Movement Disorders
Vol. 00, No. 00, 2010, pp. 000–000

Review

The Identification of Parkinson's Disease Subtypes Using Cluster Analysis: A Systematic Review

Stephanie M. van Rooden, MSc,[1]* Willem J. Heiser, PhD,[2] Joost N. Kok, PhD,[3,4] Dagmar Verbaan, PhD,[1] Jacobus J. van Hilten, MD, PhD,[1†] and Johan Marinus PhD[1†]

[1]*Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands*
[2]*Department of Psychology, Faculty of Social and Behavioral Sciences, Leiden University, The Netherlands*
[3]*Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands*
[4]*Department of Medical Statistics and Bioinformatics/Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands*

Abstract: The clinical variability between patients with Parkinson's disease (PD) may point at the existence of subtypes of the disease. Identification of subtypes is important, since a focus on homogeneous groups may enhance the chance of success of research on mechanisms of disease and may also lead to tailored treatment strategies. Cluster analysis (CA) is an objective method to classify patients into subtypes. We systematically reviewed the methodology and results of CA studies in PD to gain a better understanding of the robustness of identified subtypes. We found seven studies that fulfilled the inclusion criteria. Studies were limited by incomplete reporting and methodological limitations. Differences between studies rendered comparisons of the results difficult. However, it appeared that studies which applied a comparable design identified similar subtypes. The cluster profiles "old age-at-onset and rapid disease progression" and "young age-at-onset and slow disease progression" emerged from the majority of studies. Other cluster profiles were less consistent across studies. Future studies with a rigorous study design that is standardized with respect to the included variables, data processing, and CA technique may advance the knowledge on subtypes in PD. © 2010 Movement Disorder Society

Key words: systematic review; Parkinson's disease; subtype; subgroup; cluster analysis

INTRODUCTION

Parkinson's disease (PD) is clinically characterized by a broad spectrum of motor and nonmotor manifestations.[1] There is, however, considerable variability between patients with PD concerning the clinical phenotype, which may indicate that there are subtypes of the disease.[2] Identification of PD subtypes may be important for research on underlying disease mechanisms, since homogeneous groups of patients are more likely to share pathological and genetic features. Second, the identification of subtypes may ultimately lead to tailored management strategies.

In 2002, Foltynie et al. explored the concept of heterogeneity in PD from several perspectives, including the clinical phenotype. It appeared that on a clinical level, subtypes generally were classified according to prespecified or hypothesized criteria that were based on predominant clinical features (tremor, bradykinesia/rigidity, postural instability, cognitive impairment), age-at-onset, and rate of progression.[2] The potential bias inherent to this approach is that less obvious or unexpected patterns may be missed.

When no a priori structure of the data is known, a data-driven method like cluster analysis (CA) may be a very suitable method to study subtypes. CA can be used to explore whether individuals can be classified into groups in such a way that differences within a group of

†Both authors contributed equally
*Correspondence to: S.M. van Rooden, MSc, Department of Neurology, K5Q-92 Leiden University Medical Center, P.O. Box 9600, NL-2300 RC Leiden. E-mail: s.m.van_rooden@lumc.nl

Potential conflict of interest: None.

Received 15 December 2009; Revised 29 January 2010; Accepted 24 February 2010

Published online in Wiley InterScience(www.interscience. wiley.-com). DOI: 10.1002/mds.23116

patients are small, while the differences between groups are large.[3] In this so-called unsupervised classification, the characteristics of the subtypes arise from the data. Next to this apparent advantage, it is important to be aware that the results of CA are dependent on choices that are made in the process of analysis, such as the variables selected for analysis, the clustering technique, and the number of clusters.[2,3] In the review on heterogeneity by Foltynie et al.,[2] only one study was reported that explored the existence of clinical subtypes in PD by CA, but several other studies have been published since. In this systematic review, we evaluated which clinical subtypes of PD have been identified by CA, discuss their robustness and reflect on the methodological issues that may influence the results.

METHODS

Search Strategy

The following databases were searched on April 27th, 2009: PubMed (1949 to April 2009), EMBASE (OVID-version, 1980 to April 2009), and Web of Science (1945 to April 2009). The search consisted of the combination of the following terms: (1) Parkinson disease, and (2) heterogeneity, cluster analysis, k-means, self organizing maps, mixture models, data driven, or cluster combined with fuzzy, kernel based, or hierarchical. The search strategy was optimized for all consulted databases, taking into account the differences of database-specific technical variations. Additionally, the reference lists of all included articles were searched. The results were limited to articles in English, German, and Dutch.

Methods of Review

The selection procedure was performed by two independent reviewers (SR, JM). This assessment was not blind with respect to authors or institutions. The obtained articles were first screened by title, after which the abstracts of potentially relevant articles were reviewed. If the abstract was considered relevant, the full text of the article was studied. Studies were included if (1) the study population consisted of PD patients, (2) the existence of subtypes was evaluated by CA, and (3) the CA was based on clinical characteristics. Studies that focused on a specific domain of the disease were excluded to avoid incomparable findings.

Data Extraction

Study methods and results were abstracted by one of the authors (SR). Whenever information was incomplete or unclear, the authors of that study were contacted with the request to provide additional information. Since the outcomes of CA are dependent on the applied method, the variables included in the analysis, the characteristics of the involved sample, and the way data are processed,[2,3] all information pertaining to these issues was collected and recorded on a standard score sheet. Specifics of these issues are detailed below.

Sample Characteristics.

Characteristics of the included sample may affect the cluster structure. Therefore, it is important to know whether a sample represents a population of interest, e.g., de novo patients.[4] Additionally, the sample size is of relevance for the generalizability of the findings, since small studies will yield less precise estimates.

Variables Selected for CA.

It is important to select variables that are considered relevant in phenotyping and discriminating subtypes of the disease.[4]

Data Preprocessing.

Variables are generally measured with different units of measurement. To adjust for differences due to scaling of the measurement instrument, the variables are usually standardized, for instance by transformation into Z-score or range.

Clustering Algorithm.

CA can be performed by different techniques, of which hierarchical clustering and K-means clustering are the most common. K-means CA is a partitioning method, meaning that patients are assigned to a prespecified (K) number of clusters without a hierarchical structure. K initial clusters are formed, after which patients are assigned to the cluster they most resemble. Subsequently, cluster means are calculated after which the distance to each cluster mean is calculated for each patient. Patients are reassigned to another cluster if they are closer to that cluster mean than to the mean of the cluster they were allocated to in the previous step. In the next step cluster means are recalculated followed by calculation of the patients' distances to the cluster means. This iterative process stops when no patients need to be reassigned, and the optimal solution for the clusters is achieved.[3]

Local Optimum.

Cluster methods that involve iterative processes stop when an optimum is achieved. This optimum, however, may not be the optimal solution among all possible solutions but represent a so-called local optimum. The process of partitioning is sensitive to the starting points. To reduce the risk of ending in a local optimum, the clustering can be repeated a number of times with randomly chosen different starting points, after which the optimal solution is selected.[5]

Determination of Number of Clusters.

The validity of the cluster result is dependent on the estimation of the number of clusters.[6] In K-means CA the number of clusters has to be indicated by the investigator. This optimal number can be estimated by statistical methods, of which the Calinski and Harabasz index (pseudo F-statistic)[7] was considered most appropriate.[8] It is important to report on which statistical grounds or other rationale the choice of the number of clusters was determined.

Cluster Validation.

Validation of the results is an important step since CA methods can always generate a division in clusters, which do not necessarily represent true subtypes.[6] The results are preferably replicated in an independent sample. Other methods are cross-validation, the demonstration of stability, and face validity.[2,4,9]

Interpretation of Cluster Results.

The final goal of CA is to evaluate whether the cluster sizes and profiles are meaningful and clinically interpretable. A discriminant function analysis can be performed to evaluate which combination of variables best differentiates the subtypes. In contrast, F-values only provide insight in the magnitude of univariate differences between clusters. In post-hoc analyses, clusters can be further characterized on variables which were not included in the CA. This may provide insight in factors that play a role in the development of a specific phenotype profile.

RESULTS

The search strategy yielded 259 studies of which eight were judged eligible by at least one reviewer. The overall inter-reviewer agreement for the inclusion of the studies by reviewing titles and abstracts was 99.6% (Cohen's Kappa 0.93). Differences were reconciled by consensus, after which eight studies were considered eligible for this review. Eight other studies performed CA in PD but focused on a specific domain of the disease and were therefore excluded.[10–17] One eligible study was published as a congress abstract[18] and as an article in a peer-reviewed journal.[19] We included only the latter publication, resulting in seven included studies (Table 1). The selection process is presented in Figure 1.

Methodological Appraisal

Study characteristics are presented in Table 1 and 2.

Sample Characteristics.

All but one study analyzed >100 patients, with samples ranging from 44[23] to 176[25] patients. All studies applied validated criteria to diagnose PD.[26–28] Four studies applied additional inclusion criteria, which were based on disease duration in three[20,23,24] and on disease severity in one study.[22] One study did not provide patient characteristics for the total group.[24] Only in the study of Post et al. age-at-onset was specified, which was 65.1 (10.4) [mean (SD)].[20] In the other studies, except that of Gasparoli et al.,[24] information on mean age at onset can be obtained by subtracting disease duration from age, but the SD is unknown. Mean age-at-onset was 55.7[25], 56.6[22], 62[23], 62.2[19], and 65.8[21] years.

Variables Selected for CA.

Studies showed large variability, not only in included variables, but also in measures or measurement instruments for a similar clinical domain (Table 2 and Supporting Appendix). Six of the seven studies included motor symptoms assessed by the Unified Parkinson's Disease Rating Scale (UPDRS), although different sum and subscores were used.[19,20,22–24] Motor symptoms were combined with measures of cognition in five studies,[19,20,22,23,25] of which four also included depression[19,20,22,25] and age-at-onset.[19,20,22,25] Three of these latter four studies also included a measure of disease progression.[19,20,22] Other variables were less frequently included (Table 2).

Data Preprocessing.

In five studies scores were standardized before analysis,[19,20,22,23,25] which concerned transformation into Z-scores in three.[19,20,22] One study presented results

TABLE 1. *Methodology of cluster analyses on the spectrum of PD*

Methodological steps	Reijnders 2008[19]	Post 2008[20]	Schrag 2006[21]	Lewis 2006[22]	Dujardin 2004[23]	Gasparoli 2002[24]	Graham 1999[25]
No. of patients	346	131	124	120	44	103	176
Inclusion criteria, in addition to PD	None	De novo	None	H&Y I–III	Time since diagnosis 3 yrs	Time since diagnosis 5 yrs	None
Sample characteristics (means, SD)							
Age, yrs	70.4[a]	66.7 (10.4)	71.9 (11.0)[b]	64.4 (9.3)	66 (median)	NS	63.2 (10.2)
Disease duration, yrs	8.2	1.7 (0.9)	6.1 (4.4)[b]	7.8 (5.4)	4 (median)	NS	7.5 (6.4)
H&Y	2.7	1.8[b]	2.8[b]	2.1[b]	NS	NS	NS
Data preprocessing	Z-scores	Z-scores[b]	No[b]	Z-scores[b]	NS	[c]	NS
Clustering algorithm	K-means	K-means	K-means	K-means	K-means	[c]	K-means
Basis of the determination of the number of clusters	Non-statistical[d]	No criteria	No criteria	Non-statistical[e]	Statistical[f]	Non-statistical[g]	No criteria
Cluster validation on independent sample	Yes	No	No	No	No	No	No
Evaluation of discriminative variables	No	Yes, F-values	Yes, F-values	No	Yes, discriminant analysis	No	Yes, F-values
Post hoc analysis of variables not included in CA	Yes	Yes	Yes	Yes	Yes	Yes	Yes

PD: Parkinson's disease, H&Y; Hoehn and Yahr stage, yrs: years, SD: standard deviation, NS: Not Specified, CA: cluster analysis.

[a]Total sample is built op from two samples: $N = 224$, age 73.2 (8.4), disease duration 9.5 (5.7), H&Y 2.8 (1.0); $N = 122$, age 65.3 (10.0), disease duration 6.7 (5.0), H&Y 2.4 (0.8) (mean (SD)).

[b]Personal communication.

[c]Unknown.

[d]Based on changes in cluster distances in successive steps; face validity.

[e]Based on a clear distinction in clusters with a sufficiently large size.

[f]Pseudo F-statistic, Cubic Clustering Criterion, squared correlation ratio.

[g]Aim to find two clusters.

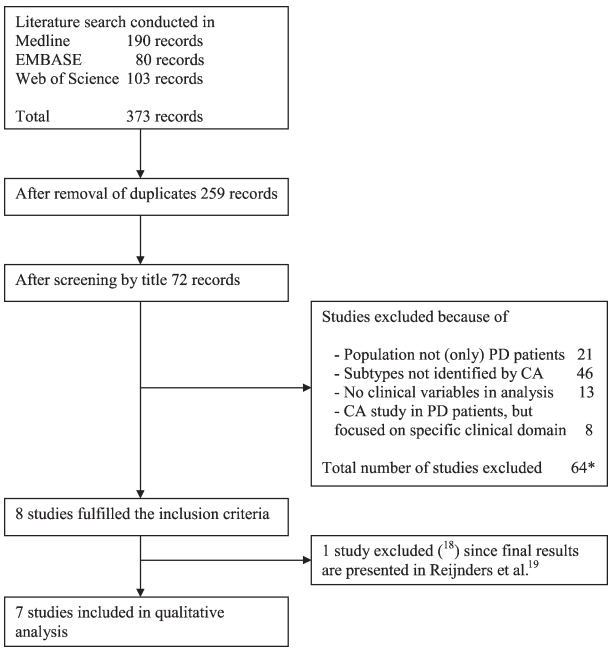

FIG. 1. Flow chart of study search and selection. PD: Parkinson's Disease; CA: cluster analysis. *Sum of studies excluded per criterion does not equal the total number of excluded studies, since some studies were excluded because of more than one exclusion criterion.

based on an analysis with unstandardized scores. However, repeating the analysis with scores transformed into Z-scores revealed similar findings (personal communication).[21]

Clustering Algorithm.

Information on the applied CA method was not reported in one study;[24] all other studies used K-means CA.

Local Optimum.

None of the studies reported how they tried to avoid local optima.

Determination of Number of Clusters.

One study determined the optimal number of clusters on statistical grounds by three different indices (Pseudo F-statistic, Cubic Clustering Criterion, squared correlation ratio).[23] One study indicated that the aim was to find two clusters.[24] Two studies evaluated two to five-cluster solutions after which the optimal number was determined based on changes in cluster distances in successive steps and face validity in one study[19] and based on a clear distinction in clusters with a sufficiently large size in the other study.[22] Two studies evaluated both two and three-cluster solutions and one a five-cluster solution without determining the optimal number of clusters.[20,21,25]

Cluster Validation.

Only the study of Reijnders et al.[19] verified the cluster solution in a second sample; based on the cluster

TABLE 2. *Variables included in the cluster analyses*

	Reijnders	Post	Schrag	Lewis	Dujardin	Gasparoli	Graham
Motor symptoms							
Onset symptom							
Cognitive impairment			[a]				
Depression							
Apathy							
Hallucinations							
Motor complications			[a]				
Time to MF/dysk, years							
Time to falls, years							
Disease progression			[a]				
Disease severity							
Disease duration, years							
Age at onset, years			[a]				
Age, years			[a]				
Medication							
ADL							

MF: motor fluctuations, dysk: dyskinesias, ADL: activities of daily living.

Grey and black marked variables were included in the cluster analysis; the variables marked in black were discriminative variables that emerged from discriminant analysis or had large F-values. Details of measures and measurement instruments are provided in the Supplementary Appendix.

[a]A cluster analysis was performed with all marked variables and with only the variables marked with[a]. Both analyses resulted in similar solutions (personal communication).

TABLE 3. *Identified cluster profiles and cluster sizes (% of total sample)*

Cluster characteristic	Reijnders	Post	Schrag	Lewis	Dujardin	Gasparoli	Graham
Old age at onset/rapid disease progression	6	40	64	17		39	21
Young age at onset/slow disease progression	29	34	36	41		61	
Intermediate onset, anxiety, depression		27					
Tremor dominant	47			17			
Non-tremor dominant	17			26			
More severe motor & cognitive impairment					36		32
Mild motor & mild cognitive impairment					59		
Motor only							47
	100	100	100	100	95[a]	100	100

[a]Sum of percentages does not equal 100, because one of the clusters was discarded.

means and covariance matrices of the first sample they evaluated the probability of a cluster membership of patients in a second sample. That study, as well as two other studies,[20,22] also evaluated post-hoc how variables that were not included in the CA differed between the clusters. Lewis et al.[22] also evaluated to which extent patients consistently grouped together in the 3, 4, and 5 cluster solutions. Four studies did not report any information about validation of the cluster solution.[21,23–25]

Interpretation of the Results

The studies identified five[25], four[19,22], three[20,23], and two[21,24] clusters. The sizes of all clusters were >5% of the total sample in all studies, except one in the study by Dujardin et al.[23] In the latter study, however, this small cluster was discarded after the analysis because authors concluded that the patients in this cluster had developed Alzheimer's disease (Table 3).

It is important to emphasize that all studies included a different set of variables in the CA (Table 2). Since both the number and nature of included variables have a prominent role in the outcomes of the CA, all following results should be considered in the context the total set of variables that was included in the CA in each study. Consequently, the findings of different studies have limited comparability.

Discriminative Variables.

Only one study performed a discriminant analysis to evaluate which variables best discriminated the subtypes.[23] Three other studies presented F-values, indicating which variables showed large differences between clusters.[20,21,25] It appeared that in two studies cognitive dysfunction, specifically executive dysfunction, best differentiated between clusters.[23,25] In two other studies, age-at-onset had large F-values,[20,21] in combination with age and axial motor symptoms in one,[20] and levodopa dose in the other study.[21] As stated above, these discriminating variables should be considered in the context of the total set of variables, which varied betweens studies (Table 2).

Cluster Profiles.

The majority of studies reported two clusters with a largely similar profile regarding age-at-onset and rate of disease progression (Tables 3 and 4).

The sizes of the clusters with the profile "Rapid disease progression and old age-at-onset" differed considerably and ranged from 6 to 64% of the total sample. The mean age-at-onset ranged from 61.0[25] to 72.9[20] years. Four studies found an association with axial impairment, either directly from the cluster profile[19,20] or through post-hoc analyses.[24,25] Three studies found an association between this profile and predominance of bradykinesia/rigidity (cluster profile[19]; post hoc analysis[24,25]). Conflicting results for this cluster profile were found with respect to the association with motor complications (sporadic[21] and frequent[24]) and cognitive impairment (unaffected[22], mildly impaired,[19] and impaired[25]).

Clusters with the profile "Slow disease progression and young age-at-onset" also differed in size ranging from 29%[19] to 61%[24] of the total sample. Mean age-at-onset ranged from 50[22] to 59.1[21] years. This cluster profile was further characterized in three studies by mild motor symptoms[20,22,24] and absence of cognitive impairment.[19,20,22] Conflicting results were reported on the association between this profile and the severity of motor complications (sporadic[24] and severe[19,22]) and depressive symptoms (mild[22] and severe[21]).

Lewis et al.[22] and Reijnders et al.[19] also distinguished a "bradykinesia/rigidity and PIGD-dominant" and a "tremor-dominant" profile. Notably, since these were the only studies that included the subdomains tremor, bradykinesia, and rigidity in the CA, this profile could not have been identified in other studies.[19,22]

TABLE 4. *Cluster characteristics and associations*

Cluster profiles	Other characteristics of the cluster profile	Cluster-associated variables not included in the CA
Rapid disease progression and old age at onset[19–22,24,25]	- motor impairment (total score)[25] - bradykinesia/rigidity[19] - bradykinesia/rigidity/tremor[20] - axial impairment[19,20]	- predominance bradykinesia/rigidity[21,25] - axial impairment[21,25] - bilateral PD signs at onset[25]
	- cognitive impairment[25] - mild cognitive impairment[19] - no cognitive impairment[22]	
	- frequent motor complications[24]	
	- sporadic motor complications[21]	- frequent symptomatic orthostasis[21] - low LDOPA dose[27,28] - short disease duration[19] - higher H&Y stage[26] - higher level of disability[26] - low level QoL (physical)[26]
Slow disease progression and young age at onset[19–22,24]	- mild motor symptoms[20,22,24]	- predominance tremor[25] - absence of gait disturbance[25]
	- no cognitive impairment[20–22]	- unilateral PD signs at onset[25]
	- severe depression[21] - mild depression[22]	
	- severe motor complications[19] - sporadic motor complications[24]	- severe motor complications[27]
	- high L-dopa dose[28]	- large proportion using DA[27] - relatively long disease duration[19] - younger age[19,28]
Tremor dominant[19,22]	- modest motor symptoms[22] - no cognitive impairment[19,22] - no depression[19,22]	- frequent tremor at onset[27] - anti-cholinergic medication[27] - relatively short disease duration[19] - lower H&Y stage[19]
Dominance of bradykinesia/rigidity, PIGD[19,22]	- cognitive impairment[19,22] - executive dysfunction[22]	- cognitive impairment[27]
	- depressive symptoms[19,22] - apathy[19]	
	- hallucinations[19]	- relatively long disease duration[19] - higher H&Y stage[19] - worse ADL[19] - worse QoL (mobility, cognition)[27]

CA: Cluster analysis, PD: Parkinson's disease, L-dopa: levodopa, H&Y: Hoehn and Yahr, QoL: quality of life, DA: dopamine agonists, PIGD: postural instability and gait disorder, ADL: activities of daily living.

The first cluster profile showed similarities with the "rapid disease progression and old age-at-onset" cluster profile whereas the latter was comparable to the "slow disease progression and young age-at-onset" cluster profile. However, each profile had specific characteristics.

DISCUSSION

We found seven studies that performed CA techniques on a combination of PD features with the aim to identify clinical subtypes in PD. The cluster profiles "old age-at-onset and rapid disease progression" and "young age-at-onset and slow disease progression" emerged from the majority of studies.[19,20–22,24,25] Two studies further distinguished a "tremor-dominant" and a "bradykinesia/rigidity and PIGD dominant" cluster profile.[19,22] Other profiles were less consistently identified. These results suggest that age-at-onset and rate of disease progression are important determinants for subtypes in PD and are related to each other. Further, several studies found that the "old age-at-onset and rapid

disease progression" cluster profile was also characterized by axial motor symptoms, bradykinesia, and rigidity, whereas "young age-at-onset and slow disease progression" was associated with mild motor and cognitive impairment. However, the results of the studies for either of these two profiles also clearly differed with respect to further characterization of the profiles and cluster sizes.

Differences in design may at least partly account for the differences in cluster results between studies, since choices in the process of CA will affect the results. However, large variability between studies in characteristics of study populations, in variables included in the CA and in measurement instruments, and in the number of clusters in the solution rendered the comparison of the results between studies difficult. All these factors may have to some extent influence on the cluster result. As a consequence, it was impossible to indicate specific explanation for differences in, for instance, cluster profiles or prevalence. This was further complicated by incomplete reporting of methodological steps in most studies. Interestingly, studies that included a largely similar set of variables in the CA found four more or less similar profiles,[19,22] which indicates that these subtypes are rather robust considering the fact that they were consistently identified despite considerable differences in samples characteristics between studies. It may be expected that application of CA in a more standardized design will increase the yield of studies on subtypes.

The study populations showed differences in age, disease duration, and severity. Information on age-at-onset and the distribution of this characteristic, which was found to be an important determinant for the cluster profiles, was lacking in the majority of studies, rendering results difficult to interpret and to compare between studies.

Additionally, we identified several issues regarding the selection of variables that were included in the CA to identify PD subtypes. First, large differences between studies were noted in the extent to which clinical domains of PD were included in the analysis. It is possible that domains essential in discriminating subtypes may have been missed. This also holds for subdomains with independent behavior with respect to associations with other variables. Second, all but two studies[23,24] included not only clinical impairments, but also variables like age-at-onset or medication. The use of mixtures of clinical and nonclinical variables in establishing distinct phenotypes is questionable and may yield phenotype profiles that are conceptually less meaningful. Third, since PD is a progressive disorder, the phenotypic expression of patients may change with increasing disease duration. Only the studies of Gasparoli et al.[24] and Dujardin et al.[23] included patients with similar disease duration; all other studies included patients with variable disease durations and thus ran the risk of identifying subtypes that reflect different stages of the disease rather than reflecting different phenotypic subtypes. It should be noted that all studies included in this review had a cross-sectional design which ruled out the possibility of detecting subtypes with different longitudinal patterns of change, that is, different disease courses. Longitudinal studies could provide important information, in addition to the phenotypic profiles that are identified by cross-sectional studies.

K-means, the CA technique applied by at least six of the seven included studies, does not indicate how many clusters are optimal, which is unfortunate since the number of clusters in the solution has consequences for the cluster profiles and sizes. Only one study dealt with this problem by calculating indices to determine the optimal number of clusters.[23] A second limitation of this clustering technique is that K-means is sensitive to outliers. In that respect, it is irrational that one study discarded one of the clusters after the analysis, since the excluded patients possibly had deviant scores and thus could have distorted the results.[23] Third, K-means is an iterative clustering method, and thus studies ran the risk of ending in a local optimum. However, none of the studies reported that they attempted to avoid local optima.

Another important observation is that only one study validated the outcomes in a second sample, an essential step to obtain insight in the robustness and generalizability of the findings.[19]

In spite of these methodological limitations and variations rendering comparisons difficult, the advantages of studying subtypes in an objective manner were also noted. When compared with studies on clinical subtypes based on prespecified criteria as described in the review of Foltynie et al.,[2] two important differences between the studies of interest in both reviews were noticed. First, the subtypes as described in the study by Foltynie et al.[2] were classified by only one dimension, i.e., one variable, whereas in the studies included in this review subtypes were classified and characterized on the basis of different dimensions. Additionally, four studies included in the present review allowed insight in the extent to which each variable contributed to the classification of the subtypes.[20,21,23,25] Second, in the studies reported in the review by Foltynie et al.[2] young age-at-onset, for example, was defined as <40

years, while the mean age-at-onset of the young onset subtypes that were found in the studies included in the present review ranged from 50 to 60 years and already showed clear differences with profiles with an old age-at-onset. Thus, researcher-based cut-off criteria may differ from mean values of clusters that are determined by CA and this may have consequences for the subtypes.

In conclusion, although CA has a great potential in identifying subtypes, the current review shows that the findings of different CA studies in PD are difficult to compare because of methodological differences between studies. These differences combined with methodological limitations in many of the studies lead to not fully conclusive results. In spite of the methodological differences, a profile characterized by higher age-at-onset and faster rate of disease progression and a profile characterized by lower age-at-onset and slower rate of disease progression, emerged from most studies. Since the identification of subtypes potentially has consequences for studies on the etiology of PD as well as for patient care, there is a need for CA studies with a rigorous design using a standardized approach. Future studies are recommended to (1) select a sample of PD patients with a preferably similar disease duration; (2) critically select a set of conceptually similar clinical variables that adequately represent the clinical spectrum of PD and are relevant in discriminating phenotypic profiles; (3) take the limitations of K-means into account or apply another CA technique that does not have these limitations; (4) critically evaluate the cluster results: Are they clinically meaningful and interpretable? Which variables discriminate best between clusters? How do the clusters differ with respect to variables that were not included in the CA?; (5) validate the results in independent samples. Studies that apply a similar design in different cohorts and take into account the aforementioned recommendations will likely increase our knowledge on subtypes in PD.

Acknowledgments: The authors would like to thank Dr. A.F.G. Leentjens, Dr. S. Lewis, Dr. B. Post, J.S.A.M. Reijnders, MA, and Dr. A. Schrag for providing additional information and J.W. Schoones, MSc, for his assistance with the literature search. This study is 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 Parkinson's disease Society.

Financial Disclosure J.J. van Hilten, MD PhD serves on the scientific advisory boards from Novartis Pharma and GlaxoSmithKline, and serves as an editorial board member of Movement Disorders. He received honoraria for speaking engagements from Medtronic and Novartis Pharma, and travel funding from Boehringer Ingelheim. All other authors report no financial disclosures.

Author Roles: Manuscript: Execution of review: SR, JM; writing of first draft of manuscript: SR; Review and critique: WH, JK, DV, JH, JM; Conception of idea: JH, JM.

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. Everitt BS, Landau S, Leese M. Cluster analysis. Fourth ed. London: Hodder Headline Group; 2001.
4. Dilts D, Khamalah J, Plotkin A. Using cluster analysis for medical resource decision making. Med Decis Making 1995;15:333–347.
5. Steinley D. Local optima in K-means clustering: what you don't know may hurt you. Psychol Methods 2003;8:294–304.
6. Xu R, Wunsch D. Survey of clustering algorithms. IEEE Trans Neural Netw 2005;16:645–678.
7. Calinski T, Harabasz J. A dendrite method for cluster analysis. Commun Stat 1974;3:1–27.
8. Milligan GW, Cooper MC. An examination of procedures for determining the number of clusters in a data set. Psychometrika 1985;50:159–179.
9. Dupuy A, Simon RM. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. J Natl Cancer Inst 2007;99:147–157.
10. Aarsland D, Bronnick K, Ehrt U, et al. Neuropsychiatric symptoms in patients with Parkinson's disease and dementia: frequency, profile and associated care giver stress. J Neurol Neurosurg Psychiatry 2007;78:36–42.
11. Bronnick K, Aarsland D, Larsen JP. Neuropsychiatric disturbances in Parkinson's disease clusters in five groups with different prevalence of dementia. Acta Psychiatr Scand 2005;112:201–207.
12. Kaiser I, Kryspin-Exner I, Brucke T, Volc D, Alesch F. Long-term effects of STN DBS on mood: psychosocial profiles remain stable in a 3-year follow-up. BMC Neurol 2008;8:43.
13. Kulisevsky J, Pagonabarraga J, Pascual-Sedano B, Garcia-Sanchez C, Gironell A. Prevalence and correlates of neuropsychiatric symptoms in Parkinson's disease without dementia. Mov Disord 2008;23:1889–1896.
14. Macht M, Schwarz R, Ellgring H. Patterns of psychological problems in Parkinson's disease. Acta Neurol Scand 2005;111:95–101.
15. Merchant H, Luciana M, Hooper C, Majestic S, Tuite P. Interval timing and Parkinson's disease: heterogeneity in temporal performance. Exp Brain Res 2008;184:233–248.
16. Thompson-Ward EC, Theodoros DG, Murdoch BE, Cahill L. The use of a miniature lip transducer system in the assessment of patients with Parkinsons disease. J Med Speech Lang Pathol 1999;7:175–179.
17. Weintraub D, Moberg PJ, Culbertson WC, Duda JE, Stern MB. Evidence for impaired encoding and retrieval memory profiles in Parkinson disease. Cogn Behav Neurol 2004;17:195–200.
18. Leentjens AFG, Reijnders J, Ehrt U, Lousberg R, Aarsland D. Clinical phenotypes of Parkinson's disease: how does psychopathology cluster with motor symptoms? Mov Disord 2007;22:S266.
19. 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.

20. Post B, Speelman JD, de Haan RJ. Clinical heterogeneity in newly diagnosed Parkinson's disease. J Neurol 2008;255:716–722.
21. Schrag A, Quinn NP, Ben Shlomo Y. Heterogeneity of Parkinson's disease. J Neurol Neurosurg Psychiatry 2006;77:275–276.
22. 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.
23. Dujardin K, Defebvre L, Duhamel A, et al. Cognitive and SPECT characteristics predict progression of Parkinson's disease in newly diagnosed patients. J Neurol 2004;251:1383–1392.
24. Gasparoli E, Delibori D, Polesello G, et al. Clinical predictors in Parkinson's disease. Neurol Sci 2002;23:S77–S78.
25. Graham JM, Sagar HJ. A data-driven approach to the study of heterogeneity in idiopathic Parkinson's disease: identification of three distinct subtypes. Mov Disord 1999;14:10–20.
26. Daniel SE, Lees AJ. Parkinson's Disease Society Brain Bank, London: overview and research. J Neural Transm Suppl 1993;39:165–172.
27. Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. Arch Neurol 1999;56:33–39.
28. 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.