RESEARCH ARTICLE

Clinical Subtypes of Parkinson's Disease

Stephanie M. van Rooden, MSc,[1]* Fabrice Colas, PhD,[2] Pablo Martínez-Martín, MD,[3] Martine Visser, PhD,[1] Dagmar Verbaan, PhD,[1] Johan Marinus, PhD,[1] Ray K. Chaudhuri, MD,[4] Joost N. Kok, PhD,[2,5] and Jacobus J. van Hilten, MD[1]

[1]*Department of Neurology, Leiden University Medical Centre, Leiden, The Netherlands*
[2]*Leiden Institute for Advanced Computer Science, Leiden University, Leiden, The Netherlands*
[3]*Alzheimer Disease Research Unit, National Centre for Epidemiology and CIBERNED, Carlos III Institute of Health, Madrid, Spain*
[4]*National Parkinson Foundation Centre of Excellence, King's College Hospital and University, Hospital Lewisham, London, United Kingdom/King's College and Institute of Psychiatry, London, United Kingdom*
[5]*Department of Medical Statistics and Bioinformatics/Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands*

ABSTRACT: The clinical heterogeneity of Parkinson's disease (PD) may point at the existence of subtypes. Because subtypes likely reflect distinct underlying etiologies, their identification may facilitate future genetic and pharmacotherapeutic studies. Aim of this study was to identify subtypes by a data-driven approach applied to a broad spectrum of motor and nonmotor features of PD. Data of motor and nonmotor PD symptoms were collected in 802 patients in two different European prevalent cohorts. A model-based cluster analysis was conducted on baseline data of 344 patients of a Dutch cohort (PROPARK). Reproducibility of these results was tested in data of the second annual assessment of the same cohort and validated in an independent Spanish cohort (ELEP) of 357 patients. The subtypes were subsequently characterized on clinical and demographic variables. Four similar PD subtypes were identified in two different populations and are largely characterized by differences in the severity of nondopaminergic features and motor complications: Subtype 1 was mildly affected in all domains, Subtype 2 was predominantly characterized by severe motor complications, Subtype 3 was affected mainly on nondopaminergic domains without prominent motor complications, while Subtype 4 was severely affected on all domains. The subtypes had largely similar mean disease durations (nonsignificant differences between three clusters) but showed considerable differences with respect to their association with demographic and clinical variables. In prevalent disease, PD subtypes are largely characterized by the severity of nondopaminergic features and motor complications and likely reflect complex interactions between disease mechanisms, treatment, aging, and gender. ©2010 *Movement* Disorder Society

Key Words: Parkinson's disease; subtypes; cluster analysis; nondopaminergic; motor complications

Parkinson's disease (PD) is generally known as a movement disorder, but there is an increasing awareness that the clinical spectrum of PD encompasses also many nonmotor domains like cognition and autonomic function.[1] Patients with PD exhibit conspicuous differences in the disease profile and progression rate.[2] This clinical heterogeneity may indicate the existence of subtypes. Identification of subtypes is important as homogeneous groups likely reflect stronger clinical, pathological, and genetic coherence, which, in turn, may facilitate our understanding of involved biological pathways. This may ultimately lead to tailored treatment strategies.

In studies on PD subtypes, patients have often been classified according to predefined criteria [e.g., young versus old age at onset (AO) or dominance of tremor versus bradykinesia/rigidity], after which other clinical variables were compared between the resulting groups.[2] Alternatively, subtypes may be identified through a data-driven approach like cluster analysis (CA), in which the profile of the subtypes arise from the data without a priori clinical assumptions.

Additional Supporting Information may be found in the online version of this article.

***Correspondence to:** Ms. Stephanie M. van Rooden, Department of Neurology, K5Q-92 Leiden University Medical Center, P.O. Box 9600, NL 2300 RC Leiden, The Netherlands; s.m.van_rooden@lumc.nl

Relevant conflicts of interest: Nothing to report.
Full financial disclosures and author roles may be found in the online version of this article.

Received: 19 April 2010; **Revised:** 2 June 2010; **Accepted:** 18 June 2010
Published online 16 November 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.23346

In a systematic review, seven studies were identified that used CA to identify subtypes in the broad clinical spectrum of PD.[3] The majority of those studies identified a subtype with "old age-at-onset and rapid disease progression" and a subtype characterized by "young age-at-onset and slow disease progression." However, the results of these studies were difficult to compare because of methodological differences. Studies focused on different PD domains, which were occasionally combined with other variables such as AO and dopaminergic therapy, resulting in conceptually unclear cluster solutions. Further, the applied CA techniques showed some limitations.[3] The aim of this study is to identify PD subtypes based on motor and nonmotor features of PD, by using a data-driven approach applied to the data of two large independent European cohorts.

Methods

Patients

Data were obtained from two European longitudinal cohorts, the PROfiling PARKinson's disease cohort (PROPARK; n = 415; www.scopa-propark.eu) and the Estudio Longitudinal de pacientes con Enfermedad de Parkinson cohort (ELEP; n = 387).[4] Data of the first and second annual assessment of the PROPARK cohort were collected between May 2003 and April 2007. Data of first annual assessment of the ELEP cohort were collected between March and December 2006.

Patients in both cohorts fulfilled the United Kingdom PD Society Brain Bank criteria for idiopathic PD.[5] The recruitment procedure of the PROPARK cohort has been described in detail elsewhere.[6] In short, patients were recruited from outpatient departments of three university and six regional hospitals and assessed at the Leiden University Medical Centre. Four equally large strata based on AO (</>50 years) and disease duration (</>10 years) were constructed across the cohort to ensure an adequate distribution of factors that are considered to be important determinants of the disease course. Patients in the ELEP cohort were assessed at 20 centers in Spain. Here, six equally large strata were constructed based on sex, AO (30–60/>60 years), and disease duration (</>5 years). In both cohorts, AO was defined as onset of first symptoms as perceived by the patient. Patients who underwent stereotactic surgery were excluded because this intervention may influence the expression of the phenotype. No other selection criteria were applied. All patients gave written informed consent. Studies were approved by the medical ethics committees of the Leiden University Medical Centre (PROPARK) and the Research Committee of the Carlos III Institute of Public Health and the Clinical Research Ethics Committee of the Hospital de la Princesa, Madrid (ELEP).

Measurement Instruments

Except for depression, patients in both cohorts were assessed with the same instruments, which have been described in more detail elsewhere.[7] The following features were assessed: motor symptoms and motor complications (MCs) (SPES/SCOPA sections Motor and Motor Complications), cognitive functioning (SCOPA-COG), autonomic symptoms (SCOPA-AUT), psychotic symptoms (SCOPA-PC, items 1–5), nighttime sleep problems and excessive daytime sleepiness (SCOPA-SLEEP), and depressive symptoms [PROPARK: Beck Depression Inventory (BDI); ELEP: Hospital Anxiety and Depression Scale (HADS)[8]]. Based on results of our previous study[9] motor features were evaluated by a tremor factor including rest and postural tremor, a factor consisting of bradykinesia and rigidity, and two axial factors, i.e., one factor comprising rise, gait, and postural instability, reflecting what is commonly known as postural-instability-gait-difficulty (PIGD), and one factor comprising freezing during on, speech, and swallowing (FOSS). Autonomic dysfunction was evaluated with the gastrointestinal, urological, and cardiovascular items (4–6, 8–16) of the SCOPA-AUT, since these items were considered most relevant for phenotyping. Higher scores reflect poorer function for all scales except the SCOPA-COG; scores of this latter scale were inversed to facilitate interpretability. Patients using antiparkinsonian medication were assessed in "on" state. For each patient, a total levodopa (L-dopa) dose equivalent (LDE) was calculated.[10]

Statistical Analysis

If 25% or more of the items of a scale was missing, this patient was excluded from statistical analyses. If, for a particular patient, less than 25% of the items of a scale were missing, missing data were imputed by the mean value of the nonmissing items of that scale of that patient.

Because severity of PD features increases with longer disease duration and, thus, may act as a potential confounder in the process of identifying distinct phenotypes, each variable was adjusted for disease duration: for each clinical feature, the residual value was obtained from a linear regression with the clinical feature as the dependent and disease duration as an independent variable. Finally, all variables were transformed into *z*-scores to obtain scale invariant outcomes.

Cluster Analysis

To identify subtypes, we performed a model-based CA on data of PROPARK year 1. Clusters were tested

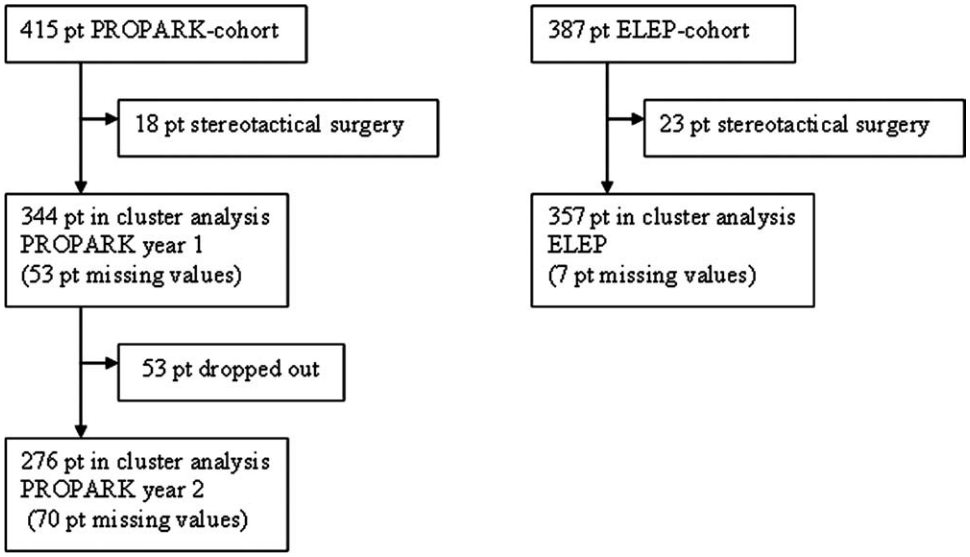

FIG. 1. Flow chart of patients; pt: patients.

for reproducibility in data of PROPARK year 2. The selected model was validated in data of the ELEP cohort by comparing the characteristics of the same model as selected in PROPARK year 1. Details of the clustering method and cluster selection criteria are provided in the supporting information (Supporting Information Appendix).

Characterization of Clusters

Cluster profiles were graphically displayed in heatmaps. Clusters were characterized for clinical-, demographic-, and disease-related variables using data of PROPARK year 1. Differences between clusters were evaluated using ANOVA and χ^2 tests, using Bonferroni correction for multiple testing. A P-value <0.05 was considered significant.

To evaluate which features best discriminated the clusters of the PROPARK and ELEP cohort, a discriminant analysis was performed with the clusters as dependent and the PD features as independent variables. A second discriminant analysis was performed on the PROPARK cohort with demographic- (age and sex) and disease-related variables (AO, disease duration, and LDE) as independent variables.

CAs and cluster visualization were performed with SubtypeDiscovery 1.11 (https://gforge.nbic.nl/projects/subtypediscover/) in R 2.7.0 (www.r-project.org). Other statistical analyses were performed using SPSS 16.0 (SPSS, Inc., Chicago, IL, USA).

Results

Data of 344 patients (1st assessment) and 276 patients (2nd assessment) of the PROPARK cohort and 357 patients of the ELEP cohort were available for analysis (Fig. 1, Table 1). A total of 48 (1%) values were imputed to replace missing values of 45 (13%) of the patients in the dataset of the first PROPARK assessment; 46 (1%) of the values were imputed to replace missing values of 38 (14%) patients in the dataset of the second PROPARK assessment. No data were missing for the ELEP cohort.

Cluster Characteristics

We selected a model with four clusters for further analysis (Supporting Information Appendix and Table; Fig. 2A, Tables 2 and 3). Cluster 1 (49%) was characterized by an overall mild severity in all clinical domains. These patients were relatively young, had a younger AO, and had lower intake of and shorter exposure to L-dopa.

Cluster 2 (13%) was characterized by severe and frequent MCs and moderately severe sleep problems and depressive symptoms. These patients had longer disease duration and higher intake of and longer exposure to dopaminergic medication than patients in other subtypes. Patients in this subtype were comparatively young and had the youngest AO, and the proportion of women was relatively large.

Cluster 3 (30%) showed intermediate severity in nondopaminergic domains, while MCs were mild and less frequent. Patients were relatively old and had a higher AO.

Cluster 4 (8%) included patients who were severely affected in most domains, although tremor was relatively mild. MCs were prominent but less severe than in Cluster 2. This cluster was characterized by relatively high age and AO, long duration of L-dopa use, and a comparatively large proportion of women.

Disease duration did not significantly differ between Clusters 1, 3, and 4. The discriminant analysis showed that motor fluctuations, PIGD, and depression best discriminated the four subtypes of the PROPARK

TABLE 1. Patient characteristics

	PROPARK cohort, year 1	PROPARK cohort, year 2	ELEP cohort
N	344	276	357
Sex, men/women (% men)	226/118 (66)	184/92 (67)	193/164 (54)
Disease duration (yr), mean (SD)	9.9 (6.2)	11.0 (6.2)	7.7 (5.8)
Age (yr), mean (SD)	60.8 (11.3)	61.5 (11.0)	66.2 (11.2)
Age at onset (yr), mean (SD)	50.8 (11.9)	50.5 (11.8)	58.5 (11.7)
H&Y, median (IQR)	2 (2–3)	3 (2–4)	2 (1–2)
Patients on L-dopa, N (%)	223 (65)	199 (72)	267 (75)
Patients on DA, N (%)	234 (68)	192 (70)	242 (68)

H&Y, Hoehn and Yahr stage; IQR, interquartile range; l-dopa, levodopa; DA, dopamine agonists.

cohort. Based on those variables, 251 (73%) patients were correctly classified. Of the demographic- and disease-related variables, AO and total LDE best discriminated the four subtypes; based on these variables, 170 (50%) patients were correctly classified. When AO was substituted by age, 167 (49%) patients were correctly classified.

Cluster Validation

The EII-4 model of the ELEP cohort was similar to the one obtained in the PROPARK cohort (Fig. 2C). Clusters 1–4 comprised 58%, 11%, 27%, and 5% of the patients, respectively. This distribution did not significantly differ from the distribution of the PROPARK subtypes ($\chi^2 = 6.21$; $df = 3$; $P = 0.102$). Discriminant analysis showed that motor fluctuations, PIGD, and autonomic dysfunction correctly classified 286 (80%) patients in the ELEP cohort. By using the same discriminative variables as in the PROPARK cohort (depression instead of autonomic dysfunction), 274 (77%) patients were correctly classified.

Discussion

Conspicuous clinical heterogeneity exists among patients with PD, which is most likely attributable to differences in the mechanisms that underlie PD and complications of dopaminergic treatment. In this study, we aimed to identify clinical subtypes in PD and found four subtypes. The strengths of this study are the use of a data-driven approach in a large number of patients who were extensively characterized on a broad array of motor and nonmotor domains. Moreover, the results were validated in an independent Spanish cohort, in which patients were assessed by similar measurement instruments as in the PROPARK cohort. The latter finding emphasizes the transcultural validity of these subtypes and underscores their robustness despite differences in sample collection and sample characteristics.

By comparing the profiles of the subtypes (Fig. 2), it appeared that two subtypes (Clusters 1 and 4) differed only by a severity gradient (benign versus malignant), regardless the clinical domain of interest. However, the profile of the other subtypes (Clusters 2 and 3) showed that, based on severity, certain clinical domains grouped together. This grouping of clinical domains is consistent with the results from an earlier study on the coherency of motor and nonmotor domains.[7] Two subtypes (Clusters 3 and 4) had prominent involvement of PIGD, cognitive impairment, autonomic dysfunction, psychosis, daytime sleepiness, and depression, which are predominantly nondopaminergic features (PND complex). In our previous study, this PND complex was associated with both disease severity and age and most likely reflects advancing disease.[7] PIGD has been previously identified as an important motor phenotype associated with cognitive decline, a higher risk for depression, and a more progressive course.[11–15] When PD is viewed from a broader perspective, PIGD may actually be the motor component of the much larger PND complex.

A tremor dominant subtype associated with a more favorable disease course[11–13] was not identified in this study. Recently, a clinicopathologic study showed that tremor is not an independent indicator of a benign disease course.[16] In addition, a longitudinal study showed that with the development of PIGD, the number of tremor predominant patients gradually decreased.[14] Hence, the prevalent character of both cohorts used in this study (mean disease durations 10 and 8 years) may explain why a tremor dominant subtype was not identified.

The two subtypes with prominent involvement of the PND complex (Clusters 3 and 4) had higher AO and age than the two other subtypes (Clusters 1 and 2). Because of the cross-sectional nature of this study and largely similar mean disease duration of the subtypes, it is impossible to unravel whether AO or age most strongly determined disease severity. In the previous CA studies, cluster profiles were characterized on higher AO and rapid disease progression.[3] However, in longitudinal studies on cognitive impairment in PD where both AO and age were taken into account, only the latter was found related to faster rate of cognitive

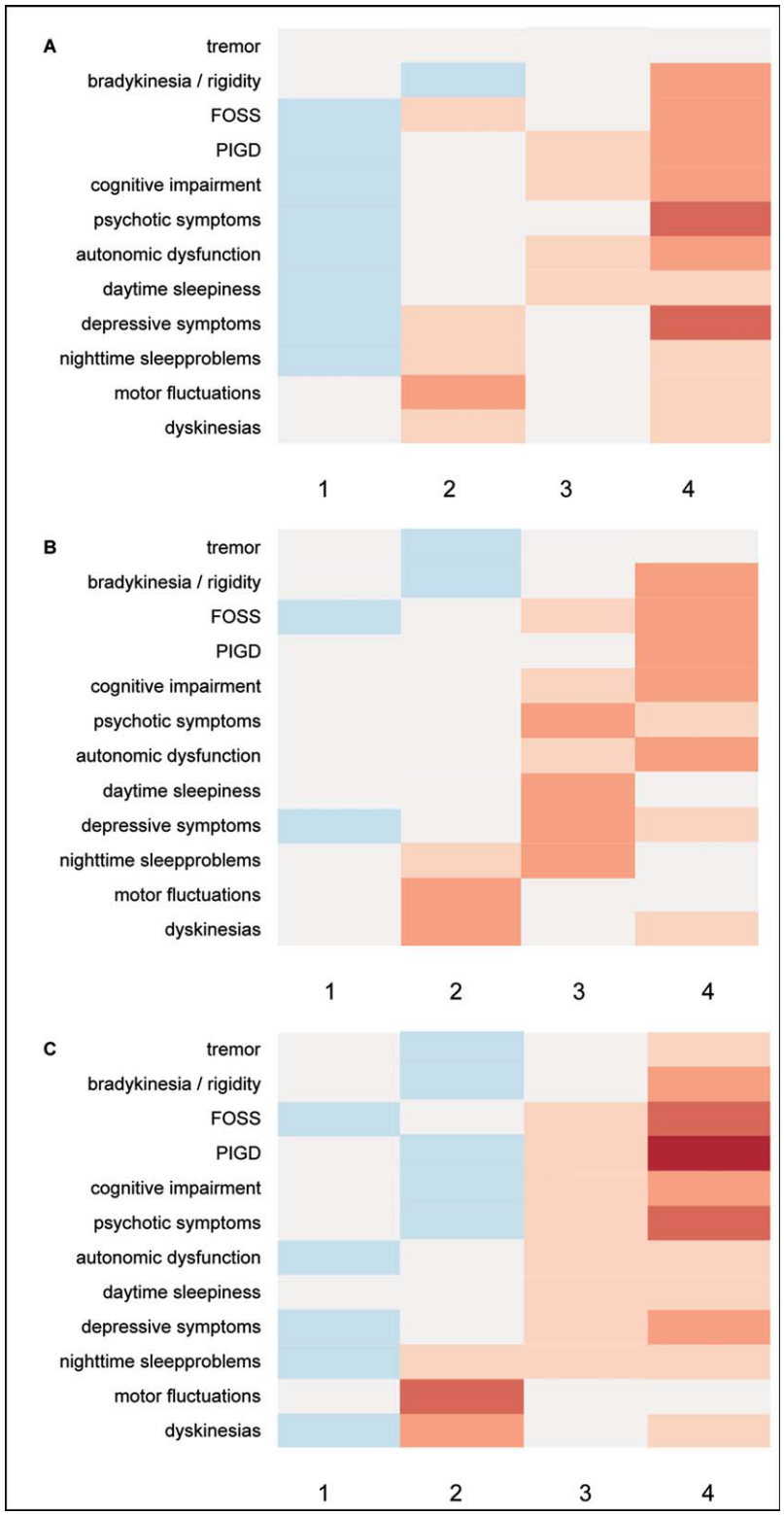

FIG. 2. Heatmaps of the EII—4 models. A, PROPARK cohort year 1; B, PROPARK cohort year 2; C, ELEP cohort. FOSS: an axial motor factor consisting of "freezing during on," "swallowing," and "speech." PIGD: an axial motor factor consisting of "rise," "gait," and "postural instability." Each column represents a cluster. The colors indicate whether the cluster mean severity of a symptom is lower (blue color) or higher (red color) than the mean of the total group.

TABLE 2. Cluster characteristics of the four clusters of PROPARK cohort year 1

Symptoms (range scores)	Cluster 1	Cluster 2	Cluster 3	Cluster 4
N (%)	169 (49)	45 (13)	104 (30)	26 (8)
Bradykinesia/rigidity (0–12)	4.7 (1.7)[a–c]	3.8 (1.7)[b–d]	5.5 (1.8)[a,c,d]	7.4 (2.0)[a,b,d]
Tremor (0–12)	3.6 (1.9)	3.1 (1.8)[b]	4.1 (2.2)[a]	3.1 (1.5)
PIGD (0–9)	1.2 (1.1)[a–c]	2.2 (1.5)[b–d]	3.1 (1.7)[a,c,d]	4.8 (1.8)[a,b,d]
Cognitive impairment (0–43)	14.7 (5.5)[b,c]	15.8 (4.8)[b,c]	19.7 (5.4)[a,c,d]	25.7 (5.1)[a,b,d]
Psychotic symptoms (0–15)	1.0 (1.1)[a–c]	2.4 (1.7)[c,d]	2.3 (1.7)[c,d]	4.8 (2.0)[a,b,d]
Autonomic dysfunction (0–36)	7.1 (4.1)[a–c]	11.3 (4.4)[b–d]	13.5 (5.1)[a,d]	15.7 (6.1)[a,d]
Daytime sleepiness (0–18)	2.9 (2.6)[a–c]	4.4 (3.9)[b,d]	7.7 (3.5)[a,d]	6.4 (3.0)[d]
Depression (0–63)	6.7 (4.0)[a–c]	13.9 (5.5)[b–d]	10.6 (4.9)[a,c,d]	21.5 (8.3)[a,b,d]
Nighttime sleep problems (0–15)	3.1 (3.1)[a–c]	7.7 (3.0)[b,d]	4.5 (3.8)[a,c,d]	7.0 (3.6)[b,d]
Motor fluctuations (0–6)	0.3 (0.8)[a,c]	2.8 (1.2)[b–d]	0.4 (0.9)[a,c]	1.2 (1.5)[a,b,d]
Dyskinesias (0–6)	0.5 (1.3)[a,c]	2.3 (1.9)[b,d]	0.4 (1.2)[a,c]	1.8 (1.8)[b,d]
FOSS (0–9)	1.5 (1.2)[a–c]	3.0 (1.9)[c,d]	2.6 (1.6)[c,d]	4.2 (1.5)[a,b,d]

Means (SD) presented for all variables. For all PD signs and symptoms: higher scores reflect more problems.
In the columns presenting data of clusters 1–4 only significant differences are indicated.
[a]Significant difference ($P < 0.05$) with cluster 2 (ANOVA).
[b]Significant difference ($P < 0.05$) with cluster 3 (ANOVA).
[c]Significant difference ($P < 0.05$) with cluster 4 (ANOVA).
[d]Significant difference ($P < 0.05$) with cluster 1 (ANOVA).
PIGD, rise, gait, postural instability; FOSS, freezing, speech, swallowing.

decline.[17,18] Compelling data from other studies suggest that advancing age is also an important determinant of clinical progression in PD, as reflected by PIGD, cognitive decline, and hallucinations.[16,19] Notably, clusters with a similar AO and age also showed remarkable differences in severity of the PND complex (Cluster 1 versus 2 and 3 versus 4); underscoring that in addition to aging, other disease-modifying factors must play a role in the progression of PD. Levy[17] proposed an appealing model in which nondopaminergic manifestations result from biologic interaction between the disease process and aging. The PND complex likely reflects advancing Lewy body pathology,[20] which has been related to the enhanced level of alpha-synuclein.[21] Interestingly, the latter is influenced by both the disease process and aging.[22] Aging may

TABLE 3. Cluster characteristics on variables that were not included in the cluster analysis (PROPARK cohort year 1)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Disease duration (yr)[a]	9.1 (6.3)[b]	12.3 (5.3)[c]	10.3 (6.4)	9.9 (4.6)
Age (yr)[a]	57.8 (10.6)[d,e]	57.8 (9.6)[d]	65.9 (10.3)[b,c]	64.7 (14.2)[c]
Age onset (yr)[a]	48.7 (11.5)[d]	45.4 (9.3)[d,e]	55.6 (10.8)[b,c]	54.9 (14.3)[b]
Sex, men/women (% men)[f]	119/50 (70)[b]	22/23 (49)[c]	73/31 (70)	12/14 (46)
First- and secnd-degree relatives with PD, N (%)[f]	42 (25)	13 (29)	22 (21)	5 (19)
Side of onset, N (%)[g]				
Left	54 (38)	14 (38)	31 (44)	3 (33)
Right	80 (56)	19 (51)	34 (49)	4 (44)
Both	9 (6)	4 (11)	5 (7)	2 (22)
H&Y, median (IQR)[f]	2 (2–2)[b,d,e]	3 (2–3)[c,e]	3 (2–3)[c,e]	4 (3–4)[b–d]
Presence of motor fluctuations (% yes)[f]	14[b,e]	96[c–e]	19[b]	42[b,c]
Presence of dyskinesias (% yes)[f]	14[b,e]	69[c,d]	14[b,e]	53[c,d]
LDE (mg)[a]	436 (379)[b]	989 (556)[c–e]	570 (375)[b]	606 (394)[b]
LDE L-dopa (mg)[a]	230 (283)[b,d,e]	585 (403)[c,d]	377 (334)[b,c]	432 (323)[c]
LDE DA (mg)[a]	205 (226)[b]	405 (247)[c–e]	193 (191)[b]	173 (210)[b]
Exposure to L-dopa (yr)[a]	2.7 (4.2)[b,d,e]	7.3 (5.2)[c,d]	4.9 (5.3)[b,c]	7.5 (4.6)[c]
Exposure to DA (yr)[a]	2.7 (3.2)[b]	5.4 (3.2)[c,d]	3.5 (3.9)[b]	4.2 (3.9)
Patients on clozapine, N (%)[f]	4 (2)	3 (7)	3 (3)	3 (12)
Patients on amantadine, N (%)[f]	46 (27)[e]	17 (38)	42 (40)	15 (58)[c]

Means (SD) presented for all variables unless stated otherwise. In the columns presenting data of clusters 1–4 only significant differences are indicated
[a]ANOVA, *post hoc*, and *t* test with Bonferroni correction for multiple testing.
[b]Significant difference ($P < 0.05$) with cluster 2.
[c]Significant difference ($P < 0.05$) with cluster 1.
[d]Significant difference ($P < 0.05$) with cluster 3.
[e]Significant difference ($P < 0.05$) with cluster 4.
[f]χ^2 test, Bonferroni correction for multiple testing.
[g]Kruskal-Wallis test.
H&Y, Hoehn and Yahr stage; IQR, interquartile range; LDE, levodopa dose equivalent; L-dopa, levodopa; DA, dopamine agonists.

influence the progression of PD through shared involvement of processes fundamental to neuronal vitality, including the maintenance of protein homeostasis and mitochondrial function.[23]

Our results further identified two subtypes with pronounced MCs (Clusters 2 and 4). Young AO, female gender, higher (cumulative) L-dopa dose, longer duration, and higher severity of the disease have been reported risk factors for MCs.[24–26] Except for disease severity, all the reported determinants were identified in Cluster 2. Conversely, in Cluster 4, only female gender and disease severity were identified determinants. Because younger AO is considered a risk factor for development of MCs [24,27] it was surprising that also a substantial number of patients with an older AO clustered in a subtype with pronounced MCs. The finding that female gender was the only common risk factor in both subtypes with MCs highlights that subtype-specific interactions between medication and disease-related variables may result in a certain susceptibility to MCs.

Subtypes with prominent MCs also exhibited more severe mood and sleep disturbances. MCs may directly affect sleep and mood, but clustering of these symptoms may also be linked by factors related to female gender, since insomnia and depression are more frequent in women.[28]

In conclusion, PD subtypes are largely characterized by the severity on two axes: the PND complex and MCs (Fig. 3). Axial function and MCs, derivatives of both axes, discriminated best between subtypes, providing further support for this classification. Our findings further show that subtype expression is based on complex interactions between disease mechanisms, treatment, aging, and gender. These findings may have consequences for epidemiologic studies and trials, which hitherto have considered PD a homogeneous disorder, since risk factors and treatment effects may be subtype-specific. The contribution of genetic factors in the subtypes was not evaluated, because a previous screening for mutations in the *Parkin*, *DJ-1*, *PINK1*, *LRRK2*, and *SNCA* gene in the PROPARK cohort showed that pathogenic variations were demonstrated in only 4% of the patients.[29] Hence, it seems unlikely that differences between clusters can be explained by these known mutations.

With the model-based CA, we have tried to avoid the disadvantages of hierarchical and k-means CA. In contrast to these methods, model-based CA is not sensitive to outliers, and the model and number of clusters is not arbitrarily chosen but guided by a fit statistic.[30–32] This method estimates different models, varying in assumptions on the distribution of the clusters, thereby resulting in more precise solutions.[31,32] And, more importantly, we were able to validate the results in an independent sample.

The recruitment strategy of PROPARK, aimed at obtaining an equal distribution of AO and disease duration, may limit generalization of our findings. However, this study aimed to identify subtypes and not their prevalence. Furthermore, despite different sample characteristics of the ELEP cohort, we identified similar subtypes. Because disease duration may influence the phenotypic expression of the disease, we aimed to correct each variable for the influence of disease duration. Small differences in mean disease duration between the subtypes nevertheless remained (range, 9.1-12.3 years), which was significant only between Clusters 1 and 2. Thus, it seems unlikely that these differences play a decisive role in the differential expression of subtypes. Nevertheless, the cross-sectional design of the study precludes the possibility to determine whether patients within clusters had followed similar disease courses. Moreover, given the prevalent nature of both cohorts, it is impossible to determine which factors at disease onset predict the subsequent development into particular subtypes. Collectively, this knowledge could be important for the development of tailored treatment strategies and, also, highlights the need for longitudinal studies on incident cases. ■

Acknowledgments: The authors thank the ELEP Group by its contribution to the database. The SCOPA-PROPARK project was funded by the Michael J. Fox Foundation (PD-subtypes program), the Prinses Beatrix Fonds (project no. WAR05-0120), the van Alkemade-Keuls Foundation, and the Dutch PD Society. The ELEP study was supported by a grant of the Intramural Research Program, Carlos III Institute of Health (EPY1271/05).

References

1. Chaudhuri KR, Healy DG, Schapira AH. Non-motor symptoms of Parkinson's disease: diagnosis and management. Lancet Neurol 2006;5:235–245.
2. Foltynie T, Brayne C, Barker RA. The heterogeneity of idiopathic Parkinson's disease. J Neurol 2002;249:138–145.
3. van Rooden SM, Heiser WJ, Kok JN, Verbaan D, Hilten JJv, Marinus J. The identification of Parkinson's disease subtypes using cluster analysis: a systematic review. Mov Disord 2010;25: 969–978.
4. ELEP group. A longitudinal study of patients with Parkinson's disease (ELEP): aims and methodology. Rev Neurol 2006;42: 360–365.
5. Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 1988;51:745–752.
6. Verbaan D, van Rooden SM, Visser M, Marinus J, van Hilten JJ. Nighttime sleep problems and daytime sleepiness in Parkinson's disease. Mov Disord 2008;23:35–41.
7. van Rooden SM, Visser M, Verbaan D, Marinus J, van Hilten JJ. Patterns of motor and non-motor features in Parkinson's disease. J Neurol Neurosurg Psychiatry 2009;80:846–850.
8. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983;67:361–370.
9. van Rooden SM, Visser M, Verbaan D, Marinus J, van Hilten JJ. Motor patterns in Parkinson's disease: a data-driven approach. Mov Disord 2009;24:1042–1047.
10. Esselink RA, de Bie RM, de Haan RJ, et al. Unilateral pallidotomy versus bilateral subthalamic nucleus stimulation in PD: a randomized trial. Neurology 2004;62:201–207.
11. Jankovic J, McDermott M, Carter J, et al. Variable expression of Parkinson's disease: a base-line analysis of the DATATOP cohort. The Parkinson Study Group. Neurology 1990;40:1529–1534.

12. 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;76:343–348.
13. 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;15: 379–382.
14. Alves G, Larsen JP, Emre M, Wentzel-Larsen T, Aarsland D. Changes in motor subtype and risk for incident dementia in Parkinson's disease. Mov Disord 2006;21:1123–1130.
15. Burn DJ, Rowan EN, Minett T, et al. Extrapyramidal features in Parkinson's disease with and without dementia and dementia with Lewy bodies: a cross-sectional comparative study. Mov Disord 2003;18:884–889.
16. Selikhova M, Williams DR, Kempster PA, Holton JL, Revesz T, Lees AJ. A clinico-pathological study of subtypes in Parkinson's disease. Brain 2009;132:2947–2957.
17. Levy G. The relationship of Parkinson disease with aging. Arch Neurol 2007;64:1242–1246.
18. Aarsland D, Kvaloy JT, Andersen K, et al. The effect of age of onset of PD on risk of dementia. J Neurol 2007;254: 38–45.
19. Kempster PA, Williams DR, Selikhova M, Holton J, Revesz T, Lees AJ. Patterns of levodopa response in Parkinson's disease: a clinico-pathological study. Brain 2007;130:2123–2128.
20. Lang AE. The progression of Parkinson disease: a hypothesis. Neurology 2007;68:948–952.
21. Kim C, Lee SJ. Controlling the mass action of alpha-synuclein in Parkinson's disease. J Neurochem 2008;107:303–316.
22. Brundin P, Li JY, Holton JL, Lindvall O, Revesz T. Research in motion: the enigma of Parkinson's disease pathology spread. Nat Rev Neurosci 2008;9:741–745.
23. van Ham TJ, Thijssen KL, Breitling R, Hofstra RM, Plasterk RH, Nollen EA. C. elegans model identifies genetic modifiers of alpha-synuclein inclusion formation during aging. PLoS Genet 2008;4: e1000027.
24. Bhidayasiri R, Truong DD. Motor complications in Parkinson disease: Clinical manifestations and management. J Neurol Sci 2008; 266:204–215.
25. Jankovic J. Motor fluctuations and dyskinesias in Parkinson's disease: clinical manifestations. Mov Disord 2005;20(suppl 11): S11–S16.
26. Paus S, Gadow F, Knapp M, Klein C, Klockgether T, Wullner U. Motor complications in patients form the German Competence Network on Parkinson's disease and the DRD3 Ser9Gly polymorphism. Mov Disord 2009;24:1080–1084.
27. Linazasoro G. New ideas on the origin of L-dopa-induced dyskinesias: age, genes and neural plasticity. Trends Pharmacol Sci 2005; 26:391–397.
28. Eichling PS, Sahni J. Menopause related sleep disorders. J Clin Sleep Med 2005;1:291–300.
29. Macedo MG, Verbaan D, Fang Y, et al. Genotypic and phenotypic characteristics of dutch patients with early onset Parkinson's disease. Mov Disord 2009;24:196–203.
30. Colas F, Meulenbelt I, Houwing-Duistermaat J, et al. A scenario implementation in R for subtype discovery examplified on chemoinformatics data. In: Margaria T,Steffen B, editors. Leveraging Applications of Formal Methods, Verification and Validation, Communications in Computer and Information Science. Berlin, Heidelberg: Springer-Verlag, 2008. p 669–683.
31. Colas F, Meulenbelt I, Houwing-Duistermaat J, Slagboom PE, Kok JN. A comparison of two methods for finding groups using heatmaps and model-based clustering. Presented at: Proceedings of the 27th Annual International Conference of the British Computer Society (SGAI), AI 2007, Cambridge 2007;119–131.
32. Fraley C, Raftery AE. How many clusters? Which clustering method? Answers via model-based cluster analysis. Computer J 1998;41:578–588.