PAPER

Heterogeneity of Parkinson's disease in the early clinical stages using a data driven approach

S J G Lewis, T Foltynie, A D Blackwell, T W Robbins, A M Owen, R A Barker

....

See end of article for authors' affiliations
....

Correspondence to:
S J G Lewis, Cambridge Centre for Brain Repair, Forvie Site, Addenbrooke's Hospital, Cambridge, CB2 2PY, UK; sjgl2@wbic.cam.ac.uk

Received 7 December 2003
In revised form 26 April 2004
Accepted 10 June 2004
....

Objective: To investigate the heterogeneity of idiopathic Parkinson's disease (PD) in a data driven manner among a cohort of patients in the early clinical stages of the disease meeting established diagnostic criteria.
Methods: Data on demographic, motor, mood, and cognitive measures were collected from 120 consecutive patients in the early stages of PD (Hoehn and Yahr I–III) attending a specialist PD research clinic. Statistical cluster analysis of the data allowed the existence of the patient subgroups generated to be explored.
Results: The analysis revealed four main subgroups: (*a*) patients with a younger disease onset; (*b*) a tremor dominant subgroup of patients; (*c*) a non-tremor dominant subgroup with significant levels of cognitive impairment and mild depression; and (*d*) a subgroup with rapid disease progression but no cognitive impairment.
Conclusions: This study complements and extends previous research by using a data driven approach to define the clinical heterogeneity of early PD. The approach adopted in this study for the identification of subgroups of patients within Parkinson's disease has important implications for generating testable hypotheses on defining the heterogeneity of this common condition and its aetiopathological basis and thus its treatment.

Idiopathic Parkinson's disease (PD) is a common condition that presents in a variety of ways suggesting clinical diversity. Clinicopathological reviews have highlighted the difficulties in clinical diagnosis of PD, with some series reporting an accuracy of less than 80%.[1] Consequent on this, the implementation of rigorous diagnostic criteria in the setting of a specialist assessment clinic greatly improves predictive value, such that over 95% of cases are accurately diagnosed,[2] although some "atypical" cases can be missed. Indeed in some early cases ^{18}F-dopa positron emission tomography imaging can be normal[3] while in advanced disease, coexistent pathologies confound the clinical presentation.[4]

Given these limitations, we elected to study patients in the early clinical stages of the disease using the United Kingdom Parkinson's Disease Society (UKPDS) Brain Bank criteria in the context of a specialist clinic to define heterogeneity. Previously attempts have used a "matched groups" approach with classifications based on predetermined patient attributes, such as age of disease onset,[5] cognitive performance,[6 7] motor phenotype,[8] the presence of depression,[9 10] disease severity,[8 11] or motor symptom laterality.[12 13] However, all such approaches suffer from the limitations arising from the prospective assumptions about the classification, namely the arbitrary division of patients based on the criteria adopted. Furthermore, inconsistencies between inclusion criteria and assessment techniques limit the comparability of results between studies, as does the small number of cases studied in some instances.

An alternative approach would be to use methodology that avoids the need for prospective definition and is capable of assessing variables in conjunction, rather than independently. One such technique is cluster analysis,[14] which seeks to divide patients into discrete clusters, such that any patient belongs to one cluster only, and the complete set of clusters contains all the patients. Each cluster should ideally have internal cohesion and external isolation[15] such that the clinical characteristics of patients within a given cluster should not vary greatly from one another but clear distinctions could be drawn between differing clusters.

This data driven approach for defining heterogeneity in PD has been used only once before[16] and identified three groups from a cohort of patients at all stages of the disease (including advanced): "motor only", "motor and cognitive", or "rapidly progressive". However, inclusion of patients with advanced disease is problematic as they may mask some of the subtle clinical variance that is seen only in the earlier stages of the disease as well as having associated comorbidity.[4] Furthermore, the methodology adopted in that study sought to derive different subgroups on the basis of all available patient data, so restricting post hoc comparisons for validation. To test the findings of their classification and circumvent problems in advanced cases, a cluster analysis where the validity of the results could be assessed was performed in 120 patients with mild to moderate PD attending a specialist research clinic. It was anticipated that this data driven methodology would identify the existence of distinct clinical subgroups in PD and by doing so generate testable hypotheses in this and other populations of patients with PD.

METHODS
Patients
A total of 120 patients with PD between Hoehn and Yahr stages I and III (77 men, 43 women; mean age 64.4 (SD 9.3) years; mean disease duration 7.8 (5.4) years) were included in this study, having been assessed consecutively in the Cambridge Centre for Brain Repair PD research clinic between January 2000 and July 2001. These patients were referred from multiple sources including general neurology

....

Abbreviations: BDI, Beck depression inventory; MMSE, Mini Mental State Examination; NART, National Adult Reading Test; PD, Parkinson's disease; PRM, pattern recognition memory; TOL, Tower of London (test); UPDRS, Unified Parkinson's Disease Rating Scale

clinics, geriatricians, and PD nurse specialists from across the region. All patients satisfied the UKPDS Brain Bank criteria[17] and were assessed by neurologists experienced in studying PD. Whilst no systematic imaging was undertaken, 80/120 patients had undergone recent brain scanning (38 computed tomography (CT) and 42 magnetic resonance imaging (MRI)) and no significant pathology was found in any case. Imaging was not performed in the remaining 40 cases, as the patients were felt to have classic idiopathic PD. Permission for the study was obtained from the local research ethical committee and all patients consented to participation. They were assessed in their ''best on'' state in a single session (lasting approximately two hours) allowing appropriate rest periods as required.

Demographics
Details of age at disease onset, disease duration, symptoms at onset, medications, motor fluctuations, L-dopa induced dyskinesia, and family history were recorded.

Clinical assessment
The patients were assessed on the mood and mentation, activities of daily living, and motor sections of the Unified Parkinson's Disease Rating Scale (I–III UPDRS).[18] The Hoehn and Yahr scale[8] was used for disease staging. Severity of depressive symptoms was derived from the Beck depression inventory (BDI)[19] with a cut-off of 8/9 indicating a degree of depression.[20]

Cognitive function
An estimate of premorbid verbal intelligence quotient (IQ) was derived using the National Adult Reading Test (NART),[21] and the patients were assessed on the Folstein Mini-Mental State Examination (MMSE).[22] Neuropsychological testing of verbal and categorical fluency (FAS 60 seconds,[23] animals 90 seconds[24]) and pattern (PRM) and spatial recognition memory (SRM) tests[25] (from the Cambridge Neuropsychological Test Automated Battery (CANTAB)) as well as performance accuracy on the Tower of London planning test (TOL)[26] were conducted. Normative data provided in the package for the CANTAB tests allowed these results to be compared with the expected performance of an age and intelligence matched healthy population. Mild impairment was defined as greater than one standard deviation below normal and significant impairment as more than two standard deviations below normal.

Quality of life
All patients completed the PDQ-39 self-assessment questionnaire,[27] which evaluates patient mobility, activities of daily living, emotional wellbeing, stigma, social support, communication, cognition, and bodily discomfort.

Statistical analysis
Non-hierarchical (k-means) cluster analysis was performed (SPSS-PC v10.0) on the 120 patients for a two, three, four, and five cluster solution and subsequent post hoc testing was undertaken to compare the patient subgroups. It should be stressed though that the results obtained from cluster analysis are highly dependent on the selection of appropriate variables and the number of subgroups sought. In the present study we selected the variables from a broad range of phenotypic features that have been previously highlighted in the literature as being significant in disease heterogeneity. These comprised the standardised values for age of disease onset, rate of disease progression, dopaminergic therapy, motor phenotype score, MMSE, NART, PRM, TOL, and BDI. Limiting the number of variables to those above allowed us to undertake a post hoc comparison confirmation of the clusters

by using independent variables linked but not included in the cluster analysis. In addition, by deliberately varying the number of subgroups sought through a range of possible solutions, the consistency of the subgroups could be confirmed.

Derivation of the variables used in the analysis
To allow some comparison of the rate of disease progression between patients with differing disease durations assessed at only a single time point, a variable was calculated by dividing the total UPDRS score for sections I–III by the disease duration (years). To assess dopaminergic therapy as an ordinal variable in the cluster analysis a scale ranging from zero to two was pre-defined. A score of zero represented no treatment with L-dopa or dopamine agonist (DA). A dopaminergic therapy score of one represented either an L-dopa dose below 1000 mg per day in isolation or combined with a DA, or DA monotherapy. Finally, a score of two was given to patients taking over 1000 mg of L-dopa with or without concomitant DA. We attempted to compare the tremor and non-tremor symptoms with the motor phenotype score[28], an approach similar to that adopted by others.[29] The score was obtained by dividing a patient's ''tremor score'' by their ''non-tremor score''. The tremor score was derived from the sum of items 16 and 20–26 on the UPDRS divided by 8 (the number of items included) and represents the degree of tremor reported in the activities of daily living section of the UPDRS, along with objective tremor at rest and with action, determined on physical examination. The non-tremor score was derived from the sum of items 5, 7, 12–15, 18, 19, and 27–44 on the UPDRS divided by 26 (the number of items included). This measure assesses speech, swallowing, ability to turn in bed, falls, freezing, and walking from the activities of daily living section as well as speech, facial expression, rigidity, bradykinesia, ability to stand, posture, gait, and postural stability determined by the motoric examination in the UPDRS.

The cognitive analysis included MMSE as a measure of global cognitive function and NART to encompass the effect of premorbid IQ. PRM, which is sensitive to temporal lobe damage,[26] was included and is generally unimpaired in the testing of patients with mild PD.[30] In contrast, the TOL is sensitive to executive impairment in PD even in the early stages of the disease[11 31] and was thus included in the analysis. The BDI was included in the analysis to determine the impact of any affective disorder in the subgroups.

Cluster analysis was repeated four times on the patient cohort to seek five, four, three, and two subgroup solutions so that the consistency of the clusters derived could be assessed. Post hoc comparisons on the subgroups generated by the cluster analysis were then performed using either a one way analysis of variance (ANOVA) or unpaired t tests for continuous variables, whilst categorical and ordinal data were analysed using the χ^2 technique.

RESULTS
Demographic data
Sex (77 men, 43 women), motor symptom laterality (54 predominantly left sided, 49 right sided, 17 no clear laterality), positive family history of PD (13 patients), and use of antidepressants, benzodiazepines, monoamine oxidase inhibitors, catechol-O-methyltransferase (COMT) inhibitors, and amantadine were not significantly associated with any of the subgroups for any clustering solution.

Clustering solutions
Results from the various clustering solutions revealed generally similar patterns for patient clusters and consistent features were observed. Four independent patient clusters

Table 1 Group characteristics for the four cluster solution

|	YO (n = 49)	T+ (n = 20)	RDP+ (n = 20)	NT+ (n = 31)
Age, years	60 (8)	66 (7)	66 (10)	69 (9)
Onset, years	50 (10)	59 (7)	62 (10)	62 (10)
Hoehn and Yahr	2 (1)	2 (1)	2 (1)	2 (1)
Unified Parkinson's Disease Rating Scale III	21 (10)	25 (12)	30 (11)	28 (13)
Duration, years	10 (6)	8 (4)	3 (1)	7 (5)
Disease progression score	4 (2)	5 (2)	13 (5)	8 (4)
L-dopa dose, mg	560 (468)	290 (264)	245 (229)	456 (580)
Motor phenotype score	1 (0.5)	2 (1)	1 (0.4)	0.6 (0.5)
Mini Mental State Examination	30 (1)	29 (1)	30 (1)	28 (1)
National Adult Reading Test	117 (6)	115 (8)	117 (6)	108 (10)
Verbal fluency (FAS)	46 (13)	40 (12)	44 (9)	33 (12)
Categorical fluency	24 (6)	21 (7)	22 (5)	17 (6)
Pattern recognition memory (max = 24)	21 (2)	18 (4)	21 (2)	17 (2)
Spatial recognition memory (max = 20)	16 (2)	14 (2)	15 (2)	14 (2)
Tower of London (max = 14)	10 (2)	10 (3)	11 (2)	5 (4)
Beck depression inventory	8 (6)	9 (5)	9 (5)	11 (6)

Values are mean (SD).
This clustering solution revealed four patient subgroups: younger onset (YO), tremor dominant (T+), non-tremor dominant (NT+), and rapid motor progression (RDP+). Group comparisons clearly demonstrated significant differences between numerous features of the disease. NT+ patients demonstrated cognitive deficits and had higher affective score ratings, while YO, T+, and RDP+ patients were distinguished by their age at onset, motor phenotype score, and rate of progression of motoric features, respectively. Similar distinctions were observed for other clustering solutions and notably the solution with five subgroups delineated the non-tremor dominant patients into those with moderate and those with severe cognitive impairment.

were readily identifiable: (*a*) patients with a younger disease onset; (*b*) a tremor dominant subgroup of patients; (*c*) a non-tremor dominant subgroup with significant levels of cognitive impairment and mild depression; and (*d*) a subgroup with rapid disease progression but no cognitive impairment. These distinctions were most clearly observed in the four cluster solution (table 1) and further validated in the five cluster solution, which served only to delineate the non-tremor dominant patients into those with moderate and those with severe cognitive impairment. It was felt that clustering solutions beyond five subgroups would have resulted in cohorts with numbers too small to allow significant conclusions to be drawn. The three cluster analysis demonstrated a younger onset, a tremor dominant,

and a non-tremor dominant cluster, and finally, the two cluster solution demonstrated a cluster with younger onset compared to the other patient cohort.

The validity of the subgrouping was explored by following the classification of patients in the various clustering solutions (fig 1). Patients generally demonstrated consistency in their subgroup between the various solutions, although a very small number of cases were classified in apparently contrasting subgroups. However, it should be noted that varying the number of subgroups generated by different clustering solutions would subtly alter the individual clinical characteristics of the cases contained within each cluster. For example, whereas in a three subgroup clustering solution the overall clinical characteristics of an individual case might best fit within the tremor dominant cohort, in a four subgroup solution the same case could be more closely aligned to a different clinical phenotype. Given that the present data showed very few instances of this occurring, such cases are unlikely to have influenced the overall clinical subgroups generated. In the three, four, and five cluster solutions, 46 patients were always classified as belonging to the younger onset subgroup, 16 were always in the tremor dominant subgroup, and 25 were always categorised as non-tremor dominant. Deliberately varying the number of clusters sought also revealed a meaningful pattern within the subgroups generated. This is clearly demonstrated for example, by the findings for the non-tremor dominant subgroup of patients who were always classified in this cohort throughout the three, four, and five cluster solutions (fig 1).

In all clustering solutions the group characterised by having a younger disease onset was found to have a slow rate of disease progression, mild motor symptoms, no cognitive impairment, and lower depression ratings. However, it is important to stress that the patients in this cohort were older than patients defined as having young onset PD[32] with its clear genetic associations. The validity of this subgroup is further supported by the data recording the motor complications of disease and usage of dopamine agonist medications, variables that were not included in generating the clustering solution. The younger onset patients showed a significant association with the switching "off" phenomenon (three clusters: $\chi^2 = 17.1$, p<0.05; four clusters: $\chi^2 = 26.1$, p<0.05; five clusters: $\chi^2 = 19.3$, p<0.05),
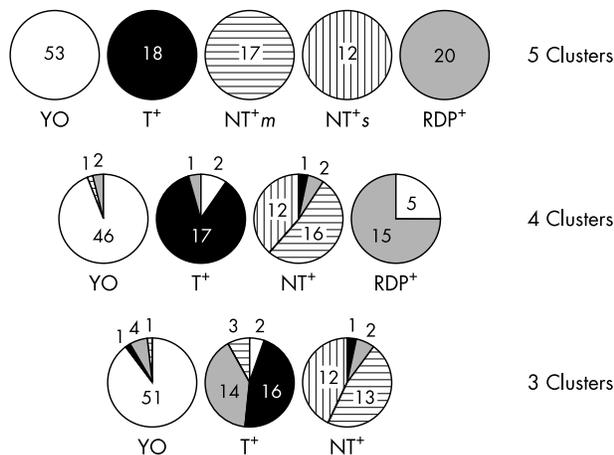
Figure 1 Patient "movement" between variable cluster solutions. The diagram shows the classification of the patient subgroups derived in the five cluster solution for the three and four cluster solutions. It is clear that most of the patients remain classified in the same subgroup between the varying cluster solutions, and this is well demonstrated by the 12 patients classified as NT+s in the five cluster solution who were classified as NT+ in the three and four subgroup clustering solutions. NT+, non-tremor dominant; NT+m, non-tremor dominant with moderate cognitive impairment; NT+s, non-tremor dominant with severe cognitive impairment; RDP+, rapid disease progression; T+, tremor dominant; YO, younger onset.

Table 2 Expected performance: subgroup characteristics

	Three clusters			Four clusters				Five clusters				
	YO	T+	NT+	YO	T+	RDP+	NT+	YO	T+	RDP+	NT+m	NT+s
PRM			†				†					††
SRM			†				†					†
TOL			††				††				†	††
BDI			*				*			*		*

*Mild depression
Only patients in the non-tremor dominant subgroup demonstrated any significant deficits (†one standard deviation below healthy controls) for pattern (PRM) and spatial recognition memory (SRM). These patients were also severely impaired on the Tower of London (TOL) task of executive function (††two standard deviations below healthy controls) and along with those patients with rapid motoric progression, also showed higher scores on the Beck depression inventory (BDI).
NT+, non-tremor dominant; NT+m, non-tremor dominant with moderate cognitive impairment; NT+s, non-tremor dominant with severe cognitive impairment; RDP+, rapid disease progression; T+, tremor dominant; YO, younger onset.

dyskinesias (three clusters: $\chi^2 = 8.55$, p<0.05), and a higher proportion of dopamine agonist use (three clusters: $\chi^2 = 8.5$, p<0.05; four clusters: $\chi^2 = 10.7$, p<0.05) compared with the other subgroups, in keeping with their earlier onset and prolonged duration of disease.

The tremor dominant subgroup was internally similar based largely on a tremor dominant motor phenotype score, but other concomitant features were consistent across the clustering solutions. These patients typically demonstrated a slow rate of disease progression, modest motor symptoms, no statistically significant cognitive impairments, and an absence of depressive symptoms (table 1). The validity of this patient subgroup is also supported by the results of variables regarding medication and presenting symptoms that were not included in the clustering solution. The tremor dominant subgroup was the only cohort to be significantly associated with the use of anticholinergic medication (four clusters: $\chi^2 = 12.4$, p<0.05; five clusters: $\chi^2 = 10.2$, p<0.05) and also demonstrated a significantly higher proportion of patients with the predominant symptom of tremor at presentation compared with the younger onset (three clusters: $\chi^2 = 4.1$, p<0.05; four clusters: $\chi^2 = 5.6$, p<0.05; five clusters: $\chi^2 = 6.1$, p<0.05) and non-tremor dominant cohorts (three clusters: $\chi^2 = 5.1$, p<0.05; four clusters: $\chi^2 = 4.16$, p<0.05; five clusters: $\chi^2 = 4.3$, p<0.05).

The non-tremor dominant subgroup also exhibited a robust pattern of results across the range of cluster solutions with characteristic motor phenotype scores, cognitive impairment most clearly demonstrated by executive dysfunction (table 2), significant ratings of depression scoring >9 on the BDI (table 2) and a generally more rapid disease progression than the younger onset and tremor dominant subgroups (table 1). Results on the PDQ-39 self-assessment questionnaire, another variable not included in the clustering solution, also added validity to the identification of this subgroup with these patients demonstrating significantly higher scores on the ratings for mobility (three clusters: $F_{2,117} = 5.6$, p<0.05; four clusters: $F_{3,116} = 4.1$, p<0.05; five clusters: $F_{4,115} = 4.0$, p<0.05) and cognitive impairment (five clusters: $F_{4,115} = 2.7$, p<0.05).

The subgroup of patients with rapid disease progression was only observed in the four and five cluster solutions and consistently demonstrated an aggressive course but no severe motor disability or cognitive impairment compared with the other subgroups (table 1). However, in the four and five cluster solutions the ratings of this subgroup on the BDI were indicative of mild depression (table 2). The L-dopa dose of these patients was significantly less than that of the younger onset cohort (four clusters: $t_{67} = 2.9$, p<0.05; five clusters: $t_{71} = 5.9$, p<0.05) and those patients identified in the five cluster solution with non-tremor dominant phenotype and severe cognitive impairment (five clusters: $t_{30} = 5.5$, p<0.05). However, the dose was similar to the doses taken by the tremor and other non-tremor dominant cohorts identified in the four and five cluster solutions.

DISCUSSION

The present study developed and confirmed the existence of distinct subgroups of patients in the early clinical stages of PD using a novel objective technique. The data driven approach robustly identified distinct cohorts of patients: *younger onset*, *tremor dominant*, *non-tremor dominant*, and *rapid disease progression*. These subgroups demonstrated consistent clinical features even when statistical clustering explored a number of different possible solutions, and when they were tested on variables not included in the original cluster analysis. The identification of these subgroups, while showing consistency in this study, and similar to that previously reported, serves more as a platform for testing hypotheses, rather than as a definitive classification system, which would require detailed clinicopathological study. Currently, no clinicopathological cohort has had sufficient data to derive these correlations[17 33] and it is hoped that this approach will ultimately form the conclusion of this study.

The issue of subtypes of PD has been addressed previously using different and less rigorous methodological approaches. Our results are in agreement with these studies, which consistently identified a cohort of PD patients with young onset of disease.[34–36] This subgroup demonstrates a slower rate of disease progression,[5 29 37 38] less cognitive impairment,[29 38 39] and a greater potential to develop motor fluctuations[4 17] (possibly as a consequence of the prolonged exposure to L-dopa) compared with subgroups with later disease onset, as also shown with our cluster analysis solution.

In contrast, the association with a predominance of tremor symptoms[29] and a positive family history[29] in younger onset patients was not seen in our study. Our failure to demonstrate any significant familial risk in the younger onset subgroup probably relates to the fact that this cohort had a mean age of 50 years at the emergence of symptoms. Genetic risks are greatest in younger patients (typically less than 30 years old) and only one of the 120 patients in our study was under the age of 30 years at onset (29.4 years). Therefore, it is unlikely that known genetic mutations, especially parkin[32] made any contribution to the subtyping, a conclusion supported by the small number of patients with a positive family history included in this study (n = 13).

Another variable that has attracted investigation in terms of heterogeneity within PD is the predominant motor

phenotype.[4] [8] [29] [38] [40–42] Many studies have made a distinction between a benign ''tremulous'' form of the disease compared with a more aggressive ''bradykinetic-rigid'' form, but these studies did not use the strict diagnostic guidelines followed in this work and thus the inclusion of non-PD cases may have affected the analyses. However, the tremor dominant subgroup identified here concurs with those of previous works which have shown this type of patient to have a slower rate of disease progression[8] [29] [38] [43] and less cognitive impairment[29] [39] [40] compared with non-tremor dominant patients. Our study did not provide evidence to support a younger age of disease onset and increased prevalence of a positive family history in the tremor dominant patients, as has been reported elsewhere,[29] suggesting that contamination of families exhibiting essential tremor was avoided.

The final subgroup identified in our analysis included patients with a rapid rate of motoric disease progression. This group may have been overlooked previously if only predetermined criteria such as motor phenotype were used. The one previous data driven approach to classification[16] identified a similar subgroup of patients with a rapid deterioration in motor and cognitive parameters. Interestingly, subsequent neuropsychological testing using a working memory paradigm on a selected cohort of patients from this subgroup has demonstrated specific deficits (SJGL, unpublished data). These patients may represent a parkinsonian syndrome other than idiopathic PD with their associated more rapid natural history[44] [45]; but given the use of the strict diagnostic criteria implemented in this study,[2] misdiagnosis is unlikely to be the explanation for this subgroup. Imprecise estimations for disease duration would also seem unlikely given that patient recall has been reported to be very accurate.[46] All patients in this study were managed by experienced neurologists and given that the L-dopa doses of this subgroup were similar to those found in the tremor and non-tremor dominant patient cohorts under treatment is not likely to have led to the artificial generation of this cluster. Furthermore they did not show deficits on a cognitive task of planning previously shown to be sensitive to dopaminergic effects.[30] The definition of this group of patients with aggressive disease is thus broadly in agreement with the earlier findings of Graham and Sagar.[16]

Previous studies have suggested differing neural pathologies for distinct motoric symptoms, such as akinesia[3] [47] [48] and tremor[49] [50] as well as the cognitive features of the disease.[28] [51–59] It is likely that different clinical subgroups probably have different pathological processes and foci,[60] [61] which in turn may have differing aetiological bases. This forms one of the future pathological aims of this study. Alternatively, the patient subgroups may reflect different genetic backgrounds with superimposed PD.[62]

Defining subgroups of patients with PD is helpful in delineating the natural history, prognosis, and therapeutic options. Recently, defining the optimal treatment for the specific motor[63–66] and cognitive symptoms has been of much interest.[67] Thus our identification of different patient subgroups may be a helpful predictor for current management strategies and may also be relevant in neurosurgical interventions such as deep brain stimulation, as well as being important in emerging novel experimental therapies.[68] [69]

Whilst this study has clearly identified subgroups of patients, the methodology is not without criticism—for example, to satisfy the diagnostic criteria for inclusion, all patients of PD will of course share some degree of homology. Furthermore, it is also clear that derivation of clusters that are significantly different from each other on the basis of the variables included in the analysis is almost inevitable. One needs to be aware that the choice and number of variables selected for inclusion as well as the number of clusters sought

can have profound effects on the results.[14] However, the variables included in the clustering solutions sought in this study were carefully selected not only to allow exploration of clinical heterogeneity but also to enable the comparison of variables not included in the cluster analysis, which validates any subgroup classification. Furthermore, the internal cohesion and external isolation of subgroups with differing clustering solutions further supported the existence of these discrete cohorts of patients with PD.

The cohort studied in this work comprised patients in the early clinical stages of the disease who were referred on a voluntary basis for participation in a specialist research clinic. As can be seen from the relatively young mean age of the participants this approach might have allowed an element of selection bias to occur although it must be stressed that these patients were recruited from a wide range of sources within the community. Furthermore, we deliberately targeted patients in the early clinical stages of disease as it is recognised that advanced disease may coexist with other neuropathological processes[4] and the loss of subtle information regarding clinical phenotype, which make the investigation of end stage disease complex and insensitive. These limitations may have potential implications when generalising the findings of this study to the more widespread PD community and highlight the need for further research.

In summary, this study reports the existence of clinical heterogeneity within the early clinical stages of PD using a data driven approach. The study identified a subgroup of PD patients with younger disease onset who have a slow rate of disease progression and no cognitive impairment; a tremor dominant subgroup of patients who are not cognitively impaired, and a non-tremor dominant subgroup with mild depression showing cognitive impairment that is most pronounced in the performance of executive functions. Finally, a subgroup of patients may exist who have a rapid rate of motor progression but no marked cognitive impairment. These subgroups obviously require validation on another independent and representative cohort of patients. A longitudinal assessment of the patients in this study, along with the long term aim of securing pathological data through post mortem analysis will also be vital. However, if confirmed, the existence of clinical subgroups is likely to be important in predicting prognosis and optimal therapy as well as reflecting differing aetiopathologies.

. . . .

Authors' affiliations
S J G Lewis, T Foltynie, R A Barker, Cambridge Centre for Brain Repair, University of Cambridge; and Department of Neurology, University of Cambridge, UK
A D Blackwell, Department of Psychiatry, University of Cambridge, UK
T W Robbins, Department of Experimental Psychology, University of Cambridge, UK
A M Owen, Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK

REFERENCES
1 **Hughes AJ**, Daniel SE, Kilford L, *et al.* Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992;**55**:181–4.

2 **Hughes AJ**, Daniel SE, Ben-Shlomo Y, *et al.* The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain* 2002;**125**:861–70.

3 **Morrish PK**, Sawle GV, Brooks DJ. Clinical and [18F]dopa PET findings in early Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1995;**59**:597–600.

4 **Hughes AJ**, Daniel SE, Blankson S, *et al.* A clinicopathologic study of 100 cases of Parkinson's disease. *Arch Neurol* 1993;**50**:140–8.

5 **Diamond SG**, Markham CH, Hoehn MM, *et al.* Effect of age at onset on progression and mortality in Parkinson's disease. *Neurology* 1989;**39**:1187–90.

6 **Lewis SJ**, Cools R, Robbins TW, *et al.* Using executive heterogeneity to explore the nature of working memory deficits in Parkinson's disease. *Neuropsychologia* 2003;**41**:645–54.

7 **Woods SP**, Troster AI. Prodromal frontal/executive dysfunction predicts incident dementia in Parkinson's disease. *J Int Neuropsychol Soc* 2003;**9**:17–24.

8 **Hoehn MM**, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967;**17**:427–42.

9 **Mayeux R**, Stern Y, Cote L, *et al.* Altered serotonin metabolism in depressed patients with Parkinson's disease. *Neurology* 1984;**34**:642–6.

10 **Santamaria J**, Tolosa E, Valles A. Parkinson's disease with depression: a possible subgroup of idiopathic parkinsonism. *Neurology* 1986;**36**:1130–3.

11 **Owen AM**, James M, Leigh PN, *et al.* Fronto-striatal cognitive deficits at different stages of Parkinson's disease. *Brain* 1992;**115**:1727–51.

12 **Direnfeld LK**, Albert ML, Volicer L, *et al.* Parkinson's disease. The possible relationship of laterality to dementia and neurochemical findings. *Arch Neurol* 1984;**41**:935–41.

13 **Tomer R**, Levin BE, Weiner WJ. Side of onset of motor symptoms influences cognition in Parkinson's disease. *Ann Neurol* 1993;**34**:579–84.

14 **Everitt BS**. *Cluster analysis.* New York: Edward Arnold, 1993.

15 **Everitt BS**. Commentary: classification and cluster analysis. *BMJ* 1995;**311**:535–6.

16 **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.

17 **Gibb WR**, Lees AJ. A comparison of clinical and pathological features of young- and old-onset Parkinson's disease. *Neurology* 1988;**38**:1402–6.

18 **Fahn S**, Elton R. Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Jenner P, *et al.*, eds. *Recent Developments in Parkinson's Disease.* New Jersey: Macmillan Health Care Information, 1987:153–63.

19 **Beck AT**, Ward CH, Mendelson M, *et al.* An inventory for measuring depression. *Arch Gen Psychiatry* 1961;**4**:561–71.

20 **Leentjens AF**, Verhey FR, Luijckx GJ, *et al.* The validity of the Beck Depression Inventory as a screening and diagnostic instrument for depression in patients with Parkinson's disease. *Mov Disord* 2000;**15**:1221–4.

21 **Nelson HE**. *National Adult Reading Test Manual.* Windsor: NFER-Nelson, 1982.

22 **Folstein MF**, Folstein SE, McHugh PR. ''Mini-mental state''. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;**12**:189–98.

23 **Benton AL**. Differential behavioural effects of frontal lobe disease. *Neuropsychologia* 1983;**6**:53–60.

24 **Goodglass H**. *Assessment of aphasia and related disorders.* Philadelphia: Lea and Febiger, 1972.

25 **Sahakian BJ**, Morris RG, Evenden JL, *et al.* A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain* 1988;**111**:695–718.

26 **Owen AM**, Sahakian BJ, Hodges JR, *et al.* Dopamine-dependent fronto-striatal planning deficits in early Parkinson's disease. *Neuropsychology* 1995;**9**:126–40.

27 **Jenkinson C**, Fitzpatrick R, Peto V, *et al.* The Parkinson's Disease Questionnaire (PDQ-39): development and validation of a Parkinson's disease summary index score. *Age Ageing* 1997;**26**:353–7.

28 **Lewis SJG**, Dove A, Robbins TW, *et al.* Cognitive impairments in early Parkinson's disease are accompanied by reductions in activity in frontostriatal neural circuitry. *J Neurosci* 2003;**23**:6351–6.

29 **Zetusky WJ**, Jankovic J, Pirozzolo FJ. The heterogeneity of Parkinson's disease: clinical and prognostic implications. *Neurology* 1985;**35**:522–6.

30 **Lange KW**, Robbins TW, Marsden CD, *et al.* L-dopa withdrawal in Parkinson's disease selectively impairs cognitive performance in tests sensitive to frontal lobe dysfunction. *Psychopharmacology (Berl)* 1992;**107**:394–404.

31 **Owen AM**, Beksinska M, James M, *et al.* Visuospatial memory deficits at different stages of Parkinson's disease. *Neuropsychologia* 1993;**31**:627–44.

32 **Lucking CB**, Durr A, Bonifati V, *et al.* Association between early-onset Parkinson's disease and mutations in the parkin gene. French Parkinson's Disease Genetics Study Group. *N Engl J Med* 2000;**342**:1560–7.

33 **Hughes AJ**, Daniel SE, Lees AJ. The clinical features of Parkinson's disease in 100 histologically proven cases. *Adv Neurol* 1993;**60**:595–9.

34 **Ludin SM**, Ludin HP. Is Parkinson's disease of early onset a separate disease entity? *J Neurol* 1989;**236**:203–7.

35 **Giovannini P**, Piccolo I, Genitrini S, *et al.* Early-onset Parkinson's disease. *Mov Disord* 1991;**6**:36–42.

36 **Dubois B**, Pillon B, Sternic N, *et al.* Age-induced cognitive disturbances in Parkinson's disease. *Neurology* 1990;**40**:38–41.

37 **Goetz CG**, Tanner CM, Stebbins GT, *et al.* Risk factors for progression in Parkinson's disease. *Neurology* 1988;**38**:1841–4.

38 **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**:1529–34.

39 **Portin R**, Rinne UK. Predictive factors for cognitive deterioration and dementia in Parkinson's disease. *Adv Neurol* 1987;**45**:413–16.

40 **Mortimer JA**, Pirozzolo FJ, Hansch EC, *et al.* Relationship of motor symptoms to intellectual deficits in Parkinson disease. *Neurology* 1982;**32**:133–7.

41 **Ebmeier KP**, Calder SA, Crawford JR, *et al.* Clinical features predicting dementia in idiopathic Parkinson's disease: a follow-up study. *Neurology* 1990;**40**:1222–4.

42 **Huber SJ**, Christy JA, Paulson GW. Cognitive heterogeneity associated with clinical subtypes of Parkinson's disease. *Neuropsychiatry, Neuropsychol Behav Neurol* 1991;**4**:147–57.

43 **Huber SJ**, Paulson GW, Shuttleworth EC. Relationship of motor symptoms, intellectual impairment, and depression in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1988;**51**:855–8.

44 **Colosimo C**, Albanese A, Hughes AJ, *et al.* Some specific clinical features differentiate multiple system atrophy (striatonigral variety) from Parkinson's disease. *Arch Neurol* 1995;**52**:294–8.

45 **Wenning GK**, Ben Shlomo Y, Magalhaes M, *et al.* Clinical features and natural history of multiple system atrophy. An analysis of 100 cases. *Brain* 1994;**117**:835–45.

46 **Reider CR**, Halter CA, Castelluccio PF, *et al.* Reliability of reported age at onset for Parkinson's disease. *Mov Disord* 2003;**18**:275–9.

47 **Playford ED**, Jenkins IH, Passingham RE, *et al.* Impaired mesial frontal and putamen activation in Parkinson's disease: a positron emission tomography study. *Ann Neurol* 1992;**32**:151–61.

48 **Nandi D**, Aziz TZ, Giladi N, *et al.* Reversal of akinesia in experimental parkinsonism by GABA antagonist microinjections in the pedunculopontine nucleus. *Brain* 2002;**125**:2418–30.

49 **Parker F**, Tzourio N, Blond S, *et al.* Evidence for a common network of brain structures involved in parkinsonian tremor and voluntary repetitive movement. *Brain Res* 1992;**584**:11–17.

50 **Doder M**, Rabiner EA, Turjanski N, *et al.* Tremor in Parkinson's disease and serotonergic dysfunction: An ^{11}C-WAY 100635 PET study. *Neurology* 2003;**60**:601–5.

51 **Cools R**, Stefanova E, Barker RA, *et al.* Dopaminergic modulation of high-level cognition in Parkinson's disease: the role of the prefrontal cortex revealed by PET. *Brain* 2002;**125**:584–94.

52 **Mattay VS**, Tessitore A, Callicott JH, *et al.* Dopaminergic modulation of cortical function in patients with Parkinson's disease. *Ann Neurol* 2002;**51**:156–64.

53 **Cash R**, Dennis T, L'Heureux R, *et al.* Parkinson's disease and dementia: norepinephrine and dopamine in locus ceruleus. *Neurology* 1987;**37**:42–6.

54 **Dubois B**, Pilon B, Lhermitte F, *et al.* Cholinergic deficiency and frontal dysfunction in Parkinson's disease. *Ann Neurol* 1990;**28**:117–21.

55 **McKeith IG**. Spectrum of Parkinson's disease, Parkinson's dementia, and Lewy body dementia. *Neurol Clin* 2000;**18**:865–902.

56 **Paulus W**, Jellinger K. The neuropathologic basis of different clinical subgroups of Parkinson's disease. *J Neuropathol Exp Neurol* 1991;**50**:743–55.

57 **Rinne JO**, Rummukainen J, Paljarvi L, *et al.* Dementia in Parkinson's disease is related to neuronal loss in the medial substantia nigra. *Ann Neurol* 1989;**26**:47–50.

58 **Marie RM**, Barre L, Dupuy B, *et al.* Relationships between striatal dopamine denervation and frontal executive tests in Parkinson's disease. *Neurosci Lett* 1999;**260**:77–80.

59 **Ito K**, Nagano-Saito A, Kato T, *et al.* Striatal and extrastriatal dysfunction in Parkinson's disease with dementia: a 6-[18F]fluoro-L-dopa PET study. *Brain* 2002;**125**:1358–65.

60 **Jellinger KA**. Post mortem studies in Parkinson's disease—is it possible to detect brain areas for specific symptoms? *J Neural Transm Suppl* 1999;**56**:1–29.

61 **Brooks DJ**. Functional imaging of Parkinson's disease: is it possible to detect brain areas for specific symptoms? *J Neural Transm Suppl* 1999;**56**:139–53.

62 **Foltynie T**, Goldberg TE, Lewis SJG, *et al.* Working memory ability in PD is associated with the COMT val108/158met polymorphism. *Mov Disord* (in press).

63 **Pogarell O**, Gasser T, van Hilten JJ, *et al.* Pramipexole in patients with Parkinson's disease and marked drug resistant tremor: a randomised, double blind, placebo controlled multicentre study. *J Neurol Neurosurg Psychiatry* 2002;**72**:713–20.

64 **Navan P**, Findley LJ, Jeffs JA, *et al.* Double-blind, single-dose, cross-over study of the effects of pramipexole, pergolide, and placebo on rest tremor and UPDRS part III in Parkinson's disease. *Mov Disord* 2003;**18**:176–80.

65 **Schrag A**, Keens J, Warner J. Ropinirole for the treatment of tremor in early Parkinson's disease. *Eur J Neurol* 2002;**9**:253–7.

66 **Marjama-Lyons J**, Koller W. Tremor-predominant Parkinson's disease. Approaches to treatment. *Drugs Aging* 2000;**16**:273–8.

67 **Aarsland D**, Laake K, Larsen JP, *et al.* Donepezil for cognitive impairment in Parkinson's disease: a randomised controlled study. *J Neurol Neurosurg Psychiatry* 2002;**72**:708–12.

68 **Barker RA**. Repairing the brain in Parkinson's disease: where next? *Mov Disord* 2002;**17**:233–41.

69 **Gill SS**, Patel NK, Hotton GR, *et al.* Direct brain infusion of glial cell line-derived neurotrophic factor in Parkinson disease. *Nat Med* 2003;**9**:589–95.