Journal of the International Neuropsychological Society (2009), **15**, 426–437.
Copyright © 2009 INS. Published by Cambridge University Press. Printed in the USA.
doi:10.1017/S1355617709090614

Cognitive decline in Parkinson's disease: A prospective longitudinal study

DINO MUSLIMOVIĆ,[1] BART POST,[1] JOHANNES D. SPEELMAN,[1] ROB J. DE HAAN,[2] AND BEN SCHMAND[1,3]

[1]Department of Neurology, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands
[2]Department of Clinical Epidemiology and Biostatistics, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands
[3]Department of Psychonomics, University of Amsterdam, Amsterdam, The Netherlands

(RECEIVED October 15, 2008; FINAL REVISION February 6, 2009; ACCEPTED February 9, 2009)

Abstract

This controlled prospective study examined the evolution and predictors of cognitive decline in Parkinson's disease (PD). Consecutive patients diagnosed at baseline with PD (n = 89), established PD (EPD) patients (n = 52) with a mean disease duration of 6.5 years, and healthy control subjects (n = 64) underwent extensive neuropsychological assessment twice, approximately 3 years apart. A standardized regression-based method, normative data, and multivariate normative comparisons were used to assess the cognitive course of PD. Cognitive performance of newly diagnosed patients decreased significantly over time, particularly on measures of psychomotor speed and attention and to a lesser extent on tests of memory, visuospatial skills, and executive functions. About 50% of the patients showed cognitive decline and 9% developed dementia. Similar results were observed in EPD patients. None of the baseline features predicted cognitive change in newly diagnosed patients, whereas age at disease onset and axial impairment (postural and gait disorders) contributed to decline in established patients. We conclude that within few years after diagnosis, PD patients show faster rate of cognitive decline than matched healthy subjects, particularly in domains of attention and psychomotor speed. Selection bias probably led to underestimation of the true extent of cognitive decline in established patients. (*JINS*, 2009, *15*, 426–437.)

Keywords: Parkinson's disease, Cognitive impairment, Natural history, Neuropsychological tests, Longitudinal design, Prognostic factors

INTRODUCTION

Cognitive functions in Parkinson's disease (PD) tend to deteriorate over time, and up to 80% of patients may eventually develop dementia (Aarsland et al., 2003; Hely et al., 2008). However, little is known about the rate of change in different cognitive abilities or the pattern of evolution of cognitive deficits. The brevity of follow-up intervals (often 1 year or less), limited assessment of cognitive function (usually just global dementia screening measures), lack of control groups, and insufficient control of practice effects hamper the ability to draw firm conclusions about the nature and extent of cognitive decline in PD (Muslimovic et al., 2007). With few exceptions (Reid et al., 1996; Williams-Gray et al., 2007), most studies have examined patients with long-standing disease. In addition, most studies have evaluated cognitive change at a group level, not accounting for individual variability in the trajectory of decline, which is to be expected given the heterogeneity of PD. Hence, it remains to be established how different cognitive domains change over time after the onset of disease, which individual patients exhibit cognitive decline, and what factors predict cognitive deterioration.

Previously, we described the profile of cognitive dysfunction in a cohort of PD patients at the time of diagnosis (Muslimovic et al., 2005). We now report the 3-year follow-up in this cohort. The aims of this study were to (1) assess the nature, magnitude, and frequency of cognitive change in PD patients compared with demographically matched healthy controls; (2) identify individual patients exhibiting cognitive decline; and (3) determine baseline clinical and demographic variables that best predict

Correspondence and reprint requests to: Dino Muslimović, Department of Neurology, Academic Medical Centre, University of Amsterdam, H2-222, PO Box 22660, 1100 DD Amsterdam, The Netherlands. E-mail: dinomuslimovic@hotmail.com

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

cognitive decline. In addition, we also report 3-year follow-up data from a cohort of more advanced PD patients in an attempt to examine the influence of disease duration on the pattern of cognitive decline. This issue has rarely been studied in a longitudinal design.

METHODS

Subjects

A total of 116 patients with newly diagnosed PD (NDPD) and 64 patients with established PD (EPD) participated in the neuropsychological assessment at baseline. These patients were part of a larger sample of PD patients who participated in a 3-year longitudinal research project investigating the course of functional decline in PD. Details of the case-finding procedure have been published elsewhere (Muslimovic et al., 2005). In brief, 133 consecutive patients with NDPD were recruited between July 2002 and April 2005 from the neurology outpatient clinics of six general hospitals in Amsterdam and surrounding area in the Netherlands. In addition, 69 patients with more advanced PD were identified from the medical records of the participating hospitals and from the Dutch Parkinson's Disease Association. A diagnosis of PD was confirmed in all patients at entry to the study by the project neurologist based on standard clinical criteria (Gelb et al., 1999). Exclusion criteria were age 85 years or older, insufficient command of the Dutch language, neurosurgical operation for relief of motor symptoms, global cognitive deterioration [Mini-Mental State Examination (MMSE) <24] (Folstein et al., 1975), and the presence of somatic illness with a life expectancy of less than a year.

Seventy spouses or friends of the patients volunteered to serve as control subjects. They were relatives or friends who accompanied the patients to the clinics or they were invited by the patients to participate in the study. None of the controls had impairment in hearing and visual acuity, history of major psychiatric disorders, head injury with loss of consciousness, cerebrovascular disease, current alcohol, substance or psychoactive medication abuse, or other central nervous system diseases that could influence cognitive performance. Written informed consent was obtained from all subjects. The study was approved by the local ethics committees of the participating hospitals.

Over 3 years, 5 patients from the NDPD group had died, 17 refused retesting, 3 were not able to participate in follow-up evaluation due to worsening health condition (stroke, renal disease, and pneumonia), and 2 could not be retested due to dementia or severe confusion. Within the EPD group, three patients had died, six refused retesting, one underwent neurosurgery and was no longer eligible, and two were not able to participate because of colon cancer with complications and severe dementia, respectively. Three control subjects refused retesting and three had moved and could not be traced anymore. Thus, 89 patients diagnosed at baseline (77%), 52 established patients (81%), and 64 healthy controls (91%) were available for follow-up. The mean test–retest interval was 37 months (interquartile range 36–38 months). Demographic and clinical characteristics are shown in Table 1.

Procedure

Clinical assessment

A detailed clinical examination was performed at baseline and annually thereafter over a 3-year period. Information about demographic variables, disease history, and drug treatment was obtained at each visit using a semistructured interview. All patients were examined by research nurses with experience in PD. Nurses were trained in administration and scoring of study instruments before starting this project and they performed assessments under supervision of the project neurologist to ensure reliability of ratings (Post et al., 2005). A clinical diagnosis of PD was revalidated by the project neurologist through review of the medical records. Only patients whose diagnosis was confirmed were included in the present analysis.

Extrapyramidal symptoms were rated in "on" phase using the motor section of the Unified Parkinson's Disease Rating Scale (UPDRS; Fahn et al., 1987). The stage of PD was determined with the Hoehn and Yahr Rating Scale (1967). Functional status was evaluated using the Schwab and England Activities of Daily Living (SE-ADL) Scale (Schwab & England, 1969), the Barthel Activities of Daily Living Index (Mahoney & Barthel, 1965), and the Functional Independence Measure (FIM; Keith et al., 1987). The FIM is divided into motor (13 items, possible score range 13–91) and cognitive (5 items, range 5–35) subscales. Higher scores on all scales indicate better functional capacity. Duration of disease was defined as the time between the appearance of the first symptom of PD as reported by the patient and the moment of assessment. Different drugs were pooled and expressed as a levodopa equivalent dose (LED; Esselink et al., 2004).[a] The Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983) was used as a screening instrument of affective symptoms.

Neuropsychological assessment

A comprehensive neuropsychological test battery was administered at baseline and 3-year follow-up. All subjects were examined by the same investigator (D.M.) on both assessments. The examination was suspended whenever a patient indicated that he or she went into "off" phase.

Six cognitive domains were evaluated: psychomotor speed, attention, language, memory, executive functions, and visuospatial skills. *Psychomotor speed* was examined using the Wechsler Adult Intelligence Scale-Revised (WAIS-R) Digit Symbol Test (Wechsler, 1981), Trail Making Test part A

[a] LED = regular levodopa dose × 1 + slow release levodopa × 0.75 + bromocriptine × 10 + apomorphine × 10 + ropinirol × 20 + pergolide × 100 + pramipexol × 100 + [regular levodopa dose + (slow-release levodopa × 0.75)] × 0.2 if taking entacapone.

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

Table 1. Demographic and clinical characteristics of NDPD patients, EPD patients, and HCS at baseline and 3-year follow-up

	Baseline				3-year follow-up			
Variable	NDPD (n = 89)	EPD (n = 52)	HCS (n = 64)	p Value	NDPD (n = 89)	EPD (n = 52)	HCS (n = 64)	p Value
Age, *M* (*SD*)	64.7 (9.7)	64.6 (6.9)	63.6 (7.5)	.73	67.7 (9.6)	67.6 (6.9)	66.7 (7.4)	.74
Gender (M/F)	47/42	28/24	34/30	.99				
Education (years), *M* (*SD*)	11.8 (2.5)	11.6 (1.9)	12.3 (2.5)	.21				
DART-IQ, *M* (*SD*)	104.0 (21.8)	106.4 (18.3)	106.3 (16.0)	.68				
MMSE, *M* (*SD*)	27.9 (1.6)	28.1 (1.7)	28.2 (1.5)	.36	27.7† (1.9)	28.4 (1.9)	28.1 (1.4)	.05
HADS, *M* (*SD*)	10.7* (7.5)	9.5* (5.4)	7.5 (4.3)	.04	10.8‡ (6.6)	11.1‡ (6.8)	6.8 (3.5)	<.001
Duration of PD (months), *M* (*SD*)	18.4 (11.1)	77.2 (23.3)		<.001	55.1 (11.5)	113.8 (23.6)		<.001
UPDRS motor section, *M* (*SD*)	16.5 (7.6)	25.2 (10.3)		<.001	24.6 (8.7)	27.0 (9.8)		.28
Hoehn and Yahr stage, *M* (*SD*)	1.7 (0.7)	2.3 (0.5)		<.001	2.5 (0.5)	2.5 (0.5)		.43
LED (mg/day), *M* (*SD*)	152.7 (143.5)	469.6 (294.0)		<.001	418.2 (275.3)	622.8 (362.0)		<.001
Dopamine agonist, *n* (%)	28 (32)	35 (67)		<.001	50 (56)	38 (73)		.03
Anticholinergics, *n* (%)	0 (0)	3 (6)		.05	1 (1)	4 (8)		.07
Antidepressants, *n* (%)	0 (0)	3 (6)		.05	6 (7)	3 (6)		.78
Benzodiazepines, *n* (%)	1 (1)	4 (8)		.07	8 (9)	3 (6)		.46
Antipsychotics, *n* (%)	0 (0)	1 (2)		.38	0 (0)	2 (4)		.14
Cholinesterase inhibitors, *n* (%)	—	—			2 (2)	0 (0)		.38
SE-ADL (%)	90.3 (5.9)	86.9 (7.6)		.004	84.7 (11.0)	81.9 (12.8)		.15
Barthel Activities of Daily Living Index, *M* (*SD*)	19.7 (1.1)	19.5 (0.9)		.08	19.2 (1.9)	19.1 (1.6)		.56
FIM, *M* (*SD*)	122.4 (4.7)	119.7 (5.9)		.002	115.8 (8.6)	114.1 (8.4)		.14
FIM-motor subscale, *M* (*SD*)	87.5 (4.7)	84.9 (5.5)		.001	82.6 (7.3)	80.3 (7.6)		.04
FIM-cognitive subscale, *M* (*SD*)	34.9 (0.3)	34.8 (0.6)		.43	33.2 (2.6)	33.8 (1.6)		.20

Note. DART-IQ, Dutch Adult Reading Test; HCS, healthy control subjects.
* $p < .05$ compared to controls.
† $p < .05$ compared to EPD group.
‡ $p < .001$ compared to controls.

(TMT-A; Reitan, 1992), and Stroop Color Word Test part A (word reading) and part B (color naming) (Stroop, 1935). *Attention* was assessed with the Trail Making Test part B (TMT-B; Reitan, 1992), Stroop Color Word Test part C (interference condition) (Stroop, 1935), and an adapted version of WAIS-R forward and backward digit span (Lindeboom & Matto, 1994). *Language* function was examined with the Boston Naming Test (BNT; Kaplan et al., 1983). *Memory* was assessed with the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964), Rivermead Behavioural Memory Test (RBMT) Logical Memory (LM) subtest (Wilson et al., 1985), Wechsler Memory Scale-III Faces Recognition subtest (Wechsler, 1997b), and Visual Association Test (Lindeboom et al., 2002). *Executive functions* were examined with the Modified Wisconsin Card Sorting Test (MWCST; Nelson, 1976), Category fluency (animals and supermarket items) (Luteijn & Barelds, 2004), Controlled Oral Word Association Test (COWAT; Benton & Hamsher, 1976), Tower of London-Drexel Test (TOL; Culbertson & Zilmer, 2001), and Wechsler Adult Intelligence Scale (Third Version) (WAIS-III) Similarities (Wechsler, 1997a). *Visuospatial and constructive abilities* were assessed with the Judgment of Line Orientation (JOLO; Benton et al., 1983), Clock Drawing Test (Royall et al., 1998), and the subtest Visuospatial Reasoning of the Groningen Intelligence Test (GIT; Luteijn & Barelds, 2004). To minimize practice effects, we used alternate forms at follow-up where available (RAVLT, RBMT, and COWAT). Due to time constraints and/or patients' fatigue or cooperation factors, not all tests were administered to all participants.

Statistical Analyses

Group differences in demographic and clinical characteristics at baseline and follow-up were analyzed with one-way analysis of variance, Kruskal–Wallis test, independent sample *t* test, Mann–Whitney test, and chi-square test where appropriate.

A standardized regression-based method (Frerichs & Tuokko, 2005; McSweeney et al., 1993) was used to establish the normative pattern and magnitude of cognitive change. This method takes into account sources of measurement error inherent in evaluating cognitive change over time, such as practice effects and regression to the mean, as well as demographic variables that may affect test performance, and initial level of performance. We conducted multiple regression analyses in the control group using baseline test scores, age, sex, and education as possible determinants to derive equations for predicting retest scores for each cognitive measure. The resulting equations were then used to obtain predicted retest scores for PD patients. The differences between the observed and the predicted retest scores were divided by the *SE* of the estimate to obtain standardized *z* scores. Negative *z* scores indicate lower than expected test performance. Paired

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

sample *t* tests were used to compare observed and predicted follow-up test scores in the two PD groups.

Additionally, standard scores were derived from available normative data for each neuropsychological measure to assess the frequency of cognitive deficits at baseline and follow-up in a more clinically meaningful manner. A test score was categorized as impaired if it was at least 2 *SD*s below the mean of the normative sample after correction for age and (if possible) for level of education. Normative data provided in test manuals were used. For some tests, normative data were taken from other sources: BNT—Marien et al. (1998); TMT-A, TMT-B, Stroop Test, RBMT, and COWAT—Schmand et al. (2003); and MWCST—Lineweaver et al. (1999). For TOL and Clock Drawing Test, standardized residual scores were used as described above. To adjust for the Flynn effect (Flynn, 1987) due to the use of the older version of Digit Symbol Test, we corrected normative data for this test by subtracting 1 *SD* from the observed scores.

A multivariate normative comparison method (Huizenga et al., 2007) was employed to identify *individual* patients exhibiting a significant decline in cognitive performance compared with healthy controls. The pattern of change in test performance of each PD patient over the follow-up interval was compared with the control group to assess whether the patient deviates from the norm on several cognitive variables simultaneously while controlling for false-positive rate due to multiple comparisons. For this analysis, regression-based *z* scores were used as the dependent variables. We hypothesized that both NDPD and EPD patients would exhibit greater cognitive change than controls ($p < .05$, one-sided).

Patients with cognitive decline as determined with the multivariate normative comparison method were contrasted with cognitively stable patients to identify baseline demographic and clinical characteristics associated with cognitive deterioration in PD. To examine the contribution of specific motor signs of PD to cognitive decline, the UPDRS motor section was subdivided into six domains (Levy et al., 2000). Facial expression was quantified with item 19. Tremor score was calculated by adding items 20 (tremor at rest) and 21 (action or postural tremor of hands). Rigidity was measured with item 22. Bradykinesia was evaluated with combination of items 24 (hand movements), 25 (rapid alternating movements of hands), 26 (leg agility), and 31 (body bradykinesia and hypokinesia). Speech was rated with item 18. Last, a score for axial impairment was obtained by adding items 27–30 (rising from a chair, posture, gait, and postural stability). Subsequently, all variables that were univariately associated with cognitive decline at $p \leq .15$ were entered into a hierarchical multivariate logistic regression analysis, with cognitive status at follow-up as the dependent variable. Per block step, the improvement of the regression model was compared using the likelihood ratio test. The goodness of fit of the model was assessed with the Hosmer–Lemeshow Test. Given the explorative nature of this analysis, we did not statistically correct for multiple comparisons (Rothman, 1990).

Furthermore, we sought to establish the proportion of patients developing dementia at follow-up. We applied clinical diagnostic criteria for dementia recently proposed by the Movement Disorders Society Task Force, which require the presence of cognitive deficits in at least two domains and functional disability resulting from cognitive decline (Emre et al., 2007). To this end, composite scores were computed for each cognitive domain by averaging regression-based *z* scores from the individual tests that contributed to the particular domain and subjecting these composite scores to a multivariate normative comparison method. Cognitive decline was defined here as decrement in two areas of cognitive function other than psychomotor speed relative to the control group. Functional impairment was operationally defined as a score of 4 or less (corresponding to minimal assistance in performing daily tasks) on at least one item from the FIM-cognitive subscale related to the following activities: language comprehension, language expression, social interaction, problem solving, and memory.

At follow-up, 6% of the test scores were missing in the NDPD and EPD groups and 1% in the control group. All analyses were repeated after imputation of these missing data. A missing value was replaced by the worst value observed within the particular patient group when a patient had attempted to perform a test but was unable to do it. In other cases, the missing value was replaced by the mean group score at follow-up. *p* Values .05 or less were considered significant, unless otherwise indicated.

RESULTS

NDPD patients without follow-up were older, had later disease onset, and performed slower on Stroop Test word reading condition at baseline than those who returned to the second evaluation (all *p* values < .05). Regarding the EPD group, patients without follow-up were older, had later onset of disease, had greater degree of axial impairment, received higher LED, and performed worse on TMT-B at baseline than those who participated in the second evaluation (all *p* values < .05).

Test–Retest Reliability and Cognitive Change in Control Group

The test–retest reliability coefficients in the control group for the 27 measures examined ranged from .34 to .89, with a median of .68. For most of the tests, the levels of stability over time were adequate given the long follow-up interval. The proportion of variance explained by baseline test scores and demographic variables in predicting expected retest scores ranged between .15 and .80. Using the standard cutoff criterion of $z = \pm 1.645$ [corresponding to a score outside the 90% confidence interval (CI), $\alpha = .10$, two-tailed], all tests were found to fit the expected distribution (i.e., 5% classified as improved and 5% as deteriorated), except BNT and MWCST number of perseverative errors, which were associated with an increased risk of misclassification (χ^2 goodness of fit, $p < .001$).

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

Cognitive Change in NDPD Group

Magnitude of cognitive change

Table 2 summarizes baseline and follow-up scores, predicted retest scores, and standardized *z* scores for NDPD patients. For 20 of the 27 measures, observed retest scores were significantly lower than predicted scores, suggesting a greater decrease in performance over time than in healthy controls. Tests that tap short-term and working memory (digit span forward and backward), planning (TOL), and spatial reasoning (GIT visuospatial task) showed no decline. Follow-up performance on BNT and MWCST perseverative errors exceeded predicted retest scores, but these findings are questionable given the poor classification accuracy of these measures in the control group.

As indicated by *z* scores, tests of psychomotor speed and attention (TMT-B) showed the greatest degree of decline. Within the memory domain, decline in performance was more prominent on measures of verbal recognition memory, immediate and delayed story recall, than on word list learning or visual memory. Tests of executive function showed small but significant rates of decline. Observed retest scores for the WAIS-III Similarities test were better than predicted. Within the visuospatial domain, tests of spatial orientation (JOLO) and constructive skills (Clock Drawing) showed modest declines.

Frequency of cognitive deficits at baseline and 3-year follow-up

Comparison with normative data revealed that impairments in the NDPD group at baseline were most frequent on Digit Symbol Test (12%), TOL (12%), TMT-B (11%), Clock Drawing Test (10%), and RBMT LM (10%). After 3 years, these deficits became even more prevalent and there was an

Table 2. Observed baseline and follow-up test scores, test–retest reliability (Pearson *r*), predicted follow-up scores, standardized *z* scores, and *p* values for observed *versus* predicted follow-up scores (paired *t* test) in NDPD patients ($n = 89$)

	Baseline score		Follow-up score			Predicted follow-up score			
Test	*M*	*SD*	*M*	*SD*	*r*	*M*	*SD*	*z* Score	*p* Value
Psychomotor speed									
WAIS-R Digit Symbol Test	39.2	11.3	33.6	14.3	.85	39.0	11.3	−1.01	<.001
TMT-A (s)	46.1	16.2	58.2	43.7	.61	36.0	12.7	−2.07	<.001
Stroop Test (word reading) (s)	48.5	10.4	53.1	13.2	.73	45.2	5.9	−1.67	<.001
Stroop Test (color naming) (s)	63.9	14.0	68.5	17.3	.78	61.3	8.9	−1.06	<.001
Attention									
TMT-B (s)	123.9	74.4	177.1	161.6	.73	142.5	85.5	−1.10	.006
Stroop Interference Test (s)	122.5	48.5	132.6	68.0	.85	113.0	21.3	−0.84	.001
Digit span forward	12.0	2.8	12.1	3.0	.71	12.0	1.6	0.03	.742
Digit span backward	8.2	2.5	8.4	3.0	.63	8.2	1.7	0.11	.414
Language									
BNT	53.6	4.3	52.5	4.9	.85	48.1	3.9	2.12	<.001
Memory									
RAVLT trials 1–5	39.7	9.8	37.4	11.7	.79	39.8	8.5	−0.40	.001
RAVLT delayed recall	7.9	2.9	7.2	3.3	.69	7.7	2.4	−0.33	.040
RAVLT recognition	28.0	1.8	27.2	2.8	.38	28.4	1.1	−1.00	<.001
RBMT LM immediate recall	14.1	5.6	13.3	6.6	.77	16.5	5.0	−0.82	<.001
RBMT LM delayed recall	10.5	5.4	10.0	6.7	.76	13.2	4.4	−0.74	<.001
WMS-III Faces immediate	32.2	4.1	33.9	4.4	.66	35.3	2.8	−0.45	.001
WMS-III Faces delayed	36.0	3.7	36.3	4.6	.56	37.6	2.6	−0.45	.005
Executive functions									
MWCST categories	4.1	1.7	4.3	1.7	.50	4.7	0.7	−0.38	.005
MWCST perseverations	5.4	4.8	5.2	4.6	.59	7.3	1.7	0.95	<.001
Animal fluency	19.7	5.6	20.2	6.1	.57	21.8	3.3	−0.45	.004
Supermarket fluency	20.3	5.2	20.8	5.7	.59	23.0	1.7	−0.58	<.001
COWAT Letter fluency	30.6	10.6	28.7	11.9	.77	32.4	9.6	−0.52	<.001
Tower of London	6.0	1.9	5.6	2.3	.27	5.9	1.1	−0.18	.252
WAIS-III Similarities	21.9	6.0	21.2	6.0	.82	19.8	5.5	0.35	.003
Visuospatial/constructive skills									
JOLO	23.6	4.5	23.2	4.6	.73	24.9	3.0	−0.79	<.001
GIT visuospatial task	10.4	3.3	9.9	3.7	.74	9.5	3.0	0.16	.248
Clock Drawing Test	12.6	1.3	12.1	1.8	.42	12.8	0.5	−0.72	<.001

Notes. WMS-III, Wechsler Memory Scale-III. *z* Score = [observed follow-up score − predicted follow-up score]/*SE* of the estimate from regression analysis. Negative *z* scores indicate lower than expected retest score. No corrections were made for multiple comparisons.

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

additional increase in the frequency of impairments on measures of psychomotor speed, selective attentional (Stroop interference) and executive skills (COWAT), and memory (RAVLT delayed recall). Whereas impairments on BNT, WAIS-III Similarities, and supermarket fluency were fairly frequent at baseline, they remained unchanged or slightly decreased at follow-up assessment. The frequency of impairment on the remaining tasks remained stable or changed only slightly. In contrast to NDPD patients, the prevalence of cognitive deficits in the control group was low at baseline, and it remained fairly stable during the follow-up interval (Table 3).

Cognitive change at the individual level

Multivariate comparisons revealed that 43 out of 89 NDPD patients (48%) exhibited decline in cognitive performance during follow-up to a degree that exceeded the 5th percentile in controls. Patients exhibiting cognitive decline were compared with cognitively stable patients with respect to baseline demographic and clinical variables (Table 4). Differences in the UPDRS motor section and the Hoehn and Yahr Scale between patients with cognitive decline and cognitively stable patients in the univariate analyses were probably due to axial impairment because other motor symptoms at baseline were comparable across the two patient groups. Therefore, the UPDRS and Hoehn and Yahr scores were not included in the multivariate regression analysis. Furthermore, age was not included in the analysis because of its strong correlation with age at PD onset. For hierarchical regression analysis, age at onset and disease duration were entered as independent variables in the first step. In the second step, axial impairment at baseline was added to the model. The results showed that adding age at onset and disease duration as predictors in the first step of the analysis significantly improved the model (likelihood ratio test: $p = .021$). The regression model appeared to adequately fit the data (Hosmer–Lemeshow Test: $\chi^2 = 11.13$; $p = .19$). Disease duration was an independent predictor of cognitive change at follow-up [odds ratio (OR) = 1.96, 95% CI: 1.02–3.74; $p = .043$]. Adding axial impairment

Table 3. Frequency of cognitive impairments at baseline and 3-year follow-up in NDPD patients, EPD patients, and HCS

	% Impairments NDPD		% Impairments EPD		% Impairments HCS	
Test	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up
Psychomotor speed						
WAIS-R Digit Symbol Test	12	20	10	19	2	0
TMT-A	6	13.5	11.5	15	0	2
Stroop Test (word reading)	3	10	8	8	3	0
Stroop Test (color naming)	2	8	4	4	2	0
Attention						
TMT-B	11	21	8	19	0	2
Stroop Interference Test	4.5	9	6	6	3	0
Language						
BNT	11	11	6	13.5	6	3
Memory						
RAVLT trials 1–5	1	3	0	2	0	0
RAVLT delayed recall	1	8	2	6	0	0
RBMT LM immediate recall	10	16	6	10	2	2
RBMT LM delayed recall	10	21	19	13.5	2	0
WMS-III Faces immediate	1	1	2	0	0	0
WMS-III Faces delayed	0	0	0	0	0	0
Visual Association Test	0	3	2	2	0	0
Executive functions						
MWCST categories	6	7	4	8	0	2
MWCST perseverations	1	0	4	4	2	0
Animal fluency	9	11	4	13.5	2	0
Supermarket fluency	11	6	10	6	2	0
COWAT Letter fluency	2	12	0	6	3	2
Tower of London Test	12	17	10	10	6	3
WAIS-III Similarities	11	11	2	2	2	3
Visuospatial/constructive skills						
JOLO	2	2	0	0	2	0
GIT visuospatial task	9	10	2	6	3	5
Clock Drawing Test	11	18	17	27	5	5

Notes. WMS-III, Wechsler Memory Scale-III; HCS, healthy control subjects. Impairment was defined as a score below 2 *SDs* of age and (if possible education) norms. No corrections were made for multiple comparisons.

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

Table 4. Baseline demographic and clinical characteristics of NDPD patients with stable cognitive performance and cognitive decline after 3 years

	Stable (n = 46)		Decline (n = 43)		
Variable	*M*	*SD*	*M*	*SD*	*p* Value
Age	62.9	9.4	66.5	9.8	.082
Gender (M/F)	21/25		26/17		.162
Education (years)	11.9	2.5	11.7	2.5	.781
DART-IQ	103.1	21.0	104.8	22.8	.718
Age at onset of PD	61.6	9.6	64.8	9.6	.122
Duration of PD (months)	16.1	7.5	20.8	13.7	.053
UPDRS motor section	15.8	8.1	17.3	6.9	.146
Tremor	2.2	2.0	2.5	2.0	.453
Bradykinesia	5.9	3.3	6.4	2.8	.355
Rigidity	3.2	2.3	2.8	2.1	.362
Axial symptoms	1.1	1.6	2.1	2.0	.009
Speech	0.4	0.5	0.5	0.5	.306
Facial expression	1.1	0.6	1.1	0.7	.963
Hoehn and Yahr Scale	1.6	0.6	1.9	0.7	.058
MMSE	28.2	1.4	27.5	1.7	.052
HADS	9.9	6.9	11.6	8.1	.342
HADS-D	4.3	3.3	5.4	4.1	.199
HADS-A	5.5	4.2	6.2	4.6	.542
LED (mg/day)	135.8	146.4	170.7	139.7	.253
SE-ADL (%)	90.6	5.3	90.0	6.5	.718
Barthel Activities of Daily Living Index	19.8	0.4	19.5	1.5	.431
FIM	123.5	3.4	121.2	5.6	.038
FIM-motor subscale	88.6	3.3	86.3	5.6	.034
FIM-cognitive subscale	34.9	0.3	34.9	0.3	.579
3-year follow-up					
SE-ADL (%)	87.8	9.9	81.4	11.2	<.001
Barthel Activities of Daily Living Index	19.4	2.1	18.9	1.7	.042
FIM	118.6	7.6	113.0	8.8	.001
FIM-motor subscale	84.5	7.2	80.6	6.9	.002
FIM-cognitive subscale	34.0	1.2	32.4	3.3	.002

Notes. DART, Dutch Adult Reading Test; HADS-D, Depression subscale; HADS-A, Anxiety subscale. No corrections were made for multiple comparisons.

at baseline as a predictor did not improve the regression model. In the final step, with all candidate predictors included in the equation, none of the independent variables turned out to be significantly associated with cognitive decline. Multicollinearity between predictors was not present ($.11 < r < .39$).

Dementia at follow-up

Multivariate normative comparison using composite domain scores revealed that 16 NDPD patients (18%) showed decline in two or more cognitive areas. Four of them (4.5% of NDPD patients) exhibited significant decline in at least two cognitive domains other than psychomotor speed and had evidence of functional disability at follow-up, defined as a score of ≤4 on at least one item on the FIM-cognitive subscale. Review of medical records revealed that out of the 122 patients from the original sample who participated in annual clinical examination over a 3-year period, an additional 5 patients (4%) were diagnosed as demented by their consultant neurologist between the last two study visits. Thus, in all, 8.5% of NDPD patients developed dementia during follow-up.

Cognitive Change in EPD Group

Comparison with either control subjects (Table 5) or normative data (Table 3) revealed deteriorating performance in the EPD group over the follow-up interval across several cognitive domains, particularly on measures of attention, psychomotor speed, and constructive skills. Analysis of cognitive change at the individual level showed that 26 out of the 52 patients (50%) displayed cognitive decline exceeding 5th percentile in controls. Following the same procedure as in the NDPD group, univariate analyses were conducted to select candidate predictors of cognitive decline. Patients with EPD who showed decline in cognitive function over the 3-year follow-up interval were older and more often male and had later onset of disease, lower scores on a global measure of cognitive status, greater

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

Table 5. Observed baseline and follow-up test scores, test–retest reliability (Pearson *r*), predicted follow-up scores, standardized *z* scores, and *p* values for observed *versus* predicted follow-up scores (paired *t* test) in EPD patients ($n = 52$)

	Baseline score		Follow-up score			Predicted follow-up score			
Test	*M*	*SD*	*M*	*SD*	*r*	*M*	*SD*	*z* Score	*p* Value
Psychomotor speed									
WAIS-R Digit Symbol Test	41.1	12.4	36.3	14.2	.87	40.9	12.4	−0.85	<.001
TMT-A (s)	50.1	21.7	53.9	27.4	.72	39.2	16.9	−1.38	<.001
Stroop Test (word reading) (s)	46.4	8.8	51.2	11.5	.75	44.0	5.0	−1.53	<.001
Stroop Test (color naming) (s)	61.9	14.0	66.4	17.5	.82	60.0	8.9	−0.94	<.001
Attention									
TMT-B (s)	122.5	77.7	186.8	189.1	.72	140.9	89.4	−1.46	.021
Stroop Interference Test (s)	116.6	55.1	123.3	62.6	.68	111.1	23.3	−0.52	.091
Digit span forward	13.6	2.8	13.0	2.6	.70	12.9	1.7	0.04	.719
Digit span backward	9.1	2.7	9.2	2.7	.58	8.7	1.6	0.26	.093
Language									
BNT	54.5	4.0	53.4	4.4	.87	48.9	3.6	2.21	<.001
Memory									
RAVLT trials 1–5	42.0	11.1	39.5	10.6	.78	41.4	9.1	−0.31	.037
RAVLT delayed recall	8.9	3.0	7.7	3.1	.66	8.4	2.7	−0.47	.026
RAVLT recognition	27.9	2.0	27.3	2.8	.47	28.3	1.1	−0.91	.003
RBMT LM immediate recall	15.1	5.2	14.8	5.0	.68	17.2	4.2	−0.61	<.001
RBMT LM delayed recall	11.0	5.0	11.3	5.1	.72	13.5	3.6	−0.50	<.001
WMS-III Faces immediate	33.2	4.2	34.6	5.1	.58	36.0	2.7	−0.46	.029
WMS-III Faces delayed	37.0	4.1	36.6	4.2	.55	38.2	2.6	−0.52	.010
Executive functions									
MWCST categories	4.4	1.6	3.8	1.8	.60	4.9	0.6	−0.82	<.001
MWCST perseverations	5.5	5.9	6.9	6.6	.50	7.4	1.7	0.20	.594
Animal fluency	20.6	5.8	21.3	7.1	.63	22.2	3.2	−0.27	.217
Supermarket fluency	19.8	5.8	21.6	6.4	.72	23.0	1.2	−0.37	.083
COWAT Letter fluency	33.7	8.7	31.9	12.3	.72	34.8	7.8	−0.42	.013
Tower of London	6.1	1.8	5.7	2.0	.35	5.8	1.1	−0.12	.492
WAIS-III Similarities	22.8	5.3	22.4	5.6	.80	20.6	4.8	0.42	.003
Visuospatial/constructive skills									
JOLO	23.5	4.3	23.9	3.9	.70	24.9	2.9	−0.43	.016
GIT visuospatial task	9.7	3.3	9.2	3.1	.74	8.8	3.0	0.18	.250
Clock Drawing Test	12.4	2.1	11.8	2.3	.64	12.7	0.8	−1.00	.001

Notes. WMS-III, Wechsler Memory Scale-III. *z* Score = [observed follow-up score − predicted follow-up score]/*SE* of the estimate from regression analysis. Negative *z* scores indicate lower than expected retest score. No corrections were made for multiple comparisons.

overall severity of motor disorder, and greater degree of axial impairment and bradykinesia at baseline assessment than those with stable cognitive performance. There were no differences between the groups with respect to education, premorbid intelligence, duration of PD, the prevalence of affective symptoms, severity of tremor, rigidity, speech impediments and facial mobility, or LED. Subsequently, the hierarchical logistic regression was performed, with age at onset and sex entered first in the model, followed by baseline axial symptoms and bradykinesia. Likelihood ratio tests showed that the model significantly improved at each step of the analysis (step 1: $p = .001$; step 2: $p = .003$). At the final step, with all independent variables entered in the equation, age at disease onset (OR = 1.20, 95% CI: 1.02–1.40; $p = .025$) and axial symptoms at baseline (OR = 1.79, 95% CI: 1.05–3.05; $p = .032$) were found to be significant predictors of cognitive decline in the EPD group. The regression model adequately fitted the data (Hosmer–Lemeshow Test: $\chi^2 = 8.52$; $p = .38$).

Two patients (4%) showed decline in two or more cognitive areas when using composite domain scores along with evidence of functional disability. Out of the 67 patients who participated in the 3-year clinical assessment, two additional patients (3%) were diagnosed as demented by their consultant neurologist. Thus, in all, 7% of EPD patients developed dementia during follow-up.

DISCUSSION

This study examined the evolution and predictors of cognitive impairment over 3 years in patients with PD. In an attempt to overcome methodological limitations of previous research in this area, our approach included a prospective follow-up of a well-defined inception cohort of patients from

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

the time of diagnosis in addition to established cases, comprehensive neuropsychological assessment, demographically matched control group, and analytic methods that control for sources of error in test–retest designs and allow evaluation of change in individual patients.

We found a significant decrease in performance on most cognitive measures in NDPD patients (Table 2). The most severe decline occurred on measures of psychomotor speed and attention. In the domains of memory, visuospatial skills, and executive function, deterioration was smaller in magnitude but still significantly greater than in controls. Comparison with normative data revealed already at baseline a higher prevalence of impairments across several cognitive measures in patients than in controls. Three years later, the frequency of deficits in PD patients further increased on tests of psychomotor speed, attention, constructive skills, verbal memory, and selective executive abilities (Table 3). Analysis of cognitive change at an individual level revealed that nearly half of the patients exhibited cognitive decline. This decline appeared to occur in a diffuse manner, but deterioration in psychomotor speed and attention seemed to be the most prominent.

An important methodological issue concerns the extent to which cognitive changes observed in our patient sample are affected by motor impairment. In this respect, it is important to note that the measures that showed the greatest change involve a timed component and motor skills. However, correlation analyses using regression-based z scores and UPDRS change scores between baseline and follow-up assessment as variables showed that within the NDPD group, an increase in motor impairment was significantly related only with the rate of decline on the Digit Symbol Test ($r = -.23$) (data not shown). Among patients with EPD, an increase in motor impairment was found to be significantly related only with changes in performance on RBMT LM immediate recall ($r = -.30$) and GIT visuospatial task ($r = -.31$), which are both nonmotor tasks (data not shown). Thus, it is seems unlikely that the cognitive changes observed in our patient sample can be accounted for in terms of motor impairment alone.

The present finding of a diffuse decline across multiple cognitive abilities is consistent with some previous (uncontrolled) studies that prospectively assessed PD patients over longer time intervals with an extensive neuropsychological test battery (Caparros-Lefebvre et al., 1995; Palazzini et al., 1995). In a recent meta-analysis of 25 longitudinal studies, we found small but significant declines in global cognitive ability, memory, and visuoconstructive skills among initially nondemented PD patients during a mean follow-up interval of 2.5 years (Muslimovic et al., 2007). Although methodological differences preclude direct comparisons, the present results confirm and extend these previous observations by demonstrating decline in performance on measures of memory, constructive skills, and executive function 3 years after diagnosis. The present results concerning domains of psychomotor speed and attention differ from our review of the literature, which indicated a rather stable pattern of performance over time in these two domains. This discrepancy may in part be due to different methods of assessment. Of the seven studies included in the meta-analysis that assessed psychomotor speed and attention, only two used the Digit Symbol Test and three employed TMT. On the other hand, the most frequently used measure was Digit Span Test. Notably, digit span forward and backward were the only two attentional measures that did not reveal change in our current patient sample. In view of our findings that the Digit Symbol Test, TMT, and Stroop Test are highly sensitive to longitudinal change, it is possible that our meta-analysis would have produced different results if more studies had employed these tests. Another important finding is that traditional measures of executive function, such as the Wisconsin Card Sorting Test, word fluency, and Tower of London, were found to exhibit a relatively smaller degree of decline in our NDPD patients than some tests of verbal memory and visuospatial abilities. The comparatively lower level of change on tests of executive function is in contrast to the observations from cross-sectional comparisons of patients at different stages of disease (Owen et al., 1992). Thus, the present longitudinal findings, in keeping with our recent review of the literature, suggest that executive function may not be the earliest or primary cognitive domain to show decline in PD.

We found that 9% of patients developed dementia 3 years after diagnosis. However, there are several issues that need to be considered when interpreting this finding. First, diagnostic criteria for dementia applied in this study (i.e., Movement Disorders Society Task Force) (Emre et al., 2007) require functional impairment resulting from cognitive dysfunction. We attempted to minimize the influence of motor impairment by investigating the functional status with the FIM-cognitive subscale, which contains questions about activities that are unlikely to be affected by motor symptoms (e.g., comprehension, social interaction, problem solving, and forgetfulness). This information was acquired by asking patients to indicate the extent to which they perform these activities independently. However, this method of assessment relying on self-report may have led to underestimation of the frequency of dementia due to possibly reduced insight into daily functioning, considering the fact that almost a half of our patients showed evidence of cognitive decline. In this respect, it is important to note that even when cognitive decline rather than dementia was used as outcome measure, it still appeared to be functionally relevant. This is reflected by our finding that patients with adverse cognitive outcome attained significantly lower scores on each of the functional scales relative to cognitively stable patients at follow-up (Table 4). The relation between the cognitive decline and the FIM-cognitive subscale remained significant even after the effects of motor impairment had been taken into account.

Another issue that might have influenced our findings is the presence of selection bias due to the loss to follow-up. Of the nine patients who developed dementia within 3 years after diagnosis, four participated in the follow-up, whereas the remaining five patients were diagnosed as demented by their neurologist. Given the high sensitivity of neuropsychological tests to detect cognitive changes, it is likely that

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

patients with milder forms of dementia were missed in the group that was lost to follow-up. Thus, the frequency of dementia probably would have been somewhat higher if these patients had been included, because they were older at baseline, and were therefore at greater risk of developing dementia (Emre et al., 2007).

Despite these potentially confounding factors, the proportion of patients with dementia in this study is similar to that in a recent report, in which 10% of 126 PD patients from a community sample developed dementia 3.5 years after diagnosis (Williams-Gray et al., 2007). However, a direct comparison of these findings is difficult because in this earlier study, the presence of cognitive impairment was inferred from the MMSE and functional disability was quantified with the SE-ADL Scale. Nonetheless, this highlights the need for a uniform operational definition of dementia in PD in future studies if progress is to be made in our understanding of this important concomitant of the disease.

The trajectory of decline in patients with more advanced PD appeared to be similar to that observed among our newly diagnosed patients. However, there are two important methodological problems that are likely to have affected this finding. First, our established sample appears to have been biased already at inclusion in the study toward relatively better functioning patients with slower disease progression. Patients who agreed to participate at baseline were younger than those who refused to take part in the study. This selection bias is a likely explanation for a relatively stable course of motor impairment in our established sample, which is in contrast to progression of motor symptoms seen among newly diagnosed patients (see UPDRS motor section and Hoehn and Yahr Scale in Table 1). Second, selection bias during the follow-up interval might have influenced our findings in this sample. Patients who were lost to neuropsychological follow-up (n = 12) were older and had later disease onset, more severe axial symptoms, and more deficits on TMT-B at baseline than those who were retested. In view of our finding that older age at disease onset and greater severity of axial symptoms were significantly associated with cognitive decline, it is likely that the current results underestimate the true extent of cognitive deterioration and dementia. It should be noted that chronological age was not included in the regression model due to its strong correlation with age at onset of PD. Therefore, chronological age may also be a potential predictor of cognitive decline in this group. Given the selection bias in our established sample, firm conclusions about the effect of disease duration longer than 3 years on cognitive course in PD cannot be drawn.

A final finding of this study is that none of the baseline demographic and clinical features appeared to be associated with cognitive change in the first few years following the PD diagnosis. In the EPD group, age at disease onset and axial impairment at baseline are identified as the main predictors of cognitive decline at follow-up. This latter finding is consistent with previous longitudinal studies (Aarsland et al., 2004; Burn et al., 2006; Williams-Gray et al., 2007) and adds further support to the notion that motor symptoms assumed to result from pathological changes in nondopaminergic brain systems (i.e., postural instability and gait difficulty) are likely to be of prognostic significance for cognitive course.

Some methodological limitations of this study should be recognized. First, 23% of newly diagnosed patients and 19% of established patients were lost during the follow-up interval, and this attrition rate was higher than that among control subjects (9%). Furthermore, some patients (6% in each group) had deceased. Because mortality is associated with dementia in PD (Levy et al., 2002), the attrition due to death in our samples may not be random, and this could have influenced our results. Second, the majority of our patients formed a clinic-based sample, and the degree to which the present findings can be generalized to the PD population at large remains to be determined. However, our inception cohort was composed by inviting every newly diagnosed patient referred to outpatient clinics, and therefore, they are likely to represent the kind of patients most specialists deal with in their practices. Third, a follow-up interval of 3 years is modest in duration, recognizing that PD tends to progress rather slowly over several years. Fourth, the available norms for several tests used in this study only permit a correction for age but not for level of education. This might have distorted somewhat the frequency of cognitive deficits in our sample. Finally, the classification of tests into cognitive domains is arbitrary to some degree. For example, many tests (Stroop Test, TMT-A, and TMT-B) that purport to assess different aspects of attention are also used as measures of executive function.

In conclusion, the present longitudinal evaluation of neuropsychological performance indicates a progressive course of cognitive impairment in PD beyond cognitive changes associated with normal aging. Cognitive decline in PD is evident across several cognitive domains but appears to be particularly prominent for psychomotor speed and attention. Axial impairment (i.e., postural instability and gait difficulty) and age at disease onset are major predictors of cognitive decline in later stages of PD. This information about the cognitive prognosis in PD may be relevant for patients and their families and could help to identify patients requiring closer attention and monitoring. In addition, these results may be useful for future trials evaluating disease-modifying drugs. Overall, our study highlights the advantage of prospectively investigating patients from an early point in the course of disease and illustrates the methodological difficulties in studying established cases.

ACKNOWLEDGMENTS

This work was funded by the Prinses Beatrix Fonds (Grant PGO 01-0138 to B.S.). B.P., J.D.S., and R.J. de H. are members of the Co-morbidity and Ageing in Rehabilitation Patients; the influence on Activities (CARPA) study group. The CARPA study group is supported by a grant from ZonMw, The Hague, The Netherlands. The information in this manuscript and the manuscript itself have never been published either electronically or in print. There are no financial or other relationships that may create a conflict of interest. We thank the patients for their willingness to participate in this research. We are grateful to the research nurses D. Standaar,

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

M. Postma, L. Daeter, and A. Gorissen. We also thank the staff from neurology departments of Medical Centre Alkmaar, Meander Medical Centre, Amersfoort, Slotervaart Hospital, Lucas Andreas hospital, Onze Lieve Vrouw Gasthuis, and Academic Medical Centre in Amsterdam for patient recruitment and logistic support. We are grateful to the Dutch Parkinson's Disease Association.

REFERENCES

Aarsland, D., Andersen, K., Larsen, J.P., Lolk, A., & Kragh-Sørensen, P. (2003). Prevalence and characteristics of dementia in Parkinson disease: An 8-year prospective study. *Archives of Neurology*, *60*, 387–392.

Aarsland, D., Andersen, K., Larsen, J.P., Perry, R., Wentzel-Larsen, T., Lolk, A., & Kragh-Sørensen, P. (2004). The rate of cognitive decline in Parkinson disease. *Archives of Neurology*, *61*, 1906–1911.

Benton, A.L. & Hamsher, K.D.S. (1976). *Multilingual Aphasia Examination*. Iowa City, IA: University of Iowa.

Benton, A.L., Hamsher, K., Varney, N., & Spreen, O. (1983). *Contributions to neuropsychological assessment—A clinical manual*. New York: Oxford University Press.

Burn, D.J., Rowan, E.N., Allan, L.M., Molloy, S., O'Brien, J.T., & McKeith, I.G. (2006). Motor subtype and cognitive decline in Parkinson's disease, Parkinson's disease with dementia, and dementia with Lewy bodies. *Journal of Neurology, Neurosurgery, and Psychiatry*, *77*, 585–589.

Caparros-Lefebvre, D., Pecheux, N., Petit, V., Duhamel, A., & Petit, H. (1995). Which factors predict cognitive decline in Parkinson's disease? *Journal of Neurology, Neurosurgery, and Psychiatry*, *58*, 51–55.

Culbertson, W.C. & Zilmer, E. (2001). *The Tower of London*^{Dx} *(TOL*^{Dx}*): Manual*. North Tonawanda, NY: Multi-Health Systems.

Emre, M., Aarsland, D., Brown, R., Burn, D.J., Duyckaerts, C., Mizuno, Y., Broe, A.G., Cummings, J., Dickson, D.W., Gauthier, S., Goldman, J., Goetz, C., Korczyn, A., Lees, A., Levy, R., Litvan, I., McKeith, I., Olanow, W., Poewe, W., Quinn, N., Sampaio, C., Tolosa, E., & Dubois, B. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Movement Disorders*, *22*, 1689–1707.

Esselink, R.A., de Bie, R.M., de Haan, R.J., Lenders, M.W., Nijssen, P.C., Staal, M.J., Smeding, H.M., Schuurman, P.R., Bosch, D.A., & Speelman, J.D. (2004). Unilateral pallidotomy versus bilateral subthalamic nucleus stimulation in PD: A randomized trial. *Neurology*, *62*, 201–207.

Fahn, S., Elton, R.I., & members of the UPDRS Development Committee. (1987). In S. Fahn, C.D. Marsden, D.B. Calne & M. Goldstein (Eds.), *Recent development in Parkinson's disease*. New York: Macmillan Health Care Information, 153–164.

Flynn, J.R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure? *Psychological Bulletin*, *101*, 171–191.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini Mental-State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Frerichs, R.J. & Tuokko, H.A. (2005). A comparison of methods for measuring cognitive change in older adults. *Archives of Clinical Neuropsychology*, *20*, 321–333.

Gelb, D.J., Oliver, E., & Gilman, S. (1999). Diagnostic criteria for Parkinson disease. *Archives of Neurology*, *56*, 33–39.

Hely, M.A., Read, W.G., Adena, M.A., Halliday, G.M., & Morris, J.G. (2008). The Sydney multicenter study of Parkinson's disease: The inevitability of dementia at 20 years. *Movement Disorders*, *23*, 837–844.

Hoehn, M.M. & Yahr, M.D. (1967). Parkinsonism: Onset, progression and mortality. *Neurology*, *17*, 427–442.

Huizenga, H.M., Smeding, H.M., Grasman, R.P., & Schmand, B. (2007). Multivariate normative comparisons. *Neuropsychologia*, *45*, 2534–2542.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia, PA: Lea & Febiger.

Keith, R.A., Granger, C.V., Hamilton, B.B., & Sherwin, F.S. (1987). The Functional Independence Measure: A new tool for rehabilitation. In M.G. Eisenberg & R.G. Grzesiak (Eds.), *Advances in clinical rehabilitation*. New York: Springer.

Levy, G., Tang, M.X., Louis, E.D., Cote, L.J., Alfaro, B., Mejia, H., Stern, Y., & Marder, K. (2000). Motor impairment in PD: Relationship to incident dementia and age. *Neurology*, *55*, 539–544.

Levy, G., Tang, M.X., Louis, E.D., Cote, L.J., Alfaro, B., Mejia, H., Stern, Y., & Marder, K. (2002). The association of incident dementia with mortality in PD. *Neurology*, *59*, 1708–1713.

Lindeboom, J. & Matto, D. (1994). Digit series and Knox cubes as concentration tests for elderly subjects [in Dutch]. *Tijdschrift Gerontologie en Geriatrie*, *25*, 63–66.

Lindeboom, J., Schmand, B., Tulner, L., Walstra, G., & Jonker, C. (2002). Visual Association Test to detect early dementia of the Alzheimer type. *Journal of Neurology, Neurosurgery, and Psychiatry*, *73*, 126–133.

Lineweaver, T.T., Bond, M.W., Thomas, R.G., & Salmon, D.P. (1999). A normative study of Nelson's (1976) modified version of the Wisconsin Card Sorting Test in healthy older adults. *Clinical Neuropsychologist*, *13*, 328–347.

Luteijn, F. & Barelds, D.P.H. (2004). *Groningen Intelligence Test 2 (GIT-2): Manual*. Amsterdam, The Netherlands: Harcourt Assessment BV.

Mahoney, F.I. & Barthel, D.W. (1965). Functional evaluation: The Barthel Index. *State Medical Journal*, *14*, 61–65.

Marien, P., Mempaey, E., Vervaet, A., Saerens, J., & de Deyn, P.P. (1998). Normative data for the Boston Naming Test in native Dutch-speaking Belgian elderly. *Brain Language*, *65*, 447–467.

McSweeney, A.J., Naugle, R.I., Chelune, G.J., & Luders, H. (1993). "T scores for change": An illustration of a regression approach to depicting change in clinical neuropsychology. *Clinical Neuropsychologist*, *7*, 300–312.

Muslimovic, D., Post, B., Speelman, J.D., & Schmand, B. (2005). Cognitive profile of patients with newly diagnosed Parkinson disease. *Neurology*, *65*, 1239–1245.

Muslimovic, D., Schmand, B., Speelman, J.D., & de Haan, R.J. (2007). Course of cognitive decline in Parkinson's disease: A meta-analysis. *Journal of the International Neuropsychological Society*, *13*, 920–932.

Nelson, H.E. (1976). A Modified Card Sorting Test sensitive to frontal lobe defects. *Cortex*, *12*, 313–324.

Owen, A.M., James, M., Leigh, P.N., Summers, B.A., Marsden, C.D., Quinn, N.P., Lange, K.W., & Robbins, T.W. (1992). Fronto-striatal cognitive deficits at different stages of Parkinson's disease. *Brain*, *115*, 1727–1751.

Palazzini, E., Soliveri, P., Filippini, G., Fetoni, V., Zappacosta, B., Scigliano, G., Monza, D., Caraceni, T., & Girotti, F. (1995). Progression of motor and cognitive impairment in Parkinson's disease. *Journal of Neurology*, *242*, 535–540.

Post, B., Merkus, M.P., de Bie, R.M., de Haan, R.J., & Speelman, J.D. (2005). Unified Parkinson's Rating Scale for motor examination: Are ratings of nurses, residents in neurology, and movement disorder specialists interchangeable? *Movement Disorders*, *20*, 1577–1584.

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press

Reid, W.G.J., Hely, M.A., Morris, J.G.L., Broe, G.A., Adena, M., O'Sullivan, D.J., & Williamson, P.M. (1996). A longitudinal study of Parkinson's disease: Clinical and neuropsychological correlates of dementia. *Journal of Clinical Neuroscience*, *3*, 327–333.

Reitan, R.M. (1992). Trail Making Test: Manual for administration and scoring. Tucson, AZ: Reitan Neuropsychological Laboratory.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris, France: Presses Universitaires de France.

Rothman, K.J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology*, *1*, 43–46.

Royall, D.R., Cordes, D.A., & Polk, M. (1998). CLOX: An executive clock drawing task. *Journal of Neurology, Neurosurgery, and Psychiatry*, *64*, 588–594.

Schmand, B., Houx, P., & de Koning, I. (2003). *Norms for Stroop Color Word Test, Trail Making Test, and story recall of Rivermead Behavioural Memory Test*. Amsterdam, The Netherlands: Neuropsychology Section of the Dutch Institute for Psychologists.

Schwab, J.F. & England, A.C. (1969). Projection technique for evaluating surgery in Parkinson's disease. In F.J. Gillingham & M.C. Donaldson (Eds.), *Third symposium on Parkinson's disease*. Edinburgh, UK: E&S Livingstone.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. New York: Psychological Corporation.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale (WAIS-III)* (3rd ed.). New York: Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale (WMS-III)* (3rd ed.). New York: Psychological Corporation.

Williams-Gray, C.H., Foltynie, T., Brayne, C.E.G., Robbins, T.W., & Barker, R.A. (2007). Evolution of cognitive dysfunction in an incident Parkinson's disease cohort. *Brain*, *130*, 1787–1798.

Wilson, B., Cockburn, J., & Baddeley, A. (1985). *Rivermead Behavioural Memory Test*. Reading, UK: Thames Valley Test Company.

Zigmond, A.S. & Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*, 361–370.

https://doi.org/10.1017/S1355617709090614 Published online by Cambridge University Press