Evolution of cognitive dysfunction in an incident Parkinson’s disease cohort

C. H. Williams-Gray,1 T. Foltynie,1 C. E. G. Brayne,2 T. W. Robbins3 and R. A. Barker1

1Cambridge Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, 2Department of Public Health and Primary Care, University of Cambridge and 3Department of Experimental Psychology, University of Cambridge, UK

Correspondence to: Caroline Williams-Gray, MRCP, Cambridge Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, Forvie Site, Robinson Way, Cambridge, CB2 2PY, UK.
E-mail: chm27@cam.ac.uk

We have previously performed detailed clinical and neuropsychological assessments in a community-based cohort of patients with newly diagnosed parkinsonism, and through analysis of a subcohort with idiopathic Parkinson’s disease (PD), we have demonstrated that cognitive dysfunction occurs even at the time of PD diagnosis and is heterogeneous. Longitudinal follow-up of the cohort has now been performed to examine the evolution of cognitive dysfunction within the early years of the disease. One hundred and eighty (79%) eligible patients from the original cohort with parkinsonism were available for re-assessment at between 3 and 5 years from their initial baseline assessments. PD diagnoses were re-validated with repeated application of the UKPDS Brain Bank criteria in order to maximize sensitivity and specificity, following which a diagnosis of idiopathic PD was confirmed in 126 patients. Thirteen out of 126 (10%) had developed dementia at a mean (SD) of 3.5 (0.7) years from diagnosis, corresponding to an annual dementia incidence of 30.0 (16.4–52.9) per 1000 person-years. A further 57% of PD patients showed evidence of cognitive impairment, with frontostriatal deficits being most common amongst the non-demented group. However, the most important clinical predictors of global cognitive decline following correction for age were neuropsychological tasks with a more posterior cortical basis, including semantic fluency and ability to copy an intersecting pentagons figure, as well as a non-tremor dominant motor phenotype at the baseline assessment. This work clarifies the profile of cognitive dysfunction in early PD and demonstrates that the dementing process in this illness is heralded by both postural and gait dysfunction and cognitive deficits with a posterior cortical basis, reflecting probable non-dopaminergic cortical Lewy body pathology. Furthermore, given that these predictors of dementia are readily measurable within just a few minutes in a clinical setting, our work may ultimately have practical implications in terms of guiding prognosis in individual patients.

Keywords: Parkinson’s disease; cognitive; dementia; incidence

Abbreviations: PD = Parkinson’s disease


Introduction

Although Parkinson’s disease (PD) is defined classically in terms of its motor symptomatology, it has become apparent in recent years that non-motor deficits form an important part of the syndrome (Chaudhuri et al., 2006). Dementia in particular is a common feature of the disease, with an estimated prevalence of between 24 and 31% (Aarsland et al., 2005b). Furthermore, dementia has a major impact on quality of life (Schrag et al., 2000) and prognosis (Nussbaum et al., 1998), as well as being an important risk factor for nursing home placement (Aarsland et al., 2000), with consequent major health economics implications (Findley et al., 2003). The underlying pathophysiology of PD dementia is thought to involve limbic and neocortical Lewy body deposition, although neurofibrillary tangles and senile plaques may play a role in some patients and dysfunction of non-dopaminergic neurotransmitter systems, in particular the cholinergic system, has also been heavily implicated (reviewed in Williams-Gray et al., 2006).

Estimates of the incidence of dementia in PD have varied considerably (Table 1). However, a significant methodological problem with the majority of previous studies is the
inclusion of prevalent patients at varying disease stages. Only one study has investigated the development of dementia in a newly diagnosed PD cohort followed longitudinally; 48% of surviving patients ultimately met criteria for dementia at 15 years, but annual incidence figures were not calculated (Hely et al., 2005). Furthermore, this study and others used biased cohorts recruited from hospital outpatient clinics rather than the community (Mayeux et al., 1990; Biggins et al., 1992; Mahieux et al., 1998; Hughes et al., 2000). Accurate determination of the true natural history of cognitive dysfunction within PD requires detailed prospective longitudinal follow-up of an incident rather than prevalent cohort which is population-representative.

A further methodological concern in published studies investigating this issue in PD is the accuracy of dementia diagnosis. The majority of studies have used DSM criteria, essentially requiring the demonstration of memory impairment and higher cortical dysfunction which interfere significantly with social or occupational functional capacity (Diagnostic and Statistical Manual of Mental Disorders, 1994). Determination of whether or not a patient fulfils such criteria is subjective, and may differ substantially from centre to centre given that there is no standardized definition of ‘impairment’. The adoption of an additional more objective criterion, such as a cut-off score on a standardized cognitive rating scale (MMSE <24 has been widely used, see Tangalos et al., 1996), circumvents this problem to some extent. However, the MMSE is also affected by lack of standardization of administration between centres, and fulfilment of the criterion of a score below 24 may be confounded by a patient’s pre-morbid IQ. Hence an alternative outcome variable based on repeated assessments and reflecting rate of cognitive decline may actually be a more useful and informative measure of cognitive dysfunction for the individual patient. We have calculated such a measure for our patients in addition to estimating dementia incidence using standard DSM-IV and MMSE-based criteria.

The cognitive dysfunction of PD encompasses a wide range of deficits, but perhaps the most well defined are executive in nature, affecting ability to plan, organize and regulate goal-directed behaviour. Such deficits are demonstrable on frontally mediated tasks of working memory, planning and attentional set shifting (Owen et al., 1992, 1995a). Performance is typically influenced by both levodopa therapy (Lange et al., 1992; Owen et al., 1995a; Cools et al., 2001) and functional polymorphisms in genes involved in dopamine regulation (Foltynie et al., 2004b), supporting the notion that executive deficits reflect dopaminergic dysfunction in frontostriatal networks. Verbal fluency deficits are also well described in PD, which in part reflect deficiencies in executive search and retrieval processes. However, a recent large meta-analysis has demonstrated that PD patients are more impaired on tests of semantic than phonemic fluency, implying that pathology in the temporal lobe might contribute to the observed fluency impairment (Henry and Crawford, 2004b).

Further evidence for posterior cortically based deficits comes from studies describing explicit memory impairment (Helkala et al., 1988; Pillon et al., 1993) and visuospatial and constructional dysfunction in PD (Girotii et al., 1988; Levin et al., 1991; Pillon et al., 1991; Aarsland et al., 2003). Cognitive deficits in PD also seem to be heterogeneous in terms of their neurochemical basis. In particular, Dubois et al. (1987) have demonstrated that low-dose scopolamine causes impairment in visual recognition memory in patients with PD but not controls, implying that some cognitive deficits in PD are acetylcholine-dependent.

In terms of when such deficits arise in PD, it is clear that dysfunction across a range of neuropsychological domains occurs even in the earliest stages (Foltynie et al., 2004a; Levin and Katzen, 2005; Muslimovic et al., 2005). Our previous work has demonstrated that cognitive impairment is heterogeneous, with deficits exclusively in frontostriatally based tasks in 12%, deficits in temporal lobe-based tasks in 8% and global deficits in 15% of a community-based cohort of 159 newly diagnosed patients (Foltynie et al., 2004a), thus adding weight to the hypothesis that the neural substrates of these neuropsychological deficits are anatomically and/or neurochemically diverse. How these deficits in early disease relate to the later development of dementia is unclear, but in this study we have investigated this relationship through following up our incident cohort.
This is of the utmost importance for improving our understanding of the neural basis of PD dementia, as well as for identifying predictive markers of dementia to enable the implementation of therapeutic and supportive strategies at a stage of disease when they are more likely to be effective.

Previous attempts to determine early neuropsychological predictors of dementia have produced varied and inconsistent results: executive deficits (Mahieux et al., 1998; Levy et al., 2002a; Janvin et al., 2005), impaired verbal fluency (Jacobs et al., 1995; Mahieux et al., 1998), visuospatial deficits (Mahieux et al., 1998) and memory and language dysfunction (Levy et al., 2002a; Hobson and Meara, 2004) have all been suggested as useful prognostic markers in longitudinal studies. Their findings are, of course, influenced by the particular selection of neuropsychological tests employed. Different ‘executive’ tests, for example, can vary considerably in terms of their cognitive demands. Furthermore, it is clear that no human neuropsychological test is uniquely associated with damage to a specific brain region. We have selected for our assessments a range of well-validated tests, including several items from the CANTAB battery (Owen et al., 1990; Robbins et al., 1994) which depend to varying degrees upon frontal, temporal and parietal lobe function. This approach enables us to characterize the profile of impairments in our cohort and estimate the type of dysfunction predominantly associated with subsequent cognitive decline and the development of dementia.

We have previously established a community-based incident PD cohort and described its clinical and neuropsychological characteristics at the time of diagnosis (Foltynie et al., 2004a). In this study, we use outcome data from follow-up assessments between 3 and 5 years later to address three principal aims. These are: first, to estimate the incidence of dementia in early PD; secondly to redefine the profile of cognitive dysfunction in the disease; and finally, to determine which baseline clinical and neuropsychological variables best predict cognitive decline.

Material and Methods

A community-based cohort of 239 patients with incident Parkinsonism was recruited within the county of Cambridgeshire, UK, over a 2-year period between December 2000 and December 2002 using multiple sources of case ascertainment. Full details of the recruitment process have been published elsewhere (Foltynie et al., 2004a). All patients underwent comprehensive clinical and neuropsychological assessments at baseline. Follow-up assessments were conducted in patients’ own homes over a 13-month period between December 2004 and December 2005 and comprised a repeat of the full baseline assessment and an evaluation of dementia status (see later).

Diagnosis of PD

One hundred and fifty-nine of the 239 recruited patients with Parkinsonism were diagnosed with PD at baseline (Foltynie et al., 2004a) using the UK Parkinson’s Disease Society (UKPDS) Brain Bank criteria (Gibb and Lees, 1988). These criteria were reapplied at follow-up to the entire cohort, irrespective of initial diagnosis, to optimize diagnostic accuracy. Only those meeting a diagnosis of PD following this re-validation process were included in subsequent analyses of dementia incidence and factors predicting cognitive decline. Specifically, criteria for PD diagnosis were:

1. Fulfilment of the UKPDS Brain Bank criteria at baseline and follow-up assessments, or
2. Parkinsonism of uncertain cause at baseline and fulfilment of UKPDS Brain Bank criteria at follow-up, or
3. Suspected drug-induced parkinsonism at baseline but progression of symptoms despite discontinuation of the drug in question for >2 years and fulfillment of UKPDS Brain Bank criteria in other respects at follow-up.

Patients with suspected PD but significant cognitive impairment at baseline, i.e. an MMSE of <24 (Folstein et al., 1975), were excluded from further analysis, thus minimizing the possibility of erroneously including patients with Dementia with Lewy Bodies in our cohort.

Clinical assessment

This included a full history of the disease and co-morbid conditions, drug history, family history of neurological disease and a standardized neurological assessment including the Unified Parkinson’s Disease Rating Scale (Fahn and Elton, 1987) together with screening for atypical features pointing to other parkinsonian diagnoses. Patients were classified in terms of motor phenotype as ‘tremor dominant’, ‘mixed’ or ‘postural instability and gait disturbance’ (PIGD) on the basis of tremor and PIGD scores derived from the motor subsection of the UPDRS, as in previous studies (Zetuksky et al., 1985; Jankovic et al., 1990). All patients completed the Beck Depression Inventory (Beck et al., 1961). Functional independence scores were assessed using the Schwab and England scale (Schwab and England, 1969).

Doses of dopaminergic medication were converted to equivalent levodopa doses using the following formula, developed from those previously used in the literature (Broksky et al., 2003). Equivalent levodopa dose = [levodopa × (1.2 if COMT inhibitor) × (1.2 if 10 mg selegiline OR × 1.1 if 5 mg selegiline)] + [pramipexole × 400] + [ropinirole × 40] + [cabergoline × 160] + [pergolide × 200] + [bromocriptine × 10] + [lisuride × 160], all doses in mg.

Neuropsychological assessment

Patients underwent a battery of standardized, previously validated cognitive tests at both baseline (Foltynie et al., 2004a) and follow-up. These included the following: the 30-item Mini-Mental state examination (MMSE, Folstein et al., 1975); the National Adult Reading Test (NART, a measure of verbal IQ, Nelson and O’Connell, 1978); a test of phonemic fluency for words starting with the letters F, A and S for 1 min each (Benton, 1968); a test of semantic fluency for animals in a 90-s period (Goodglass, 1972); selected computerized neuropsychological tests from the CANTAB battery including Pattern and Spatial Recognition Memory (PRM and SRM; Sahakian et al., 1988) and the ‘one-touch’ Tower of London (TOL; Owen et al., 1995a). In addition, the interlocking pentagon copying item within the MMSE was graded using a 0–2 rating scale modified from Ala et al. (2001) with 2 points indicating that all 10 angles were present and the 2 pentagons
were intersecting, 1 point indicating that two intersecting figures were present, one with 5 angles and 0 indicating a less-acceptable copy.

The neuropsychological battery was selected to probe frontal, temporal and parietal lobe function, although the individual tests cannot be fully segregated in terms of their neuroanatomical basis. In terms of verbal fluency, phonemic and semantic tests are thought to load differentially on frontal and temporal processing, with both tests being impaired to a similar extent by frontal lobe damage but semantic fluency being predominantly impaired by temporal damage (Henry and Crawford, 2004a). CANTAB PRM is sensitive to temporal lobe, but not frontal lobe lesions, whereas SRM shows the opposite pattern of sensitivity (Owen et al., 1995b). Performance on the TOL, a test of planning and working memory, is also impaired in patients with frontal lobe damage (Owen et al., 1990), and functional imaging studies have confirmed that the test reliably activates the prefrontal as well as parietal cortex (e.g. (Baker et al., 1996; Lazeron et al., 2000). Further evidence supporting the idea that the CANTAB tests depend on different neural circuits comes from levodopa withdrawal studies and cross-sectional comparisons of medicated and unmedicated patients. These studies suggest that levodopa has a beneficial effect on performance on the TOL and SRM tests (Lange et al., 1992; Owen et al., 1995a; Swainson et al., 2000), but does not influence performance on the PRM test (Lange et al., 1992), supporting the concept that deficits on the latter arise from posterior cortical circuitry rather than dopaminergic, frontostriatal circuitry.

Tasks involving figure copying are used to detect constructional apraxia, a common feature of both PD dementia and Dementia with Lewy Bodies (DLB) (Aarsland et al., 2003; Cormack et al., 2004). It is widely recognized that visuospatial and constructional deficits follow focal parietal lobe lesions (Di Renzi, 1997) and functional imaging studies have confirmed that picture copying in particular activates the parietal cortex (Makuuchi et al., 2003). We selected the MMSE pentagon copying item as a constructional test for this study because it is widely utilized in a clinical setting and has been shown to be highly sensitive for distinguishing Lewy Body disease from Alzheimer’s disease (Ala et al., 2001).

A diagnosis of dementia at follow-up was made on the basis of an MMSE of less than 24 and fulfillment of DSM-IV criteria for dementia. Although the DSM-IV criteria rely on subjective judgements, their application was standardized as much as possible. Specifically, criterion A, requiring memory impairment plus impairment of at least one other ‘higher cortical function’ was evaluated using pre-defined cut-off values on our neuropsychological test battery (see later). Criterion B, relating to impairment of occupational and social functioning, required a functional independence score on the Schwab and England scale of less than or equal to 60% (denoting inability to perform certain activities of daily living) although some subjective judgement was required to determine whether this disability was attributable to cognitive rather than motor impairment.

Data analysis
Dementia incidence was calculated using the person-years method, i.e. by dividing the number of cases of dementia by the total number of ‘at risk’ person-years of follow-up. For non-demented patients, the number of ‘at risk’ years was simply the time period between assessments in years. For cases of incident dementia, time of dementia onset was assumed to be the midpoint of the interval between assessments, hence the number of ‘at risk’ years was calculated by halving the time interval between assessments (Aarsland et al., 2001). This analysis was restricted to those patients with a ‘definite’ diagnosis of PD, i.e. those whose diagnosis had been re-validated at follow-up. We did not attempt to adjust for mortality given that it was impossible to retrospectively determine a patient’s cognitive status in the period immediately prior to their death with any certainty.

Amongst those not meeting criteria for dementia, the proportion of patients scoring below a specified cut-off value was calculated for each test to allow the profile of mild cognitive impairments at follow-up to be determined. Similar cognitive profiling was performed retrospectively for baseline assessments for comparison with follow-up data. For the CANTAB tests, pre-defined cut-offs for impairment were adopted, as in our previous studies (Lewis et al., 2003; Foltynie et al., 2004a), i.e. number correct <14/20 for the SRM, <16/20 for the PRM and <8/14 for the TOL. These cut-off scores were defined as 1 SD below normative means in age and IQ-matched samples of healthy controls. Normative data for the PRM and SRM were provided by Cambridge Cognition, UK and for the ‘one-touch’ TOL these data were collected locally (Sahakian et al., unpublished data). Phonemic fluency impairment was defined as a score below a cut-off of 1 SD below the mean in age-matched controls, i.e. <25 words in 3 min (Tombaugh et al., 1999). The semantic fluency cut-off of 16 words in 90 s was derived from normative data (Tombaugh et al., 1999) in a similar fashion, adjusted by a factor of 1.2 to account for the fact that our animal-naming scores were collected over 90 rather than 60 s (which is associated with a mean score increment of 20% according to our unpublished data). Pentagon copying was considered to be impaired if a score of less than 2 was obtained.

In order to accurately identify predictors of cognitive decline and dementia, we used a rigorous and systematic approach of data encompassing two complementary outcome measures. In the first stage of analysis, change in MMSE per year, calculated by dividing the difference in MMSE scores at baseline and follow-up by the time interval between assessments, was adopted as an outcome variable. This continuous variable was selected rather than development of dementia per se in order to optimize power, given that the number of patients developing dementia at 3.5 years was small. Furthermore, as previously discussed, this variable could be argued to be more clinically relevant to individuals than meeting criteria for dementia. Simple bivariate analyses were performed to identify potential predictors of cognitive decline. Non-categorical clinical and neuropsychological variables from the baseline assessment were dichotomized at the median. Between-group comparisons were made using Student’s t tests or one-way ANOVA. In the second stage of analysis, baseline variables significantly associated with cognitive decline in the bivariate analyses (P ≤ 0.05) were entered into a multivariate regression analysis, again using change in MMSE per year as the dependent variable. A backward stepwise method was employed such that all potential predictor variables were entered into the model, and non-significant predictors (P > 0.10) were removed in a stepwise fashion. In the third stage of analysis, we sought to confirm an association between identified predictors of cognitive decline and the alternative outcome variable of dementia. Bivariate analyses were used, with the calculation of relative risks where appropriate. All analyses were conducted using SPSS version 11.5.
Results

Re-validation of diagnoses

Figure 1 summarizes the number and outcome of follow-up assessments conducted at a mean (SD) time from diagnosis of 3.5 (0.7) years. Ten patients from the original group with PD diagnosed according to the UKPDS Brain Bank criteria (‘PDBB’) had an MMSE of <24 at baseline. Retrospective application of DSM-IV criteria confirmed that they were demented at presentation, hence this subgroup were excluded from subsequent analyses. A further 14 patients from the ‘PDBB’ group died prior to the follow-up assessment. These patients were significantly older (\(P = 0.01\)) and had lower Schwab and England scores (\(P = 0.02\)) indicating more functional impairment than those assessed, but they did not differ in terms of gender or baseline UPDRS, MMSE or Beck depression scores (Supplementary Table 1). Those lost to follow-up from the ‘PDBB’ group (n = 7) did not differ significantly from those assessed on any of these variables (Supplementary Table 1). Re-validation of diagnoses through repeated application of the UKPDS Brain Bank criteria resulted in the exclusion of 12 patients from the original PD group, thus the criteria had a positive predictive value for a ‘definite’ PD diagnosis at follow-up of 90%. Of those with undefined parkinsonism at baseline, six fulfilled UKPDS Brain Bank criteria at follow-up, and a further six from the ‘drug-induced’ group also met diagnostic criteria for PD (all had progressive symptoms for >2 years post discontinuation of the drug in question), corresponding to negative predictive values of 50 and 66.6% for the criteria in these groups respectively. No patients with baseline diagnoses of Essential Tremor, Dystonic Tremor, Corticobasal Degeneration, Multiple System Atrophy, Dementia with Lewy Bodies or Vascular Parkinsonism met PD diagnostic criteria at follow-up, indicating a negative predictive value for the criteria of 100% in these patients. One hundred and twenty-six patients in total were included in the ‘PDBB’ group at 3.5 years and are included in our subsequent analyses.

Dementia incidence

Thirteen (10%) out of 126 patients met DSM-IV criteria for dementia at the follow-up assessment. Eleven of these had MMSE scores below 24. The remaining two had MMSE scores equal to 24, on the borderline of the arbitrary cut-off, hence it was deemed appropriate to also include them within the dementia group. There were no additional patients with MMSE-defined dementia who did not meet DSM-IV criteria. The estimated annual incidence of dementia was 30.0 per 1000 person-years of observation (95% confidence intervals 16.4–52.9 per 1000 person years). Exclusion of the 12 patients whose PD diagnosis was not established until the follow-up assessment did not alter the dementia incidence estimate (one case of dementia was diagnosed amongst this group).

---

**Fig. 1** Flow-chart summarizing the re-validation of diagnoses in a cohort of 239 patients with incident parkinsonism based on the outcome of follow-up assessments conducted at a mean (SD) time of 3.5 (0.7) years from baseline. ‘PDBB’ indicates a diagnosis of PD according to the UKPDS Brain Bank criteria. ‘Other’ parkinsonism indicates a diagnosis of Dystonic Tremor, Corticobasal Degeneration, Multiple System Atrophy, Dementia with Lewy Bodies or Vascular Parkinsonism. ‘Drug-induced’ indicates parkinsonism in the context of recent exposure to a neuroleptic drug; this diagnosis was revised at follow-up only if parkinsonism had progressed despite omission of the drug in question for a period of >2 years. ‘Unspecified’ indicates evidence of parkinsonism but with no clear diagnosis. ‘ET’ indicates essential tremor. ‘Patient unable to complete assessment due to recent major cerebrovascular event.'
Profile of cognitive impairment

At baseline, 62% of patients were impaired on at least one neuropsychological test. The overall proportion of patients with cognitive dysfunction was similar at 3.5 years, although 10% now met criteria for dementia. Amongst non-demented patients, deficits on SRM and TOL, which depend at least partially on frontal processing, were more common at the second assessment, whereas deficits on the more temporal lobe-based PRM were less common (Fig. 2). The subgroup with dementia exhibited a global pattern of cognitive deficits, with 12/13 being impaired on pentagon copying, 11/13 being impaired on phonemic fluency, and all being impaired on semantic fluency. Nine of 13 were unable to complete the CANTAB battery due to inability to comprehend the instructions for the tasks.

Risk factors for cognitive decline

Rate of change of MMSE ranged from +1.4 to −6.8 points per year (mean −0.3 ± 0.1). No other relevant co-morbid conditions were identified to account for the most extreme rates of cognitive decline. There was a clear relationship between change in MMSE and age, with a particular susceptibility to cognitive decline over 70 years (Fig. 3). Bivariate comparisons of baseline demographic, clinical and neuropsychological variables versus rate of cognitive decline are shown in Table 2. In addition to older age, a non-tremor dominant motor phenotype, a higher UPDRS motor score, and below average performance on tests of semantic fluency, pentagon copying, SRM and TOL were associated with a more rapid rate of cognitive decline (P < 0.05) and were therefore selected for inclusion in a multivariate analysis.

Multivariate analysis revealed that a non-tremor dominant motor phenotype, poor semantic fluency and inaccurate pentagon copying were the most significant predictors of cognitive decline, independently of age (Table 3). Inclusion of estimated IQ (NART) as a potential confounding factor in the multivariate model did not significantly change the outcome. Cross-group comparisons suggested that these predictor variables were associated with dementia risk per se in addition to an increased rate of cognitive decline (Table 4). Patients with a non-tremor-dominant phenotype were 4.1 times more likely to develop dementia than tremor-dominant patients, and those who were unable to accurately copy intersecting pentagons were 5.2 times more likely to dement. Verbal fluency deficits were stratified to differentiate between semantic and phonemic impairments. Whereas isolated semantic and global deficits were associated with a significantly increased
dementia risk (RR 9.5 and 9.4 respectively, reference group unimpaired), isolated phonemic deficits were not, thus implicating the semantic component of the fluency task in particular in predicting dementia.

**Discussion**

In a population-based cohort of 126 patients with incident PD followed up for a mean of 3.5 years, we have shown that dementia occurs at an estimated incidence rate of 30.0 per 1000 person-years. In addition to the 10% of patients meeting criteria for dementia at follow-up, a further 57% showed some degree of impairment on neuropsychological testing, hence two-third of the cohorts were cognitively impaired at 3.5 years. Through a systematic approach of analysis we have identified a number of clinical predictors of cognitive decline independent of age, including a non-tremor-dominant motor phenotype, impaired semantic fluency and impaired pentagon copying at presentation.

The major strengths of this longitudinal study lie in the nature of the cohort: it is incident, rather than cross-sectional, thus allowing us to monitor the evolution of cognitive dysfunction throughout the course of the disease, and it is community-based, thus representative of the true spectrum of PD within the population. These unique features are likely to account for the disparity between our estimate of dementia incidence in PD and previous estimates (Table 1). Our incidence figure, although lower than in other studies, is still notably higher than the dementia incidence rate for UK population at a comparable age: the MRC Cognitive Function and Ageing study (Matthews and Brayne, 2005) estimates a dementia incidence of 10.3 (95% CI 6.2–19.9) at age 70–74 years, the mean age of our cohort. It is possible that dementia incidence might increase with disease duration in PD, thus the low incidence in our cohort might reflect early disease stage. However, the population-based Rotterdam study reported no increase in hazard ratio for dementia with increasing disease duration in 166 PD dementia cases.

**Table 2** Bivariate comparisons of baseline demographic, clinical and neuropsychological variables versus rate of cognitive decline (change in MMSE per year) using Student’s t test (2 categories) or ANOVA (>2 categories)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Change in MMSE/year</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;72</td>
<td>−0.01</td>
<td>0.001</td>
</tr>
<tr>
<td>≥72</td>
<td>−0.61</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>−0.35</td>
<td>0.59</td>
</tr>
<tr>
<td>female</td>
<td>−0.25</td>
<td></td>
</tr>
<tr>
<td>Motor phenotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tremor dominant</td>
<td>−0.01</td>
<td>0.002</td>
</tr>
<tr>
<td>mixed/PIGD</td>
<td>−0.53</td>
<td></td>
</tr>
<tr>
<td>UPDRS motor score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25b</td>
<td>−0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>≥25</td>
<td>−0.52</td>
<td></td>
</tr>
<tr>
<td>Equivalent levodopa dose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>−0.28</td>
<td>0.32</td>
</tr>
<tr>
<td>1–250</td>
<td>−0.35</td>
<td></td>
</tr>
<tr>
<td>251–500</td>
<td>−0.64</td>
<td></td>
</tr>
<tr>
<td>501–750</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>751–1000</td>
<td>−0.14</td>
<td></td>
</tr>
<tr>
<td>NARTc (IQ)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;112</td>
<td>−0.31</td>
<td>0.76</td>
</tr>
<tr>
<td>≥112</td>
<td>−0.26</td>
<td></td>
</tr>
<tr>
<td>Phonemic fluency (FAS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;34</td>
<td>−0.49</td>
<td>0.06</td>
</tr>
<tr>
<td>≥34</td>
<td>−0.13</td>
<td></td>
</tr>
<tr>
<td>Semantic fluency (animals)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>−0.62</td>
<td>0.001</td>
</tr>
<tr>
<td>≥20</td>
<td>−0.01</td>
<td></td>
</tr>
<tr>
<td>Pattern recognition memory score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;19</td>
<td>−0.46</td>
<td>0.17</td>
</tr>
<tr>
<td>≥19</td>
<td>−0.19</td>
<td></td>
</tr>
<tr>
<td>Spatial recognition memory score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;15</td>
<td>−0.57</td>
<td>0.05</td>
</tr>
<tr>
<td>≥15</td>
<td>−0.13</td>
<td></td>
</tr>
<tr>
<td>Tower of London score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;10</td>
<td>−0.62</td>
<td>0.01</td>
</tr>
<tr>
<td>≥10</td>
<td>−0.07</td>
<td></td>
</tr>
<tr>
<td>Pentagon copying score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>−1.21</td>
<td>0.03</td>
</tr>
<tr>
<td>1</td>
<td>−0.52</td>
<td>1.82</td>
</tr>
<tr>
<td>2</td>
<td>−0.21</td>
<td>0.85</td>
</tr>
<tr>
<td>Beck depression score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;7</td>
<td>−0.26</td>
<td>0.09</td>
</tr>
<tr>
<td>≥7</td>
<td>−0.35</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Continuous variables are dichotomized at the median, with the exception of levodopa dose, which is stratified into five subgroups.

aPreliminary analyses suggested similar rates of cognitive decline in PIGD and mixed subgroups, hence these were combined into a single subgroup for analysis.

bUnified Parkinson’s Disease Rating Scale.

cNational Adult Reading Test.

**Table 3** Multivariate regression model with change in MMSE per year as the dependent variable; $F = 6.79$, $P < 0.001$, $R^2 = 0.195$

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$ coefficient</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>−0.76</td>
<td>0.04</td>
</tr>
<tr>
<td>Age ≥72</td>
<td>−0.38</td>
<td>0.05</td>
</tr>
<tr>
<td>Non-tremor-dominant motor phenotype</td>
<td>−0.35</td>
<td>0.07</td>
</tr>
<tr>
<td>Pentagon copying score</td>
<td>−0.35</td>
<td>0.05</td>
</tr>
<tr>
<td>Semantic fluency &lt;0.20</td>
<td>−0.38</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Note: ref = reference category for comparison.

Verbal fluency deficits defined as performance <1 SD below mean (cphonemic score <25 and semantic score >16; asemantic score <16 and phonemic score ≥25; dsemantic score <16 and phonemic score ≥25).

although the majority of these were established rather than incident cases hence subject to recruitment bias (de Lau et al., 2005). Further longitudinal follow-up of our incident cohort should resolve this question, allowing estimation of disease duration-specific incidence rates for dementia in PD.

A further strength of this study is the rigorous process applied to define PD cases. We used the UKPDS Brain Bank criteria (Gibb and Lees, 1988), estimated in a clinicopathological study to have a diagnostic accuracy of 90% (Hughes et al., 2001). Furthermore, given that diagnoses made at a single time point are likely to be less accurate than those made on the basis of longitudinal data, we followed up our entire cohort of parkinsonian patients to 3.5 years, and reapplied the same diagnostic criteria to all patients for a second time. Through this process we identified 12 patients with false positive PD diagnoses at presentation (10%) and a further 12 patients with false negative diagnoses (40% of unspecified/’drug-induced’ group), thus we were able to optimize specificity and sensitivity of diagnosis at follow-up. Although this process resulted in the inclusion of 12 patients in the final cohort who were ‘officially’ diagnosed with PD at a later time than the other 114, all 12 had significant parkinsonian symptoms at baseline, and comparison of mean baseline UPDRS motor scores in this group versus the 114 meeting UKPDS Brain Bank criteria at both assessments revealed no significant difference (27.0 versus 25.4, respectively, P = 0.97, Mann–Whitney U test). Furthermore, given that the association between clinical signs and pathological stage is unclear in PD, time of diagnosis is a somewhat arbitrary measure.

One concern with all longitudinal studies of this type is attrition, which potentially introduces bias. In our study 21% of 229 eligible patients were not available for follow-up, principally due to death (15%), whilst 7% were uncontrollable or withdrew consent. We suspect that bias due to attrition was minimal amongst the subcohort with ‘definite’ PD at baseline, the seven patients lost to follow-up did not differ clinically from those assessed. The 14 patients who died within this group were not surprisingly significantly older and had more functional impairment, probably reflecting their greater comorbidity, but notably did not differ from those assessed in terms of their baseline MMSE scores (Supplementary Table 1).

Amongst the group with possible or ‘drug-induced’ parkinsonism at baseline, there was an attrition rate of 23%. However, through applying the expected negative predictive values of the UKPDS Brain Bank criteria in these groups of 50 and 66.6%, respectively, we anticipate that this degree of attrition would have led to the loss of only four true PD cases from the cohort.

Studies describing cognitive dysfunction in PD tend to focus on the development of dementia, although it is important to establish the frequency of milder degrees of cognitive impairment, which also impact significantly on quality of life (Frank et al., 2006). The cognitive profile of this cohort has previously been described at baseline, when performance on three tests, namely the MMSE, PRM and TOL, was used to divide patients into subgroups with suspected frontostriatal impairment, temporal lobe type impairment, global impairment and no impairment (Foltynie et al., 2004a). Preliminary analysis of our follow-up data, however, revealed that these subgroups were uninformative in terms of predicting cognitive outcome. This is not particularly surprising given that first, numbers within each subgroup were small, and secondly, the tests used to determine subgroups were chosen arbitrarily, without the benefit of longitudinal data, with the purpose of simply establishing that the cohort was
cognitively heterogeneous. Here, we adopted an alternative approach to describe the profile of cognitive dysfunction, based on the proportion of patients impaired on each test. This profiling (Fig. 2) reveals little change in the overall proportion of patients with cognitive impairment at follow-up versus baseline (67 versus 62%) but illustrates a change in the pattern of impairments: dementia develops in 10% and amongst the remaining non-demented patients, frontostriatally based impairments become more frequent whilst deficits on the more temporal lobe-based PRM become less frequent with time. This may indicate the beginnings of a segregation of demented and non-demented patients in terms of their cognitive profiles, but further longitudinal data needs to be collected to explore this possibility. Alternative neuropsychological tools would be necessary to adequately probe the pattern of cognitive deficits amongst the demented group given that 9/13 were unable to complete the CANTAB PRM, SRM and TOL. These particular tests were primarily chosen to probe deficits amongst those with mild cognitive impairment and may require a threshold level of attentional capacity and short-term recall which some of our demented patients were unable to meet.

One of our key aims was to determine early clinical predictors of cognitive decline in PD. Age has consistently been demonstrated to be an important predictor of both dementia (Mahieux et al., 1998; Aarsland et al., 2001; Levy et al., 2002b; Hobson and Meara, 2004) and cognitive decline (Aarsland et al., 2004), and our study is in agreement with this finding. We have also identified three further predictors which influence rate of cognitive decline independently of age, namely motor phenotype, semantic fluency and pentagon copying performance. In addition, each of these predictors significantly increased the risk of meeting dementia criteria at 3.5 years, although the confidence intervals for the relative risk estimates were wide as a consequence of the small numbers with dementia at follow-up. Nonetheless, these RR estimates provide a preliminary indication of the value of these variables in predicting dementia. The predictors are particularly informative in combination, thus 62.5% (5/8) of patients with all three identified risk factors developed dementia versus none of the 39 patients with no such risk factors.

The association between motor symptom type and the development of cognitive impairment in PD has been examined previously. Our own cross-sectional cluster analysis of 120 PD patients identified a non-tremor dominant subgroup with significant cognitive impairment and a cognitively unimpaired tremor-dominant group (Lewis et al., 2005). In longitudinal studies, axial and speech impairments have been associated with incident dementia (Levy et al., 2000) and a more rapid rate of cognitive decline (Aarsland et al., 2004), as has PIGD motor phenotype (Burn et al., 2006). Furthermore, recent evidence suggests a temporal relationship between the development of PIGD symptoms and occurrence of dementia (Alves et al., 2006). Hence our results, suggesting an association between non-tremor-dominant phenotypes and increased rate of cognitive decline, which still approaches significance in a multivariate model with correction for age, are broadly in keeping with those in the literature. It is possible that axial symptoms and PD dementia have overlapping aetiopathologies, with distinct loci of dysfunction different from those underlying tremor-dominant PD. Specifically, the postural instability of PD tends to be refractory to dopaminergic therapy, and may relate to dysfunction within the cholinergic system (Burn et al., 2003) which also seems to play an important role in the dementia of PD (Whitehouse et al., 1983; Perry et al., 1985; Hilker et al., 2005).

Verbal fluency impairment is well described in PD (Henry and Crawford, 2004b), and several other authors have identified fluency deficits as predictors of later-occurring dementia (Jacobs et al., 1995; Mahieux et al., 1998; Levy et al., 2002a), though phonemic rather than semantic fluency has typically been implicated (Mahieux et al., 1998; Levy et al., 2002a). There is clearly a degree of overlap in terms of the underlying neurobiological basis of phonemic and semantic fluency deficits, with both relying on frontally based executive strategies. However, a difference in relative performance on the two tasks may be useful in terms of neuroanatomical localization, given that a disproportionate impairment of semantic fluency is seen in Alzheimer’s disease and semantic dementia (Hodges et al., 1992), in contrast to more prominent phonemic fluency deficits in subcortical dementias (Green, 2000). Furthermore, lesion studies indicate that temporal lobe damage produces significantly greater deficits in semantic compared to phonemic fluency, whereas frontal lobe lesions produce more comparable semantic and phonemic deficits (Henry and Crawford, 2004a). These observations probably reflect a reliance of semantic fluency on semantic memory, whose neural substrate is widely accepted to lie within the temporal lobe. Indeed, a functional imaging study has confirmed activation of the left medial temporal lobe during a category fluency task in contrast to a number listing task (Pihlajamaki et al., 2000). Hence our finding that semantic but not phonemic fluency is a useful predictor of cognitive decline and dementia in PD has interesting implications with regard to the site of pathology in PD dementia, implying a posterior cortical rather than frontostriatal basis.

However, some studies have suggested that the relationship between fluency performance and neuroanatomical site of pathology is more complex [reviewed in (Troyer et al., 1998a)]. Hence the number of words generated within a specific time may not be the most informative measure in this respect, but rather ‘clustering’ (i.e. producing clusters of semantically or phonemically related words) and ‘switching’ (i.e. shifting from one subcategory of words to another) may correlate more specifically with lesion site. Specifically, frontal lobe patients tend to be impaired on
switching during phonemic fluency tasks, whereas temporal lobe patients tend to generate smaller clusters of words during semantic fluency tasks (Troyer et al., 1998a). The same pattern is claimed by these authors for PD dementia and Alzheimer’s disease, respectively (Troyer et al., 1998b), implying that PD dementia is a frontal rather than temporal lobe syndrome, in contrast to our findings. This discrepancy may relate to the definition of ‘dementia’ in their study; this was based on Dementia Rating Scale score alone, with no reference to functional disability, which is generally recognized as a key component of the dementia syndrome. Their PD dementia group may actually be comparable to our non-demented but cognitively impaired group, in whom frontally based deficits certainly do appear to be more prominent. Whilst further investigation of the prognostic significance of clustering and switching performance is warranted, one clear advantage of the more simplistic measure of ‘number of words generated’ is its ease of use in general clinical practice. Furthermore, irrespective of its underlying neural basis, this measure does appear to predict cognitive decline in PD.

The apparent utility of pentagon copying, as a measure of visuospatial and constructional ability, as another predictor of cognitive decline in PD is not unexpected given that such deficits are prominent in DLB (Ala et al., 2001; Tiraboschi et al., 2006) and PD dementia (Cormack et al., 2004). Other authors have similarly reported that constructional deficits predict dementia in PD (Mahieux et al., 1998) and it is widely accepted that these deficits reflect parietal lobe dysfunction (Di Renzi, 1997; Makuuchi et al., 2003). These clinical observations, together with recent post-mortem data (Aarsland et al., 2005a), provide strong support for the hypothesis that the dementia of PD has a posterior cortical basis. Although Lewy body deposition seems the most likely aetiological factor, the extent to which posterior cortically based deficits are influenced by subcortical dopaminergic and cholinergic systems remains unresolved given that these cortical areas receive innervation from both systems (Berger et al., 1991; Hurd et al., 2001).

In conclusion, we have established through prospective follow-up of an incident population-based cohort that two-thirds of patients with early PD develop cognitive deficits within 3.5 (+/−0.7) years from disease onset. These deficits include both frontospatially based and more posterior cortical impairments, and 10% develop a global dementia. Furthermore, we have demonstrated that age, a non-tremor-dominant motor phenotype, and poor performance on two simple neuropsychological tests, namely semantic fluency and pentagon copying, are useful predictors of global cognitive decline. These predictors are particularly valuable in that they are all readily measurable within just a few minutes in an outpatient clinic setting. Thus this unique prospective study of an incident PD cohort using a rigorous process of assessment and analysis not only adds to our understanding of the cognitive heterogeneity of PD and the basis of PD dementia, but its findings have practical prognostic implications for clinical practice.

Supplementary material
Supplementary material is available at BRAIN online.

Acknowledgements
We would like to dedicate this paper to the memory of Imogen Rose Barker, who inspired us to do so much, and who was sadly killed on February 24, 2007 aged 15 years. The work was supported by grants from the Medical Research Council and the Parkinson’s Disease Society. C.H.W.G. is a Patrick Berthoud Clinical Research Fellow, and holds a Raymond and Beverly Sackler Studentship.

References
Burn DJ, Rowan EN, Allan LM, Molloy S, O’Brien JT, McKeith IG. Motor subtype and cognitive decline in Parkinson’s disease, Parkinson’s disease...
Cognitive dysfunction in Parkinson’s disease


