

Identifying the Clusters within Nonmotor Manifestations in Early Parkinson's Disease by Using Unsupervised Cluster Analysis

Hui-Jun Yang¹, Young Eun Kim², Ji Young Yun³, Han-Joon Kim², Beom Seok Jeon^{2*}

1 Department of Neurology, Ulsan University Hospital, Ulsan, Korea, **2** Department of Neurology and Movement Disorder Center, Parkinson Study Group, Neuroscience Research Institute, Seoul National University Hospital, Seoul, Korea, **3** Department of Neurology, Ewha Womans University Mokdong Hospital, Seoul, Korea

Abstract

Background: Classical and data-driven classifications of Parkinson's disease (PD) are based primarily on motor symptoms, with little attention being paid to the clustering of nonmotor manifestations.

Methods: Clinical data on demographic, motor and nonmotor features, including the Korean version of the sniffin' stick (KVSS) test results, and responses to the screening questionnaire of the nonmotor features were collected from 56 PD patients with disease onset within 3 years. Nonmotor subgroups were classified using unsupervised hierarchical cluster analysis (HCA). In addition to unsupervised HCA, we performed a cross-sectional analysis comparing the performance on the KVSS olfactory test with other nonmotor manifestations of the patients.

Results: Forty-nine patients (87.5%) had hyposmia based on the KVSS test. HCA suggested three nonmotor clusters for all PD patients and two nonmotor clusters in *de novo* PD patients, without *a priori* assumptions about the relatedness. In the cross-sectional analysis, dream-enactment behavior was more prevalent in patients with lower olfactory scores, implying impaired olfactory function ($P=0.029$ for all PD patients; $P=0.046$ for *de novo* PD patients).

Conclusion: We propose the existence of different clusters of nonmotor manifestations in early PD by using unsupervised hierarchical clustering. To our knowledge, this study is the first to report the identification of nonmotor subgroups based on unsupervised HCA of multiple nonmotor manifestations in the early stage of the disease.

Citation: Yang H-J, Kim YE, Yun JY, Kim H-J, Jeon BS (2014) Identifying the Clusters within Nonmotor Manifestations in Early Parkinson's Disease by Using Unsupervised Cluster Analysis. PLoS ONE 9(3): e91906. doi:10.1371/journal.pone.0091906

Editor: Oscar Arias-Carrion, Hospital General Dr. Manuel Gea González, Mexico

Received: July 23, 2013; **Accepted:** February 18, 2014; **Published:** March 18, 2014

Copyright: © 2014 Yang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by grants of the Korean Health Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (A101273 and A110875). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: brain@snu.ac.kr

Introduction

Parkinson's disease (PD) is a clinically and pathologically heterogeneous disease [1]. This heterogeneity is thought to indicate different subtypes of PD and many studies have sought to elucidate them [2,3].

Classically, patients are divided by their motor phenotype into tremor-dominant (TD) or postural instability-gait disturbance (PIGD) subgroups, or classified by onset age as juvenile, young, or late onset PD [1]. In recent years, cluster analysis has been introduced as an objective data-driven grouping method [2,4]. Several cluster analysis studies have shown that the age of onset [4–9], rate of disease progression [5–10], and motor phenotype [6,9–10] are the major dimensions of PD subgroup classification.

The classical and data-driven approaches have focused primarily on the motor symptoms of PD. Some of the nonmotor features, such as depression [8] or cognitive decline [4,10] were associated with previous classifications based mainly on motor phenotypes. Although recent results from two independent European cohorts have shown that the severity of nonmotor symptoms as well as motor complications are important factors in the characterization

of PD subtypes [11], little attention has been paid to the clustering of nonmotor manifestations themselves [12,13].

In the current study, we investigated the existence of different subgroups of nonmotor manifestations by using unsupervised hierarchical cluster analysis (HCA). To our knowledge, this is the first identification of nonmotor clusters using unsupervised HCA on a range of nonmotor features in early stage PD.

Methods

Subjects

Between July 2007 and January 2008, 119 consecutive patients with idiopathic PD who met the diagnostic criteria of the United Kingdom Brain Bank were referred to our movement disorder center and clinically followed up. At referral, 56 of these patients who were within 3 years of motor symptom onset underwent olfactory function testing using the Korean version of the sniffin' stick (KVSS) test, the first olfactory function test to use odorants familiar to Koreans [14]; moreover, these patients were personally interviewed by a trained member of the movement disorders clinic by utilizing a screening questionnaire consisting of 8 nonmotor

items in Korean, as part of the routine clinical evaluation of patients with PD. Patients were asked about the presence or absence of each nonmotor feature (insomnia, orthostatic dizziness, depression, excessive daytime sleepiness, urinary symptom, memory disturbance, and dream-enactment behavior) mentioned in each item.

We retrospectively performed a systematic review of the hospital electronic medical records to collect clinical data on current age, age at onset of PD, gender, Hoehn and Yahr (H-Y) stage, dopaminergic drugs in a levodopa-equivalent daily dose (LEDD, mg/day), presence of constipation and responses to the screening questionnaire of the nonmotor features. Constipation was defined as having fewer than three bowel movements per week. Patients had to be followed for at least 1 year to be included. None of the *de novo* PD patients, a subset of the all PD samples, had previously taken antiparkinsonian medication. This retrospective study was approved by the Institutional Review Board (IRB) at the Seoul National University Hospital (H-1109-122-379). Requirement for informed consent was waived for this retrospective analysis of clinical data. The IRB also approved the consent procedure. All data used in this study was analyzed anonymously.

Smell testing

The KVSS test is a modified version of the “Sniffin’ Stick” test [14]. Its validity and reliability have been demonstrated in comparison with the Cross-cultural Smell Identification Test (CC-SIT) [15]. KVSS I is a rapid screening test and KVSS II is a comprehensive test that involves three subsets: threshold, discrimination, and identification. The olfactory threshold was defined as the mean concentration at which the pen containing *n*-butanol was differentiated correctly four times from two blank pens. Olfactory discrimination was assessed using triplets of odorant pens in which two pens have identical odors and the other has a different odor, and the patients identify the pen with a different odorant. For olfactory identification, all 16 different odorants familiar to Koreans were presented in felt-tip pens, with the patients choosing one of four odor items [14]. The scores ranged from 0 to 16 in all three subsets. The sum of the threshold, discrimination, and identification subset scores is presented as the composite threshold-discrimination-identification (TDI) score.

Statistical analysis

Statistical analyses were performed with the open-source statistical software R version 2.15.1 (<http://www.r-project.org>) and IBM SPSS statistics version 19.0 (IBM, Somers, NY).

Unsupervised HCA was performed and the clustering results are shown using a dendrogram. The main advantage of an unsupervised hierarchical approach is that it can be applied when the optimum number of clusters is not known in advance [16]. We used Yule’s *Q* as a measure of similarity for asymmetric binary variables. The dissimilarity between clusters was calculated by the most common method, unweighted pair group method with arithmetic mean (UPGMA) also known as between-group average method [17]. Different methods of measuring similarity (Jaccard’s coefficient, Dice’s coefficient and Yule’s *Q*) and different measures of intergroup distance (average linkage within-groups method, and UPGMA) were used to demonstrate the stability of the clustering [16,17]. Multiscale bootstrap resampling was performed with the R package pvclust version 1.2–2 to compute the confidence of the hierarchical clustering with 1000 bootstrapped samples [18,19]. Clusters with approximately unbiased probability value (AU *P* value) >95% were considered significant [19].

The cross-sectional analysis between the two olfactory groups based on the performance in the KVSS test was performed

independently with unsupervised cluster analysis. The patients were divided into high- and low-scoring groups by using the KVSS II median TDI score along previous literatures [20]. The chi-square test and Fisher’s exact test were used to determine the relationship between olfactory dysfunction and other nonmotor manifestations. Either Student’s *t*-test or Mann-Whitney *U* test was used to analyze clinical differences between the two olfactory groups. The results were considered statistically significant at $P < 0.05$.

Results

Table 1 shows demographic and clinical characteristics of the subjects. The 56 patients included 28 men and 28 women (age range, 46–81 years). No subject had motor fluctuation or levodopa-induced dyskinesia. Twenty-seven patients were newly diagnosed *de novo* PD patients. The KVSS tests were well accepted by all patients. Forty-nine patients (87.5%) had hyposmia based on the reported criteria [14,15]. For all of the PD patients, the cluster analysis of nine nonmotor features gave three clusters without *a priori* assumptions about relatedness. Figure 1A shows the corresponding dendrogram based on UPGMA distance. Cluster 1 included hyposmia, dream-enactment behavior, and constipation. Cluster 2 comprised memory disturbance and orthostatic dizziness. Cluster 3 contained urinary symptoms and excessive daytime sleepiness.

HCA in the *de novo* PD group revealed two main clusters (Fig. 1B). Cluster 1 was defined by hyposmia, dream-enactment behavior, and constipation. The larger cluster 2 was defined by depression, insomnia, memory disturbance, orthostatic dizziness, excessive daytime sleepiness, and urinary symptoms. The clustering stability was assessed by comparing the results of different methods of measuring similarity and different measures of intergroup dissimilarity. We obtained similar clustering results and concluded that the group structure was stable.

In cross-sectional investigation, the patients were split into two olfactory groups based on the median TDI score [20]. Gender distribution, mean current age, disease duration, LEDD, and H-Y stage did not differ between the two olfactory groups (Table 1). Dream-enactment behavior was more prevalent in patients with lower TDI scores, which imply impaired olfactory function ($P = 0.029$ for all PD patients; $P = 0.046$ for *de novo* PD patients; Fig. 2). There were no significant differences in other nonmotor symptoms (Figs. 2A and 2B).

Discussion

In this study, 87.5% of patients were hyposmic as assessed by the KVSS olfactory test, which was in line with results from previous olfactory function studies in PD patients [18,20,21]. Unsupervised cluster analysis suggests three nonmotor clusters for the entire group of PD patients and two clusters for the *de novo* PD patients. This clustering result and the cross-sectional investigation result are in agreement with results of previous studies demonstrating close relationships between olfactory dysfunction and rapid eye movement sleep behavior disorders (RBDs) [22]. Notably, our study suggests that depression is independent of olfactory dysfunction. This is a somewhat striking result, since depression, together with hyposmia, is a well-known predated nonmotor manifestation [23]. Our clustering result concurs with those of several works showing that olfactory dysfunction did not correlate with depression [20,24,25]. The discrepancy with the findings of Berendse *et al.*, which addressed a potential correlation of hyposmia with depression, might partly depend on disease duration or the use of different olfactory function tests [26].

Table 1. Demographics and clinical profiles of the PD patients.

Characteristics	All PD patients (n = 56)						de novo PD patients (n = 27)							
	Total		TDI top (n = 28)		TDI bottom (n = 28)	P	Total		TDI top (n = 14)		TDI bottom (n = 13)	P		
Age (years)	64.0	(8.5)	62.1	(9.9)	65.9	(6.5)	.101	64.5	(7.7)	64.4	(8.2)	64.7	(7.4)	.912
Gender (% male)	50.0		42.9		57.1		.285	48.1		42.9		53.8		.568
Age at PD onset (years)	62.6	(8.8)	60.7	(10.5)	64.5	(6.4)	.103	63.6	(8.0)	63.4	(8.5)	63.7	(7.8)	.934
Duration of PD (years)	1.6	(0.9)	1.5	(1.0)	1.6	(1.0)	.871	1.3	(0.8)	1.2	(0.7)	1.4	(0.9)	.450
Hoehn and Yahr stage	2.0	(0.7)	2.0	(0.7)	2.1	(0.7)	.618	2.1	(0.7)	2.1	(0.7)	2.0	(0.7)	.939
I (%)	16	(28.6)	8	(28.6)	8	(28.6)		6	(22.2)	3	(21.4)	3	(23.1)	
II (%)	30	(53.6)	15	(53.6)	15	(53.6)		16	(59.2)	8	(42.9)	8	(38.5)	
III (%)	10	(17.9)	5	(17.9)	5	(17.9)		5	(18.5)	3	(21.4)	2	(15.4)	
LEDD, mg/day	149.6	(231.1)	101.2	(188.1)	198.1	(261.8)	.118							
KVSS I	5.0	(1.5)	5.7	(0.7)	4.1	(1.4)	<0.001	5.3	(1.2)	5.9	(0.7)	4.7	(1.3)	.009
KVSS II thresholds	4.1	(3.5)	6.2	(3.3)	2.0	(2.1)	<0.001	4.6	(3.8)	7.1	(3.5)	1.8	(1.5)	<0.001
KVSS II discrimination	8.0	(2.6)	9.6	(1.9)	6.4	(2.3)	<0.001	8.3	(3.3)	9.9	(2.4)	6.6	(3.3)	.009
KVSS II identification	8.2	(2.7)	10.1	(2.2)	6.3	(1.6)	<0.001	8.7	(3.1)	10.7	(2.7)	6.6	(1.9)	<0.001
TDI score	20.1	(6.6)	25.5	(3.4)	14.7	(4.1)	<0.001	21.3	(7.4)	27.1	(2.4)	15.0	(5.6)	<0.001

Data are shown as means (standard deviation) unless otherwise indicated. P-values are from Student's t-test for continuous measures or Mann-Whitney U test for ordinal data. PD, Parkinson's disease; LEDD, levodopa equivalent daily dose; KVSS, Korean version of the sniffin' stick; TDI, threshold-discrimination-identification. doi:10.1371/journal.pone.0091906.t001

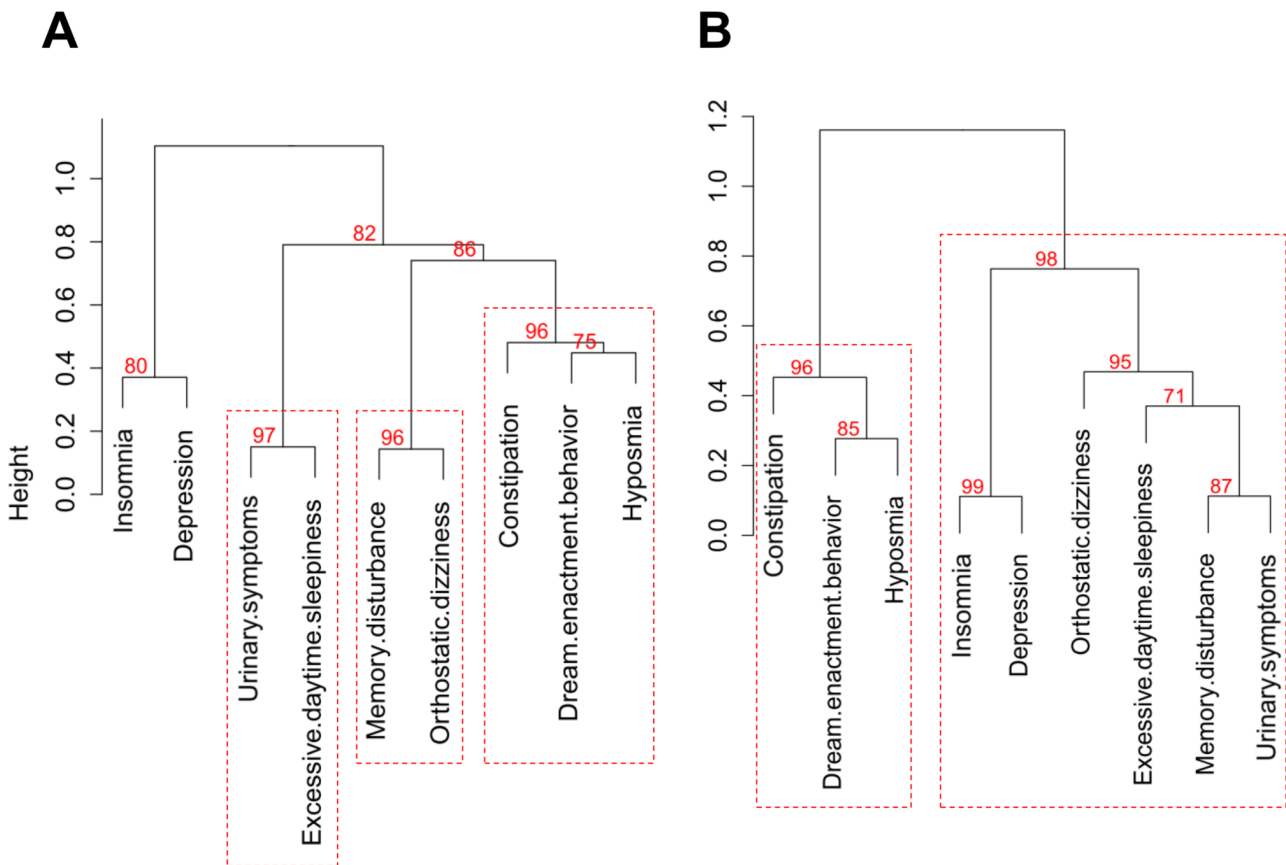


Figure 1. Dendrogram of the hierarchical cluster analysis and multiscale bootstrap resampling results of nonmotor features of Parkinson's disease (PD). The clusters with an approximately unbiased (AU, red) P-value > 95% are highlighted by the red dashed rectangles. (A) all PD patients. (B) de novo PD patients. doi:10.1371/journal.pone.0091906.g001

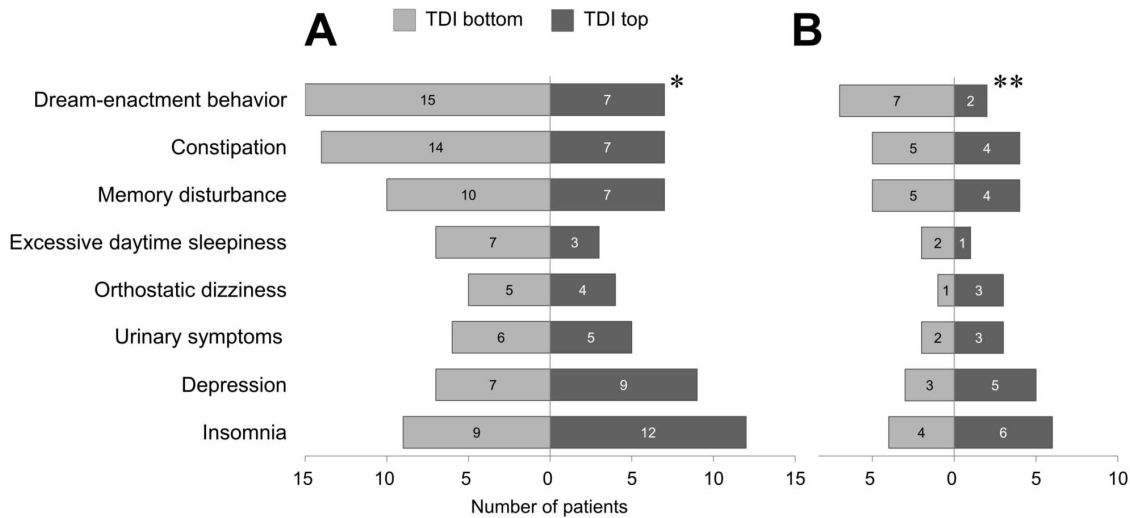


Figure 2. Group differences in nonmotor manifestations based on the KVSS II TDI score (* $P=0.029$; ** $P=0.046$, chi-square test or Fisher's exact test, as appropriate). (A) all PD patients. (B) *de novo* PD patients.
doi:10.1371/journal.pone.0091906.g002

Some nonmotor manifestations, such as hyposmia, RBDs, constipation, and depression, can develop during the prodromal period of PD and precede the onset of classical PD motor symptoms [27,28]. While patients in the premotor stage of PD exhibit no motor symptoms, they do demonstrate various prodromal nonmotor symptoms [23,29]. Conventional PD classification based on motor phenotype (tremor-dominant or PIDG) cannot be applied to these premotor PD; therefore, classifications that are wholly based on nonmotor symptoms are needed. The finding of this study can have implications for developing the premotor PD classification based on nonmotor features. Our research subjects, patients with newly diagnosed *de novo* PD and patients with relatively early stages of PD (with disease onset within 3 years), may be the closest diagnostic PD group to represent ideal premotor PD populations. Our study demonstrates the existence of different nonmotor symptom clusters in early stage PD patients and support the idea of a premotor PD subtype based on nonmotor manifestations [30–32].

While this cluster analysis gave interesting results, the clinical relevance of the complicated relationships among nonmotor manifestations is yet to be determined. A nonmotor feature could be caused by coexisting conditions, for example, daytime sleepiness due to nocturia. It is also possible that some symptoms are caused by PD medication, such as levodopa or dopamine agonists [33]. In addition, the patterning of nonmotor manifestations seen in our results implicates common neuropathological and neurochemical processes underlying PD [13,34,35]. The underlying alpha-synuclein pathology has been proposed as a shared

mechanism for hyposmia, RBD, and constipation [22,35], which define the first cluster of *de novo* PD patients (Fig. 1B). Other nonmotor features that comprise the second cluster of *de novo* PD patients, such as depression, memory disturbance and autonomic dysfunction, may be early disease manifestations that reflect underlying non-dopaminergic deficit [23,36].

One limitation of this study is its retrospective design with a relatively small sample size, which potentially influences the generalizability of our results. Another limitation is the use of the screening questionnaire of the nonmotor features that has not been validated. The lack of replication is another caveat of this study. Although we used multiscale bootstrap resampling to validate confidence in the clusters, a larger study should be conducted to verify our results [2,18].

This preliminary analysis examining the presentation of multiple nonmotor features suggests that it is possible to identify structures based on the profiles of nonmotor manifestations. Future data-driven replications with larger independent populations are required, and these will serve to increase our understanding of associations between various nonmotor features of PD.

Author Contributions

conceived and designed the experiments: HY YK JY HK BJ. performed the experiments: HY YK JY BJ. analyzed the data: HY YK JY. contributed reagents/materials/analysis tools: HY HK BJ. wrote the manuscript: HY HK BJ.

References

- Foltnie T, Brayne C, Barker RA (2002) The heterogeneity of idiopathic Parkinson's disease. *J Neurol* 249: 138–145.
- van Rooden SM, Heiser WJ, Kok JN, Verbaan D, van Hilten JJ, et al. (2010) The identification of Parkinson's disease subtypes using cluster analysis: a systematic review. *Mov Disord* 25: 969–978.
- Rajput AH, Voll A, Rajput ML, Robinson CA, Rajput A (2009) Course in Parkinson disease subtypes: A 39-year clinicopathologic study. *Neurology* 73: 206–212.
- Graham JM, Sagar HJ (1999) A data-driven approach to the study of heterogeneity in idiopathic Parkinson's disease: identification of three distinct subtypes. *Mov Disord* 14: 10–20.
- Gasparoli E, Delibori D, Polesello G, Santelli L, Ermani M, et al. (2002) Clinical predictors in Parkinson's disease. *Neurol Sci* 23 Suppl 2: S77–78.
- Lewis SJ, Foltnie T, Blackwell AD, Robbins TW, Owen AM, et al. (2005) Heterogeneity of Parkinson's disease in the early clinical stages using a data driven approach. *J Neurol Neurosurg Psychiatry* 76: 343–348.
- Schrag A, Quinn NP, Ben-Shlomo Y (2006) Heterogeneity of Parkinson's disease. *J Neurol Neurosurg Psychiatry* 77: 275–276.
- Post B, Speelman JD, de Haan RJ (2008) Clinical heterogeneity in newly diagnosed Parkinson's disease. *J Neurol* 255: 716–722.
- Reijnders JS, Ehart U, Lousberg R, Aarsland D, Leentjens AF (2009) The association between motor subtypes and psychopathology in Parkinson's disease. *Parkinsonism Relat Disord* 15: 379–382.
- Dujardin K, Defebvre L, Duhamel A, Lecoiffé P, Rogelet P, et al. (2004) Cognitive and SPECT characteristics predict progression of Parkinson's disease in newly diagnosed patients. *J Neurol* 251: 1383–1392.

11. van Rooden SM, Colas F, Martinez-Martin P, Visser M, Verbaan D, et al. (2011) Clinical subtypes of Parkinson's disease. *Mov Disord* 26: 51–58.
12. Perez Lloret S, Rossi M, Merello M, Rascol O, Cardinali DP (2011) Nonmotor symptoms groups in Parkinson's disease patients: results of a pilot, exploratory study. *Parkinsons Dis* 2011:473579. doi:10.4061/2011/473579
13. van Rooden SM, Visser M, Verbaan D, Marinus J, van Hilten JJ (2009) Patterns of motor and non-motor features in Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 80: 846–850.
14. Hong SC, Yoo YS, Kim ES, Kim SC, Park SH, et al. (1999) Development of KVSS test (Korean version of Sniffin' sticks test). *Korean J Otolaryngol* 42: 855–860.
15. Cho JH, Jeong YS, Lee YJ, Hong SC, Yoon JH, et al. (2009) The Korean version of the Sniffin' stick (KVSS) test and its validity in comparison with the cross-cultural smell identification test (CC-SIT). *Auris Nasus Larynx* 36: 280–286.
16. Řezanková H (2009) Cluster analysis and categorical data. *Statistika* 46: 216–232.
17. Quinette P, Guillery-Girard B, Dayan J, de la Sayette V, Marquis S, et al. (2006) What does transient global amnesia really mean? Review of the literature and thorough study of 142 cases. *Brain* 129: 1640–1658.
18. Baba T, Takeda A, Kikuchi A, Nishio Y, Hosokai Y, et al. (2011) Association of olfactory dysfunction and brain. *Metabolism in Parkinson's disease*. *Mov Disord* 26: 621–628.
19. Suzuki R, Shimodaira H (2006) Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22: 1540–1542.
20. Morley JF, Weintraub D, Mamikonyan E, Moberg PJ, Siderowf AD, et al. (2011) Olfactory dysfunction is associated with neuropsychiatric manifestations in Parkinson's disease. *Mov Disord* 26: 2051–2057.
21. Haehner A, Boesveldt S, Berendse HW, Mackay-Sim A, Fleischmann J, et al. (2009) Prevalence of smell loss in Parkinson's disease - a multicenter study. *Parkinsonism Relat Disord* 15: 490–494.
22. Stiasny-Kolster K, Doerr Y, Moller JC, Hoffken H, Behr TM, et al. (2005) Combination of 'idiopathic' REM sleep behaviour disorder and olfactory dysfunction as possible indicator for alpha-synucleinopathy demonstrated by dopamine transporter FP-CIT-SPECT. *Brain* 128: 126–137.
23. Postuma RB, Aarsland D, Barone P, Burn DJ, Hawkes CH, et al. (2011) Identifying prodromal Parkinson's disease: pre-motor disorders in Parkinson's disease. *Mov Disord* 27: 617–626.
24. Morley JF, Duda JE (2011) Neuropsychological correlates of olfactory dysfunction in Parkinson's disease. *J Neurol Sci* 310: 228–230.
25. Cramer CK, Friedman JH, Amick MM (2010) Olfaction and apathy in Parkinson's disease. *Parkinsonism Relat Disord* 16: 124–126.
26. Berendse HW, Roos DS, Rajmakers P, Doty RL (2011) Motor and non-motor correlates of olfactory dysfunction in Parkinson's disease. *J Neurol Sci* 310: 21–24.
27. Siderowf A, Lang AE (2012) Pre-motor Parkinson's disease: concepts and definitions. *Mov Disord* 27: 608–616.
28. Tolosa E, Gaig C, Santamaria J, Compta Y (2009) Diagnosis and the premotor phase of Parkinson disease. *Neurology* 72 Suppl 7: S12–20.
29. Chen H, Burton EA, Ross GW, Huang X, Savica R, et al. (2013) Research on the premotor symptoms of Parkinson's Disease: clinical and etiological implications. *Environ Health Perspect* 121: 1245–1252.
30. Liepelt-Scarfone I, Gauss K, Maetzler W, Müller K, Bormann C, et al. (2013) Evaluation of progression markers in the premotor phase of Parkinson's disease: the progression markers in the premotor phase study. *Neuroepidemiology* 41: 174–182.
31. Siderowf A, Jennings D, Eberly S, Oakes D, Hawkins K, et al. (2012) The PARS Investigators. Impaired olfaction and other prodromal features in the Parkinson At-Risk Syndrome study. *Mov Disord* 27: 406–412.
32. Berg D, Godau J, Seppi K, Behnke S, Liepelt-Scarfone I, et al. (2013) The PRIPS study: screening battery for subjects at risk for Parkinson's disease. *Eur J Neurol* 20: 102–108.
33. Kim HJ, Park SY, Cho YJ, Hong KS, Cho JY, et al. (2009) Nonmotor symptoms in de novo Parkinson disease before and after dopaminergic treatment. *J Neurol Sci* 287: 200–204.
34. Lerner A, Bagic A (2008) Olfactory pathogenesis of idiopathic Parkinson disease revisited. *Mov Disord* 23: 1076–1084.
35. Ravina B, Aarsland D (2013) Many roads lead to Parkinson's disease. *Mov Disord* 28: 122–123.
36. Goldstein DS, Sewell L, Holmes C (2010) Association of anosmia with autonomic failure in Parkinson disease. *Neurology* 74: 245–251.