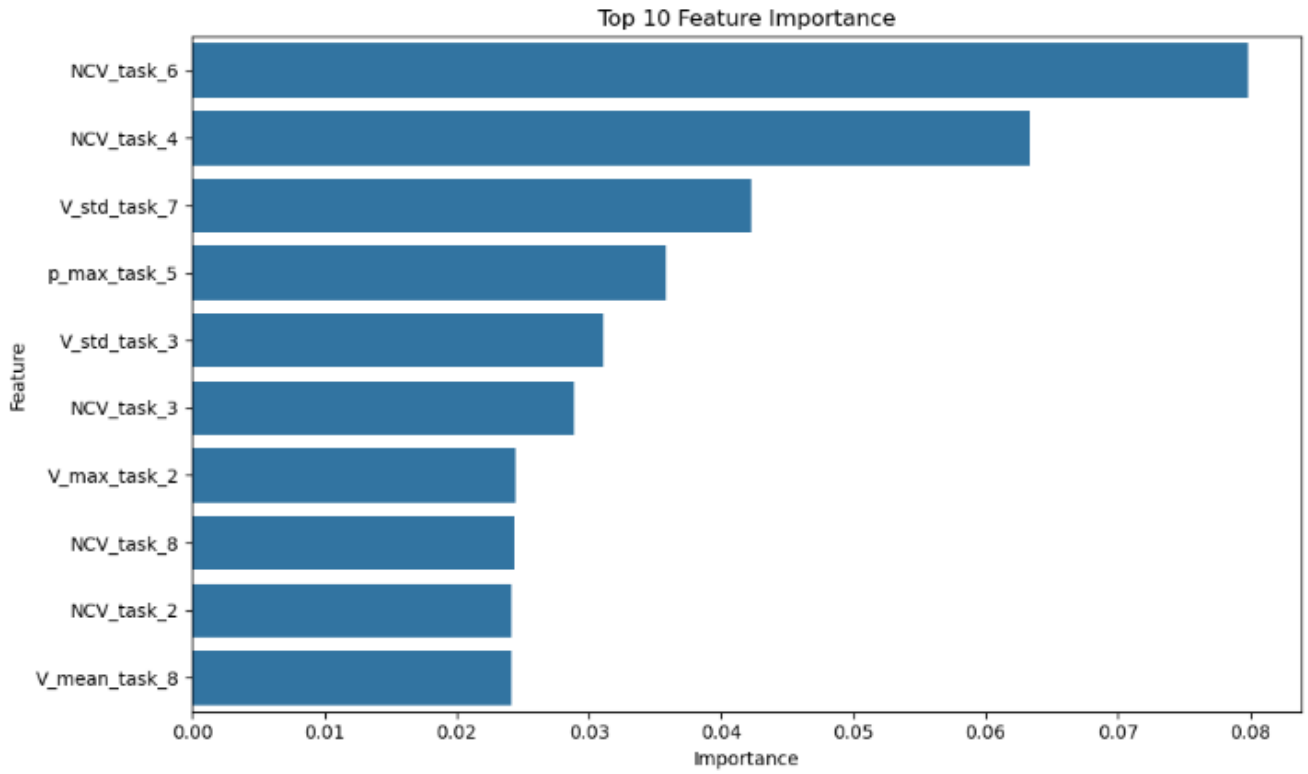

Figure 4 Feature importance for XGBoost

Economic Perspective

Aggregate Economic Data

This section outlines the data sources and variables used to estimate the economic burden associated with PD in Czechia, with a focus on productivity losses associated with the disease. The data was obtained for the years 1996 – 2018.

Prevalence of Parkinson's Disease

Data on the prevalence of PD were obtained from the Global Burden of Disease (GBD) database maintained by the Institute for Health Metrics and Evaluation (IHME) (Institute for Health Metrics and Evaluation, 2021). The GBD database provides age-standardised prevalence rates by country, gender and year, allowing for time-series analysis and demographic breakdowns.

Labor Productivity

Labour productivity, defined as the gross domestic product (GDP) per hour worked in constant 2017 international dollars (PPP-adjusted), was sourced from Our World in Data using data from the Penn World Table (Feenstra, R. C., Inklaar, R. and Timmer, MP, 2015). This metric serves as a proxy for the economic contribution of an average working hour and is used to estimate the value of lost productivity due to early retirement or labour force withdrawal among PD patients.

Average Wage

To obtain productivity estimates and approximate potential wage losses, average gross monthly wages in Czechia (reported in CZK) were collected from the Czech Statistical Office (Czech Statistical Office, 2024).

Currency Conversion

To convert wage data from Czech crowns (CZK) to U.S. dollars (USD), historical annual average exchange rates were retrieved from the Federal Reserve Economic Data (FRED) database (Federal Reserve Bank of St. Louis, 2025). This ensures consistency in monetary units across productivity and wage-related estimates.

Government Spending and Retirement Age

The average effective retirement age for men and women was sourced from Our World in Data, based on OECD statistics (Our World in Data, 2024). These figures are used to determine the number of productive years lost per PD patient, under the assumption that individuals would have remained in the labour force until the average retirement age in the absence of illness. I obtained government spending data per sector in Czechia from the OECD data repository (OECD, 2024).

Experimental Design

To estimate the total productivity loss (TPL) associated with PD in Czechia, I combine data from multiple sources, including PD prevalence estimates from the GBD database, labour productivity and retirement age data, wage and employment statistics and historical exchange rates. Prevalence pertains to the number of the Czech population found to be affected by PD (Sonja von Campenhausen et al., 2011).

I estimate the **individual productivity loss** associated with early labour market exit due to PD. This is computed by subtracting the estimated average retirement age of PD patients from the national average retirement age. This difference in productive years is then multiplied by the average number of annual working hours and output per hour worked (measured in constant 2017 USD, PPP-adjusted) to obtain an estimate of productivity lost per individual. To account for heterogeneity in labour force engagement, the calculations allow for variation in both working hours and retirement ages. This allows me to obtain a range of personal productivity loss estimates given in a specific year, rather than a fixed value. Lastly, I multiply the personal productivity loss by the total number of PD working-age patients. For future calculations, I take the minimum and maximum estimates of total annual productivity loss. I derive **wage-based income loss** as a complementary estimate to productivity loss. Finally, I contextualise the estimated economic burden by comparing minimum and maximum estimates of TPL with other public expenditure categories expressed as a share of national GDP, such as defence, housing and environmental protection. This comparative lens allows for an assessment of the relative scale of the economic impact of PD on national productivity.

Economic Results

I assessed the economic burden of PD in Czechia through productivity loss due to early labour market exit among working-age patients. The TPL due to PD (Figure 5) increased substantially between 1995 and 2018. The estimated **minimum annual loss** rose from approximately USD 26 million in 1995 to over USD 45 million in 2017, while the **maximum loss** increased from USD 32 million to over USD 54 million.

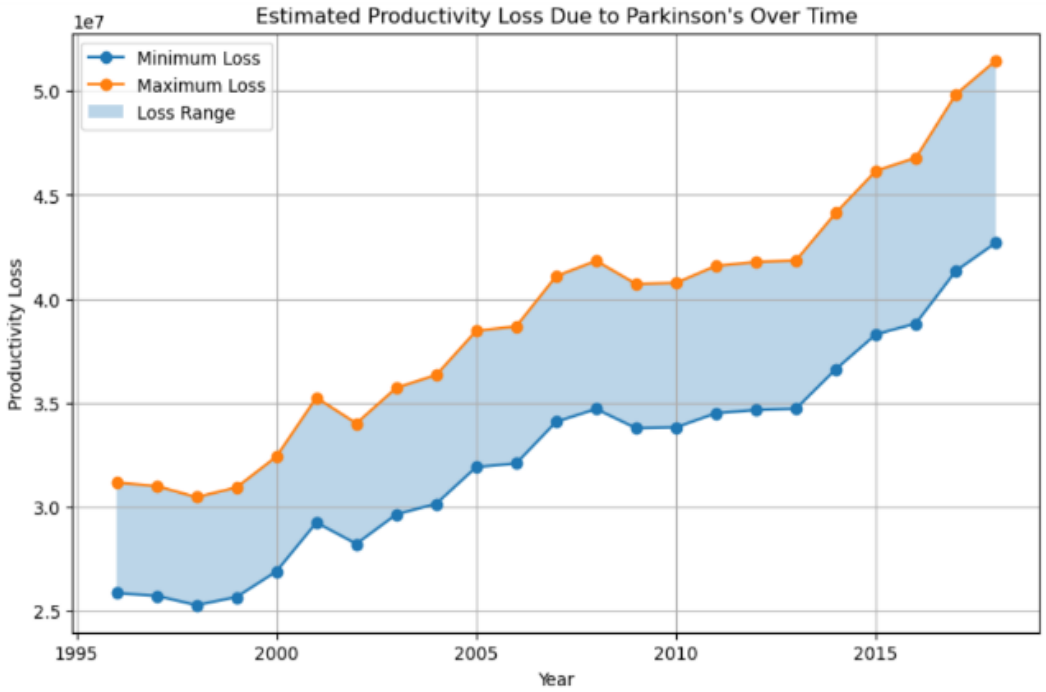

Figure 5: Estimated Productivity Loss due to PD over time

I also expressed the burden of productivity loss relative to national GDP (Figure 6). The share peaked in the early 2000s, reaching over 0.04% of GDP, before reaching a consistent decline, with below 0.02% in 2015. This downward trend can be explained by an assumption that GDP rose at a higher pace than PD prevalence, thus exerting downward pressure on PD productivity loss as a percentage of GDP.

Figure 6: Estimated Productivity Loss due to PD as % of GDP over time

To understand the relative size of PD productivity loss and its potential constraint on the national budget (Figure 7), I compare the economic impact of PD-related productivity loss with public expenditures in key sectors – environmental protection, housing and community amenities spending and defence. Productivity loss as % of GDP due to PD is considerably lower than any of the sectors shown in the graph, indicating a need for more comparable public spending data that requires fewer government resources. Still, one can observe that the TPL as % of GDP takes approximately less than one promil of the government spending on respective sectors.

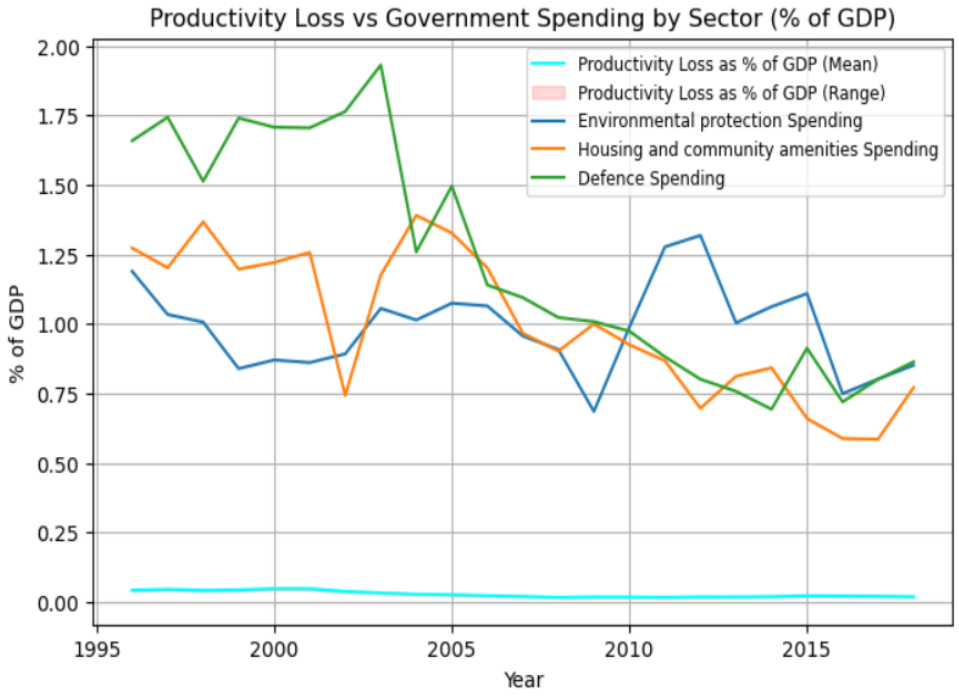

Figure 7: Estimated productivity loss due to PD as % of GDP compared to government spending on other sectors

I plot total personal income loss due to PD to understand total foregone earnings as a result of the premature labour market exit due to PD. Figure 8 shows that total personal income loss due to PD was at 2 million USD in 1995, after which it increased at a steady pace until around 2009. Although after this period, personal income loss was oscillating, potentially reflecting smaller economic shocks of recession and growth, in the long term, when these fluctuations are ironed out, it shows an increasing trend in total earnings loss.

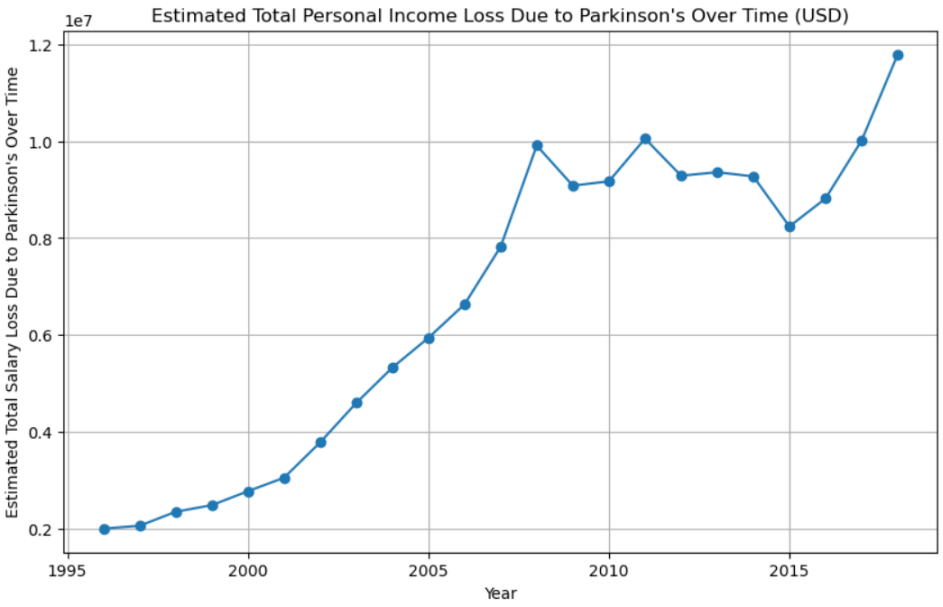

Figure 8: Total personal income loss due to PD over time

Discussion

Handwriting Perspective

PD is an extremely complex disorder with a wide range of symptoms that might vary across patients. Although PD is not comprised of a single nosological entity, a combination of symptoms such as tremors, rigidity and slowness of movement are common indicators of such neurodegenerative disease. Alongside traditional clinical methods, aforementioned motor symptoms can be reflected through behavioural properties, such as disturbance in hand motor movements. Given that handwriting is a sophisticated motor task requiring muscle coordination, it has the potential as a "biomarker" indicating early signs of PD. Although the precise correlation between certain parameters and PD symptoms is unknown,

this study's findings indicate that velocity and features derived from velocity, pressure and NCV can help discriminate between PD and controls. From a clinical perspective, kinematic properties are influenced by a variety of clinical factors, including tremors, muscle rigidity, stiffness and variation in movement speed.

In contrast, the pressure characteristics can offer extra details that the kinematic features are unable to record. The features that represent a differential power to diagnose patients with PD involve those extracted from velocity and pressure, as well as NCV. PD patients have lower velocity features (v_max, v_mean, and v_std) compared to HC. Lower mean and maximum velocity indicate slowness of movement characteristics for PD individuals. This is consistent with previous research (Nackaerts et al., 2017) that showed that individuals with PD exhibited slower handwriting compared to controls. Besides bradykinesia, patients with PD often experience rigidity, which can lead to more uniform and less variable writing speeds, resulting in a smaller standard deviation of velocity for tasks 3 and 8. Patients adopt a rigid and uniform pace to compensate for their reduced ability to make smooth and fluid adjustments. The lower standard deviation reflects a restricted range of motion or a mechanical, constrained movement pattern as they struggle to complete the task. Interestingly, although handwriting movements recorded during drawing of Archimedean spiral are common clinical and handwriting indicators, only tasks 2, 3 and 8 provide statistically significant features, indicating that in this thesis, the writing of cursive letter 'l' and a sentence provide more discriminative potential.

According to the feature importance as recorded by XGBoost, NCV is consistently important across 5 out of 10 features in Figure 4, which might reflect attempts to adjust for tremors or difficulty in maintaining a steady hand, resulting in numerous small, unintentional velocity shifts. Since NCV captures abrupt changes in velocity, it may indirectly reflect increased instantaneous acceleration, suggesting that both features could serve as

manifestations of PD tremor. Lower variability in velocity during writing tasks likely reflects compensatory strategies or constraints due to impaired fine motor control.

Although pressure features seem to exhibit less differential potential, it is nonetheless important to reflect on it as the maximum pressure for task 5 is the fourth feature according to feature importance (Figure 4). This task consists of writing a word, which might reflect the importance of pressure in tasks that require continuous writing of different letters. Whilst authors identify task 8 as the most promising task (Drotár et al., 2016), probably because it involves writing a full sentence, which allows for capturing the temporal dimension, task 5 is not much different from task 8 in terms of fine motor movements it requires. Still, task 8 is the most significant task when combined with velocity parameters, which is in line with the previous findings (Drotár et al., 2016) that velocity best differentiates between HC and PD when associated with temporal dimension.

Despite the statistical significance of these findings, their synthesis into a binary classifier for the purposes of pre-screening large populations remains challenging, particularly if one prioritises the early stages of the disease and potential efficiency gains in the public health system. The accuracy of 68% for Logistic Regression and Random Forest is significantly above what would be correctly predicted by chance. Classifying PD patients from handwriting at scale is possible, but much more complex architectures are needed to achieve promising results. While the promise of light- weight, effective tools like XGBoost has delivered in some areas of neurology (Inoue et al., 2020), it is not enough for such a complex task like motor function classification, meaning that high-complexity and resource-intensive solutions like CNNs are a necessary, but hopefully reasonable, tradeoff in search of data-driven solutions to PD diagnosis.

The findings demonstrate that various handwriting characteristics, such as velocity,

pressure and NCV can be effectively utilised in the early detection of PD. However, two limitations must be acknowledged when interpreting these results. This study focused on PD and HC in a single country, Czechia. This constrains the application of findings on a more general level, as population-specific characteristics, except available metadata, have not been accounted for in the analysis.

Insufficient precision questions whether the proposed CAD methodology can pre-screen patients accurately before doctors can do so. The question to what degree the proposed model is able to differentiate between different diseases with Parkinsonian symptoms, such as normal pressure hydrocephalus (NPH), remains insufficiently explored. Whilst the current statistical method produced features with differential potential between PD and HC, the question remains whether these features exclusively reflect the specifics of PD I analysed.

Lastly, although the proposed method resulted in features reflecting the current motor symptoms associated with PD, they do not account for the progression of motor disturbances over time. Additionally, there is a possibility that sometimes the model fails to capture differences between less rigorous handwriting styles and PD motor symptoms, as the previous record of handwriting of participants was not included in the data.

Economic Perspective

This part of the study was an evaluation and comparison of the economic burden of working-age individuals with PD in the Czechia. My results show that in the period between 1996 and 2018, economic burden attributed to PD, expressed through total productivity and personal income loss, has increased significantly. However, despite its absolute increase, when expressed as percentage of GDP, productivity loss declined. This is most likely driven by the fact that over time, the economy, hence GDP, grew at a faster pace than the prevalence of PD patients.

The analysis confirmed that premature retirement due to PD implies a non-negligible and increasing economic burden, expressed through absolute and relative TPL, as well as through total personal earnings loss. This finding contributes to the previous literature and stresses an implicit economic burden associated with disease progression. My findings are supported by literature across other countries not included as part of this study. Research from the US found a PD patients "were more than 5 times as likely to report that their health kept them from work or normal activities" (Rubenstein et al., 1997). Another study found significant personal costs associated with early retirement due to PD, where later age of diagnosis was associated with lower loss in earnings (Johnson et al., 2011), indicating that delayed progression of symptoms reduces the time to workforce exit due to PD.

Conclusion

Handwriting changes show a significant potential as discriminative indicators for neurodegenerative disease detection. This makes handwriting analysis a valuable foundation

for developing CAD tools to assist clinicians at the point of pre-screening. Given the well-documented handwriting impairments in Parkinsonian patients, such tools could play a particularly important role in supporting early identification of PD.

This thesis explored the predictive power of handwriting features using PaHaW dataset, which includes various handwriting tasks performed by both PD and HC. The statistical analysis identified a variety of velocity parameters, as well as pressure and NCV as the most significant features to differentiate between the two groups. The fact that task 8 combined with velocity features, is the most common after significance tests shows that velocity parameters and variability are particularly important when observed within task that requires a temporal dimension. The predictive models Logistic Regression and Random Forest revealed accuracy of 68%. Although this accuracy does not yet offer sufficient grounds for using discriminative potential of handwriting as a reliable tool for early PD detection overall, recall higher than 70% suggests that the model exhibits a sufficient performance when it comes to merely detecting patients with PD. Thus, for a handwriting-based screening system to be useful in correctly excluding HC, further analysis is required.

While prior work has achieved strong results on the PaHaW dataset (Drotár et al., 2016), these studies often did not account for disease severity. In contrast, this study focused exclusively on patients in the early to mild stages, making its primary contribution an investigation into the feasibility of using dynamic handwriting analysis to support early-stage PD diagnosis.

The two main limitations of this thesis are the small sample size and a lack of temporal dimension in the data collection, restricting the generalizability of the results. The former highlights the ongoing need for large, high-quality benchmark datasets in this field. The latter suggests that, although the method captures differences in handwriting between PD and

controls, it is unable to detect a temporal deterioration in handwriting movements that may vary with severity of the disease. Future research should aim to address these limitations and explore advanced classification strategies that combine temporal with static handwriting features to enhance predictive performance. Additional insights may also emerge by collecting handwriting data from PD individuals over time to validate the association of PD severity with motor deterioration.

The economic analysis of PD prevalence and TPL demonstrated that, in Czechia, PD results in non-negligible and increasing nominal economic burden, despite the relative decline over the period between 1996 and 2018. Future CAD interventions that alleviate the burden of PD- associated symptoms and delay progression, like handwriting based pre-screening, could reduce the economic burden.

Bibliography

Aarsland, D., Andersen, K., Larsen, J. P., & Lolk, A. (2003). Prevalence and Characteristics of Dementia in Parkinson Disease. *Archives of Neurology*, *60*(3), 387. https://doi.org/10.1001/archneur.60.3.387

Agentschap Innoveren & Ondernemen. (2021). *EU Funding Overview*. Eufundingoverview.be. https://eufundingoverview.be/funding/european-social-fund-esf

Alty, J., Cosgrove, J., Thorpe, D., & Kempster, P. (2017). How to use pen and paper tasks to aid tremor diagnosis in the clinic. *Practical Neurology*, *17*(6), 456–463. https://doi.org/10.1136/practneurol-2017-001719

Andlin-Sobocki, P., Jönsson, B., Wittchen, H.-U., Olesen, J., Gulacsi, L., Knapp, M., & Moscarelli, M. (2005). Costs of Disorders of the Brain in Europe Peer review panel. *EUROPEAN JOURNAL of NEUROLOGY*, *12*. https://www.fens.org/wp-content/uploads/2020/11/Cost-of-Disorders-of-the-Brain-in-Europe-EJN-June-2005.pdf#page=80

Bisio, A., Pedullà, L., Bonzano, L., Tacchino, A., Brichetto, G., & Bove, M. (2017). The kinematics of handwriting movements as expression of cognitive and sensorimotor impairments in people with multiple sclerosis. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-18066-7

Błaszczyk, J. W. (1998). Motor deficiency in Parkinson's disease. *Acta Neurobiologiae Experimentalis*, *58*(1), 79–93. https://doi.org/10.55782/ane-1998-1262

Bonney, M.-A. (2010). Understanding and Assessing Handwriting Difficulty: Perspectives from the Literature. *Australian Occupational Therapy Journal*, *39*(3), 7–15. https://doi.org/10.1111/j.1440-1630.1992.tb01751.x

Brooks, D. J. (2010). Imaging Approaches to Parkinson Disease. *Journal of Nuclear Medicine*, *51*(4), 596–609. https://doi.org/10.2967/jnumed.108.059998

Calne, D. B., Snow, B. J., & Lee, C. (1992). Criteria for diagnosing Parkinson's disease. *Annals of Neurology*, *32*(S1), S125–S127. https://doi.org/10.1002/ana.410320721

Chahine, L. M., & Stern, M. B. (2011). Diagnostic markers for Parkinson's disease. *Current Opinion in Neurology*, *24*(4), 309–317. https://doi.org/10.1097/wco.0b013e3283461723

Czech Statistical Office. (2024). *Employees and wages | Statistics*. Statistics. https://csu.gov.cz/employees-and-wages?1_pocet=10&1_start=0&pocet=10&start=0&1_skupiny=11&1_vlastnostiVystupu=12&1_razeni=-datumVydani&skupiny=11&vlastnostiVystupu=15&pouzeVydane=true&razeni=-datumVydani#data-and-time-series

Dahodwala, N., Li, P., Jahnke, J., Ladage, V. P., Pettit, A. R., Kandukuri, P. L., Bao, Y., Zamudio, J., Jalundhwala, Y. J., & Doshi, J. A. (2020). Burden of Parkinson's Disease by Severity: Health Care Costs in the U.S. Medicare Population. *Movement Disorders*, *36*(1), 133–142. https://doi.org/10.1002/mds.28265

Deuschl, G., Beghi, E., Fazekas, F., Varga, T., Christoforidi, K. A., Sipido, E., Bassetti, C. L., Vos, T., & Feigin, V. L. (2020). The burden of neurological diseases in Europe: an analysis for the Global Burden of Disease Study 2017. *The Lancet Public Health*, *5*(10), e551–e567. https://doi.org/10.1016/s2468-2667(20)30190-0

Diaz, M., Ferrer, M. A., Impedovo, D., Pirlo, G., & Vessio, G. (2019). Dynamically enhanced static handwriting representation for Parkinson's disease detection. *Pattern*

Recognition Letters, *128*, 204–210. https://doi.org/10.1016/j.patrec.2019.08.018

Dorsey, E. R., Sherer, T., Okun, M. S., & Bloem, B. R. (2018). The Emerging Evidence of the Parkinson Pandemic. *Journal of Parkinson's Disease*, *8*(S1), S3–S8. https://doi.org/10.3233/jpd-181474

Dorsey, E. R., Elbaz, A., Nichols, E., Abd-Allah, F., Abdelalim, A., Adsuar, J. C., Ansha, M. G., Brayne, C., Choi, J.-Y. J., Collado-Mateo, D., Dahodwala, N., Do, H. P., Edessa, D., Endres, M., Fereshtehnejad, S.-M., Foreman, K. J., Gankpe, F. G., Gupta, R., Hankey, G. J., & Hay, S. I. (2018). Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Neurology, 17(11), 939–953. https://doi.org/10.1016/s1474-4422(18)30295-3

Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., & Faundez-Zanuy, M. (2014). Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease. *Computer Methods and Programs in Biomedicine*, *117*(3), 405–411. https://doi.org/10.1016/j.cmpb.2014.08.007

Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., & Faundez-Zanuy, M. (2016). Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. *Artificial Intelligence in Medicine*, *67*, 39–46. https://doi.org/10.1016/j.artmed.2016.01.004

Europe PMC. (2016). *Europe PMC*. Europepmc.org. https://europepmc.org/article/med/20297871

Federal Reserve Bank of St. Louis. (2025). *FRED*. Stlouisfed.org; Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/tags/series?t=currency%3Bczech+republic

Feenstra, R. C., Inklaar, R. and Timmer, MP. (2015). *"The Next Generation of the Penn World Table."* American Economic Review, 105(10), 3150-3182.

Feigin, V. L., Nichols, E., Alam, T., Bannick, M. S., Beghi, E., Blake, N., Culpepper, W. J., Dorsey, E. R., Elbaz, A., Ellenbogen, R. G., Fisher, J. L., Fitzmaurice, C., Giussani, G., Glennie, L., James, S. L., Johnson, C. O., Kassebaum, N. J., Logroscino, G., Marin, B., & Mountjoy-Venning, W. C. (2019). Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology*, *18*(5), 459–480. https://doi.org/10.1016/s1474-4422(18)30499-x

Gubbay, S. S., & de Klerk, N. H. (1995). A study and review of developmental dysgraphia in relation to acquired dysgraphia. *Brain and Development*, *17*(1), 1–8. https://doi.org/10.1016/0387-7604(94)00110-j

Hirschauer, T. J., Adeli, H., & Buford, J. A. (2015). Computer-Aided Diagnosis of Parkinson's Disease Using Enhanced Probabilistic Neural Network. *Journal of Medical Systems*, *39*(11). https://doi.org/10.1007/s10916-015-0353-9

Impedovo, D., Pirlo, G., & Vessio, G. (2018). Dynamic Handwriting Analysis for Supporting Earlier Parkinson's Disease Diagnosis. *Information*, *9*(10), 247. https://doi.org/10.3390/info9100247

Infocommunications Journal. (2024, March). *Enhancing Parkinson's Disease Recognition through Multimodal Analysis of Archimedean Spiral Drawings*. Infocommunications Journal. https://doi.org/10.36244/ICJ.2024.1.8

Inoue, T., Ichikawa, D., Ueno, T., Cheong, M., Inoue, T., Whetstone, W. D., Endo, T., Nizuma, K., & Tominaga, T. (2020). XGBoost, a Machine Learning Method, Predicts Neurological Recovery in Patients with Cervical Spinal Cord Injury. *Neurotrauma Reports*, *1*(1), 8–16. https://doi.org/10.1089/neur.2020.0009

Institute for Health Metrics and Evaluation. (2021). *Global Health Data Exchange (GHDx)*. Institute for Health Metrics and Evaluation; University of Washington. https://vizhub.healthdata.org/gbd-results/

Islam, Md. Ariful., Majumder, H., Hussein, Md. Alomgeer., Hossain, K. M., & Miah, Md. Sohel. (2024). A review of machine learning and deep learning algorithms for Parkinson's disease detection using handwriting and voice datasets. *Heliyon*, *10*(3). https://doi.org/10.1016/j.heliyon.2024.e25469

Johnson, S. J., Davis, M., Kaltenboeck, A., Birnbaum, H. G., Grubb, E., Tarrants, M., & Siderowf, A. (2011). Early Retirement and Income Loss in Patients with Early and Advanced Parkinson's Disease. *Applied Health Economics and Health Policy*, *9*(6), 367–376. https://doi.org/10.2165/11596900-000000000-00000

K. Ray Chaudhuri, Azulay, J.-P., Odin, P., Lindvall, S., Domingos, J., Alobaidi, A., Kandukuri, P. L., Chaudhari, V. S., Juan Carlos Parra, Yamazaki, T., Oddsdottir, J., Wright, J., & Martinez-Martin, P. (2024). Economic Burden of Parkinson's Disease: A Multinational, Real-World, Cost-of-Illness Study. *Drugs - Real World Outcomes*, *11*. https://doi.org/10.1007/s40801-023-00410-1

Kalia, L. V., & Lang, A. E. (2015). Parkinson's disease. *The Lancet*, *386*(9996), 896–912. https://doi.org/10.1016/s0140-6736(14)61393-3

Larner, A. J. (2016). A Dictionary of Neurological Signs. In *Springer eBooks*. Springer Nature. https://doi.org/10.1007/978-3-319-29821-4

Leggio, L., Vivarelli, S., L'Episcopo, F., Tirolo, C., Caniglia, S., Testa, N., Marchetti, B., & Iraci, N. (2017). microRNAs in Parkinson's Disease: From Pathogenesis to Novel Diagnostic and Therapeutic Approaches. *International Journal of Molecular Sciences*, *18*(12). https://doi.org/10.3390/ijms18122698

Letanneux, A., Danna, J., Velay, J.-L., Viallet, F., & Pinto, S. (2014). From micrographia to Parkinson's disease dysgraphia. *Movement Disorders*, *29*(12), 1467–1475. https://doi.org/10.1002/mds.25990

Li, S., Jia, C., Li, T., & Le, W. (2021). Hot Topics in Recent Parkinson's Disease Research: Where We are and Where We Should Go. *Neuroscience Bulletin*, *386*(9996). https://doi.org/10.1007/s12264-021-00749-x

Margolin, D., & Wing, A. M. (1983). *Agraphia and micrographia: Clinical manifestations of motor programming and performance disorders*. *54*(1-3), 263–283. https://doi.org/10.1016/0001-6918(83)90039-2

Momina Moetesum, Siddiqi, I., Javed, F., & Uzma Masroor. (2020). Dynamic Handwriting Analysis for Parkinson's Disease Identification using C-BiGRU Model. *IEEE*, 115–120. https://doi.org/10.1109/icfhr2020.2020.00031

Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. (2003). The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations. *Movement Disorders*, *18*(7), 738–750. https://doi.org/10.1002/mds.10473

Nackaerts, E., Heremans, E., Smits-Engelsman, B. C. M., Broeder, S., Vandenberghe, W., Bergmans, B., & Nieuwboer, A. (2017). Validity and reliability of a new tool to evaluate handwriting difficulties in Parkinson's disease. *PLOS ONE*, *12*(3), e0173157. https://doi.org/10.1371/journal.pone.0173157

OECD. (2024). *Data*. OECD. https://www.oecd.org/en/data.html

Olesen, J., Gustavsson, A., Svensson, M., Wittchen, H.-U. ., & Jönsson, B. (2011). The economic cost of brain disorders in Europe. *European Journal of Neurology*, *19*(1), 155–162. https://doi.org/10.1111/j.1468-1331.2011.03590.x

Our World in Data. (2024). Our World in Data. https://ourworldindata.org/

Pereira, C. R., Pereira, D. R., Silva, F. A., Masieiro, J. P., Weber, S. A. T., Hook, C., & Papa, J. P. (2016). A new computer vision-based approach to aid the diagnosis of Parkinson's disease. *Computer Methods and Programs in Biomedicine*, *136*, 79–88. https://doi.org/10.1016/j.cmpb.2016.08.005

Phillips, J. G., Stelmach, G. E., & Teasdale, N. (1991). What can indices of handwriting quality

tell us about Parkinsonian handwriting? *Human Movement Science*, *10*(2-3), 301–314. https://doi.org/10.1016/0167-9457(91)90009-m

Poon, C., Gorji, N., Latt, M., Tsoi, K., Choi, B., Loy, C., & Poon, S. (2019). *Derivation and analysis of dynamic handwriting features as clinical markers of Parkinson's disease*.

Population, total - Czech Republic | Data. (n.d.). Data.worldbank.org. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=CZ

Rajput, A. H., Rozdilsky, B., & Rajput, A. (1991). Accuracy of Clinical Diagnosis in Parkinsonism — A Prospective Study. *Canadian Journal of Neurological Sciences / Journal Canadien Des Sciences Neurologiques*, *18*(3), 275–278. https://doi.org/10.1017/s0317167100031814

Rosenblum, S., Samuel, M., Zlotnik, S., Erikh, I., & Schlesinger, I. (2013). Handwriting as an objective tool for Parkinson's disease diagnosis. *Journal of Neurology*, *260*(9), 2357–2361. https://doi.org/10.1007/s00415-013-6996-x

Rosqvist, K., Schrag, A., & Odin, P. (2022). Caregiver Burden and Quality of Life in Late Stage Parkinson's Disease. *Brain Sciences*, *12*(1), 111. https://doi.org/10.3390/brainsci12010111

Rubenstein, L. M., Chrischilles, E. A., & Voelker, M. D. (1997). The Impact of Parkinson??s Disease on Health Status, Health Expenditures, and Productivity. *PharmacoEconomics*, *12*(4), 486–498. https://doi.org/10.2165/00019053-199712040-00006

Santos, J., Irantzu Pallarès, & Ventura, S. (2022). Is a cure for Parkinson's disease hiding inside us? *Trends in Biochemical Sciences*, *47*(8), 641–644. https://doi.org/10.1016/j.tibs.2022.02.001

Schapira, A. H. V., Chaudhuri, K. R., & Jenner, P. (2017). Non-motor features of Parkinson disease. *Nature Reviews Neuroscience*, *18*(7), 435–450. https://doi.org/10.1038/nrn.2017.62

Senatore, R., & Marcelli, A. (2019). A paradigm for emulating the early learning stage of handwriting: Performance comparison between healthy controls and Parkinson's disease patients in drawing loop shapes. *Human Movement Science*, *65*, 89–101. https://doi.org/10.1016/j.humov.2018.04.007

Sitek, E. J., Narozanska, E., Barczak, A., Jasinska-Myga, B., Michał Harciarek, Małgorzata Chodakowska-Zebrowska, Kubiak, M., Wieczorek, D., Seweryna Konieczna, Rademakers, R., Baker, M., Mariusz Berdynski, Bogna Brockhuis, Barcikowska, M., Cezary Zekanowski, Heilman, K. M., Wszolek, Z. K., & Slawek, J. (2012). Agraphia in patients with frontotemporal dementia and parkinsonism linked to chromosome 17 with P301L*MAPT*mutation: dysexecutive, aphasic, apraxic or spatial phenomenon?. *Neurocase*, *20*(1), 69–86. https://doi.org/10.1080/13554794.2012.732087

Sonja von Campenhausen, Winter, Y., Alfredo, A., Sampaio, C., Evžen Růžička, Barone, P., Poewe, W., Guekht, A., Mateus, C., Pfeiffer, K. P., Berger, K., J Skoupá, Bötzel, K., S. Geiger-Gritsch, Siebert, U., Balzer-Geldsetzer, M., Oertel, W. H., Dodel, R., & Reese, J. P. (2011). Costs of illness and care in Parkinson's Disease: An evaluation in six countries. *European Neuropsychopharmacology*, *21*(2), 180–191. https://doi.org/10.1016/j.euroneuro.2010.08.002

Tolosa, E., Garrido, A., Scholz, S. W., & Poewe, W. (2021). Challenges in the diagnosis of Parkinson's disease. *The Lancet Neurology*, *20*(5), 385–397. https://doi.org/10.1016/S1474-4422(21)00030-2

U.S. DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT. (2023). *govinfo | U.S. Government Publishing Office*. Govinfo.gov. https://www.govinfo.gov/

Winter, Y., von Campenhausen, S., Brozova, H., Skoupa, J., Reese, J. P., Bötzel, K., Eggert, K., Oertel, W. H., Dodel, R., & Ruzicka, E. (2010). Costs of Parkinson's disease in Eastern Europe: A Czech cohort study. *Parkinsonism & Related Disorders*, *16*(1), 51–

56. https://doi.org/10.1016/j.parkreldis.2009.07.005
Wong-Yu, I. S. K., Ren, L., & Mak, M. K. Y. (2021). Impaired hand function and its association with self-perceived hand functional ability and quality of life in Parkinson's disease. *American Journal of Physical Medicine & Rehabilitation, Publish Ahead of Print*. https://doi.org/10.1097/phm.0000000000001923
Wu, Y., Le, W., & Jankovic, J. (2011). Preclinical Biomarkers of Parkinson Disease. *Archives graof Neurology*, *68*(1). https://doi.org/10.1001/archneurol.2010.321
Yang, W., Hamilton, J. L., Kopil, C., Beck, J. C., Tanner, C. M., Albin, R. L., Ray Dorsey, E., Dahodwala, N., Cintina, I., Hogan, P., & Thompson, T. (2020). Current and projected future economic burden of Parkinson's disease in the U.S. *Npj Parkinson's Disease*, *6*(1). https://doi.org/10.1038/s41531-020-0117-1
Zhong, M., Peppard, R., Velakoulis, D., & Evans, A. H. (2015). The relationship between specific cognitive defects and burden of care in Parkinson's disease. *International Psychogeriatrics*, *28*(2), 275–281. https://doi.org/10.1017/s1041610215001593