OPEN ACCESS

Research Article

Data-driven stratification of Parkinson's disease patients based on the progression of motor and cognitive disease markers

Datengetriebene Stratifizierung von Patienten mit Parkinson-Krankheit anhand von Verlaufsdaten motorischer und kognitiver Kennzahlen der Erkrankung

Abstract

Parkinson's disease (PD) is a progressive neurodegenerative movement disorder with a complex set of motor and non-motor symptoms and a diverse disease progression. Subtyping PD patients is required for personalized therapies but stratification approaches based on intermediate phenotypes such as clinical assessment scores lack reproducibility and stability, which is at least partially due to the broad spectrum of methods that can be applied during different steps of data processing. We propose a novel approach that considers the progression of detailed clinical assessment scores in different domains over a period of five years. Furthermore, we confirm the robustness of our subtypes with comparisons to subtypes that emerge when using different data pre-processing or another clustering algorithm. Three subtypes were found with differentiable symptoms: The *motor-dominant* subtype has the fastest progression and is most severely affected in daily life, closely followed by the *sleep-dominant non-tremor* subtype. The *mild-motor* subtype, in contrast, is characterized by moderate progression. These subtypes emerge from their progression pattern rather than from a snapshot during one time point. Hence we advocate for stratification approaches for PD subtyping that take longitudinal data over several years into account.

Keywords: Parkinson's disease (PD), PPMI, stratification, subtypes, biomarker, machine learning, clustering

Erenik Krasniqi[1,2]
Wendelin Schramm[3]
Alexandra Reichenbach[3]

1 Heilbronn University, Heilbronn, Germany
2 University of Heidelberg, Germany
3 Medical Informatics & Center for Machine Learning, Heilbronn University, Heilbronn, Germany

Zusammenfassung

Die Parkinson-Krankheit ist eine fortschreitende, neurodegenerative Erkrankung, die sich durch komplexe motorische und nicht-motorische Symptome sowie einen vielfältigen Krankheitsverlauf auszeichnet. Subtypisierung der Patienten ist für personalisierte Therapien notwendig, jedoch fehlt es an Stratifizierungsansätzen, die auf Zwischenphänotypen wie z.B. klinischen Tests aufsetzen, an Reproduzierbarkeit und Stabilität. Dies bedingt sich teilweise durch die vielen methodischen Möglichkeiten bei der Datenprozessierung. Wir schlagen einen neuen Ansatz vor, bei dem die Entwicklung detaillierter klinischer Kennwerte aus unterschiedlichen Domänen über einen Zeitraum von fünf Jahren betrachtet wird. Die Robustheit der so erhaltenen Subtypen untermauern wir mit Vergleichen zu Subtypen, die wir mit abweichender Datenprozessierung oder einem anderen Clustering-Algorithmus gewonnen hätten. Wir finden hier drei Subtypen mit differenzierbarer Symptomatik: Der *motorisch-dominante* Subtyp ist gekennzeichnet durch den raschesten Verfall und ist im täglichen Leben am stärksten betroffen, eng gefolgt vom *Schlaf-dominanten non-Tremor* Subtyp. Im Gegensatz dazu ist der Krankheitsverlauf des *mild-motorischen* Subtyps eher moderat. Diese Subtypen

erwachsen aus den Verläufen ihrer komplexen Symptomatik und nicht aus Gruppenunterschieden während eines einzelnen Zeitpunkts. Deswegen plädieren wir dafür, für die Subtypisierung von Parkinson-Patienten Längsschnittdaten mehrerer Jahre zu verwenden.

Schlüsselwörter: Parkinson-Krankheit, PPMI, Stratifizierung, Subtypen, Biomarker, maschinelles Lernen, Clustering

Introduction

Parkinson's disease (PD) is a progressive neurodegenerative movement disorder with a complex set of motor and non-motor symptoms and a diverse disease progression. Bradykinesia, resting tremor, rigidity, and postural instability represent its cardinal symptoms [1], [2], [3]. However, the complexity of the disease manifests in a broad spectrum of symptoms, including non-motor symptoms in early disease progression and an overall heterogeneous progression of symptoms [2], [4]. Based on these disease features, it may be possible to stratify PD patients into subgroups with distinct disease courses [5]. Successful stratification of PD patients is crucial for better prediction of the individual's disease course and development of individualized therapy.

Motor and non-motor features have different underlying pathologies. Motor symptoms are mainly caused by degeneration of striatal dopaminergic neurons of the substantia nigra pars compacta in the basal ganglia, with the presence of intracytoplasmic α-synoclein protein or Lewy bodies (LB) [6], [7], [8]. Presence of LBs beyond the brainstem has been confirmed, explaining the heterogeneous characteristics of PD especially regarding non-motor features [9]. Non-motor symptoms such as sleep disorders, cognitive impairments, olfactory loss, and constipation often occur before manifestation of motor features concluding these to be pre-indicators.

Genetic mutations in 18 chromosomal regions are identified as genetic PD risk factors [10]. Mutations located in SNCA, LRRK2, MAPT, and GBA have the greatest genetic impact on developing PD. SNCA mutations can cause dysregulation of LBs, LRRK2 mutations mediate neuronal toxicity, and mutations in MAPT can cause PD related dementia [9], [11], [12], [13], [14], [15]. GBA mutations cause a wide spectrum of symptoms and are found in 8%–14% of PD autopsies [10], [16].

Previous PD subtyping studies are based on data obtained on a single time point, on the difference between baseline and follow-up measurements, or on longitudinal data without preserving the temporal structure in the data [17], [18], [19], [20], [21], [22]. However, the PD subtypes found in these studies lack reproducibility [23], [24] and stability over time [25], [26], [27], [28], [29]. To circumvent the problem of instability over time we hypothesize that PD subtypes do not necessarily differ for a snapshot of disease symptoms, obtained via e.g. clinical test scores, markers from biospecimens, or characteristics derived from neuroimaging, but that PD subtypes mainly differ in their progression of these disease symptoms.

We analysed data from 237 de novo, unmedicated PD patients from the Parkinson's Progression Markers Initiative (PPMI) cohort [30] who have completed yearly assessments of 14 disease markers in the motor, neuropsychological, cognitive, and sleep disorder domain over a five-year period. The progression of the disease is modelled with polynomial regression and the regression coefficients were used for cluster analysis, which stratifies the heterogenous group of patients in more homogenous subgroups, i.e. the PD subtypes. Regression coefficients have been used for prediction of the disease progression of PD patients [31] but for PD patient stratification this is a novel approach. Different pre-processing pipelines and clustering algorithms were compared in order to evaluate the robustness of the obtained PD subtypes.

Methods

Data

Data used in the preparation of this article were obtained from the PPMI database, downloaded in March 2019. For up-to-date information on the study, visit http://www.ppmi-info.org/. PPMI provides subject records of 18 different clinical (sub-)assessments, of which 12 are used in this study (Table 1). The Movement Disorder Society Unified Parkinson Rating Scale (MDS-UPDRS) is an update of the original Unified Parkinson's Disease Rating Scale, dividing it into four distinct parts and strengthening the non-motor features [32]. MDS-UPDRS is especially created for measuring the longitudinal disease course by assessing different disease characteristics. The subscales of the first three parts were used in this study to obtain more detailed characteristics than from the total score alone. The fourth part was excluded because of too many missing values. Additionally, Postural Instability/Gait Difficulty (PIGD) and Tremor Dominance (TD) scores were calculated from the corresponding MDS-UPDRS items [33]. Initially meant to be disease subtyping classifiers, these scores were shown to be more valuable as indicators for the disease progression [25], [26], [34]. The items of MDS-UPDRS Part II and III that were used to determine PIGD and TD were excluded from the total scores of these parts, yielding corrected scores. Apart from motor assessments, we included three neuropsychological, four cognitive, and two sleep disorder tests. Five further assessments were excluded because they were obtained at screening but not at later visits. In summary, we used all 12 clinical sub-assessments for which sufficient longitudinal data was available, added the two calculated

Table 1: Chosen clinical assessment scores at baseline (n=237) with their respective ranges. Note that the scores for MDS-UPDRS Part II and Part III were corrected for the TD and PIGD scores, respectively, which are calculated from parts of these assessments. Categorization derived from Marek and colleagues [30].

Category	Clinical (sub-)assessment	Mean ± SD	Score range [best, worst]
Motor	MDS-UPDRS Part II corrected [32]	4.0 ± 3.0	[0, 44]
	MDS-UPDRS Part III corrected [32]	19.3 ± 8.3	[0, 92]
	TD score (calculated) [33]	4.3 ± 1.0	[0, 44]
	PIGD score (calculated) [33]	4.7 ± 2.0	[0, 20]
Autonomic testing	MDS-UPDRS Part I [32]	4.0 ± 3.0	[0, 52]
Neuro/behavior	Geriatric Depression Scale (GDS) [50]	5.2 ± 1.6	[0, 15]
	State-Trait Anxiety Inventory (STAI) [51]	93.0 ± 8.3	[0, 160]
	Questionnaire for Impulsive-Compulsive Disorders (QUIP) [52]	1.4 ± 0.8	[0, 12]
Cognitive testing	Hopkins Verbal Learning Test [53]	184.1 ± 34.6	the higher, the better
	Benton Judgment of Line Orientation [54]	13.0 ± 2.0	[30, 0]
	Letter Number Sequencing [55]	10.8 ± 2.5	[21, 0]
	Modified Schwab and England ADL [56]	93.5 ± 5.9	[100, 0]
Sleep disorders	Epworth Sleepiness Scale [57]	6.8 ± 3.3	[0, 24]
	REM Sleep Disorder [58]	5.2 ± 1.6	[0, 13]

Table 2: Demographic and genetic characteristics of included patients at baseline (n=237). Eleven patients are missing some demographic data at baseline and are therefore not included in this part of the table. Some patients have multiple ethnicities, therefore the total count of ethnicities exceeds the number of patients.

	Mean ± SD/count	Range/percentage
Age	60.6 ± 9.7	[34, 85]
Years of education	15.5 ± 2.8	[5, 26]
Disease duration [months]	6.5 ± 6.6	[1, 35]
Female/male	70/156	31.0/69.0
Ethnicity: Afro American/Asian/ Caucasian/Hispanic/Native American/ not specified	2/4/230/7/2/5	0.8/1.6/91.3/2.8/0.8/2.0
Handedness: right/left	205/25	89.1/10.9
LRRK2 mutation: y/n	75/162	31.7/68.3
MAPT mutation: y/n	80/157	33.8/66.2
SNCA mutation: y/n	227/10	95.8/4.2
GBA mutation: y/n	25/212	10.6/89.4

scores PIGD and TD and corrected the sub-scores from the underlying assessments accordingly.

From the 454 PD patients of the PPMI cohort, we included the ones with yearly collections of the chosen 12 clinical assessments over a period of five years. Therefore, each of the 237 patients in our sample (Table 2) has a complete set of the 14 scores (Table 1) at baseline and 12, 24, 36, 48, and 60 months after baseline.

Pre-processing, feature extraction and selection

Each of the 14 scores was min-max normalized from 0 to 1 because this procedure yields data ranges that are comparable between scores and since the scores have natural ranges, outliers are not a problem here. To capture the progression of the disease over time and reduce the dimensionality of the data, we transformed the data into a time-series over five years from which we derived three regression coefficient sets with polynomial regression (Figure 1a). These regression coefficients, especially the higher order "slopes" β_1, β_2, and β_3, serve as numerical indicators of disease progression over the five years. Since we could not assume progression in a linear fashion, we included further possibilities with the higher order regressions. The first set includes intercepts and slopes from the linear regression of each of the 14 scores with time in months as predictor (28 coefficients total), the second set includes the three coefficients from the 14 quadratic regressions over time (42 coefficients total), and the third set includes the four coefficients from the 14 cubic regressions over time (56 coefficients total).

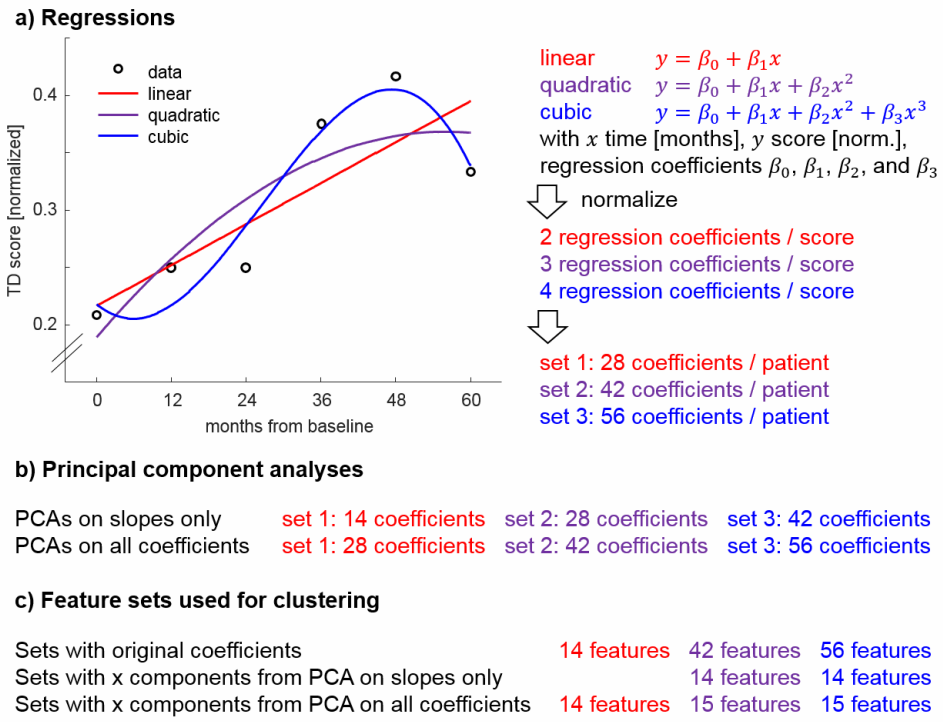

Figure 1: Feature construction. The regressions (a) are illustrated exemplarily on the TD score of one patient. The three regression types are specified by the accompanying equations and color coded throughout the whole figure. Six PCAs were carried out with the data as illustrated in b). The procedure yielded the 8 feature sets that are illustrated in c).

The regression coefficients were min-max normalized from –1 to 1 to preserve the direction of the change.
To investigate the variance distributions of the regression coefficients and reduce the number of features for clustering, principal component analyses (PCA) were performed on all regression coefficient sets (Figure 1b). Since our main focus was on the progression of the disease, we analyzed the slopes in addition to the whole set of regression parameters (Figure 1c, Figure 2).

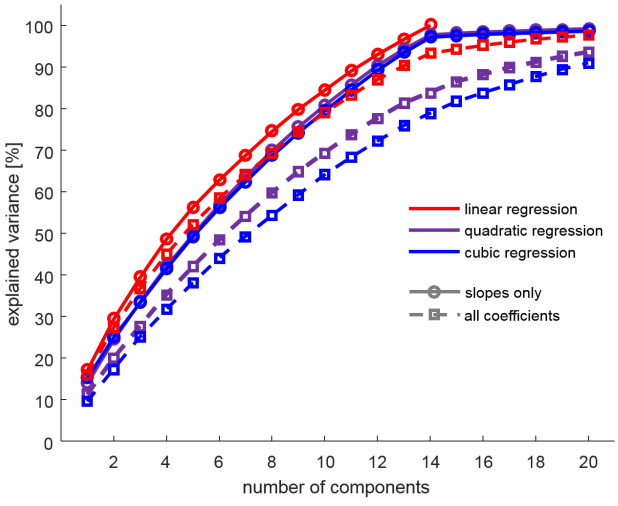

Figure 2: Explained variances of the PCAs conducted on the three complete data sets (all coefficients) and the reduced data sets (slopes only).

Using only the slopes (Figure 2, circles), it is clear that the first 14 PCs are sufficient to explain nearly the complete variance. The two sets with more than 14 coefficients experience a sharp drop in additional explained variance thereafter. The PCA on all linear coefficients shows a similar progression, although the drop after the 14th component is not as sharp. For these data sets, we therefore included the first 14 components in the feature sets for the clustering, except for the linear slopes where we used only the original data, since the PCA did not provide a dimensionality reduction here. The explained variances rise slower for all quadratic and cubic coefficients, therefore we used the first 15 components to account for this but still staying in an order similar to the other data sets. In addition to these reduced feature sets, we considered all regression coefficients from the quadratic and cubic regression as feature sets for the clustering, yielding eight feature sets for cluster analyses (Figure 1c).
Pre-processing, clustering, and statistical analyses were conducted in Python 3.7 using pandas [35], ScyPy [36], scikit-learn [37], NumPy [38], pinguin [39], and MATLAB (The MathWorks, Natick, MA, USA) using the MATLAB engine for Python.

Clustering algorithms and model evaluation

We use the two simplest and most common clustering algorithms *k*-means and hierarchical clustering in order to evaluate the stability of PD subtypes across clustering algorithms. *k*-means is the most commonly used algorithm overall, which is also extensively used on various health-related data [40], [41], including PD subtyping [20], [21],

[42]. For PD subtyping, hierarchical clustering is the most common clustering method [17], [18], [34]. This allows us to compare our stratification results with previous studies yet still directly comparing two algorithms.
For both algorithms we used the simplest configuration: Euclidean distances as similarity measure and for the hierarchical clustering Ward's criterion as linkage criterion. Both algorithms were performed on all eight feature sets.
For finding the best feature set and an optimal number of clusters, we calculated the explained variance for each clustering model from k=2 to k=20 clusters, and supplemented these data with the within-cluster sum of squares (WCSS) for each model. Both measures can easily compare models with similar complexity and both provide an idea about the optimal number of clusters using the "elbow method" [43]. For the hierarchical clustering, we additionally used the dendrograms for finding the optimal number of clusters.
To evaluate the robustness of our final cluster solution, we compared cluster memberships for this model with similar models. Similar models were the ones either obtained by the other algorithm, or with one more or less number of clusters, or using another feature set.

Statistical analyses

To describe patient subtypes from the clusters, we analysed the progression of the original assessments for the patients groups with fixed-effects ANOVAs with time (0, 12, 24, 36, 48, 60 months) as within-subjects factor and patient group as between-subjects factor. Patient groups were additionally evaluated by demographic characteristics and mutation frequencies of PD-related genes using ANOVAs or X^2-tests, respectively. Significance level was set to α=.05, uncorrected, since we merely used the tests for group description and not for group differentiation. Bonferroni correction was used for post-hoc tests within each assessment.
Furthermore, we explore whether any of the biospecimens collected for the PPMI cohort might be used as a biomarker for predicting the patient subtype from an early stage on. 89 of the specimens had data for at least half of the patients in our sample and were therefore included in the analysis. From these specimens, 41 were derived from cerebrospinal fluid, 10 from RNA, 32 from plasma, 2 from serum, 3 from urine, and one from whole blood. We conducted fixed-effects ANOVAs with the patient group as between-subjects factor on those specimens from the first visit. Significance level was set to α=.05, corrected for the number of specimens, i.e. α=.00056.

Results

Model selection

Since k-means clustering yielded slightly better values for the explained variances and both k-means and hierarchical clustering provide a rather similar picture, we demonstrate the model selection based on the measures obtained from k-means clustering (Figure 3). Both measures clearly show that the models based on the feature sets with the linear regression coefficients outperform the other models. Taken both measures together, the feature set containing only the linear slopes is the model to choose. Besides, this model has two advantages for interpretation: It is based directly on the regression slopes and not on PCA transformed data, which can be more directly interpreted. It also takes only the progression of the disease into account, i.e. is not biased by the patients' states at the time they entered the study like the models that are based on all regression coefficients, i.e. including the intercepts.
For obtaining the optimal number of clusters neither measure provides a clear "elbow" in its course (Figure 3a,b). One might detect a slight bending at three or four clusters but this is debatable. A closer look on the changes in the measures with increasing number of clusters (Figure 3c,d) suggests indeed a noticeable drop in improvement after three clusters. The dendrogram from the hierarchical clustering on the linear regression slopes confirms this observation (Figure 4a). Here, the distance measure clearly suggest three clusters. Therefore, the clustering model chosen for subtyping the PD patients is based on k-means clustering on the linear regression slopes with k=3.
A comparison of cluster assignments between the models obtained with k-means and hierarchical clustering (Figure 4b) reveals rather robust clusters between algorithms. 185 out of 237 patients (78%) were assigned to the same clusters. Comparing the assignment to the three clusters with the models based on two and four clusters (Figure 4c-d), respectively, reveals rather robust cluster assignments as well. The group splitting is also comparable to the hierarchical clustering (Figure 4a). Compared with the cluster assignments of the next best feature set, the first 14 PCs of all linear coefficients, there is only one patient with a deviating cluster assignment. Taken together, the clustering model we base our subtyping on is quite robust against the clustering algorithm, the number of clusters, and the feature set used.

Cluster group demographics and genetics

The patients were distributed rather evenly across the groups (Table 3). The groups differ neither in age, years of education, disease duration at baseline, nor handedness. There is, however, a difference in the distribution of gender (Table 3). Women are, relative to the ratio in the patient sample, under-represented in the first but over-represented in the third group. The four most prominent gene mutations associated with PD, LRR2, MAPT, SNCA, and GBA, were uniformly distributed across groups.

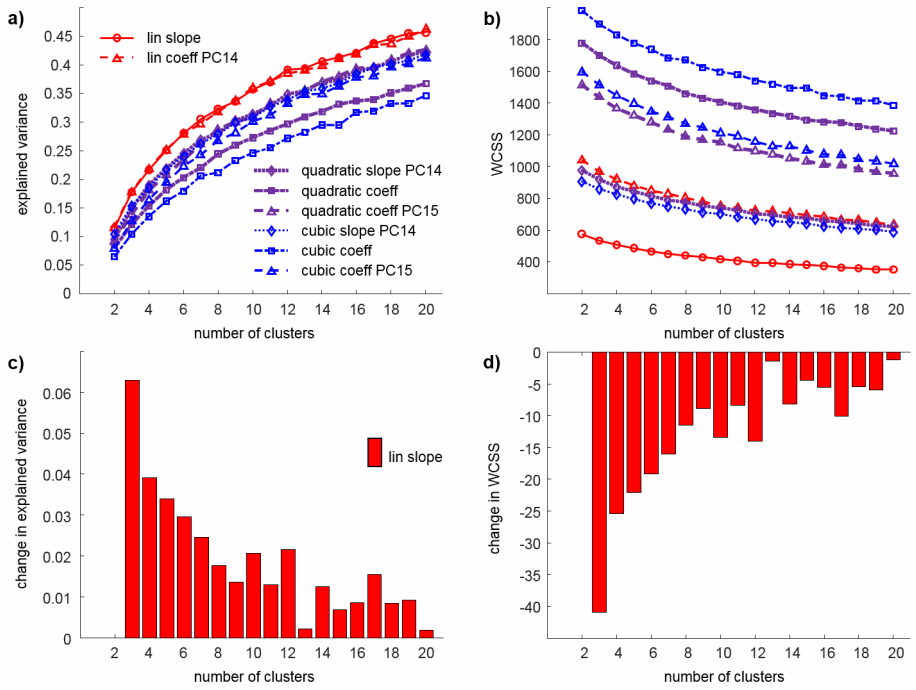

Figure 3: Evaluation of the clusters obtained from k-means clustering. The explained variance (a) indicates how much of the variance in the data can be explained by the assignment of patients to the clusters and should therefore be maximized. The within-cluster sum of squares (WCSS, b) quantifies the deviation of the cluster members from each cluster centroid and should be minimized to obtain compact clusters with large inter-cluster distances [59]. The changes in explained variance (c) and WCSS (d) with each additional cluster for the models based on the linear slopes provide a more detailed view on the fit of each cluster number.

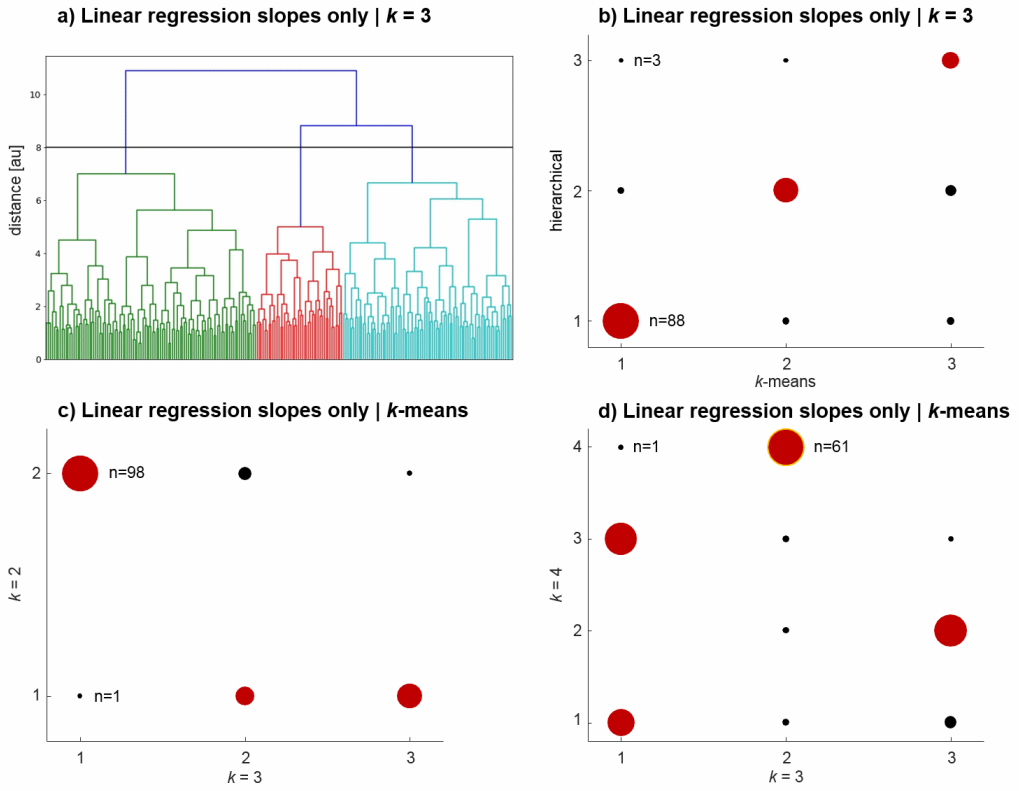

Figure 4: Model comparisons to clustering approaches similar to the chosen model. Panel a shows the corresponding dendrogram from hierarchical clustering and panel b compares the cluster assignment between the two clustering algorithms. Panels c and d compare to the cluster models with *k*=2 and *k*=4, correspondingly. The size of the discs correspond to the number of patients, the smallest and largest discs are labeled with the respective number of patients for reference, and the red discs highlight the corresponding clusters between the two compared models.

Table 3: Demographic and genetic characteristics of patients in the groups (n=237) with a test for group differences. Eleven patients are missing demographic data and are therefore not included in this part of the table.

Characteristics	Group 1	Group 2	Group 3	F/X^2	df	p
Number of subjects (%)	99 (42%)	71 (30%)	67 (28%)			
Age	61.8 ± 10.7	60.4 ± 9.0	59.2 ± 8.8	1.465	2, 223	.233
Years of education	15.2 ± 3.1	15.8 ± 3.0	15.6 ± 2.4	0.888	2, 223	.413
Disease duration at baseline [months]	7.12 ± 7.3	6.3 ± 7.1	5.8 ± 4.7	0.791	2, 223	.455
Handedness (R/L)	71/20	62/9	53/4	1.473	2	.479
Female/male (% female)	11/88 (12%)	23/48 (32%)	26/37 (41%)	6.701	2	.035
LRRK2 mutation (y/n)	30/69	22/49	24/43	0.212	2	.900
MAPT mutation (y/n)	33/66	25/46	22/45	0.064	2	.969
SNCA mutation (y/n)	99/0	69/2	59/8	0.733	2	.693
GBA mutation (y/n)	10/89	8/63	7/60	0.051	2	.975

Notation: mean ± SD or count

Patient subtyping

Analysis of the progression of the assessment scores in the different groups (Table 4) reveals a detailed profile of the progression in the different domains for each group and can therefore be used as a description for the subtypes of the disease.

Motor symptoms (Table 4, Figure 5) worsen overall and the groups demonstrate distinguishable progressions over all sub-domains. Group 1 demonstrates the steepest decline in all motor functions, their overall performance in the MDS-UPDRS Part II corr. (Figure 5a) is sign. worse than the performance of group 3 (t_{164}=3.136; p=.002), and their overall performance measured by the TD score is sign. worse than the performance of group 2 (t_{168}=3.888; p<.001) and 3 (t_{164}=3.006; p=.003). For the two other groups it is a mixed picture. They have similar performance in the MDS-UPDRS Part III corr. and the PIGD score. However, group 3 stays rather stable for MDS-UPDRS Part II corr. and worsens only mildly in the TD score. In contrast, group 2 gets sign. worse than group 3 in the MDS-UPDRS Part II corr. but sign. better in the TD score. The tremor dominance of group 2 even decreases sign. after five years compared to baseline (t_{70}=3.126; p=.003). Taken together, group 1 experiences a steep increase in all motor symptoms and group 3 shows only mild deteriorations in the motor domain. The symptoms of group 2 worsen for the ones assessed with the MDS-UPDRS Part II corr. but the tremor symptoms improve slightly.

The autonomy of the patients (MDS-UPDRS Part I) decreases in general and the groups demonstrate distinguishable progressions (Table 4, Figure 6a). The autonomy of group 1 and 2 is steeply decreasing, with group 2 exhibiting least autonomy, which is overall even sign. worse than the autonomy of group 3 (t_{136}=3.119; p=.002). Group 3 remains stable.

The neuropsychological assessments (Table 4, Figure 6b-d) cannot distinguish between the groups and remain rather stable over time. Only the scores of the Questionnaire for Impulsive-Compulsive Disorders show an overall effect of time and for this assessment as well as for the State-Trait Anxiety Inventory we find a differential effect of time on the three groups. However, the assessments in this domain fail to be good descriptors for the groups. Cognitive changes over time are present for all assessments in this domain (Table 4, Figure 7). In the Hopkins Verbal Learning Test the patients become slightly better, which might be a learning effect. The verbal working memory (Letter Number Sequencing) and the ability of performing the activities of daily life (Modified Schwab and England) are in general decreasing. The differential progression for the groups on the Schwab and England scores sets group 1 apart from the other two groups, which are rather similar. Group 1 experiences steady decrease in abilities while the two other groups have a steep loss of function in the first year but stabilize afterwards. After four years, the performance of group 1 has worsened sign. Taken together, group 1 experiences some loss of verbal working memory and substantial decrease in the ability to perform daily activities, group 2 also experiences some loss of verbal working memory and a moderate decrease in the ability to perform daily activities, and group 3 shows only a moderate decrease in the ability to perform daily activities.

Sleep problems (Table 4, Figure 8) worsen in general and the groups demonstrate distinguishable progressions over both sub-domains. Group 2 shows the steepest increase in problems on the Epworth Sleepiness Scale and their overall score is higher than the score of group 1 (t_{168}=2.216; p=.028; n.s. corr. but sign. in year four and five) and sign. higher than the score of group 3 (t_{136}=3.413; p=.001). With respect to REM Sleep Disorder, groups 1 and 2 show rather similar moderate progression, both having sign. higher overall scores than group 3 (group 1: t_{164}=2.569; p=.011; group 2: t_{136}=2.685; p=.008). Taken together, group 1 is characterized by a moderate increase on the Epworth Sleepiness Scale and in REM Sleep Disorder, group 2 experiences a steep increase on the Epworth Sleepiness Scale and a moderate increase in REM Sleep Disorder, and the sleep of group 3 remains stable.

Table 4: Group statistics for the mixed ANOVAs on the 14 assessment scores

	Group (*df*=2, 234)	Time (*df*=5, 1170)	Group x Time (*df*=10,1170)
Motor			
MDS-UPDRS Part II corr.	4.850; .009	60.833; <.001	21.570; <.001
MDS-UPDRS Part III corr.	0.936; .394	60.778; <.001	15.061; <.001
TD score	9.223; <.001	7.416; <.001	9.778; <.001
PIGD score	1.093; .337	29.961; <.001	5.761; <.001
Autonomic testing			
MDS-UPDRS Part I	4.429; .013	50.236; <.001	14.973; <.001
Neuro/behaviour			
Geriatric Depression Scale	2.098; .125	1.099; .359	1.695; .077
State-Trait Anxiety Inventory	0.772; .463	2.043; .070	1.962; .034
Quest. for Impulsive-Compulsive Dis.	1.560: .212	6.864; <.001	2.345; .001
Cognitive testing			
Hopkins Verbal Learning Test	1.676; .189	6.723; <.001	3.662; <.001
Benton Judgment of Line Orientation	2.789; .064	8.091; <.001	2.608; .004
Letter Number Sequencing	0.249; .780	5.068; <.001	1.783; .059
Modified Schwab and England ADL	0.880; .416	60.953; <.001	10.120; <.001
Sleep disorders			
Epworth Sleepiness Scale	5.499; .005	21.490; <.001	6.943; <.001
REM Sleep Disorder	4.084; .018	12.205; <.001	2.367; .009

Notation: *F*-values; *p*-values (uncorrected). Non-significant cells are shaded.

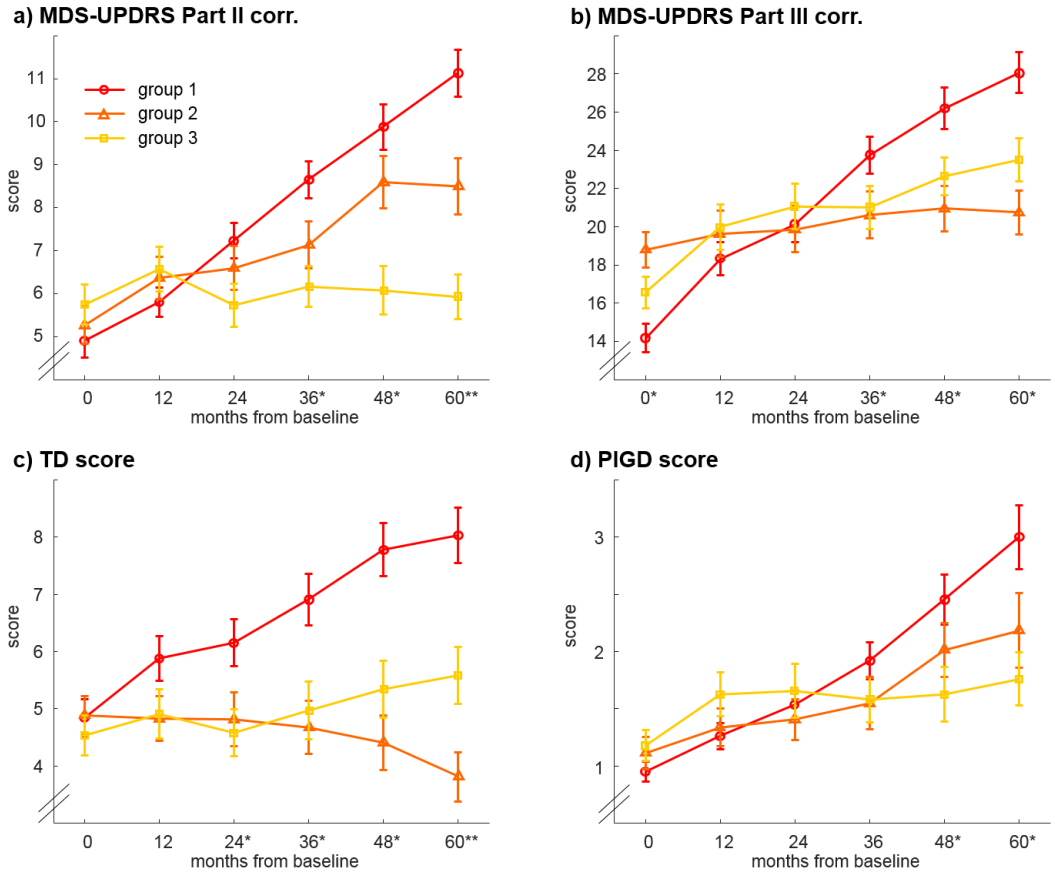

Figure 5: Development of the motor scores for the three groups. The higher the scores the worse the motor symptoms. Error bars denote SEM.

* at least two groups differ sign., ** all three groups differ sign. (only corrected for multiple comparisons within each score).

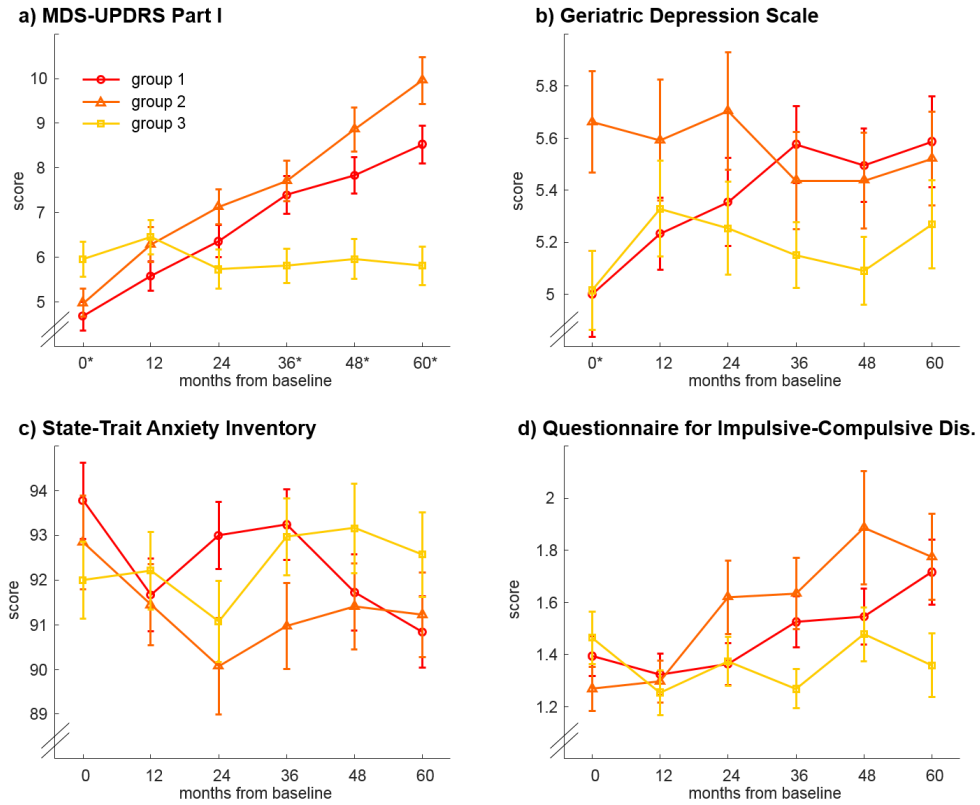

Figure 6: Development of the autonomy (a) and neuropsychological tests (b-d) for the three groups. Error bars denote SEM.
* at least two groups differ sign. (only corrected for multiple comparisons within each score).

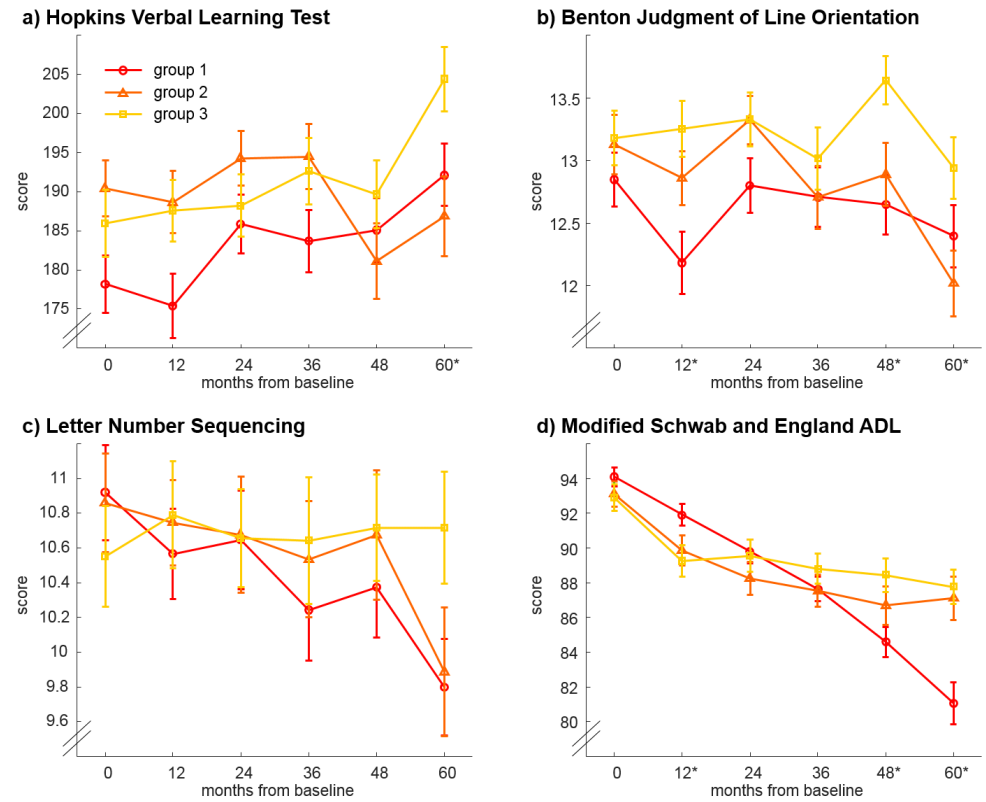

Figure 7: Development of the cognitive scores for the three groups. The higher the scores the better the cognitive abilities. Error bars denote SEM.
* at least two groups differ sign. (only corrected for multiple comparisons within each score).

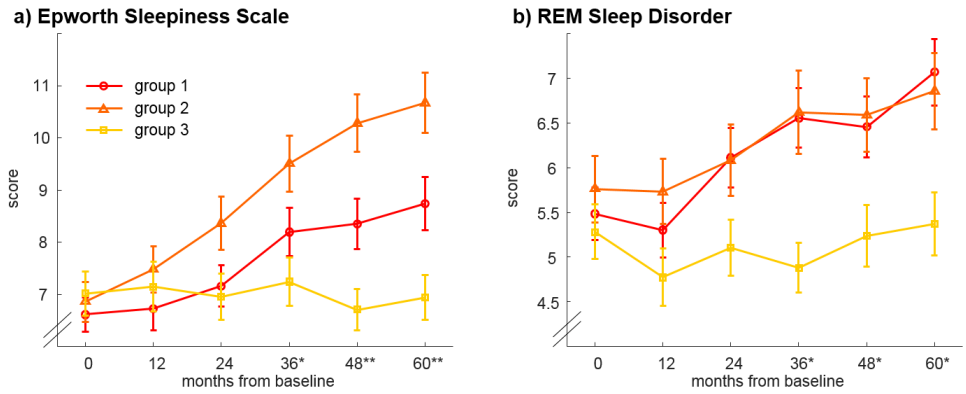

Figure 8: Development of sleep disorders for the three groups. The higher the scores the worse the sleep problems. Error bars denote SEM.

* at least two groups differ sign., ** all three groups differ sign. (only corrected for multiple comparisons within each score).

In summary, we identified three PD subtypes and can describe them in different domains. Subtype 1 (*motor-dominant*) is characterized by severe decrease in all motor domains and the ability to perform activities of daily life. These core symptoms are accompanied by mild decrease in verbal working memory and mild increase in sleep problems. Subtype 2 (*sleep-dominant non-tremor*) is characterized by severe increase in sleep problems and a shift in loss of motor function from tremor dominant to other motor symptoms. In this subtype, we also see severe decrease in daily autonomy. The core symptoms are accompanied by mild decrease in verbal working memory. Subtype 3 (*mild-motor*) experiences the least increase in symptoms with mild increase in loss of motor function and in impairments in the ability to perform activities of daily life.

Biomarker exploration

For the 89 biospecimens analyzed, only 11 ANOVAs had a probability of error $p<.05$ but none of them survived the correction for multiple comparisons (Table 5).

Table 5: Group statistics for the ANOVAs on the biospecimens. Only tests with $p<.05$ (uncorrected) are included.

	F	*df*	*p*
Cerebrospinal fluid			
C18 GlcCer	3.600	2, 197	.029
NDNA B2M CN	3.360	2, 158	.037
Plasma			
C16 SM	4.473	2, 202	.013
C20 GL2	3.118	2, 202	.046
C20 SM	3.836	2, 178	.023
C22 GL2	5.492	2, 202	.005
C22 SM	4.251	2, 197	.016
Total SM	5.003	2, 202	.008
RNA			
ALDH1A1 (rep 1)	3.266	2, 170	.041
SNCA-3UTR-1	5.726	2, 119	.004
SNCA-3UTR-2	5.397	2, 119	.006

Discussion

This study stratified de-novo PD patients based on the progression of 14 disease markers in the motor, neuropsychological, cognitive, and sleep disorder domain over a five-year period. We found three subtypes of PD patients that differ in the course of loss of function. We termed the first subtype *motor-dominant* since the core characteristics were a steep increase in all motor symptoms, accompanied by a loss of daily life autonomy. The second subtype was termed *sleep-dominant non-tremor* since the core characteristics were a severe increase in sleep problems and a shift in loss of motor function from tremor dominant to other motor symptoms, also accompanied by loss of daily life autonomy. The third subtype, *mild-motor*, is characterized by only mild increase in loss of motor function accompanied by moderate loss in daily life autonomy.

Most markers in our study cannot differentiate between the subtypes in general or at baseline but the differences emerge over the five-year period: for nine out of 14 scores we did not find main effects of group but for eleven out of 14 we found interactions between group and time, and only three out of 14 scores differentiate between subtypes at baseline. This suggests that the subtypes genuinely reflect different courses of the disease. Additionally, the patients in the groups do not differ by most demographic variables, especially not by age or disease duration at baseline. Thus, it is unlikely that the differences in our subtypes stem from patients being in different disease stages. The distribution of gender, however, was uneven for the three subgroups. Men were, relative to the gender-ratio in the sample, over-represented in the *motor-dominant* subtype and under-represented in the *mild-motor* subtype. Studies have previously found gender-differences in symptom severity but there is no clear picture yet [44]. Most common in clinical practice is subtyping PD into Tremor Dominance (TD) and Postural Instability/Gait Difficulty (PIGD) [33], characteristics easily quantified by the respective scores calculated from items of the MDS-UPDRS Parts II and III. However, recent studies question the classification power of this approach, as there is a

high between group fluctuation and most subjects initially classified as PIGD tend to switch to TD with progressing disease course, suggesting PIGD and TD may be different disease stages rather than subtypes with distinct disease courses [25], [26]. Regarding the subtypes of this study, we find some discriminative power in those two scores but they do not set themselves apart from the other assessment scores used. Early data-driven clustering studies focussed mainly on the motor symptoms and also identified subtypes characterized as tremor-dominant vs. non-tremor dominant [45], [46] but the more commonly division was characterized as "old age at onset/rapid disease progression" vs. "young age at onset/slow disease progression" [19].

While our subtypes are also characterized by different rates of overall progression, we do not find a difference for the age at onset.

More recent clustering approaches included a variety of non-motor assessments and agree on a "mild" [20], [21], "mild-motor predominant" [17] or "mainly motor/slow progression" [18] subtype comparable to the *mild-motor* subtype of this study. The studies following up on the assessments in their subtypes found the slowest progression of symptoms in this subtype [17], [18], which is in line with the progression in the present study. The other extreme, the *motor-dominant* subtype with fast progression, has been identified in other studies as well as "severe" [20] or "diffuse/malignant" [17], [18] subtype with the fastest progression. The "intermediate" subtype [17], [18] is the most diverse amongst studies with a differentiation between "*motor-dominant*" vs. "*non-motor dominant*" [20] or some motor symptoms paired with different non-motor symptoms [21]. In the present study, the most prominent feature of this group is the continued low tremor dominance paired with a steep increase in sleep symptom, hence the term *sleep-dominant non-tremor*. Even though there are some similarities between the subtypes identified in previous studies and the ones from this study, it is important to keep in mind that the subtypes from this study are based on the progression over a five-year period while the other studies derive the subtypes from baseline data. It would be especially interesting to see how well the subtypes agree to the ones found by the studies based on the same patient population [17], [21].

Lack of reproducibility [23], [24] and stability over time [25], [26], [27], [28], [29] questioned the existence of reliable subtypes in PD [24], [26]. We addressed some of the problems that were identified as possible culprits for the heterogeneity in subtyping approaches [24]. Clustering on global composite scores [17], [42] considers the one-dimensional scale of severity but disregards distinctive developments in different domains. Therefore, we used sub-scores in different domains. Analysing only a snapshot of measures [20], [21], the difference between two time-points [17], [18], or using longitudinal data without preserving the temporal structure [22] neglects the complexity of the time course the disease can take. We covered a period of five years with six measures and describe the progression with the coefficients of polynomial regression models, therewith preserving the temporal structure in the data. Because commonly used missing value imputation [17], [18], [42] can negatively affect clustering outcomes we discarded patients with missing values. Furthermore, the pre-processing steps used and the choice of clustering algorithm add to the heterogeneity. We evaluated our cluster assignment against slightly different approaches. The assignment is rather robust against some pre-processing steps such as selecting the 14 slopes from the linear regression *vs*. using the 14 first components from a PCA on all linear coefficients. It is also rather robust against the choice of clustering algorithm, namely *k*-means *vs*. hierarchical clustering. There are altogether 22% of patients that shift between groups when using another clustering algorithm, the largest group (8%) is assigned to group 3 from *k*-means and to group 2 from hierarchical clustering. The overall pattern of assessments for the groups, however, remains rather stable. With these evaluations, we successfully replicated our results by using different methods. However, the plethora of methods for clustering and data pre-processing for clustering is vast. In this study, we have combined a couple of commonly used methods to explore the stability of the models that are based on our novel features that characterize disease progression. However, we can by no means claim to have exhaustively explored their capabilities. One important limitation of our study is the inclusion of only about half the patients enrolled in the PPMI study and the selection of assessments used for stratification. Our strict inclusion regimen not only tremendously limited the number of patients but also might have introduced a selection bias. The next important step is to validate our clusters on another, ideally larger, dataset in order to ensure reproducibility and stability of the subtypes we found. Furthermore, stratification approaches might benefit from inclusion of more intermediate phenotypes of the disease such as metrics of brain anatomy and function.

In order to clinically utilize the subtypes that we describe for the prediction of disease progression or for personalized treatment, it is necessary to predict the subtype for a de novo patient. While most of the assessments we used for clustering cannot separate the subtypes at baseline, the motor assessment MDS-UPDRS Part III corr. differentiates between the *motor-dominant* and *sleep-dominant non-tremor* subtypes at baseline, the autonomy assessment MDS-UPDRS Part I differentiates between the *motor-dominant* and the *mild-motor* subtypes at baseline, and the Geriatric Depression Scale sets the *sleep-dominant non-tremor* subtype apart from the *motor-dominant* and the *mild-motor* subtypes at baseline. A combination of these assessments might therefore be used as an early indicator for subtyping since the majority of patients used in this study were diagnosed with PD for less than a year when they entered the study. However, the subtypes cannot be distinguished by genetic markers, at least not by mutations in the genes that are most commonly associated with PD: LRRK2, MAPT, SNCA, or

GBA. This is in line with previous studies showing that clinical symptoms do not differ between patients with and without LKKR2 mutations [47], that variants of MAPT and SNCA mutations are not associated with performance in cognitive tests [48], and that patients with GBA mutations do not show neuropathological differences from patients without this mutation [49]. Therefore, we tried to address this issue with exploratory analyses of biospecimens from cerebrospinal fluid, RNA, plasma, serum, urine, and whole blood taken at baseline. However, these markers cannot stratify the patients into the subtypes either, at least not with the conservative approach used in this study. Some of these markers, however, show promise and should be investigated further. The RNA markers SNCA-3UTS-1 and -2 were amongst those biospecimens as well as the plasma markers C22 GL2 and total SM. Furthermore, other genetic and physiological markers such as LB concentration and brain atrophy should be investigated to shed light on the physiological underpinnings of the subtypes and can hopefully one day be used as biomarker for prediction of disease progression or personalized treatment.

Conclusion

The current study set out to evaluate the approach of clustering PD patients based on longitudinal clinical assessments in different domains for finding stable PD subtypes. The results reveal three subtypes differing in the five-year progression of their symptoms in various motor and non-motor domains. The subtypes are robust against some methodological variations and demonstrate stability over time. Our results demonstrate that this approach shows promise and should be pursued further.

Notes

Funding

PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including MJFF, AbbVie, Avid Radiopharmaceuticals, Biogen Idec, Bristol-Myers Squibb, Covance, Eli Lilly & Co., F. Hoffman-La Roche, Ltd., GE Healthcare, Genentech, GlaxoSmithKline, Lundbeck, Merck, Meso-Scale, Piramal, Pfizer, and UCB.
This research was partially funded by the Federal Ministry of Education and Research (01IS17067; https://www.bmbf.de/en/index.html).

Competing interests

The authors declare that they have no competing interests.

References

1. DeMaagd G, Philip A. Parkinson's Disease and Its Management: Part 1: Disease Entity, Risk Factors, Pathophysiology, Clinical Presentation, and Diagnosis. P T. 2015 Aug;40(8):504-32.
2. Foltynie T, Brayne C, Barker RA. The heterogeneity of idiopathic Parkinson's disease. J Neurol. 2002 Feb;249(2):138-45. DOI: 10.1007/pl00007856
3. Obeso JA, Stamelou M, Goetz CG, Poewe W, Lang AE, Weintraub D, Burn D, Halliday GM, Bezard E, Przedborski S, Lehericy S, Brooks DJ, Rothwell JC, Hallett M, DeLong MR, Marras C, Tanner CM, Ross GW, Langston JW, Klein C, Bonifati V, Jankovic J, Lozano AM, Deuschl G, Bergman H, Tolosa E, Rodriguez-Violante M, Fahn S, Postuma RB, Berg D, Marek K, Standaert DG, Surmeier DJ, Olanow CW, Kordower JH, Calabresi P, Schapira AHV, Stoessl AJ. Past, present, and future of Parkinson's disease: A special essay on the 200th Anniversary of the Shaking Palsy. Mov Disord. 2017 Sep;32(9):1264-310. DOI: 10.1002/mds.27115
4. Berg D, Postuma RB, Bloem B, Chan P, Dubois B, Gasser T, Goetz CG, Halliday GM, Hardy J, Lang AE, Litvan I, Marek K, Obeso J, Oertel W, Olanow CW, Poewe W, Stern M, Deuschl G. Time to redefine PD? Introductory statement of the MDS Task Force on the definition of Parkinson's disease. Mov Disord. 2014 Apr;29(4):454-62. DOI: 10.1002/mds.25844
5. Sieber BA, Landis S, Koroshetz W, Bateman R, Siderowf A, Galpern WR, Dunlop J, Finkbeiner S, Sutherland M, Wang H, Lee VM, Orr HT, Gwinn K, Ludwig K, Taylor A, Torborg C, Montine TJ; Parkinson's Disease 2014: Advancing Research, Improving Lives Conference Organizing Committee. Prioritized research recommendations from the National Institute of Neurological Disorders and Stroke Parkinson's Disease 2014 conference. Ann Neurol. 2014 Oct;76(4):469-72. DOI: 10.1002/ana.24261
6. Bernheimer H, Birkmayer W, Hornykiewicz O, Jellinger K, Seitelberger F. Brain dopamine and the syndromes of Parkinson and Huntington. Clinical, morphological and neurochemical correlations. J Neurol Sci. 1973 Dec;20(4):415-55. DOI: 10.1016/0022-510x(73)90175-5
7. Forno LS. Neuropathology of Parkinson's disease. J Neuropathol Exp Neurol. 1996 Mar;55(3):259-72. DOI: 10.1097/00005072-199603000-00001
8. Jellinger KA. Pathology of Parkinson's disease. Changes other than the nigrostriatal pathway. Mol Chem Neuropathol. 1991 Jun;14(3):153-97. DOI: 10.1007/BF03159935
9. Moore DJ, West AB, Dawson VL, Dawson TM. Molecular pathophysiology of Parkinson's disease. Annu Rev Neurosci. 2005;28:57-87. DOI: 10.1146/annurev.neuro.28.061604.135718
10. Klein C, Westenberger A. Genetics of Parkinson's disease. Cold Spring Harb Perspect Med. 2012 Jan;2(1):a008888. DOI: 10.1101/cshperspect.a008888
11. Gilks WP, Abou-Sleiman PM, Gandhi S, Jain S, Singleton A, Lees AJ, Shaw K, Bhatia KP, Bonifati V, Quinn NP, Lynch J, Healy DG, Holton JL, Revesz T, Wood NW. A common LRRK2 mutation in idiopathic Parkinson's disease. Lancet. 2005 Jan 29-Feb 4;365(9457):415-6. DOI: 10.1016/S0140-6736(05)17830-1
12. Hutton M, Lendon CL, Rizzu P, Baker M, Froelich S, Houlden H, Pickering-Brown S, Chakraverty S, Isaacs A, Grover A, Hackett J, Adamson J, Lincoln S, Dickson D, Davies P, Petersen RC, Stevens M, de Graaff E, Wauters E, van Baren J, Hillebrand M, Joosse M, Kwon JM, Nowotny P, Che LK, Norton J, Morris JC, Reed LA, Trojanowski J, Basun H, Lannfelt L, Neystat M, Fahn S, Dark F, Tannenberg T, Dodd PR, Hayward N, Kwok JB, Schofield PR, Andreadis A, Snowden J, Craufurd D, Neary D, Owen F, Oostra BA, Hardy J, Goate A, van Swieten J, Mann D, Lynch T, Heutink P. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. Nature. 1998 Jun;393(6686):702-5. DOI: 10.1038/31508

13. Latourelle JC, Beste MT, Hadzi TC, Miller RE, Oppenheim JN, Valko MP, Wuest DM, Church BW, Khalil IG, Hayete B, Venuto CS. Large-scale identification of clinical and genetic predictors of motor progression in patients with newly diagnosed Parkinson's disease: a longitudinal cohort study and validation. Lancet Neurol. 2017 Nov;16(11):908-16. DOI: 10.1016/S1474-4422(17)30328-9

14. Satake W, Nakabayashi Y, Mizuta I, Hirota Y, Ito C, Kubo M, Kawaguchi T, Tsunoda T, Watanabe M, Takeda A, Tomiyama H, Nakashima K, Hasegawa K, Obata F, Yoshikawa T, Kawakami H, Sakoda S, Yamamoto M, Hattori N, Murata M, Nakamura Y, Toda T. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. Nat Genet. 2009 Dec;41(12):1303-7. DOI: 10.1038/ng.485

15. Smith WW, Pei Z, Jiang H, Dawson VL, Dawson TM, Ross CA. Kinase activity of mutant LRRK2 mediates neuronal toxicity. Nat Neurosci. 2006 Oct;9(10):1231-3. DOI: 10.1038/nn1776

16. Goker-Alpan O, Schiffmann R, LaMarca ME, Nussbaum RL, McInerney-Leo A, Sidransky E. Parkinsonism among Gaucher disease carriers. J Med Genet. 2004 Dec;41(12):937-40. DOI: 10.1136/jmg.2004.024455

17. Fereshtehnejad SM, Zeighami Y, Dagher A, Postuma RB. Clinical criteria for subtyping Parkinson's disease: biomarkers and longitudinal progression. Brain. 2017 Jul;140(7):1959-76. DOI: 10.1093/brain/awx118

18. Fereshtehnejad SM, Romenets SR, Anang JB, Latreille V, Gagnon JF, Postuma RB. New Clinical Subtypes of Parkinson Disease and Their Longitudinal Progression: A Prospective Cohort Comparison With Other Phenotypes. JAMA Neurol. 2015 Aug;72(8):863-73. DOI: 10.1001/jamaneurol.2015.0703

19. van Rooden SM, Colas F, Martínez-Martín P, Visser M, Verbaan D, Marinus J, Chaudhuri RK, Kok JN, van Hilten JJ. Clinical subtypes of Parkinson's disease. Mov Disord. 2011 Jan;26(1):51-8. DOI: 10.1002/mds.23346

20. Mu J, Chaudhuri KR, Bielza C, de Pedro-Cuesta J, Larrañaga P, Martinez-Martin P. Parkinson's Disease Subtypes Identified from Cluster Analysis of Motor and Non-motor Symptoms. Front Aging Neurosci. 2017;9:301. DOI: 10.3389/fnagi.2017.00301

21. Erro R, Picillo M, Vitale C, Palladino R, Amboni M, Moccia M, Pellecchia MT, Barone P. Clinical clusters and dopaminergic dysfunction in de-novo Parkinson disease. Parkinsonism Relat Disord. 2016 Jul;28:137-40. DOI: 10.1016/j.parkreldis.2016.04.026

22. Vavougios GD, Doskas T, Kormas C, Krogfelt KA, Zarogiannis SG, Stefanis L. Identification of a prospective early motor progression cluster of Parkinson's disease: Data from the PPMI study. J Neurol Sci. 2018 Apr;387:103-8. DOI: 10.1016/j.jns.2018.01.025

23. Mestre TA, Eberly S, Tanner C, Grimes D, Lang AE, Oakes D, Marras C. Reproducibility of data-driven Parkinson's disease subtypes for clinical research. Parkinsonism Relat Disord. 2018 Nov;56:102-6. DOI: 10.1016/j.parkreldis.2018.07.009

24. Qian E, Huang Y. Subtyping of Parkinson's Disease – Where Are We Up To? Aging Dis. 2019 Oct;10(5):1130-9. DOI: 10.14336/AD.2019.0112

25. Nutt JG. Motor subtype in Parkinson's disease: Different disorders or different stages of disease? Mov Disord. 2016 Jul;31(7):957-61. DOI: 10.1002/mds.26657

26. Simuni T, Caspell-Garcia C, Coffey C, Lasch S, Tanner C, Marek K; PPMI Investigators. How stable are Parkinson's disease subtypes in de novo patients: Analysis of the PPMI cohort? Parkinsonism Relat Disord. 2016 Jul;28:62-7. DOI: 10.1016/j.parkreldis.2016.04.027

27. Eisinger RS, Hess CW, Martinez-Ramirez D, Almeida L, Foote KD, Okun MS, Gunduz A. Motor subtype changes in early Parkinson's disease. Parkinsonism Relat Disord. 2017 Oct;43:67-72. DOI: 10.1016/j.parkreldis.2017.07.018

28. Aleksovski D, Miljkovic D, Bravi D, Antonini A. Disease progression in Parkinson subtypes: the PPMI dataset. Neurol Sci. 2018 Nov;39(11):1971-6. DOI: 10.1007/s10072-018-3522-z

29. Erro R, Picillo M, Amboni M, Savastano R, Scannapieco S, Cuoco S, Santangelo G, Vitale C, Pellecchia MT, Barone P. Comparing postural instability and gait disorder and akinetic-rigid subtyping of Parkinson disease and their stability over time. Eur J Neurol. 2019 Sep;26(9):1212-8. DOI: 10.1111/ene.13968

30. Marek K, Jennings D, Lasch S, Siderowf A, Tanner C, Simuni T, Coffey C, Kieburtz K, Flagg E, Chowdhury S, Poewe W, Mollenhauer B; Paracelsus-Elena Klinik, Sherer T, Frasier M, Meunier C, Rudolph A, Casaceli C, Seibyl J, Mendick S, Schuff N, Zhang Y, Toga A, Crawford K, Ansbach A, De Blasio P, Piovella M, Trojanowski J, Shaw L, Singleton A, Hawkins K, Eberling J, Brooks D, Russell D, Leary L, Factor S, Sommerfeld B, Hogarth P, Pighetti E, Williams K, Standaert D, Guthrie S, Hauser R, Delgado H, JankovicJ, Hunter C, Stern M, Tran B, Leverenz J, Baca M, FrankS, Thomas CA, RichardI, Deeley C, Rees L, Sprenger F, Lang E, Shill H, Obradov S, Fernandez H, Winters A, Berg D, Gauss K, Galasko D, Fontaine D, Mari Z, Gerstenhaber M, Brooks D, Malloy S, Barone P, Longo K, Comery T, Ravina B, Grachev I, Gallagher K, Collins M, Widnell KL, Ostrowizki S, Fontoura P, Ho T, Luthman J, van der Brug M, Reith AD, Taylor P; The Parkinson Progression Marker Initiative (PPMI). The Parkinson Progression Marker Initiative (PPMI). Prog Neurobiol. 2011 Dec;95(4):629-35. DOI: 10.1016/j.pneurobio.2011.09.005

31. Tsanas A, Little MA, McSharry PE, Ramig LO. Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests. IEEE Trans Biomed Eng. 2010 Apr;57(4):884-93. DOI: 10.1109/TBME.2009.2036000

32. Goetz CG, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stebbins GT, Stern MB, Tilley BC, Dodel R, Dubois B, Holloway R, Jankovic J, Kulisevsky J, Lang AE, Lees A, Leurgans S, LeWitt PA, Nyenhuis D, Olanow CW, Rascol O, Schrag A, Teresi JA, Van Hilten JJ, LaPelle N. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan. Mov Disord. 2007 Jan;22(1):41-7. DOI: 10.1002/mds.21198

33. Stebbins GT, Goetz CG, Burn DJ, Jankovic J, Khoo TK, Tilley BC. How to identify tremor dominant and postural instability/gait difficulty groups with the movement disorder society unified Parkinson's disease rating scale: comparison with the unified Parkinson's disease rating scale. Mov Disord. 2013 May;28(5):668-70. DOI: 10.1002/mds.25383

34. Fereshtehnejad SM, Postuma RB. Subtypes of Parkinson's Disease: What Do They Tell Us About Disease Progression? Curr Neurol Neurosci Rep. 2017 Apr;17(4):34. DOI: 10.1007/s11910-017-0738-x

35. McKinney W. Data Structures for Statistical Computing in Python. In: Proceedings of the 9th Python in Science Conference; 2010 Jun 28-Jul 3; Austin, Texas. p. 56-61.

36. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P; SciPy 1.0 Contributors. Author Correction: SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020 Mar;17(3):352. DOI: 10.1038/s41592-020-0772-5

37. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12(85):2825-30.

38. van der Walt S, Colbert SC, Varoquaux G. The NumPy Array: A Structure for Efficient Numerical Computation. Comput Sci Eng. 2011;13(2):22-30. DOI: 10.1109/MCSE.2011.37

39. Vallat R. Pingouin: statistics in Python. J Open Source Softw. 2018;3(31):1026. DOI: 10.21105/joss.01026

40. Jothi N, Rashid NAA, Husain W. Data Mining in Healthcare – A Review. Procedia Computer Science. 2015;72:306-13. DOI: 10.1016/j.procs.2015.12.145

41. Tomar D, Agarwal S. A survey on Data Mining approaches for Healthcare. International Journal of Bio-Science and Bio-Technology. 2013;5(5):241-66. DOI: 10.14257/ijbsbt.2013.5.5.25

42. Zhang X, Chou J, Liang J, Xiao C, Zhao Y, Sarva H, Henchcliffe C, Wang F. Data-Driven Subtyping of Parkinson's Disease Using Longitudinal Clinical Records: A Cohort Study. Sci Rep. 2019 Jan;9(1):797. DOI: 10.1038/s41598-018-37545-z

43. Yuan C, Yang H. Research on K-value selection method of K-means clustering algorithm. J (Basel). 2019;2(2):226-35. DOI: 10.3390/j2020016

44. Miller IN, Cronin-Golomb A. Gender differences in Parkinson's disease: clinical characteristics and cognition. Mov Disord. 2010 Dec;25(16):2695-703. DOI: 10.1002/mds.23388

45. Reijnders JS, Ehrt U, Lousberg R, Aarsland D, Leentjens AF. The association between motor subtypes and psychopathology in Parkinson's disease. Parkinsonism Relat Disord. 2009 Jun;15(5):379-82. DOI: 10.1016/j.parkreldis.2008.09.003

46. Lewis SJ, Foltynie T, Blackwell AD, Robbins TW, Owen AM, Barker RA. Heterogeneity of Parkinson's disease in the early clinical stages using a data driven approach. J Neurol Neurosurg Psychiatry. 2005 Mar;76(3):343-8. DOI: 10.1136/jnnp.2003.033530

47. Healy DG, Falchi M, O'Sullivan SS, Bonifati V, Durr A, Bressman S, Brice A, Aasly J, Zabetian CP, Goldwurm S, Ferreira JJ, Tolosa E, Kay DM, Klein C, Williams DR, Marras C, Lang AE, Wszolek ZK, Berciano J, Schapira AH, Lynch T, Bhatia KP, Gasser T, Lees AJ, Wood NW; International LRRK2 Consortium. Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. Lancet Neurol. 2008 Jul;7(7):583-90. DOI: 10.1016/S1474-4422(08)70117-0

48. Mata IF, Leverenz JB, Weintraub D, Trojanowski JQ, Hurtig HI, Van Deerlin VM, Ritz B, Rausch R, Rhodes SL, Factor SA, Wood-Siverio C, Quinn JF, Chung KA, Peterson AL, Espay AJ, Revilla FJ, Devoto J, Hu SC, Cholerton BA, Wan JY, Montine TJ, Edwards KL, Zabetian CP. APOE, MAPT, and SNCA genes and cognitive performance in Parkinson disease. JAMA Neurol. 2014 Nov;71(11):1405-12. DOI: 10.1001/jamaneurol.2014.1455

49. Adler CH, Beach TG, Shill HA, Caviness JN, Driver-Dunckley E, Sabbagh MN, Patel A, Sue LI, Serrano G, Jacobson SA, Davis K, Belden CM, Dugger BN, Paciga SA, Winslow AR, Hirst WD, Hentz JG. GBA mutations in Parkinson disease: earlier death but similar neuropathological features. Eur J Neurol. 2017 Nov;24(11):1363-8. DOI: 10.1111/ene.13395

50. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. Clinical Gerontologist: The Journal of Aging and Mental Health. 1986;5(1-2):165-73. DOI: 10.1300/J018v05n01_09

51. Spielberger CD, Sydeman SJ, Owen AE, Marsh BJ. Measuring anxiety and anger with the State-Trait Anxiety Inventory (STAI) and the State-Trait Anger Expression Inventory (STAXI). In: Maruish ME, editor. The use of psychological testing for treatment planning and outcomes assessment. Lawrence Erlbaum Associates Publishers; 1999. p. 993–1021.

52. Weintraub D, Mamikonyan E, Papay K, Shea JA, Xie SX, Siderowf A. Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease-Rating Scale. Mov Disord. 2012 Feb;27(2):242-7. DOI: 10.1002/mds.24023

53. Zeinoun P, Farran N, Khoury SJ, Darwish H. Development, psychometric properties, and pilot norms of the first Arabic indigenous memory test: The Verbal Memory Arabic Test (VMAT). J Clin Exp Neuropsychol. 2020 Jul;42(5):505-15. DOI: 10.1080/13803395.2020.1773408

54. Irani F. Judgment of line orientation. In: Kreutzer JS, DeLuca J, Caplan B, editors. Encyclopedia of Clinical Neuropsychology. New York: Springer; 2011. p. 1372-4. DOI: 10.1007/978-0-387-79948-3_1376

55. Wechsler D. The measurement and appraisal of adult intelligence. Baltimore: Williams & Wilkins; 1958. DOI: 10.1037/11167-000

56. Schwab RS, England AC. Projection technique for evaluating surgery in Parkinson's disease. In: Gillingham FJ, Donaldson MC, editors. Third Symposium on Parkinson's Disease. Edinburgh: E & S Livingston; 1969. p. 152–7.

57. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. 1991 Dec;14(6):540-5. DOI: 10.1093/sleep/14.6.540

58. Stiasny-Kolster K, Mayer G, Schäfer S, Möller JC, Heinzel-Gutenbrunner M, Oertel WH. The REM sleep behavior disorder screening questionnaire – a new diagnostic instrument. Mov Disord. 2007 Dec;22(16):2386-93. DOI: 10.1002/mds.21740

59. Witten DM, Tibshirani R. A framework for feature selection in clustering. J Am Stat Assoc. 2010 Jun;105(490):713-26. DOI: 10.1198/jasa.2010.tm09415

Corresponding author:
Alexandra Reichenbach
Medical Informatics & Center for Machine Learning, Heilbronn University, Max-Planck-Str. 39, 74081 Heilbronn, Germany
alexandra.reichenbach@hs-heilbronn.de

Please cite as
Krasniqi E, Schramm W, Reichenbach A. Data-driven stratification of Parkinson's disease patients based on the progression of motor and cognitive disease markers. GMS Med Inform Biom Epidemiol. 2021;17(1):Doc04.
DOI: 10.3205/mibe000218, URN: urn:nbn:de:0183-mibe0002185

This article is freely available from
https://www.egms.de/en/journals/mibe/2021-17/mibe000218.shtml

Published: *2021-05-31*