P-7

Parkinson`s Disease: A Classification of Long-Term Clinical Course Using Clustering Algorithms

Tae Soo Lee[1], Dougho Park[2,3], Su Yun Lee[4], Jong Hun Kim[5*], Hyoung Seop Kim[1*]

[1]Department of Rehabilitation Medicine, National Health Insurance Service Ilsan Hospital, Goyang, Republic of Korea
[2]Department of Rehabilitation Medicine, Pohang Stroke and Spine Hospital, Pohang, Republic of Korea
[3]Department of Medical Science and Engineering, Pohang University of Science and Technology, Republic of Korea
[4]Department of Neurology, Pohang Stroke and Spine Hospital, Pohang, Republic of Korea
[5]Department of Neurology, National Health Insurance Service Ilsan Hospital, Goyang, Republic of Korea

Objective

It remains challenging to subtype Parkinson's disease (PD) despite it having distinguishable disease courses. This retrospective cohort study aims to describe and cluster the natural course of Parkinson's disease (PD) with respect to functional disability and mortality.

Materials and Methods

Using the Korean National Health Insurance Service database, which contains the social support registry database for patients with PD, the initial sample comprised 31,167 patients with PD as the primary diagnosis. Subsequently, we extracted data regarding patients newly diagnosed with PD in 2009 and followed them up until the end of 2018. Functional disability was assessed based on the long-term care insurance (LTCI) and National Disability Registry data. All-cause mortality during the observation period was measured. The optimal number of clusters was determined using the 'NbClust' R software package and confirmed with the Hubert index, which resulting in six clusters. (Figure. 1) The baseline characteristics of these clusters were then analyzed using a divisive hierarchical tree for auto-clustering.

Results

We included 2,944 eligible patients. Table 1. shows the baseline characteristics according to the auto-clustering groups. The surviving patients were followed up for 113.7 ± 3.3 months. Among the patients who died, patients with and without disability registration were followed up for 61.4 ± 30.1 and 43.2 ± 32.0 months, respectively. The cumulative survival rate was highest in cluster 1 and decreased from cluster 1 to cluster 6 ($p<0.001$) (Figure. 2). All surviving patients were included in clusters 1 and 2, while patients who died during the observation period were included in clusters 3–6. Higher LTCI grade was associated with a significantly lower long-term mortality risk ($p < 0.001$) (Table 1.)

Table 1. Baseline characteristics according to the auto-clustering groups
Abbreviations: LTCI, long-term care insurance; NHIP, National Health Insurance Premium.

Variables	Cluster 1 (n = 750)	Cluster 2 (n = 979)	Cluster 3 (n = 425)	Cluster 4 (n = 547)	Cluster 5 (n = 187)	Cluster 6 (n = 56)	p-value
Age, years	69.3 ± 7.9	72.5 ± 7.3	76.1 ± 6.7	76.6 ± 6.7	75.0 ± 7.0	75.3 ± 8.3	<0.001
Male, n (%)	220 (29.3)	336 (34.3)	171 (40.2)	260 (47.5)	110 (58.8)	30 (53.6)	<0.001
Residence areas, n (%)							0.231
Metropolitan Seoul	163 (21.7)	212 (21.7)	82 (19.3)	96 (17.6)	30 (16.0)	10 (17.9)	
Metropolitan	118 (15.7)	182 (18.6)	74 (17.4)	110 (20.1)	38 (20.3)	10 (17.9)	
City	328 (43.7)	429 (43.8)	190 (44.7)	245 (44.8)	89 (47.6)	32 (57.1)	
County	141 (18.8)	156 (15.9)	79 (18.6)	96 (17.6)	30 (16.0)	4 (7.1)	
Health insurance types, n (%)							<0.001
Self-employed	225 (30.0)	278 (28.4)	120 (28.2)	144 (26.3)	53 (28.3)	14 (25.0)	
Employed	444 (59.2)	603 (61.6)	236 (55.5)	303 (55.4)	116 (62.0)	32 (57.1)	
Medical-aid	81 (10.8)	98 (10.0)	69 (16.2)	100 (18.3)	18 (9.6)	10 (17.9)	
NHIP levels, n (%)							<0.001
Medical-aid	81 (10.8)	98 (10.0)	69 (16.2)	100 (18.3)	18 (9.6)	10 (17.9)	
1st quartile	94 (12.5)	110 (11.2)	44 (10.4)	53 (9.7)	21 (11.2)	7 (12.5)	
2nd quartile	99 (13.2)	121 (12.4)	48 (11.3)	57 (10.4)	24 (12.8)	7 (12.5)	
3rd quartile	164 (21.9)	181 (18.5)	60 (14.1)	90 (16.5)	41 (21.9)	10 (17.9)	
4th quartile	312 (41.6)	469 (47.9)	204 (48.0)	247 (45.2)	83 (44.4)	22 (39.3)	
Comorbidities, n (%)							
Hypertension	556 (74.1)	750 (76.6)	344 (80.9)	432 (79.0)	144 (77.0)	40 (71.4)	0.092
Diabetes	303 (40.4)	432 (44.1)	205 (48.2)	279 (51.0)	91 (48.7)	28 (50.0)	0.003
Dyslipidemia	387 (51.6)	518 (52.9)	191 (44.9)	242 (44.2)	96 (51.3)	23 (41.1)	0.004
Ischemic heart diseases	63 (8.4)	80 (8.2)	38 (8.9)	61 (11.2)	19 (10.2)	7 (12.5)	0.381
Atrial fibrillation	34 (4.5)	42 (4.3)	16 (3.8)	39 (7.1)	12 (6.4)	4 (7.1)	0.097
Chronic kidney disease	13 (1.7)	11 (1.1)	14 (3.3)	34 (6.2)	9 (4.8)	6 (10.7)	<0.001
Stroke	355 (47.3)	528 (53.9)	227 (53.4)	300 (54.8)	111 (59.4)	33 (58.9)	0.012
Cancer	50 (6.7)	74 (7.6)	27 (6.4)	50 (9.1)	22 (11.8)	6 (10.7)	0.118
LTCI grade, n (%)							<0.001
1	5 (2.0)	61 (7.3)	27 (8.5)	37 (11.3)	46 (27.7)	0 (0.0)	
2	17 (6.9)	120 (14.4)	53 (16.6)	89 (27.2)	35 (21.1)	0 (0.0)	
3	73 (29.4)	553 (66.1)	219 (68.7)	201 (61.5)	85 (51.2)	0 (0.0)	
4	118 (47.6)	80 (9.6)	17 (5.3)	0 (0.0)	0 (0.0)	0 (0.0)	
5	32 (12.9)	20 (2.4)	3 (0.9)	0 (0.0)	0 (0.0)	0 (0.0)	
Others	3 (1.2)	2 (0.2)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
National Disability Registration grade, n (%)							<0.001
1	3 (0.4)	71 (7.3)	1 (0.2)	0 (0.0)	27 (14.4)	0 (0.0)	
2	4 (0.5)	97 (9.9)	0 (0.0)	0 (0.0)	44 (23.5)	0 (0.0)	
3	11 (1.5)	116 (11.8)	7 (1.6)	0 (0.0)	38 (20.3)	0 (0.0)	
4	14 (1.9)	81 (8.3)	1 (0.2)	0 (0.0)	11 (5.9)	0 (0.0)	
5	23 (3.1)	57 (5.8)	1 (0.2)	0 (0.0)	5 (2.7)	0 (0.0)	
6	5 (0.7)	27 (2.8)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
Out of grade	690 (90.2)	530 (54.1)	415 (97.6)	547 (100)	62 (33.2)	56 (100)	
Months to death[a]	113.7 ± 3.2	99.8 ± 19.8	72.0 ± 10.7	31.1 ± 13.8	25.3 ± 14.7	4.2 ± 3.0	<0.001
Months to LTCI registration[a]	106.6 ± 12.9	45.9 ± 36.6	38.4 ± 28.8	15.7 ± 14.7	11.1 ± 10.2	4.2 ± 3.0	<0.001
Months to National Disability Registration[a]	111.9 ± 8.0	69.1 ± 42.6	70.9 ± 13.2	31.1 ± 13.8	9.8 ± 7.9	4.2 ± 3.0	<0.001

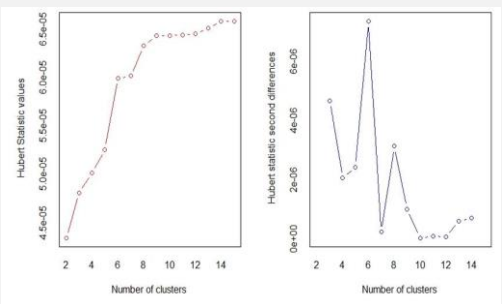

Fig. 1 Hubert statistic plots indicate that the optimal number of clusters was six.

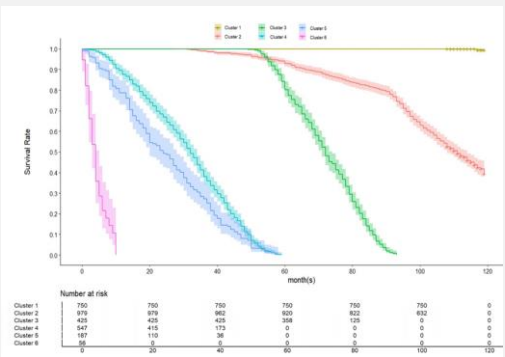

Fig. 2 Cumulative survival rate according to the auto-clustering groups.

Conclusion

Using longitudinal data from the social support registry and mortality data obtained from the NHIS database, we described six PD clinical course clusters and confirmed that PD progression is heterogeneous with respect to disability and mortality. Our findings suggest that the clinical course of patients with PD can be categorized into subtypes, enabling the establishment of personalized long-term management strategies for each patient.