Heterogeneity among patients with Parkinson's disease: Cluster analysis and genetic association ☆

CrossMark

Ling-Yan Ma [a,e], Piu Chan [b,c,d], Zhu-Qin Gu [b,c], Fang-Fei Li [a,e], Tao Feng [a,d,e,*]

[a] Center for Neurodegenerative disease, Department of Neurology, Beijing Tiantan Hospital, Capital Medical University, #6 Tian Tan Xi Li Street, Beijing 100050, China
[b] Department of Neurobiology and Neurology, Xuanwu Hospital, Capital Medical University, #45 Changchun Street, Beijing 100053, China.
[c] Key Laboratory on Neurodegenerative Disease of Ministry of Education and Key Laboratory on Parkinson's Disease of Beijing, #45 Changchun Street, Beijing 100053, China
[d] Parkinson's Disease Center, Beijing Institute for Brain Disorders, #10 You'an Men Wai Xi Tou Tiao, Beijing 100069, China
[e] China National Clinical Research Center for Neurological Diseases, #6 Tian Tan Xi Li Street, Beijing 100050, China

ARTICLE INFO

ABSTRACT

The clinical heterogeneity of Parkinson's disease (PD) reveals the presence of several PD subtypes. The objectives of this study were to identify PD subtypes using cluster analysis (CA) and to determine the association between the subtypes and the polymorphisms in LRRK2 (G2385R and R1628P) and GBA (L444P) genes. A k-means CA of demographics, disease progression, motor and non-motor symptoms was performed from 1,510 Chinese PD patients from the Chinese National Consortium on Neurodegenerative Diseases. Pearson correlation analysis was performed to eliminate uninformative characteristics. Blood samples from 852 patients were obtained for genetic analysis of LRRK2 and GBA. Genotypic associations between various subtypes and genetic variants were examined using chi-square test. We identified four different subtypes: subtype 1 was non-tremor dominant (NTD, $n = 469$; 31.1%); subtype 2 had a rapid disease progression with late onset (RDP-LO, $n = 67$; 4.4%); subtype 3 had benign pure motor characteristics (BPM, $n = 778$; 51.5%) without non-motor disturbances; and subtype 4 was tremor dominant with slow disease progression (TD-SP, $n = 196$; 13.0%). Subtypes 1, 2, and 4 had similar mean age of onset. No associations were identified between polymorphisms in LRRK2 (R1628P) and GBA (L444P) genes and the four subtypes ($P > 0.05$).

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Parkinson's disease (PD) is a heterogeneous neurodegenerative condition with different clinical phenotypes, genetics, pathology, brain imaging characteristics and disease duration [1,2]. The clinical variability among PD patients suggests the existence of disease subtypes. It is crucial to identify all PD subtypes because different groups may have different pathological and genetic features which may also lead to tailored therapeutic treatments.

Several PD subtypes had been identified mainly though experience based on predefined criteria including age of onset and motor symptoms [3,4]. According to the age of onset, PD patients can be divided into juvenile, young, and late groups [3,4]. Motor symptoms are classified into "hypokinetic rigid," "tremor dominant," and "postural instability–gait disorders" [3]. Apart from empirical classification, several studies have relied on statistical analysis to identify the different PD subtypes, among which cluster analysis (CA) is the most commonly used. CA is a data-driven approach in which the profiles of the subtypes arise solely from characteristics of the data without predefined criteria. The cluster profiles "old age of onset and rapid disease progression" and "young age of onset and slow disease progression" emerged from several CA studies [5–9]. However, due to methodological differences, other cluster profiles were less consistent among studies.

Currently, it is unknown why there are variations in PD phenotypes; however, it has been proposed the different PD subtypes emerge from interactions between genetic and non-genetic factors [10–12]. Several monogenetic forms of PD have been identified, with gene-encoding leucine-rich repeat kinase 2 (LRRK2) and glucocerebrosidase (GBA) being well characterized. Compared with non-mutation carriers, patients carrying the GBA L444P mutation have an earlier age of onset and more non-motor symptoms, while LRRK2-associated PD is sometimes related to late onset with typical symptoms [13,14]. In order to find potential factors associating with PD subtypes, we hypothesize that genetic component (mutations in LRRK2 and GBA genes) may in part contribute to the different PD subtypes [14,15]. Therefore, the

objectives of our study were (1) to identify PD subtypes based on motor and non-motor features by using CA [16] in a large number of Chinese PD patients and (2) to evaluate the association between subtypes and variants in *LRRK2* (G2385R and R1628P) and *GBA* (L444P) genes.

2. Methods

2.1. Patients and clinical features

In this study, 1,510 patients who met the United Kingdom PD Society Brain Bank criteria [17] and had complete clinical information were enrolled from a PD cohort of the Chinese National Consortium on Neurodegenerative Diseases (CNCPD), established by the Chinese Parkinson Study Group (CPSG), which dated from 2007 and contained over 5200 neurodegenerative disease cases till now. CPSG is a multi-collaborative group comprising 42 clinical centers in which patients with neurodegenerative diseases are enrolled in the cohort. The 42 centers were located in 16 provinces and 3 Municipalities (Beijing, Shanghai and Tianjin), enrolling patients from most areas of China. PD was diagnosed by movement disorder specialists. Patients who underwent stereotactic surgery were excluded. Among the 1,510 PD patients, 852 patients from 22 hospitals provided blood samples for genetic testing of *LRRK2* (G2385R and R1628P) and *GBA* (L444P) genes. Table 1 shows the demographic and baseline clinical characteristics. There is no statistic difference in demographic and clinical features between the 852 patients who gave samples and the 658 who did not (see Table 1 in supplemental materials). Written informed consent forms were obtained from all patients. The study was approved by the local ethics committees of the participating hospitals.

2.2. Evaluations

2.2.1. Motor

The Unified Parkinson's Disease Rating Scale I-IV (UPDRS I-IV) [18] was applied to evaluate mood and mentation, activity of daily living (ADL), motor disability, and therapy L-dopa complications. PD stage was determined using the Hoehn and Yahr rating scale (HY stage) [19]. Age of onset and disease duration was recorded at baseline. PD progression rate was calculated as UPDRS III/disease duration [5,20].

Several parameters were used for motor evaluation. The "Tremor score" was derived as the mean of the sub-scores of UPDRS III items 20 and 21 (rest and action/postural tremor), while a "hypokinesia score" was defined as the mean of the sub-scores of UPDRS III items 23–26 and 31 (finger tap, hand grip, rapid alternating movements, leg tap, and bradykinesia). UPDRS III item 22 was also used for the "rigidity score," and items 27–30 (arise, posture, and gait stability) were used for the "postural instability/gait disorder (PIGD) score" [21]. The mean value of items 16, 20, and 21 provided the "total tremor score," while "total non-tremor score" was obtained from the mean score of items 5, 7, 12–15, 18, 19, and 22–31 [22]. Motor phenotype was defined as ratio of total tremor to non-tremor score.

2.2.2. Non-motor

The mini mental state examination (MMSE) was used to assess global cognitive dysfunction [23]. The severity of depression symptoms was derived from the Hamilton rating scale for depression (HAMD) [24]. All patients completed the Pittsburgh Sleep Quality Index (PSQI), which evaluates sleep quality and disturbances [25]. Constipation was diagnosed according to the Rome diagnostic criteria (1 = yes, 0 = no) [26].

2.3. Statistical analysis

Patient demographic and clinical characteristics were summarized using descriptive statistics. Variables were standardized prior to CA because the clinical phenotypes were measured on different scales and with different score ranges. The analyzed data included age of disease onset, disease progression, stage (HY), motor evaluating scores (UPDRS III, tremor, hypokinesia, rigidity, PIGD), and non-motor evaluating scores (cognition/MMSE, depression/HAMD, sleep disorder/PSQI, constipation scores) (Table 2). Considering the quality of cluster results can be improved by eliminating uninformative variables, we performed Pearson correlation analysis to ensure that no variables in the core variable set were correlated at ≥ 0.7.

A *k*-means CA was performed with three, four, and five clusters following Pearson correlation analyses. Inter-cluster consistency agrees well with each other if these *k*-means-derived clusters could share large

Table 1
Patient demographics and clinical characteristics.

Clinical variables	Value* (*n* = 1510)
Age in years	66.7 (10.5; 24–91)
Age at onset in year	57.6 (10.9;12–85)
Gender (M:F)	1.5:1
Disease duration in months	63.9 (51.9; 2.0–535.2)
UPDRS -I	2.3 (2.2; 0–14)
UPDRS II	11.8 (6.7; 0–47)
UPDRS III	26.3 (14.8; 1–103)
UPDRS IV	3.6 (4.9; 0–29)
Tremor score	0.7 (0.6; 0–3.6)
Rigidity score	1.1 (0.9; 0–4)
Hypokinesia score	1.2 (0.7; 0.1–4)
PIGD	1.0 (0.7; 0–4)
Total tremor score/total non-tremor score	1.0 (1.0; 0–15.2)
Progression rate	1.2 (1.7; 0.1–17.9)
MMSE	26.8 (3.4; 10–30)
HAMD	10.0 (9.1; 0–66)
PSQI	6.5 (4.3; 0–19)
Constipation	0.6 (0.5; 0–1)

UPDRS: Unified Parkinson's Disease Rating Scale; MMSE: Mini Mental State Examination; HAMD: Hamilton rating scale for depression; PSQI: Pittsburgh Sleep Quality Index.
 * The values in each cell represent: mean (standard deviation (SD), range) expect for the gender column.

Table 2
Characteristics of the four clusters (post hoc analysis).

	Cluster 1:NTD	Cluster 2:RPD-LO	Cluster 3:BPM	Cluster 4:TD-SP	*P*-value
N (%)	469 (31.1%)	67 (4.4%)	778 (51.5%)	196 (13.0%)	
Age (years)	68.2 (9.6)	67.9 (8.4)	67.4 (9.3)	67.6 (9.6)	<0.001
Age at onset (years)	57.4 (11.0)	63.9 (9.7)	57.3 (10.7)	57.0 (11.1)	<0.001
Disease duration (months)	86.0 (58.2)	7.1 (3.6)	57.7 (42.5)	54.8 (54.7)	<0.001
HY stage	2.7 (0.7)	1.9 (0.6)	1.8 (0.6)	1.5 (0.5)	<0.001
UPDRS I	3.5 (2.6)	2.6 (2.1)	1.8 (1.7)	1.1 (1.4)	<0.001
UPDRS II	17.8 (7.0)	11.7 (5.0)	9.6 (4.1)	6.4 (3.3)	<0.001
UPDRS III	42.3 (13.1)	28.7 (10.5)	19.7 (7.8)	13.4 (7.0)	<0.001
UPDRS IV	4.6 (4.7)	3.2 (5.0)	3.2 (5.1)	2.5 (4.7)	<0.001
Tremor	1.0 (0.7)	0.7 (0.5)	0.4 (0.4)	0.7 (0.4)	<0.001
Rigidity	1.8 (0.8)	1.3 (0.8)	0.8 (0.6)	0.3 (0.4)	<0.001
Hypokinesia	1.8 (0.7)	1.2 (0.6)	0.9 (0.4)	0.5(0.3)	<0.001
PIGD	1.7 (0.8)	1.0 (0.5)	0.7 (0.4)	0.3(0.3)	<0.001
Motor phenotype score	0.7 (0.5)	0.8 (0.5)	0.7 (0.5)	2.7 (1.7)	<0.001
Progression rate	0.8 (0.6)	4.8 (2.4)	0.5 (0.5)	0.4 (0.4)	<0.001
MMSE	24.9 (4.1)	25.5 (4.3)	27.7 (2.4)	27.8 (2.5)	<0.001
HAMD	15.9 (10.6)	10.7 (8.5)	7.6 (6.9)	4.9 (5.7)	<0.001
PSQI	9.0 (4.2)	5.9(4.0)	5.6 (3.8)	3.9 (3.0)	<0.001
Constipation	0.8 (0.4)	0.5 (0.5)	0.5 (0.5)	0.4 (0.5)	<0.001

Data are expressed as mean (SD).
NTD: non-tremor dominant; RDP-LO: rapid disease progression with late onset; BPM: benign pure motor; TD-SP: tremor dominant with slow progression.
HY stage: Hoehn and Yahr stage; MMSE: Minimum Mental State Examination; HAMD: Hamilton Depression Scale; PSQI: Pittsburgh Sleep Quality Index;
Variables included in CA are marked in bold.