University of Alberta

Longitudinal performance of Neuropsychological Assessments in Parkinson's Disease.

by

Taim Abdullah Muayqil

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

In

Clinical Epidemiology

Public Health Sciences

Fall 2012
Edmonton, Alberta

Dedication

To my mother Aziza, my father Abdullah, my wife Noha, and my son Abdullah.

All have made me into the person I am today.

Abstract

Parkinson's disease (PD) involves the motor system and can lead to dementia. Dementia incidence among 102 non-demented participants (52 PD and 50 controls) followed prospectively over three years was assessed. Cognition was measured annually using the Clock Drawing Test (CDT), Mini-Mental State Exam (MMSE), Frontal Assessment Battery (FAB) and Dementia Rating Scale (DRS). Mixed-effects model was used to determine the significance of the change in neuropsychological tests (NPT) over time. Incidence of dementia in the PD group was 41.67%. The FAB and DRS showed significant. In the PD subgroup, the copy-CDT (CCDT), FAB and DRS showed decline in patients with incipient dementia (PDID). The FAB and DRS scores correlated in the primary and subgroup analyses. In the subgroup, the CCDT correlated with FAB and DRS. The NPT were useful in discriminating PD from control patients with early cognitive impairments and discriminating PDID from PDND.

Acknowledgement

This thesis is the end result of numerous contributions from many dedicated colleagues. This would not of been possible without their commitment.

I would like to specifically thank my advisors over the past several years Dr. R. Camicioli, Dr. A. Senthilselvan, Dr. T. Jeerakathil and Dr. S. Kalra. Their vast combined experiences have been most valuable.

Jennifer Sabino worked hard accumulating data and organizing patient visits. I thank her for her efforts and for showing me how to administer various rating scales to patients. Dr. Nancy Fisher and Sandra Bradford deserve thanks and recognition for their collection and interpretation of neuropsychological data. I am grateful for the patience and time they took to ensure the quality of the results. I wish to also acknowledge those who contributed to the study over the years prior to my involvement, and unfortunately did not have a chance to meet them, Sheri foster, Thomas Bouchard, and Tracy Greene.

I especially would like to thank my wife and son for their support and their patience with my involvement in this project.

Most importantly, I thank God “Allah” for making everything possible.

Contents

Index of Tables

Index of Figures

List of abbreviations

AD	Alzheimer's Disease
ADCS-ADL	Alzheimer's Disease Cooperative Study-Activities of Daily Living Inventory
ADL	Activities of Daily Living
CCDR	Caregiver-Derived Clinical Dementia Rating Scale
CCDT	Copy clock drawing test
CDR	Clinical Dementia Rating Scale
CDT	Clock drawing test
CIRS	Cumulative Illness Rating Scale
DRS	Dementia Rating Scale
DSM-IV	Diagnostic Statistical Manual fourth edition.
FAB	Frontal Assessment battery
GDS	Geriatric Depression Scale
HIS	Hachinski Ischemia Score
H & Y	Hoehn and Yahr Scale
LBD	Lewy Body Dementia
MCI	Mild Cognitive Impairment
MMSE	Mini-Mental State Exam
MoCA	Montreal Cognitive Assessment
NPT	Neuropsychological Tests
PD	Parkinson's Disease
PDD	Parkinson's Disease Dementia
PDID	Parkinson's Disease Incipient Dementia
PDND	Parkinson's Disease Non-Dementia
UPDRS	Unified Parkinson's Disease Rating Scale
UPDRS-III	Unified Parkinson's Disease Rating Scale-Motor Subscale

1. Introduction

Parkinson's disease, first described by James Parkinson [1] is a neurodegenerative disorder that typically involves resting tremor, bradykinesia, rigidity and postural instability. [2] The exact cause of the disease is not yet known, however environmental factors are suspected but not unequivocally confirmed. [2] A smaller proportion of PD patients will have a genetic mutation, particularly if disease onset is before 50 years of age. [2] Degeneration is primarily in the pars compacta of the substantia nigra, affecting the dopaminergic neurons. Cell loss is not restricted to this region and has also been described in other brainstem regions, as well as in the cerebral cortex. [2] Such pathological changes can lead to non-motor features, most importantly cognitive decline, which culminates in dementia, a state where cognitive impairment affects functional abilities.

The prevalence of dementia among PD patients has been estimated to be 24-31% and Parkinson's Disease Dementia (PDD) contributes to 3-4% of all causes of dementia. [3] In a cohort of PD patients followed up for 8 years, 78% developed dementia. [4] PD patients with dementia live shorter than those with out PDD. [5] Not unlike Alzheimer's disease, PDD can be preceded by a mild cognitive impairment (MCI) phase. Whereas MCI in AD is usually of an amnestic type, that of PD tends to be non-amnestic, usually involving executive function and visuo-constructive impairments, and to a lesser degree, memory. [6]

Patients with PD are usually older and at risk for other causes of dementia such as AD and vascular dementia. The burden of dementia on the health care system is continuously growing. [7, 8] It is important that dementia is detected early among patients at risk. Recognition of cognitive changes early and monitoring its progression is relevant for the patients' and families' knowledge, life planning, and identifying appropriate management strategies. It is also important to recognize whether cognitive deterioration is from PD or secondary to a different condition such as AD, particularly since therapeutic strategies may differ. There are many neuropsychological assessment tools available to assist in detecting and tracking cognitive dysfunction. Each has different strengths and

weaknesses and different abilities for detecting impairments in specific cognitive domains.

Thesis Statement: Given the impact dementia has on individuals and families dealing with PD and on the health care system as a whole, and considering that there are no reliable prognostic tests in predicting the development of dementia in PD patients, I will explore the utility of commonly used clinical neuropsychological tests in determining PD patients who are at risk for dementia and if the performance on these tests decline with progression of the disease over time.

2. Hypothesis & Objectives

The hypothesis is that patients with PD will show greater decline in cognitive function over time when compared to age-matched elderly controls. Also, within the PD subgroup, incipient dementia patients are expected to have worse performance on scales at baseline and throughout the observation period as compared to patients who remain non-demented.

The first objective of this study was to assess the occurrence of dementia in non-demented PD patients in comparison to age and sex matched non-demented controls in patients over 65 years of age, followed over 3 years. The second objective was to perform a longitudinal assessment of non-demented elderly participants' performance on the Clock Drawing Test (CDT), Mini-Mental State Exam (MMSE), Frontal Assessment Battery (FAB) and Dementia Rating Scale (DRS), over a period of three years in order to identify differences in performance between patients with and without PD, and to identify which tools would help in monitoring and predicting decline. In addition, a similar analysis was conducted in the non-demented PD subgroup, in order to observe how these measure changed in those who went on to develop dementia (i.e.: incipient dementia (PDID)) and those who did not develop dementia (PDND). Lastly, we examined correlations between tests that changed significantly in longitudinal analyses.

3. Background

3.1. Cognitive impairment and Dementia in Parkinson's Disease

3.1.1. Clinical profile

Aarsland et al. conducted a multicenter study that involved pooling of multiple cohorts of patients with PD in order to better assess the profile of cognitive deficits in non-demented PD patients. [6] Of the 1346 patients included in the analysis, 347 (25.8%) had MCI. Memory impairment was the most commonly involved domain in 13.3% followed by visuospatial impairment in 11%, followed by executive and attentive impairments in about 10.1%. However, the most common MCI subtype was that of a single domain non-amnestic type which was found in 152 patients (11.3%) and comprised of visuospatial impairment in 54.8% and executive and attention deficits in 52%. [6] Rodrı´guez-Ferreiro et al. additionaly found language impairment in PDD. [9] A group of non-demented PD patients that were age and education matched with a control group of 42 healthy seniors. The mean MMSE score was significantly lower in PD patients but scored > 26 in both groups. In addition to executive and attention difficulties, semantic language difficulties were identified. [9] Muslimovic et al. studied 115 consecutive newly diagnosed PD patients with MMSE ≥ 24 from a multiple neurology outpatient clinics and 70 healthy controls. [10] Detailed neuropsychological assessments of multiple cognitive domains were performed. In this study, 24% of their PD patients showed cognitive impairments, and 4% of the controls did. In the PD group, the main domains that were impaired were memory, attention and executive function, particularly when complex measures were used. Cognitive impairments were not as well appreciated in both groups when simple methods of assessment were used. Interestingly, impairment on the CDT was common, but this was not based on standardized scores; impairment was thought to be consistent with the executive function requirements of the task. [10] In addition to determining predictive factors of cognitive decline in PD patients, the follow up of the cohort 3 years later had the goal of comparing cognitive decline in PD patients with the demographically matched controls to identify individual patients with cognitive decline. The study sample included

up to 89 new-onset PD patients, 52 known PD patients and 64 controls. [11] In the newly diagnosed PD patients, attention and psychomotor speed decline was the most prominent cognitive finding in comparison to controls. Significant but not substantial declines were also observed for memory impairment, executive dysfunction, and visual-spatial deficits. All of these deficits were present at an individual level when compared to healthy controls and the deficits progressed after 3 years of follow up. Spatial orientation tests, such as the Judgement of Line Orientation (JOLO) and CDT showed changes that had modest rate of decline.

Cognitive impairment in PD also takes a toll on patients' daily activities. Rosenthal et alexamined 111 PD patients of varying cognitive status with a mean age of 72.8 years, recruited from a movement disorder center. [12] Care-givers provided information on the function of patients which was measured by the Alzheimer's Disease Cooperative Study-Activities of Daily Living Inventory (ADCS-ADL). [13] Cognition was assessed using the DRS and the MMSE. There was an association between disability and cognitive impairment, with the odds ratio (OR) for scoring below 65 on the ADCS-ADL being 1.19 (95% CI 1.03-1.36) for each drop in the DRS by one point. Even after the exclusion of demented patients, there was a relationship between the ADL and cognitive performance as measured by the DRS. It seemed to be related to the limitation in instrumental ADL rather than basic ADL, even after adjusting for mood and motor deficits. [12] The authors concluded that cognitive impairment, even if mild, affects PD patients' daily function. [12]

3.1.2. Disease process

Executive dysfunction is major finding in PD patients' cognitive profiles. Executive dysfunction is an encompassing term for many frontal cognitive functions such as planning and implementation, task shifting, working memory, and inhibitory responses and all require some degree of conceptualization. [14] Executive functions are nicely described as a process where cognitive abilities are incorporated during situations where routine behavior would not provide an adequate response. [15] The frontal lobes of the brain are most consistently associated with executive function, with the prefrontal cortex playing a most prominent role. [14, 15] Caudate, precuneate, and middle temporal gyri as well as left lateral orbito-frontal cortex and dorsolateral prefrontal cortex grey matter volumes have been found to be decreased. [16, 17] In PD, it is believed that dopamine deficiency affecting the frontal-striatal circuits are partially responsible for cognitive deficits, particularly executive dysfunction. More recently, deficiencies in the serotonergic, noradrenergic, and cholinergic systems have also been implicated.[16]Motor features of PD respond better to dopaminergic therapy than cognitive features, due to varying degrees of dopaminergic neuron loss, usually more so in the putamen. [15] In the caudate, dopamine depletion is most pronounced in the regions that are connected with the dorsolateral aspects of the frontal lobe, and relatively spared in the regions connecting with the ventral regions of the frontal lobe. [15] Memory difficulties in PD are generally attributed to retrieval defects secondary to executive dysfunction. However, left medial temporal lobe atrophy has been associated with memory deficits in PD. [16, 18] Lewy bodies are believed to be the main pathological substrate, [19] usually involving limbic and cortical brain regions. In severe cases of advanced PD dementia, particularly with prevalent cortical cognitive deficits, the pattern of cognitive impairment may not be readily distinguished from Lewy Body Dementia (LBD) or AD. As noted, PDD, like LBD, have prominent executive and visuoconstructive impairments early on that may not be distinguishable on neuropsychological assessments, while AD patients tend to have more pronounced memory deficits in the early stages. [20, 21] Interestingly, it has been suggested that involvement of the frontal cortex in PDD may occur later in the course compared to

other dementing illnesses, as it has been found that mildly demented patients with LBD have more executive function impairment than those with mild PDD. [22]

3.2. Neuropsychological Assessment Tools

3.2.1. The Clock Drawing Test

The clock drawing test (CDT) has for a long time been incorporated in cognitive assessment. Early documentation of this test is found in literature from the 1950s where it was by Dr. MacDonald Critchley in patients with parietal lobe lesions and neglect. [23] Its use has expanded significantly since then and is now a familiar bedside cognitive test among physicians; it is also helpful in predicting performance in certain tasks such as driving. [24] There have been many studies looking at the performance of clock drawing among dementia patients of various etiologies; most studies included patients with Alzheimer's Disease (AD). Some studies have included subcortical dementias such as Huntington's disease, [25] Vascular cognitive impairment, [26] and Parkinson's disease (PD). [27, 28] Others have compared the performance of CDT among different dementia groups. [29-31] Recent studies are asking if CDT is useful in detecting various types of MCI and predicting to dementia. [26, 32-34] The instructions and implementation of the CDT varies among investigators. Some provide patients with a pre-drawn circle while others ask the patient to draw one. The time requested to be illustrated is also quite different between authors; however, a majority prefer "10 past 11" as it forces the patient to attend to both hemi-fields of the clock and mentally convert the "10" into a "2". It is important that instructions be specific about a time rather than open ended. [35] There has been no mutually agreed upon standard on how to grade a clock and the literature has over a dozen different clock scales. This has been rightfully met with criticism, [36] and this variability may be an indicator of how much useful information can actually be obtained from the CDT. One of the conclusions of a meta-analysis [37] of various scoring techniques at that time was that the more complicated clock scoring techniques did not offer added benefit over the simpler ones. The type of scoring method used to detect abnormalities did not differ much when assessing driving abilities in another study. [24]

A recent study that examined the prevalence of dementia in Parkinson disease using multiple tools, [27] and the CDT was found more sensitive at identifying patients with dementia than the Min-Mental State Exam (MMSE). [27] In one study, CDT was significantly abnormal in PD compared to a control group.[38]Another study found it to be useful for detection of early cognitive impairment in PD. [39]

3.2.2. The Frontal Assessment Battery

The Frontal Assessment Battery (FAB) [40] is a more recently devised bedside cognitive assessment tool that focuses on detecting impairments in frontal lobe cognitive functions. Dubois et al. studied 42 normal controls and 121 patients with frontal lobe dysfunction, 24 had mild PD. [40] For concurrent validity the battery was correlated with the Wisconsin card sorting task and the Dementia Rating Scale (DRS). Significant correlation was found with both tests and the FAB was able to discriminate between normal and impaired individuals. Two raters examined a subgroup of 17 patients independently with excellent Kappa inter-rater reliability ($k = 0.87$, $p < 0.001$). The total obtainable score is 18 and the test requires approximately 10 minutes to implement, the subcomponents of the test include: conceptualization, by inquiring about object similarities; mental flexibility, by a phonemic verbal fluency task; programming, by a motor series task (Luria Task); sensitivity to interference, by a task that involves conflicting instructions; inhibitory control, by performing the Go-No Go task; environmental autonomy, or assessing for prehensile behavior by looking for a grasp reflex. [40]

Kaszas et al. assessed the accuracy of the FAB in 73 PD patients to detect PDD. [41] In the study, 22 of the 73 had PDD, and the area under the receiver operator characteristic (ROC) curve was found to be 0.78 (95% CI 0.64–0.9). A score of 12 was determined to be 66% sensitive and 72% specific.[41] A correlation between education level and performance on the FAB has also been reported. [42, 43] In a study that screened 75 patients with early PD and 45 healthy controls for frontal deficits, PD patients had significantly lower FAB and MMSE scores than the controls.[44] With the FAB, significant differences

were found in the similarities, motor series and conflicting instructions subscores. On the MMSE, recall and visuoconstructive abilities differed between PD and controls. The FAB scores were not related to motor symptoms or dopaminergic dosages. [44]. Lima et al. aimed to assess the usefulness of the FAB as a screen for executive dysfunction in PD patients and to assess how it correlated with other tests of executive dysfunction. They compared 122 normal subjects and 50 patients with PD whose Hoehn & Yahr score was 3-4, indicating advanced PD. In addition PD patients had 3 years less education than the control group. The FAB scores in normal participants were lower among those with lower education and higher age. Similar findings in the PD group, except education did not significantly correlate. The authors concluded good concurrent validity for the FAB as the scores correlated with other tests of executive functioning. [45] Frontal lobe dysfunction may not be the sole neural correlate of impaired performance on the FAB. [46] Matsui et al. performed SPECT scans on 30 patients with PD (who had MMSE scores > 23 and did not meet DSM 4 dementia criteria). Patients with low scores on the FAB (11 or less) had decreased perfusion in the left supramarginal gyrus and left inferior parietal lobule compared to those in the higher scoring group (12 or more). [46]

3.2.3. The Dementia Rating Scale

The Dementia Rating Scale (DRS) has been established longer than the FAB [47] and has been validated in Alzheimer's disease. [48] It has been widely used in evaluating PD patients, including those who undergo Deep Brain Stimulation. [41] The test assesses attention, language, verbal memory, visuospatial and executive functions. Subcomponents of the test include: attention by testing digit span and following commands; initiation and perseveration by verbal semantic fluency tasks, performing alternating movements and alternating graphic designs; construction by copying images; conceptualization by testing for similarities between images and verbal cues; memory and attention by testing short term verbal and visual memory recall and orientation. It is scored out of a total of 144 points. Smith et al. examined the psychometric properties of the DRS in normal controls, patients with MCI, and patients with various dementia types

over an average follow up of 3.9 years The DRS total score declined by > 10 points in less than 5% of normal individuals but in more than 60% of patients with dementia. The DRS was also found to be predictive of requiring institutionalization for patients with decline beyond two standard deviations. Patients who scored 100 on the DRS, were 2.5 times more likely to die than those who scored 135. The median survival for those scoring < 100 was 3.7 years. [49] The DRS was more accurate than the MMSE in identifying deficits in mildly impaired patients. [12] When comparing the DRS to FAB and MMSE, Kaszas et al. found the DRS to have the highest sensitivity and specificity for identifying PDD. The area under ROC curve was 0.925 (95% CI: 0.85–1) with a sensitivity of 90% and a specificity of 98%, using the cut-off score of 125 points. [41] Aarsland et al. studied the DRS in various disorders with a goal to determine the similarity of the cognitive profiles between PDD and LBD, and compare to a known subcortical dementia Progressive Supranuclear Palsy (PSP) and a known cortical dementia, AD. [22] The pattern of cognitive impairment was similar in severely demented patients with LBD and PDD. Among those with mild to moderate degree of dementia, four out of the five subscores were similar. PDD patients scored worse on initiation/perseveration subscales and better on memory subscales when compared to AD patients. Poalo et al. found a similar difference when looking at 50 AD and PD patients both with a DRS < 131. Most of the patients suffered from mild to moderate degrees of dementia with an average DRS of 104, and the AD patients performed worse on the memory subscores wile the PDD were worse on the construction subscore. [50] Brown et al. also found that most variability was in the initiation/perseveration, conceptualization and memory subscores. [51] The DRS is believed to have better predictive validity in comparison to other screening tools due to its broader coverage of cognitive domains. [51]

Matteau et al. conducted a study which aimed to determine if the DRS can identify MCI in PD patients and differentiate it from amnestic MCI (aMCI). [52] They performed a cross sectional study that included participants over 50 years of age. The groups consisted of healthy controls, aMCI group, AD, MCI with PD (MCI-PD), PD and PDD. They found that, after correcting for age and education, the MMSE significantly correlated with DRS. The DRS did not correlate with age, PD duration or Hoehn and Yahr stages in the PD groups. Both MCI groups, and both dementia groups performed similarly on the

DRS. Patients with dementia scored significantly lower than healthy controls and a cut off score of 132 or less accurately discriminated dementia patients from healthy controls. Although worse performance was found among AD patients on memory subscales in comparison to PDD patients, the AD patients did not do better than PDD patients in the subscale of initiation/perseveration, which is in contrast to the 2003 study by Aarsland et al. [22, 52] With regard to MCI, patients with aMCI and MCI-PD both scored lower than healthy controls, while still being normal on the MMSE. The aMCI and MCI-PD performed similarly on the subscales. Another study that compared 40 PD patients with 34 REM sleep behavior disorder patients, found the sensitivity and specificity of the DRS was 72% and 86% for detecting MCI in PD patients wen the cut-of score was 138. No reliable cut-off score could be determined for the MMSE. [53] A recent systematic review found that the performance changes on the DRS, which usually takes 25-45 minutes to administer, depended on age and education. [54] In a study that had a goal of establishing a cut off on the DRS that would help discriminate between PDD and PDND, 92 PD patients were studied, and 35 of them with PDD. The total DRS score best variable to discriminate PDD from PDND after correcting for covariates such as age, education and the motor subscale of the UPDRS was 123. Memory, initiation/perseveration, and conceptualization subscales also helped in discriminating PDD from PDND. [55] The ROC curve analysis for the diagnostic accuracy of the DRS for detecting PDD among patients with PD was 91.3% (95% CI 0.85–0.97). There was a 42% increase in odds for PDD for every one-point decrease in the DRS. The DRS had a sensitivity and specificity of 92.6% & 91.4%. [55]

3.2.4. The Mini-Mental State Exam

The Mini-Mental State exam [56] is scored out of 30 points and it is the most familiar cognitive assessment scale.[57] It has questions on orientation, attention & concentration, immediate and short-term memory, language, visual spatial ability and attention. The test requires 5-10 minutes to complete. It has good internal consistency and inter-rater reliability.[56, 57] A large component of the test score relies on orientation and memory, thus

it may miss detecting impairments in patients where other cognitive domains are primarily involved such as in PD, vascular dementia or frontotemporal dementia. [57]

In the original study by Folstein et al., the MMSE was performed on two participant groups. The first consisted of specifically selected patients that included 29 patients with various dementia syndromes, 40 with an affective disorder and 63 normal elderly patients. The second group consisted of unselected patients that were collected according to consecutive admissions. Nine had dementia, 45 with an affective disorder, 24 with schizophrenia, 32 with drug abuse, and 27 with neurosis. In the first group, dementia patients mean score was 9.7, which was markedly lower than controls 27.6 and psychiatric patients 19 – 25.1. A similar discrepancy was found among the unselected patients of the second group, with dementia patients scoring around 12.2 and patients with psychiatric disorders scoring 24.6 - 27.6. The inter-rater reliability was also considered good (Pearson coefficient 0.827) and the MMSE scores correlated with the Wechsler Adult Intelligence Scale; Performance IQ (Pearson $r = 0.660$ ($p < 0.001$)) with Verbal IQ (Pearson $r = 0.776$ ($p < 0.0001$)). The MMSE has been shown have a low sensitivity to detecting cognitive changes in patients with PD compared to other bed side cognitive tasks which include broader testing of cognitive domains. [58, 59] In a systematic review that assessed neuropsychological scales in PD, limitations for the use of the MMSE were identified, such as floor and ceiling effects, and difficulty determining a reliable cut-off due to its dependency on age and education. The authors concluded it had a low sensitivity for detecting PDD. [54] Sarra Nazem et al. looked at a convenience sample of 100 PD patients and found that 52% of patients considered normal by MMSE were cognitively impaired according to the Montreal Cognitive Assessment scale (MoCA). [59] A cross sectional study by Williams et al. included 108 PD patients, and found the total UPDRS score to be correlated with the MMSE and DRS scores. As did they correlate with the motor subcomponent of the UPDRS in a multivariable regression analysis. It was also found that right-sided symptoms related to the DRS (trend for MMSE P = 0.054), bradykinesia to the MMSE (trend for DRS P = 0.05), and axial symptoms (significant for both scales) were related to cognitive function. Contrary to most studies that have looked at the MMSE in PDD, Kaszas et al found that for the MMSE, the area

under the curve was 0.87 (95% CI: 0.82–0.99) and determined that a cut-off value 26 points had a sensitivity and specificity of 80% and 74% which was superior to the FAB.[41]

4. Methods

4.1. Subjects

In our study, 52 cognitively normal PD patients and 50 normal age matched controls PD patients were identified from the movement disorder clinic at Glenrose Rehabilitation Hospital in Edmonton, Alberta, the Parkinson's society or referred by community neurologists. Patients were diagnosed by an experienced recruiting neurologist according to standard criteria, specifically; subjects were required to have a movement disorder of insidious onset with two signs among tremor, rigidity or bradykinesia, in the absence of atypical clinical features. [60] The ability to ambulate independently, speak English, and have a reliable informant were required for participation in the study.

The control group subjects were recruited through advertising in local senior centers or were contacts of enrolled patients. They were matched by sex and age (+/- 5 years) with PD patients. In both groups, those with previously known neurological disorders, memory or other cognitive concerns of any etiology (eg: degenerative, systemic condition or drug), unstable medical conditions, or previous clinical stroke were excluded. Presence of Parkinsonism was an exclusion for the control group only. Current anti-cholinergic medications were the only exclusionary anti-parkinsonian medications. Stable, controlled hypertension, diabetes, osteoarthritis, and treated hypothyroidism are examples of common allowed conditions. Patients attended three visits for physical and cognitive assessments: at enrollment, 18 months and 36 months. Telephone follow-ups were conducted at 6-month intervals in between visits.

Demographic information was collected with regard to sex, age, level of education, smoking and alcohol history, and medical comorbidities. Information on the ventricular volumes of the patients, their risk of falls, and performance on other neuropsychological tests have been published. [61-63]

4.2. Interventions

The CDT was administered and scored according to the method described by Rouleau et al. [25, 64] We chose the Rouleau method because it has both quantitative and qualitative components. The numerical score is based on a brief 10-point scale. Patients were given a blank A4 size white paper and a pencil. They were asked to first draw the clock to command. The instructions were: "Draw a face of a clock, including all the numbers and the hands with the time set to ten past eleven". The instructions were repeated as necessary if not understood on initial attempts. After completion they were provided with a correct illustration of a clock with time set and asked to copy it. Two raters blinded to the individuals independently rated the CDT and a third rater analyzed the discrepant ratings. A trained research assistant administered the MMSE, and DRS according to their respective manuals and a neurologist administered the FAB. Parkinsonism was measured using the Unified Parkinson's Disease Rating Scale (UPDRS) [65] and Hoehn and Yahr (H&Y) [66] scale. Participants were also assessed for depression with the Geriatric Depression Scale (short version) (GDS), [67] for burden from other medical conditions with the Cumulative Illness Rating scale (CIRS), [68] and cerebral vascular disease burden with the Hachinski Ischemic Score (HIS). [69]

All participants had a full neurological history and exam, in addition to the above-mentioned tests, taken at each of the three annual visits. At baseline, participants had blood tested for potentially reversible causes of cognitive impairment, which included complete blood count, glucose, urea, creatinine, electrolytes, liver enzymes, thyroid stimulating hormone, vitamin B12 and folate. Assessments over the three years were performed at approximately the same time between 9 am and noon.

4.3. Outcomes

The main outcomes were the incidence of dementia in both groups, and a change over time in a cognitive assessment scale indicating worsening performance. The Clinical Dementia Rating Scale (CDR) and caregiver derived CDR (CCDR) [70] were obtained from the patient and an informant respectively, and collected via separate interviews. The CCDR was favored over the CDR as the method to determine progression over time because it relies on information provided by a reliable informant other than the patient, limiting misinformation provided by a cognitively impaired patient. Dementia was considered present if there was impairment in at least two cognitive domains with functional impairment secondary to cognitive decline, and impaired memory was not considered a requisite, modified from DSM-IV criteria. [71] This classification was based on all available information available to the examiners (neurologist and research assistant).

4.4. Sample size

To examine the incidence of dementia among patients with PD compared to a control sample, an incidence rate for dementia in PD patients (5-10% per year, 25-50% after 5 years) and for controls (1-2% per year, 5-10% after 5 years) [72] was used to determine sample size. An estimation was done using the smaller values of 25% for PD and 5% for controls; this was considered conservative as the population was expected to have a higher incidence given it was comprised exclusively of patients 65 years of age and older. Therefore, 50 subjects per group (patients and controls) would be sufficient to discern a difference in incidence rate at 18 months with a power of 70% assuming a one-sided alpha (type 1 error rate) of 0.05.

4.5. Statistical analysis

Exploratory statistical tests to determine differences at different time points with T-tests and χ^2 tests used when appropriate. Because the groups were age and sex matched by +/-

5 years, multivariate regression correcting for these two variables was used to determine significance of differences between the PD and control group. The difference in incidence of dementia between PD and control groups, was examined using the Chi-square test (χ^2) and odds ratio was calculated using logistic regression. Kaplan-Meier estimates were used to determine the occurrence of cognitive decline or dementia with failure defined as an increase in the CCDR by 0.5, and the hazard ratio was determined. Mixed effects models, specifically assessing for an interaction between group membership and with time variable, was used to assess the CDT, MMSE, FAB, and DRS between the two groups (PD and controls) and, in a separate set of analyses, between the two subgroups (PDID and PDND). Age, sex and education were the main confounders and we chose the Hoehn and Yahr (H&Y) as the scale for severity of parkinsonism in the subgroup analyses. It became apparent that the H&Y was a more robust discriminator of the PD subgroups than the UPDRS-III (motor subscale of UPDRS). Tests with results that were significantly different between the PDID and PDND groups were correlated with Pearson correlation. STATA 11.2 software was used to process the results.

5. Results

5.1. Parkinson's Disease versus controls

5.1.1. Baseline characteristics

Baseline characteristics and performance on neuropsychological tests at baseline, 18 months and 36 months for PD and control groups are shown in Table 1. Mean age, and mean years of education and distribution of sex were not statistically different between the groups. At each assessment time point, the control group had significantly higher scores on the SCDT and the CCDT than the PD group. In the qualitative analysis of the SCDT, spatial/planning errors were occurred significantly more in the PD group at each time point. Both groups tended to draw normal sized clocks, but PD patients were more likely to have graphic errors. Conceptual errors occurred more often in the PD group in the second and last assessments. Stimulus bound errors and perseverative errors were infrequent (Table 2.). In the copy condition, spatial/planning errors were still more frequent in the PD group, while conceptual, stimulus bound, and perseverative errors were very rare. There was no difference in the number of normal sized clocks drawn by each group. Graphic difficulties became more significant as time progressed in the PD group (Table 3). The MMSE was also lower in the PD group but the difference was only significant at the 36 month assessment. The FAB and DRS were significantly lower at all three assessment points (Table 1., Figure 2.). The correlation between of the examiners' ratings on the assessment scales was consistent with the previous literature. [25] It was lower on some components of the qualitative CDT such as the assessment of the degree of graphic difficulties 55.1%, the presence of stimulus bound errors 36.9%, conceptual errors 52.9% and perseverative errors 53.5%. Correlation was over 90% in determining the presence of spatial/planning errors and clock size.

.

5.1.2. Dementia incidence

The incidence of dementia at 36 months was available for 98 participants, 48 PD patients and 50 controls. From the entire cohort, 25 (25.1%) participants developed dementia, 20 (41.67%) were PD patients and 5 (10%) were controls (P <0.0001). The odds ratio (OR) for developing dementia in the PD group was 6.43 (95% CI 2.17 - 19.07, P = 0.001), using univariate logistic regression. Multivariate regression adjusting for sex, baseline age and years of education did not markedly alter the OR for the development of dementia, which was 7.01 (95% CI: 2.2 - 22.3, P = 0.001), (Hosmer-Lemeshow goodness-of-fit $\chi^2 = 10.35$, P = 0.24). The OR for the risk of dementia with every one-year increase in age was 1.14 (95% CI: 1.02 - 1.27, P = 0.021).

The CCDR results at 6-month intervals are shown (Table 1). Kaplan-Meier survival estimates, with failure defined as the time when the CCDR score increased by 0.5, are demonstrated (Figure 1) and show that cognitive decline occurred more frequently in the PD group and the occurrence was significantly increasing with time (P < 0.0001). Cox proportional hazard analysis was set with failure defined as an increase in the CCDR by 0.5 from baseline assessment. In the PD group, 34 patients had an increase by at least 0.5 in the CCDR above their initial assessment, and an increase occurred in 13 of the controls. This yielded a hazard ratio of 3.6 (P < 0.0001) in the PD group for developing dementia over the 36 month observation period. None of the patients had a CCDR of 1 at baseline, while six PD patients had a CCDR of 0.5 at baseline. Two of these PD patients remained at 0.5 at subsequent assessments, 1 dropped out before the 6 months assessment and the remaining four increased by at least 0.5. Out of the five that remained in the study 4 of them met the study's criteria for dementia at the final 36-month assessment.

5.1.3. Dropouts and deaths

From the PD group, two died prior to the 6-month telephone assessment, one prior to the 24-month telephone assessment and one prior to the 30-month telephone assessment.

One PD patient dropped out prior to the 6-month telephone assessment, and one prior to the 36-month assessment. In the control group, one died prior to the 18-month assessment and one died prior to the 36-month assessment. One control dropped out prior to the 6-month telephone assessment.

5.1.4. Mixed effects analysis of the quantitative CDT

Mixed effect model was used to determine if the numerical CDT scores varied between the two groups at baseline (t = 0) and if the difference between the two groups varied over time after controlling for age, sex and education. With regard to the SCDT, and despite it being significantly lower in the PD group at each visit (Table 1), there was no significant interaction between the two groups and time (P = 0.85). In addition, the PD group had a statistically significant lower score on the SCDT at baseline in comparison to the control group P < 0.0001). There was also a statistically significant association between SCDT and age with a lower SCDT score for every one-year increase in age (P < 0.0001). The same pattern was observed in the CCDT. The interaction between the group and the time of assessment was not significant (P = 0.18), while it was significantly lower at each visit in the PD group (Table 1). In the final model that included the above-mentioned variables, only age was significantly associated with the CCDT: every 1-point decrease in the CDT score was associated with a 1-year increase of age (P < 0.0001). The difference in score at the baseline between PD and control groups was just above the significance level (P = 0.056) (Appendix 1).

5.1.5. Mixed effects analysis of FAB

The FAB was evaluated controlling for sex, age, and education. The group difference was slightly non-significant (P = 0.066); however, there was a significant difference with each visit (P = 0.038). The interaction between the patient group and time was also significant supporting a greater decline in performance over time in the PD group (P =

0.001). Age also showed a significant association ($P < 0.001$). Males performed slightly better then females ($P = 0.031$) (Appendix 1).

5.1.6. Mixed effects analysis of DRS

The DRS was evaluated controlling for sex, age and education. The interaction between patient group and time of assessment was significant indicating a greater change over time in the DRS ($P = 0.046$) in the PD group. The scores were also significantly lower with every one-year increase in age ($P < 0.0001$) (Appendix 1).

5.1.7. Mixed effects analysis of MMSE

The MMSE was evaluated with the confounders of sex, age, and education. There was a significant decrease in MMSE scores at each visit ($P = 0.034$); however, the interaction between to two groups and time was not statistically significant ($P = 0.059$). Every 1-year increase in age was significantly associated with a decrease in the MMSE -($P < 0.0001$) (Appendix 1).

Table 1. Baseline, 18-month, and 36-month results of demographic and

	PD Mean(SD)	Control Mean(SD)	Multivariate regression correcting for sex and age P-Value
Age at enrollment	71.5 (4.6)	71.6 (4.8)	0.996
Education (years)	13.9 (3)	15.1 (3.5)	0.091
Sex	F: 30, M: 22	F: 29, M: 21	0.97*
SCDT 0	7.7 (2.4)	9 (1.3)	0.001
SCDT 18	7.5 (2.4)	8.9 (1.7)	0.001
SCDT 36	7.3 (2.5)	8.7 (1.6)	0.001
Copy CDT 0	8.7 (1.5)	9.2 (0.9)	0.049
Copy CDT 18	8.3 (1.5)	9.1 (1.2)	0.003
Copy CDT 36	7.96 (2.1)	8.9 (0.9)	0.004
CCDR at time 0	0.058(0.022)	0.03(0.12)	0.325
CCDR at time 6	0.085(0.19)	0.04(0.17)	0.211
CCDR at time 12	0.13(0.24)	0.031(0.16)	0.026
CCDR at time 18	0.23(0.55)	0.03(0.12)	0.016
CCDR at time 24	0.24(0.37)	0.1(0.23)	0.026
CCDR at time 30	0.34(0.57)	0.13(0.26)	0.022
CCDR at time 36	0.52(0.71)	0.11(0.23)	<0.0001
UPDRS 0	28.54(14.28)	3.34(3.92)	<0.0001
UPDRS 18	28.08(15.37)	3.46(4.29)	<0.0001
UPDRS 36	32.98(17.38)	4.98(5.31)	<0.0001
UPDRS-III 0	17.65(8.78)	2.14(2.96)	<0.0001
UPDRS-III 18	17.27(9.89)	1.98(2.91)	<0.0001
UPDRS-III 36	20.69(10.99)	2.66(3.25)	<0.0001
MMSE at time 0	28 (1.8)	28.4 (1.6)	0.189
MMSE at time 18	26.9(3.2)	27.8(1.7)	0.081
MMSE at time 36	26.5(3.5)	27.7(1.8)	0.045

FAB at time 0	14.4 (2.2)	15.2 (1.7)	0.023
FAB at time 18	14.19(2.68)	15.62(1.7)	0.001
FAB at time 36	13.73(3.3)	15.83(1.89)	<0.0001
DRS at time 0	136.7 (4.9)	138.5 (3.7)	0.030
DRS at time 18	135.6(8.8)	138.9(3.4)	0.014
DRS at time 36	135.8(3.7)	138.98(3.7)	0.023
Hoehn-Yahr 0	2.25(0.65)	0(0)	<0.0001
Hoehn-Yahr 18	2.19(12.7)	0(0)	<0.0001
Hoehn-Yahr 36	2.5(0.94)	0(0)	<0.0001
PD Duration at enrollment	8.75(4.4)	NA	
GDS 0	2.2(2.8)	0.74(1.07)	0.001
GDS 6	1.92(1.36)	0.62(0.99)	<0.0001
GDS 12	2.12(2.3)	0.84(1.17)	<0.0001
GDS 18	2.3(2.4)	0.69(0.99)	<0.0001
GDS 24	1.91(1.84)	0.92(1.3)	0.003
GDS 30	2.09(2.3)	1(1.6)	0.003
GDS 36	2.3(2.2)	1.06(2.12)	0.007
HIS 0	0.77(1.13)	0.4(0.73)	0.052
HIS 18	1.27(1.8)	0.79(1.2)	0.13
HIS 36	1.4(1.63)	1.4(1.7)	0.9
CIRS 0	19.4(2.5)	18.4(2.9)	0.052
CIRS 18	20.42(3.02)	19.42(18.63)	0.09
CIRS 36	20.3(3.7)	19.5(4.02)	0.34

neuropsychological variables for PD and control patients.

*Indicates χ^2 values.

Table 2. The *spontaneous* condition of the clock drawing test in PD and controls, qualitative assessment of its components

	PD (%)	Controls (%)	P Value
Presence of spatial/planning errors			
0	30 (57.7)	13 (26.53)	0.002
18	27 (56.2)	16 (34.04)	0.024
36	27 (60)	14 (30.4)	0.004
Presence of conceptual errors			
0	14 (26.92)	7 (14)	0.11
18	11 (22.92)	4 (8.33)	0.049
36	12 (26.67)	4 (8.51)	0.022
Presence of stimulus bound errors			
0	6 (11.54)	7 (14)	0.71
18	5 (10.42)	1 (2.08)	0.092
36	7 (15.56)	1 (2.13)	0.022
Presence of perseverative errors			
0	1 (1.92)	1 (2)	0.98
18	1 (2.08)	0	0.32
36	1 (2.22)	0	0.3
Presence of normal size clock			
0	42 (84)	46 (88.46)	0.62
18	40 (83.3)	43 (89.58)	0.37
36	35 (77.78)	41 (87.23)	0.23
Absence of graphic difficulties			
0	19 (36.54)	33 (66)	0.003
18	16 (33.33)	23 (47.92)	0.15
36	14 (31.11)	30 (63.83)	0.002

Table 3. The *copy* condition of the clock drawing test in PD and controls, qualitative assessment of its components

	PD (%)	Controls (%)	P Value
Presence of spatial/planning errors			
0	21 (41.2)	19 (38.8)	0.74
18	27 (57.5)	13 (27.08)	0.003
36	26 (57.78)	14 (29.79)	0.007
Presence of conceptual errors			
0	2 (3.92)	0	0.16
18	3 (6.38)	1 (2.08)	0.296
36	3 (6.67)	0	0.072
Presence of stimulus bound errors			
0	0	0	NA
18	0	1(2.08)	0.32
36	0	0	NA
Presence of perseverative errors			
0	0	0	NA
18	0	0	NA
36	0	0	NA
Presence of normal size clock			
0	49 (96.08)	50 (100)	0.157
18	46 (97.87)	47 (100)	0.31
36	42 (93.33)	47 (100)	0.144
Absence of graphic difficulties			
0	25 (49.02)	33 (66)	0.084
18	14 (29.79)	24 (50)	0.044
36	20 (44.44)	34 (72.34)	0.007

NA: Not applicable

Figure 1. Kaplan-Meier estimates for cognitive decline in the cohort of PD patients and controls. An event was defined as an increase in the CCDR by 0.5

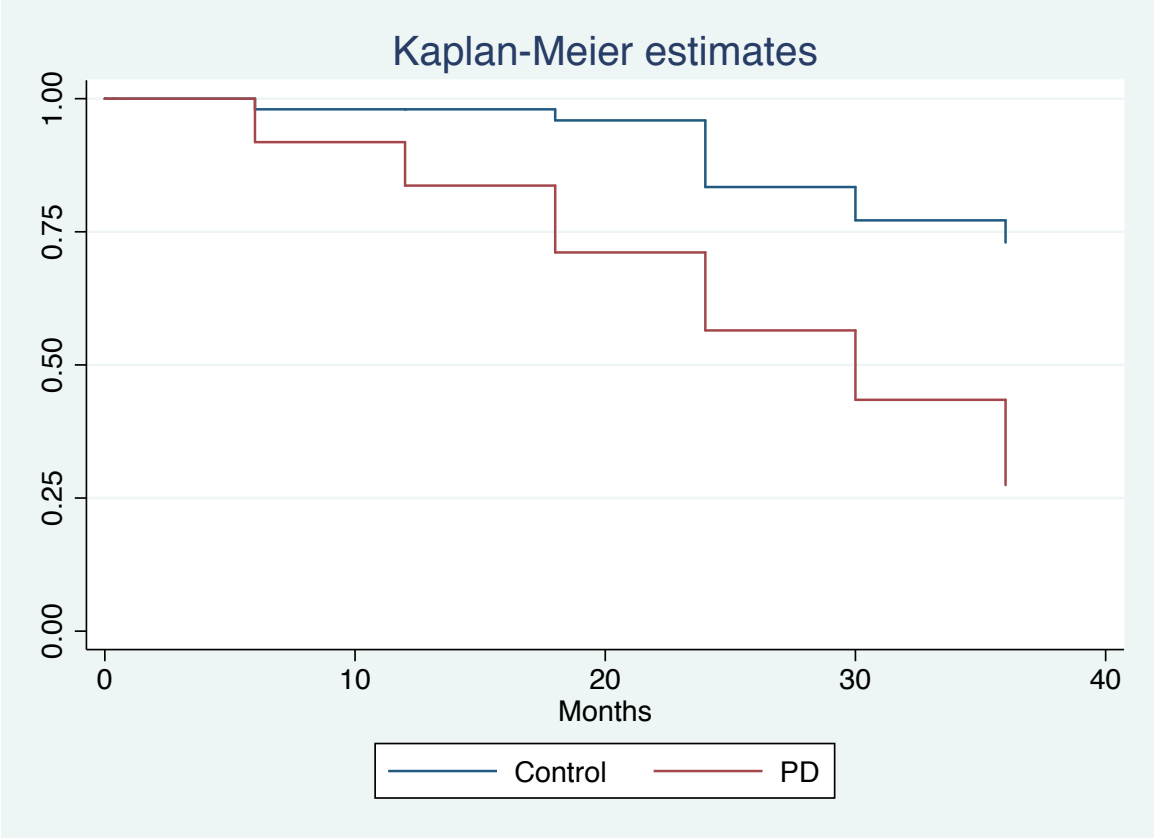
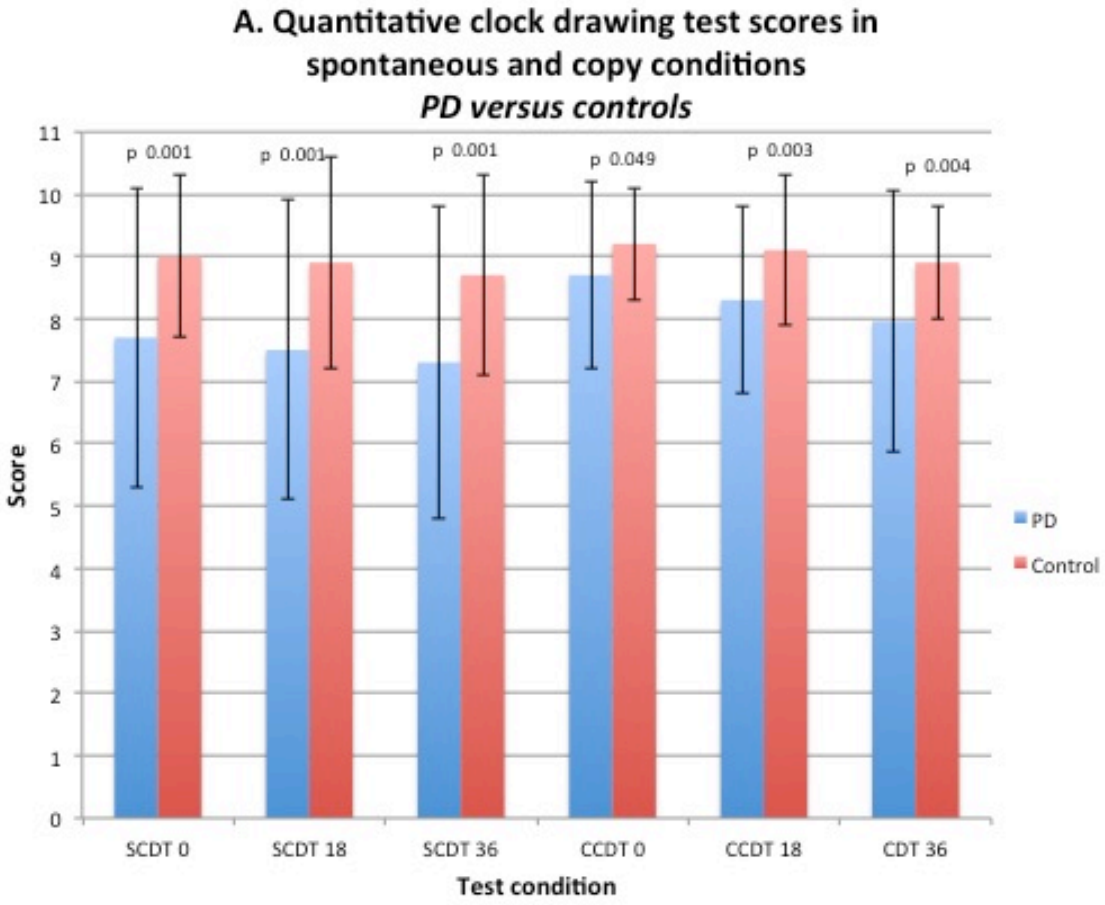
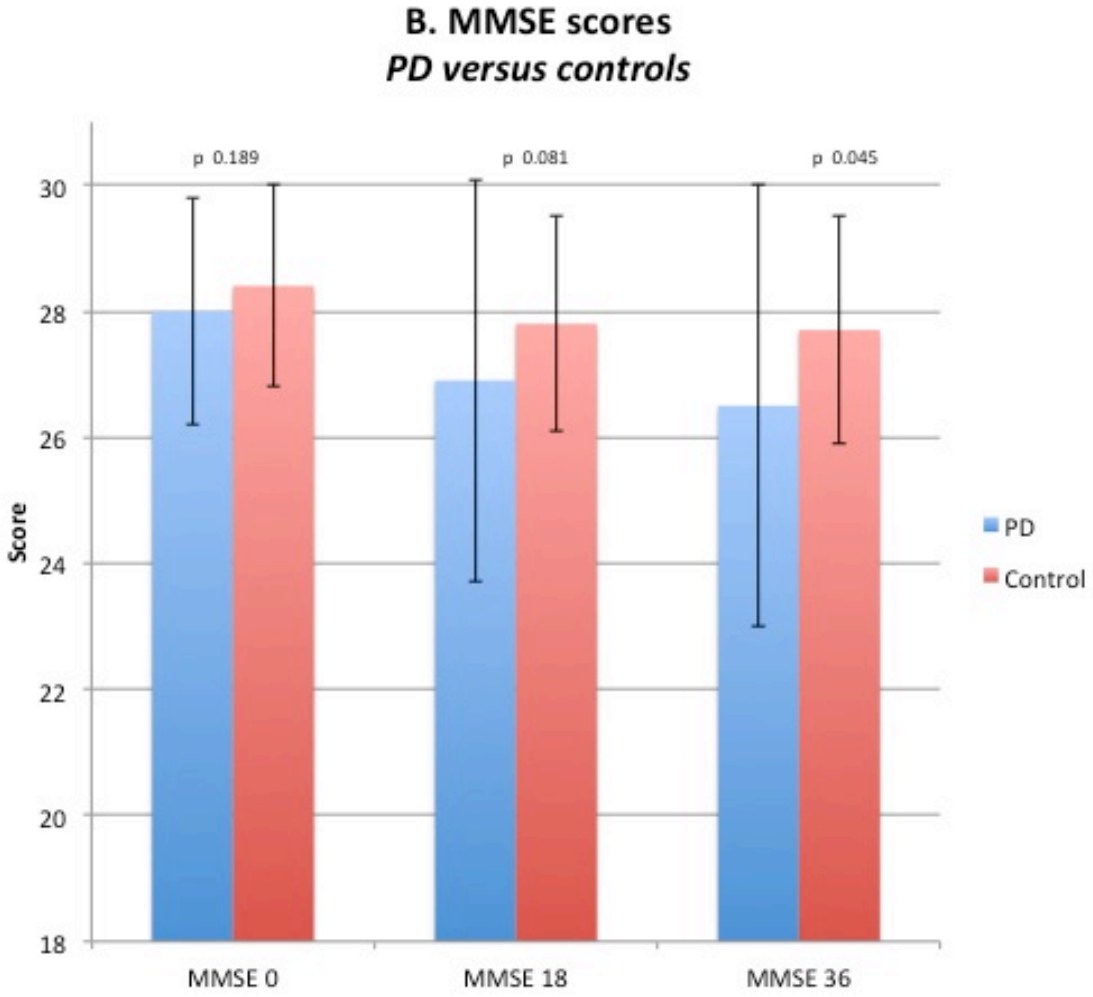
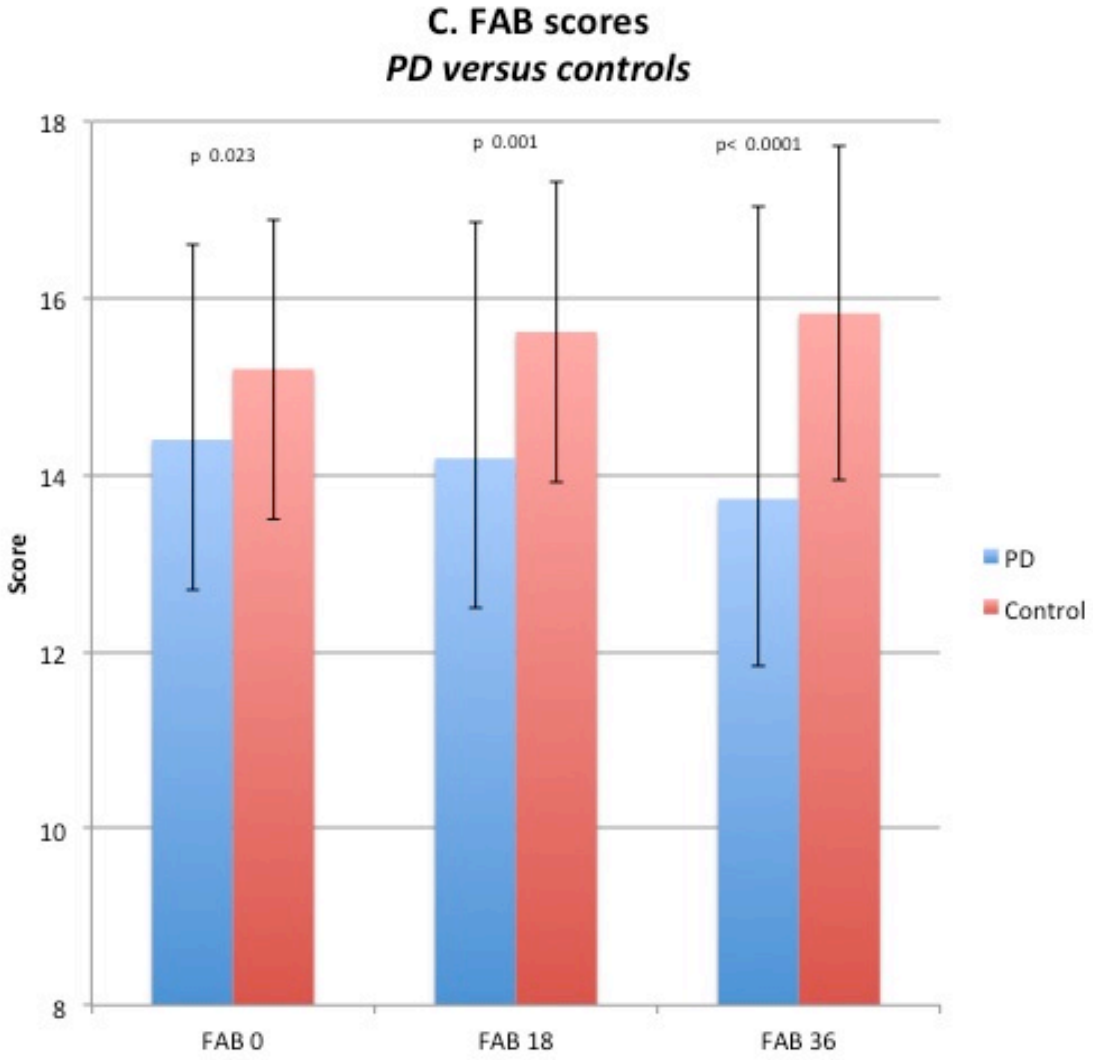
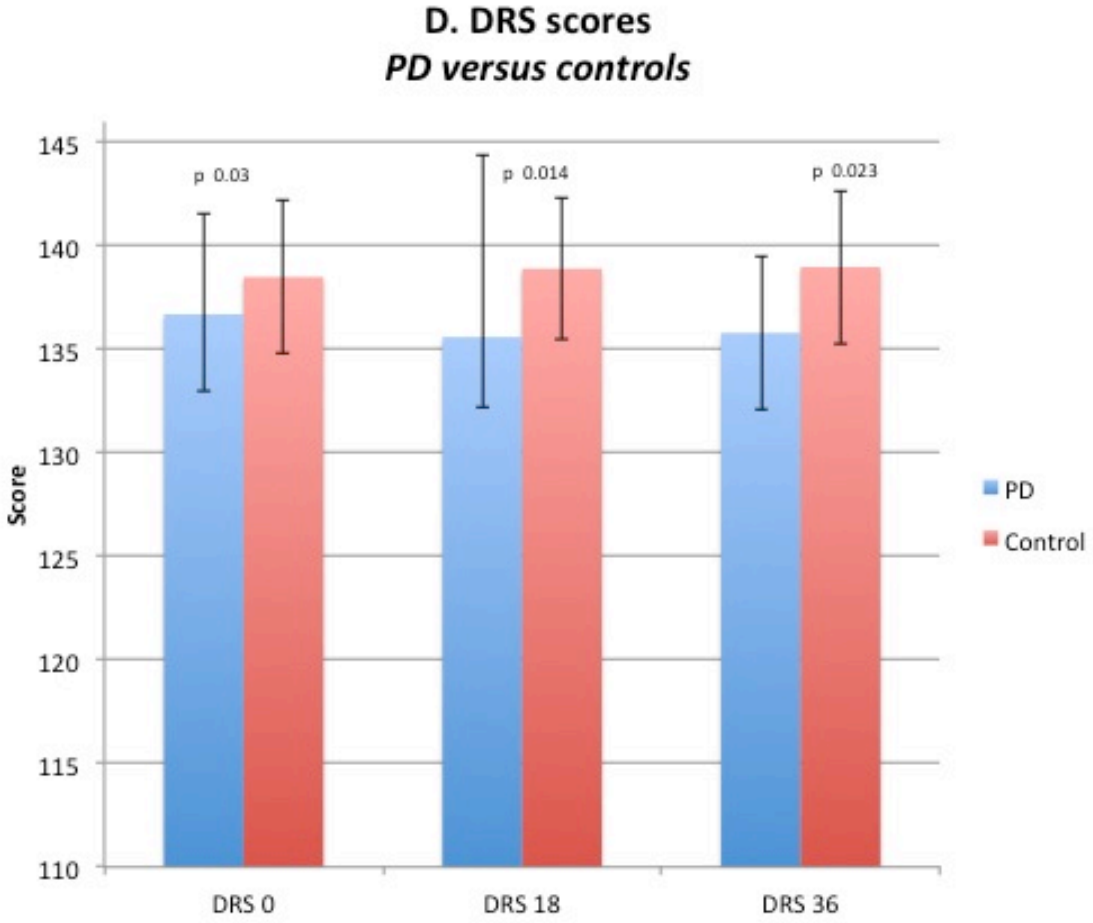


Figure 2. Mean scores of the neuropsychological scales at 0, 18, and 36 months assessments for controls and PD patients. A) Quantitative clock drawing test scores in spontaneous and copy conditions. B) Mini-mental State Exam. C) Frontal Assessment Battery. D) Dementia Rating Scale. Error bars indicate standard deviations.




B. MMSE scores
PD versus controls
p 0.189
p 0.081
p 0.045
30
28
26
24
22
20
18
Score
MMSE 0
MMSE 18
MMSE 36
PD
Control


C. FAB scores
PD versus controls
18
16
14
12
10
8
Score
p 0.023
p 0.001
p< 0.0001
FAB 0
FAB 18
FAB 36
PD
Control


D. DRS scores
PD versus controls
Score
145
140
135
130
125
120
115
110
p 0.03
p 0.014
p 0.023
DRS 0
DRS 18
DRS 36
PD
Control

5.2. Subgroup analysis: Parkinson's Disease Incipient Dementia versus Parkinson's Disease Non-Dementia

5.2.1. Baseline characteristics

Baseline characteristics and performance on neuropsychological tests at baseline, 18 months and 36 months for PDID and PDND groups are shown in Table 4. The mixed effects model included the variables age, sex, education, and time.
The SCDT was not significantly different at 36 months between PDID and PDND groups, so it was not analyzed further. The CCDT was significantly lower in the PDID group at the 18 and 36 month time points. The qualitative analysis of the spontaneous CDT showed that there was no significant difference between the groups at 36 months for the number of spatial/planning errors or stimulus bound errors, despite being significantly different at the two prior assessments. Conceptual errors were significantly more frequent in the PDID group at each time point. There was no consistent statistical difference between the two groups with regard to presence of perseverative errors, size of the clock, and graphic difficulties (table 5). The qualitative analysis of the copy CDT showed that there was no significant difference between the groups at 0 and 36 months for the number of spatial/planning errors or conceptual errors, despite being significant at the 18-month assessment. Perseverative and stimulus bound errors did not occur. There was no consistent statistical difference between the two groups with regard to presence of perseverative errors, size of the clock, and graphic difficulties (table 6). The MMSE, FAB, and DRS were statistically significantly lower in the PDID group at all time points (table 4., figure 3.).

5.2.2. Mixed effects analysis of the CCDT

In the copy condition the interaction between the group and time was significant (P = 0.040) suggesting a decline in CCDT as time progressed in the PDID group. The H&Y

score was also included in the above model and was associated with a decline in the CCDT (P = 0.001) (Appendix 2).

5.2.3. Mixed effects analysis of the MMSE

Including the variables of sex age, education and H&Y, the results showed that by looking at the interaction between dementia group and time, the decline over time in the PDID group was not statistically significant (P = 0.145). The model showed that the MMSE score was significantly lower in patients for every one-point increase in the H&Y score (P < 0.0001) (Appendix 2).

5.2.4. Mixed effects analysis of the FAB

Including the same confounding variables showed that the interaction between time of assessment and dementia group was statistically significant with more decline in performance over time in the PDID group (P < 0.0001). There was also a decline for every one-year increase in age (P = 0.001). Additionally, for every one-unit increase in the H&Y scale there was a statistically significant decline in FAB (P < 0.014) (Appendix 2).

5.2.5. Mixed effects analysis of the DRS

Analysis was done including the variables of sex, age, education and H&Y. The interaction denoting change over time was significant for PDID group (P < 0.0001). The scores were also lower for every one-unit increase in the H&Y scale (P = 0.009) (Appendix 2).

Table 4. Baseline, 18-month, and 36-month results of demographic and neuropsychological variables for PDID and PDND patients.

	PDID Mean (SD)	PDND Mean (SD)	T-Score, P value
Age at enrollment	73.99(5.1)	69.81(3.7)	0.002
Education (years)	13(2.2)	14.63(3.4)	0.153
Sex	F: 11, M: 9	F: 16, M: 12	0.883*
SCDT 0	6.7 (2.6)	8.5 (1.7)	0.001
SCDT 18	6.4 (2.9)	8.4 (1.6)	0.004
SCDT 36	6.4 (2.7)	7.9 (2.1)	0.059
Copy time 0	8.3 (1.8)	9.1 (1.1)	0.028
Copy time 18	7.2 (1.6)	9 (0.96)	<0.0001
Copy time 36	6.5 (2.4)	8.9 (1)	<0.0001
UPDRS 0	34.29(16.8)	23.61(9.5)	0.056
UPDRS 18	35.1(18.5)	23.1(10.4)	0.006
UPDRS 36	43.33(20.1)	26.07(11.12)	0.001
UPDRS-III 0	18.15(9.55)	15.43(6.76)	0.253
UPDRS-III 18	20.85(10.79)	14.71(8.48)	0.033
UPDRS-III 36	25.89(19.75)	17.22(8.6)	0.008
CCDR at time 0	0.104(0.21)	0.02(0.09)	0.069
CCDR time 6	0.21(0.25)	0(0)	<0.0001
CCDR time 12	0.24(0.31)	0.054(0.16)	0.010
CCDR at time 18	0.48(0.77)	0.054(0.16)	0.007
CCDR time 24	0.34(0.5)	0.18(0.2)	0.143
CCDR time 30	0.61(0.78)	0.16(0.27)	0.007
CCDR at time 36	1.02(0.88)	0.185(0.25)	<0.0001
MMSE at time 0	27.08(1.77)	28.75(1.35)	0.001
MMSE at time 18	25.25(4.32)	28(1.28)	0.003

MMSE at time 36	24.67(4.75)	27.78(1.5)	0.003
FAB at time 0	13.5(2.1)	15.18(1.9)	0.014
FAB at time 18	12.8(2.5)	15.18(2.3)	0.002
FAB at time 36	11(3.12)	15.56(1.83)	<0.0001
DRS at time 0	134.13(4.8)	138.96(3.8)	0.002
DRS at time 18	131.1(12.25)	138.68(2.9)	0.003
DRS at time 36	129.1(12.43)	139.52(3.3)	<0.0001
Hoehn-Yahr 0	2.52(0.62)	2.02(0.6)	0.016
Hoehn-Yahr 18	2.73(0.95)	1.8(0.6)	<0.0001
Hoehn-Yahr 36	3(1.16)	2.13(0.55)	0.002
PD Duration at enrollment	9.6(4.42)	8.02(4.33)	0.126
GDS 0	2.55(3.14)	1.39(1.55)	0.14
GDS 6	2.4(2.3)	1.54(1.57)	0.16
GDS 12	2.6(2.7)	1.7(2)	0.226
GDS 18	3(2.8)	1.8(2)	0.1
GDS 24	2.05(2.2)	1.82(1.6)	0.7
GDS 30	2.8(2.8)	1.6(1.9)	0.11
GDS 36	2.7(2.5)	2(2)	0.33
HIS 0	0.75(1.2)	0.71(0.98)	0.9
HIS 18	1.55(1.9)	1.07(1.7)	0.38
HIS 36	1.44(1.6)	1.37(1.5)	0.9
CIRS 0	19.9(2.97)	18.93(2.7)	0.25
CIRS 18	21.65(3.3)	19.5(2.5)	0.022
CIRS 36	21.6(4.6)	19.5(2.8)	0.09

*Indicates χ^2 values.

Table 5. The *spontaneous* condition of the clock drawing test in PDID and PDND, qualitative assessment of its components

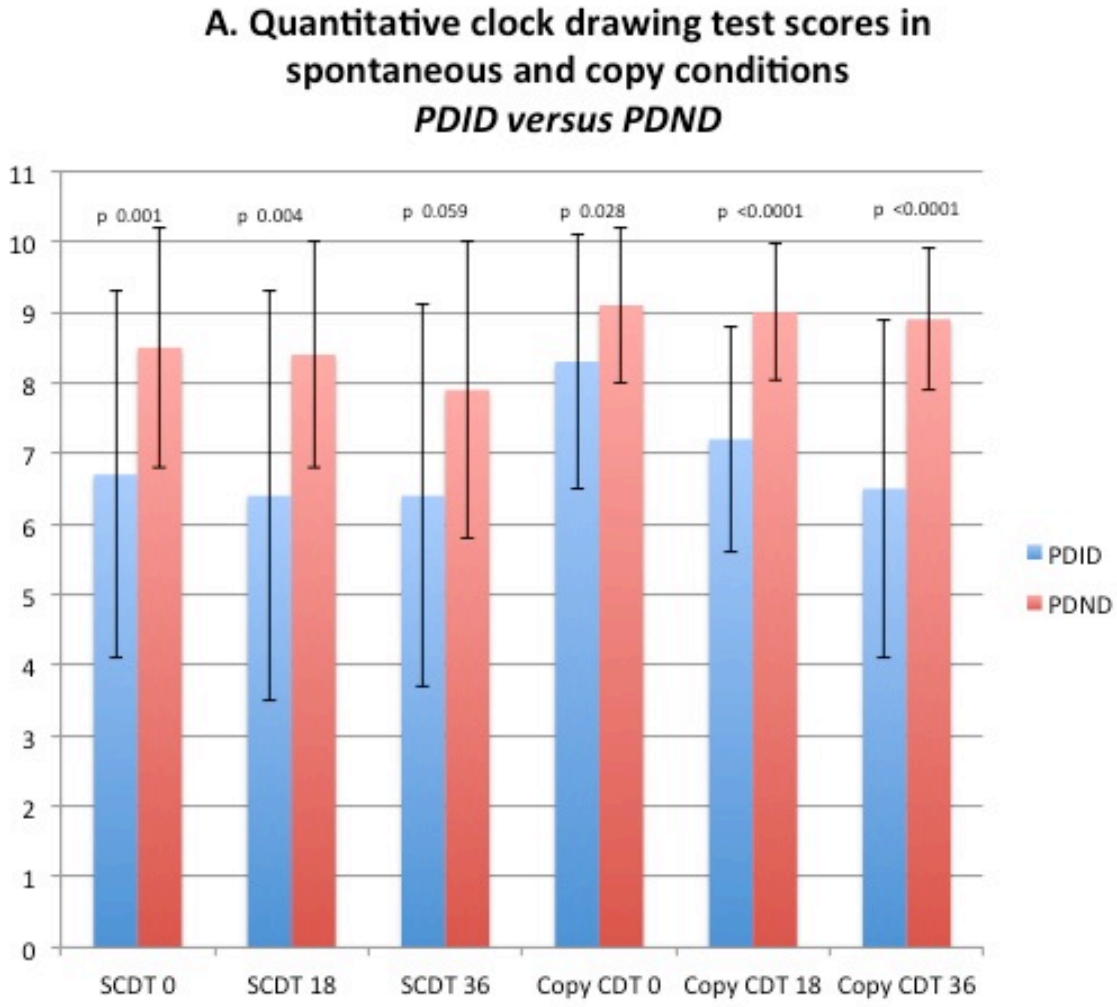
	PDID (%)	PDND (%)	P Value
Presence of spatial/planning errors			
0	15 (75)	13 (46.43)	0.048
18	16 (80)	11 (39.29)	0.005
36	13 (72.22)	14 (51.85)	0.172
Presence of conceptual errors			
0	11 (55)	2 (7.14)	<0.0001
18	8 (40)	3 (10.71)	0.017
36	8 (44)	4 (14.81)	0.028
Presence of stimulus bound errors			
0	5 (25)	1 (3.57)	0.027
18	4 (20)	1 (3.57)	0.066
36	4 (22.22)	3 (11.11)	0.314
Presence of perseverative errors			
0	1 (5)	0	0.232
18	1 (5)	0	0.232
36	1 (5.56)	0	0.215
Presence of normal size clock			
0	17 (85)	25 (89.29)	0.658
18	17 (85)	23 (82.14)	0.793
36	13 (72.22)	22 (81.48)	0.464
Absence of graphic difficulties			
0	4 (20)	13 (46.43)	0.059
18	2 (10)	14 (50)	0.004
36	9 (50)	13 (48.15)	0.088

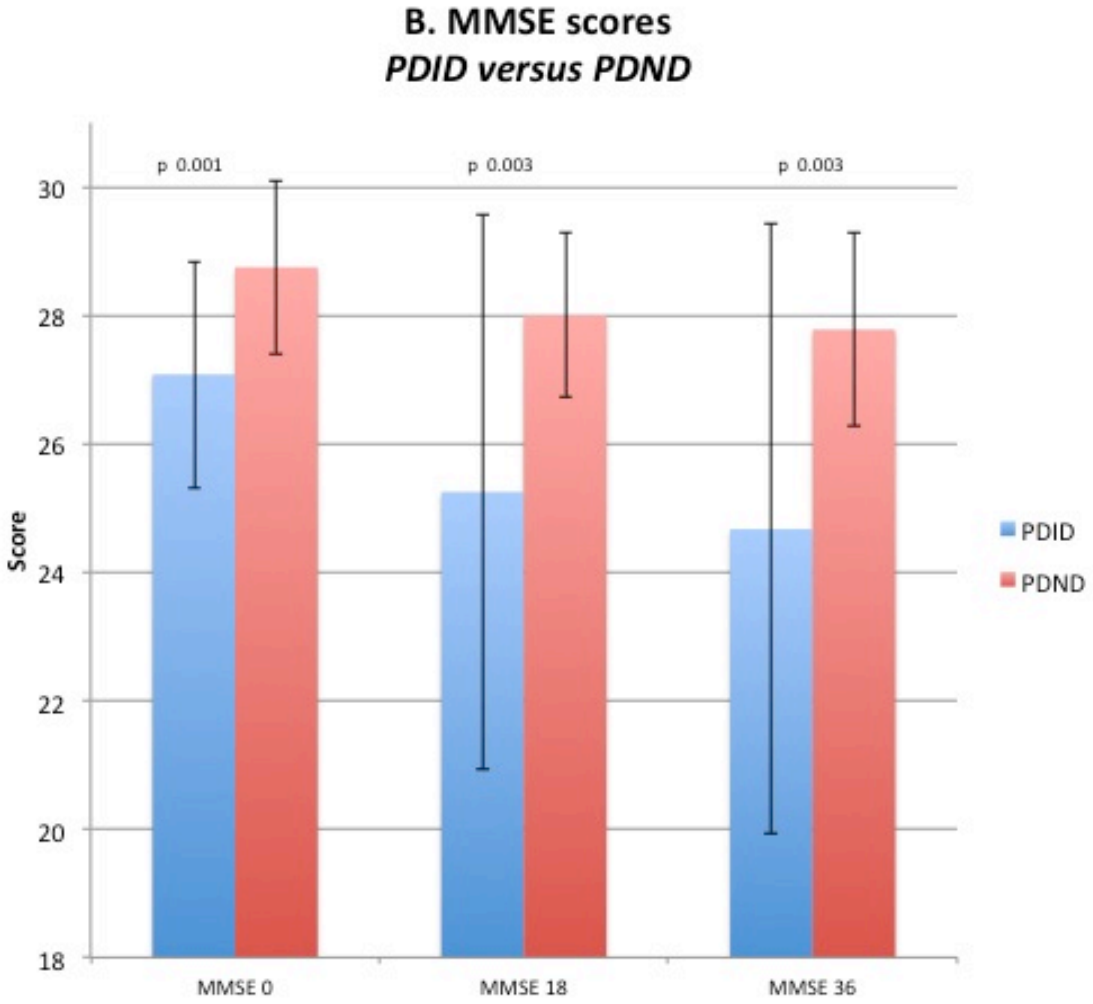
Table 6. The *copy* condition of the clock drawing test in PDID and PDND, qualitative assessment of its components

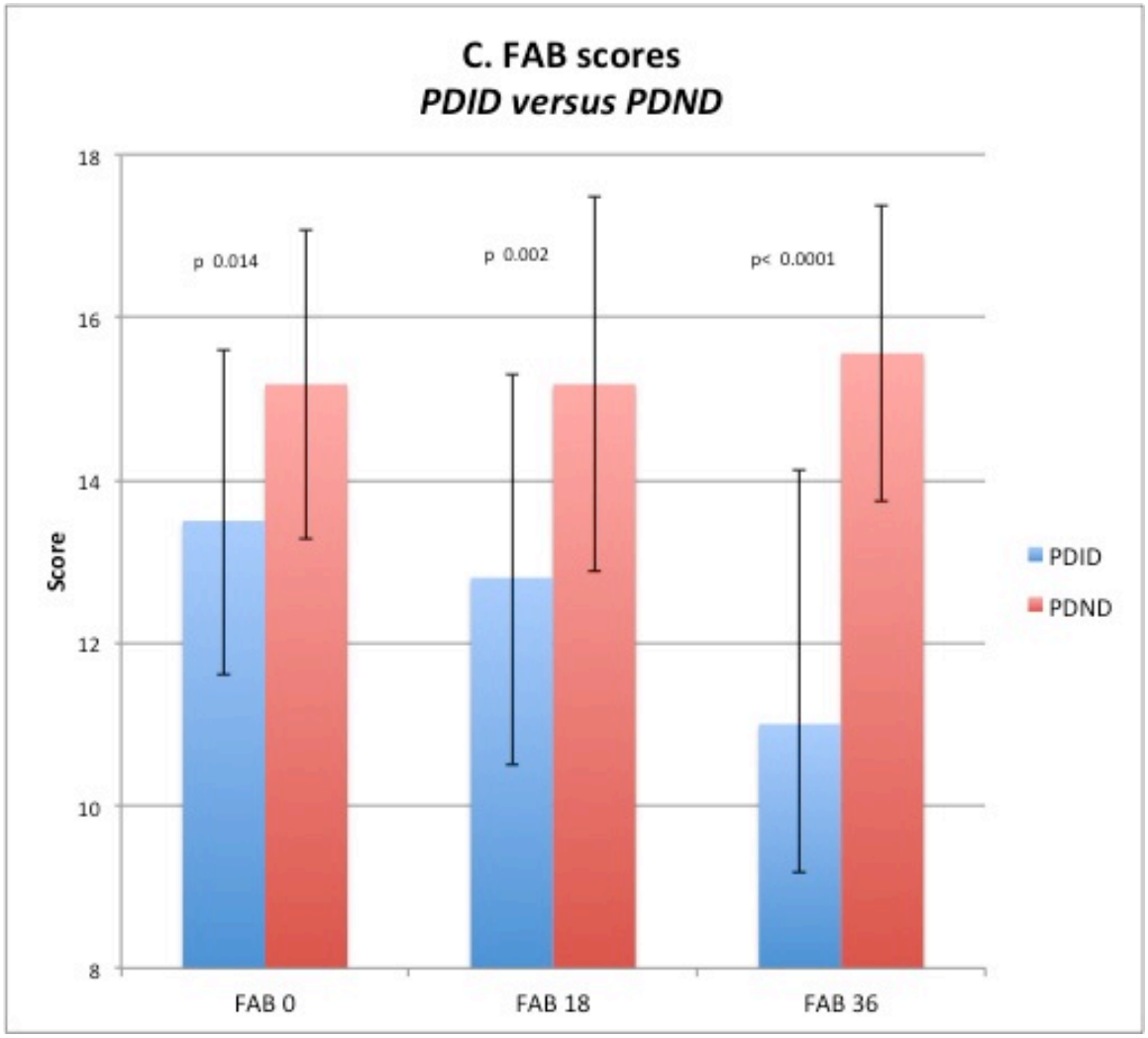
	PDID (%)	PDND (%)	P Value
Presence of spatial/planning errors			
0	12 (63.16)	8 (28.57)	0.019
18	14 (73.68)	13 (46.43)	0.064
36	8 (47.06)	14 (53.85)	0.663
Presence of conceptual errors			
0	2 (10.53)	0	0.079
18	3 (15.79)	0	0.030
36	0	2 (7.69)	0.242
Presence of stimulus bound errors			
0	0	0	NA
18	0	0	NA
36	0	0	NA
Presence of perseverative errors			
0	0	0	NA
18	0	0	NA
36	0	0	NA
Presence of normal size clock			
0	17 (89.47)	28 (100)	0.079
18	18 (94.74)	28 (100)	0.22
36	17 (100)	26 (100)	NA
Absence of graphic difficulties			
0	6 (31.58)	17 (60.71)	0.0499
18	4 (21.05)	10 (35.71)	0.281
36	11 (64.71)	14 (53.85)	0.48

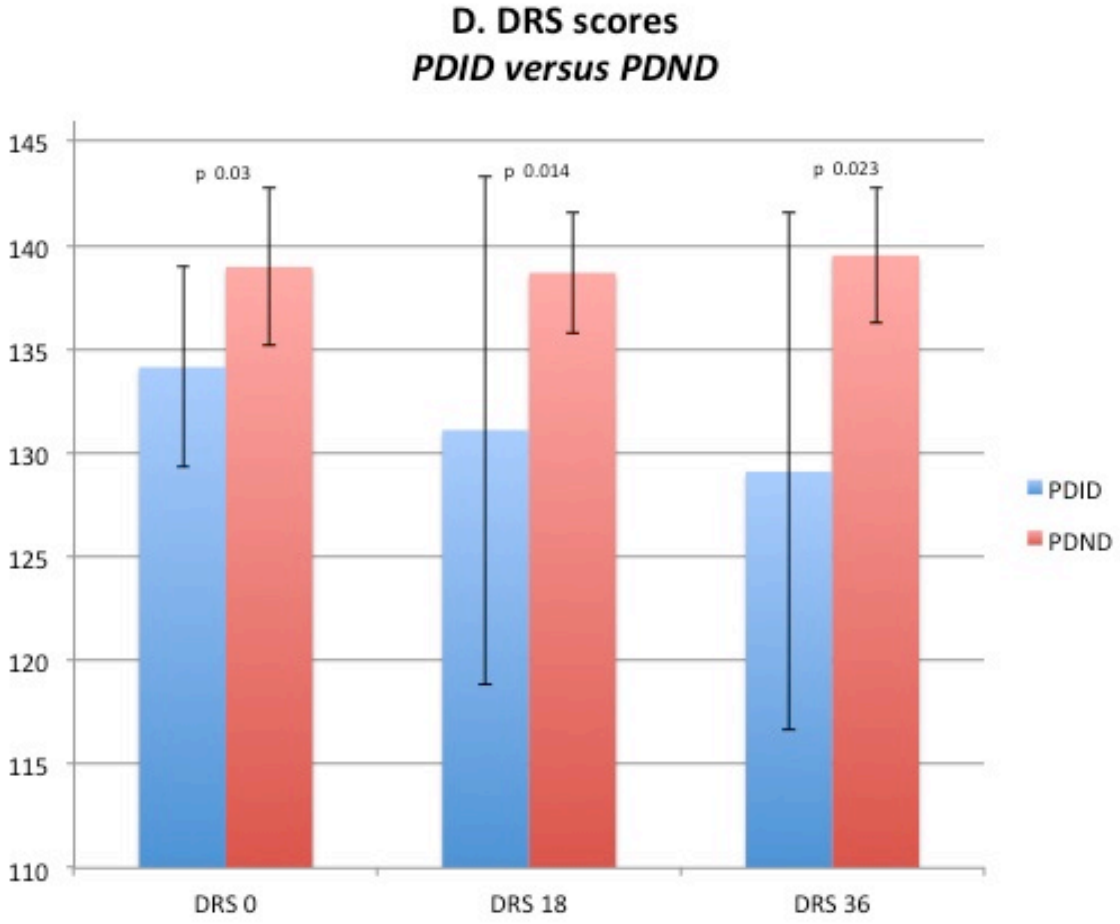
NA: Not applicable

Figure 3. Mean scores of the neuropsychological scales at 0, 18, and 36 months assessments for PDND and PDID patients. A) Quantitative clock drawing test scores in spontaneous and copy conditions. B) Mini-mental State Exam. C) Frontal Assessment Battery. D) Dementia Rating Scale.




B. MMSE scores
PDID versus PDND
p 0.001
p 0.003
p 0.003
30
28
26
24
22
20
18
Score
MMSE 0
MMSE 18
MMSE 36
PDID
PDND


C. FAB scores
PDID versus PDND
18
16
14
12
10
8
Score
p 0.014
p 0.002
p< 0.0001
FAB 0
FAB 18
FAB 36
PDID
PDND


D. DRS scores
PDID versus PDND
145
140
135
130
125
120
115
110
p 0.03
p 0.014
p 0.023
DRS 0
DRS 18
DRS 36
PDID
PDND

5.3. Correlation analysis

The results of the baseline, 18-month, and 36-month assessments were correlated with one another if the results showed significant change over time.

Given that only the FAB and DRS showed decline over time that was statistically significant in the PD and control groups, a correlation analysis was done between the two tests at each assessment and showed significant positive correlations (Table 7) & (Figure 4). A correlation analysis was not performed for the remaining tests that were not significantly different with controls over time.

In the PD subgroup, correlation was done among the tests that showed statistically significant change over time: CCDT, FAB, & DRS (Table 7) (Figures 5-11). The FAB correlation with the DRS was stronger in the subgroup comparisons $r = 0.61$ ($P < 0.0001$) at each time point. There were also positive and significant correlations between the CCDT and FAB and the CCDT and DRS at each time point (Table 7). Significance was not consistently maintained at 18 months assessment when identified outliers were excluded from the analysis (Table 7) & (Figures 4-11).

Table 7. Correlation of tests that showed significant changes over time.

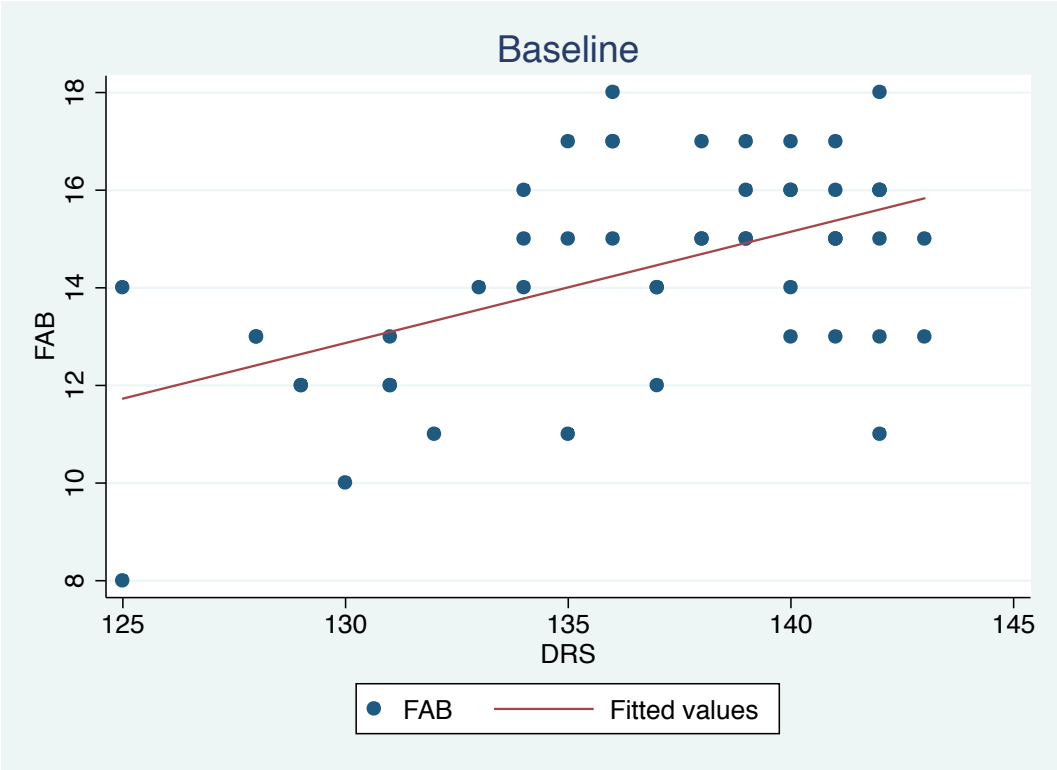
		Baseline (r, P Value)	18 months (r, P Value)	18 months (Outliers excluded) (r, P Value)	36 months (r, P Value)
Control	FAB & DRS	r = 0.39, P = 0.006	r = 0.33, P < 0.0024	N.A.	r = 0.48, P = 0.0007
PD	FAB & DRS	r = 0.52, P = 0.0001	r = 0.60, P < 0.0001	r = 0.60, P < 0.0001	r = 0.69, P < 0.0001
PDID	FAB & DRS	r = 0.47, P = 0.042	r = 0.54, P = 0.018	r = 0.5, P = 0.034	r = 0.65, P = 0.0082
	CCDT & FAB	r = 0.55, P = 0.015	r = 0.43, P = 0.06	r = 0.26, P = 0.3	r = 0.39, P = 0.113
	CCDT & DRS	r = 0.47, P = 0.042	r = 0.54, P = 0.018	r = 0.09, P = 0.74	r = 0.65, P = 0.0082
PDND	FAB & DRS	r = 0.52, P = 0.004	r = 0.47, P = 0.012	r = 0.133, P = 0.5	r = -0.04, P = 0.855
	CCDT & FAB	r = 0.68, P = 0.0001	r = 0.29, P = 0.13	r = 0.044, P = 0.83	r = 0.045, P = 0.825
	CCDT & DRS	r = 0.32, P = 0.098	r = 0.1, P = 0.625	r = 0.21, P = 0.3*	r = -0.14, P = 0.49

N.A.: Not applicable.

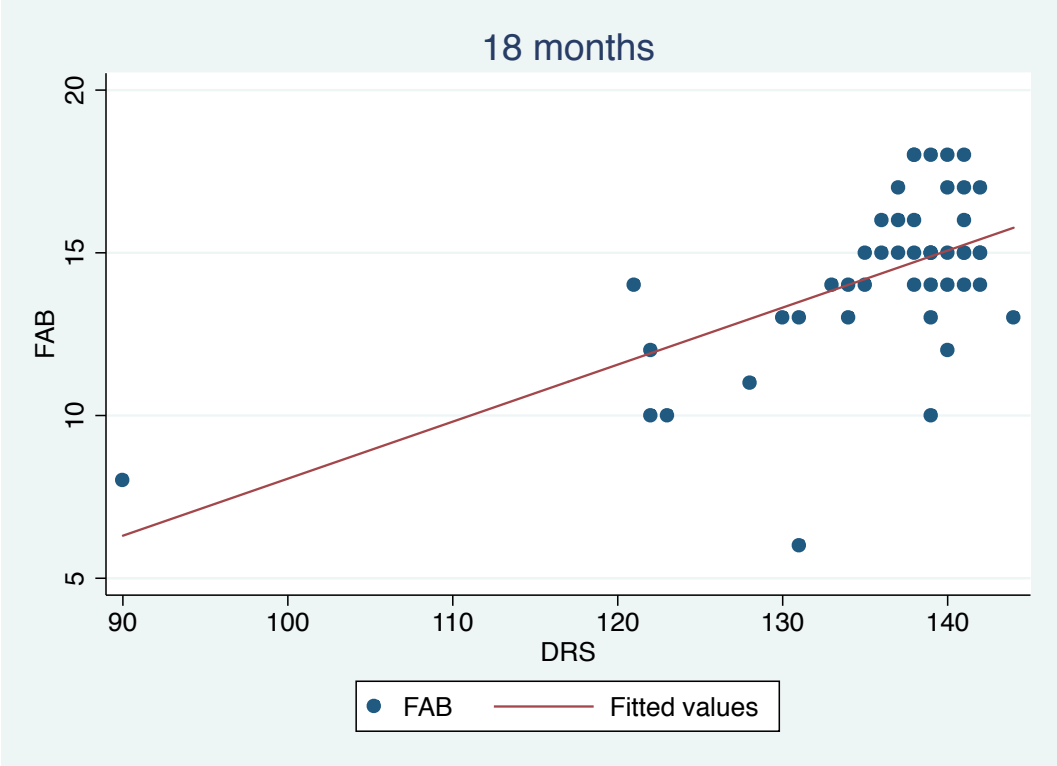
* Two outliers identified and excluded.

Figure 4. Entire PD group: Correlation of FAB with DRS at A) 0 months, B) 18 months, B1) 18 months excluding outlier, and C) 36 months assessments.

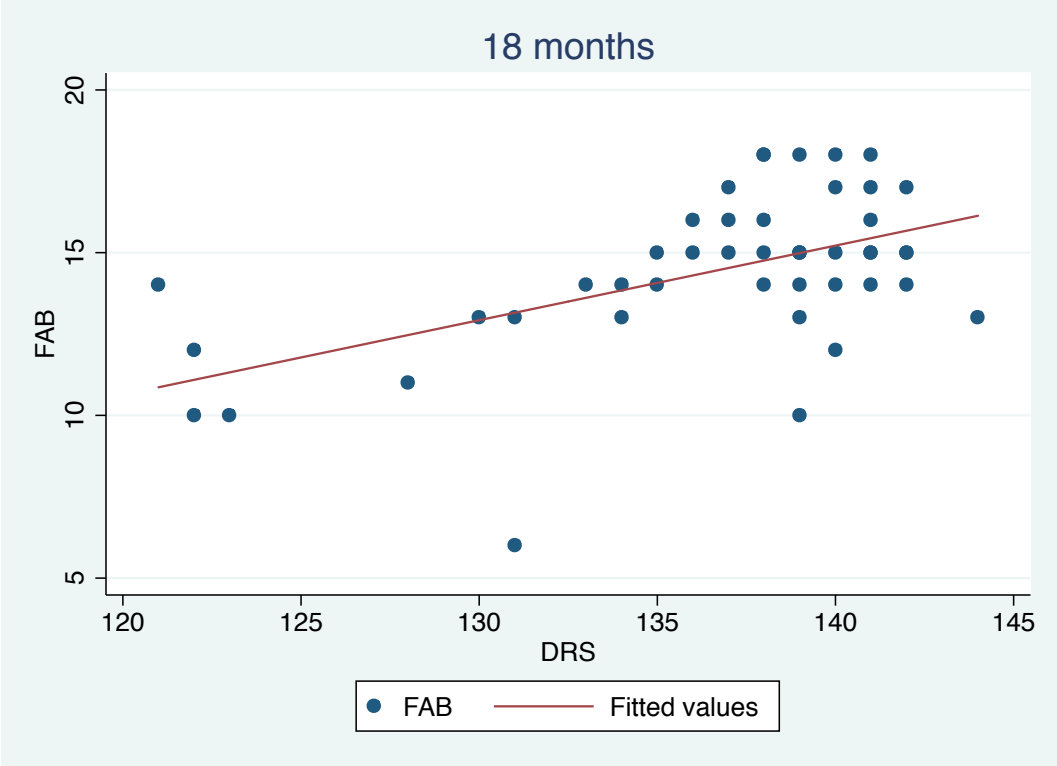
A)



B)


18 months
FAB
DRS
5
10
15
20
90
100
110
120
130
140
FAB
Fitted values

B1)


18 months
FAB
DRS
5
10
15
20
120
125
130
135
140
145
FAB
Fitted values

C)

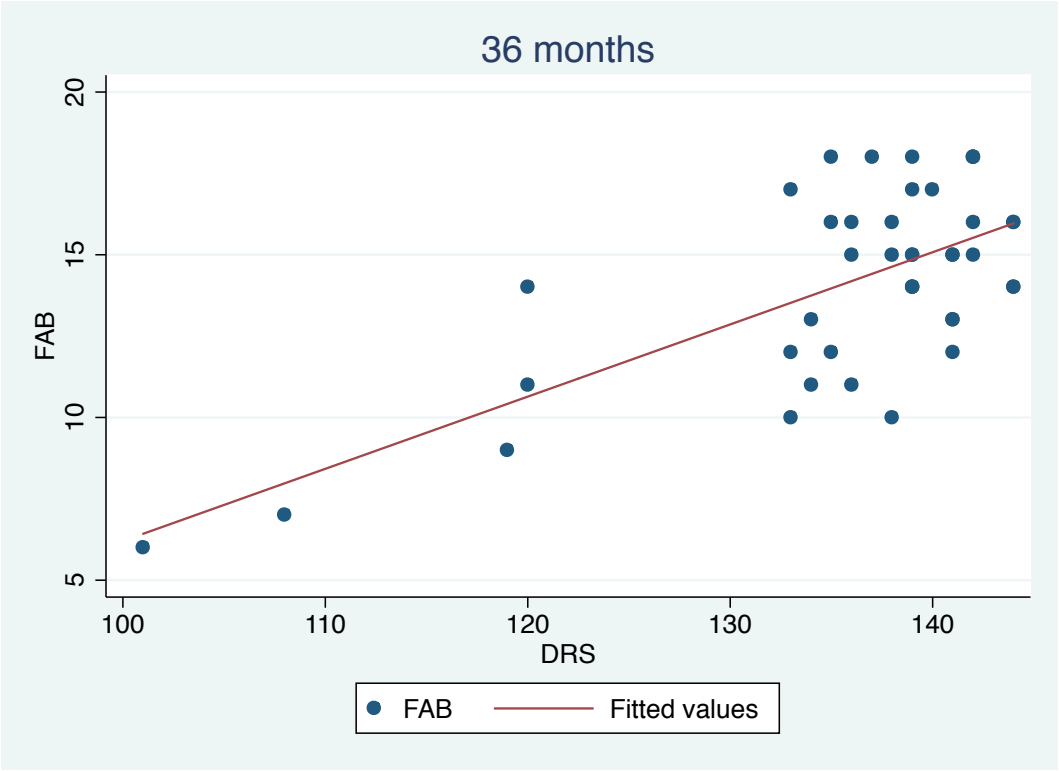
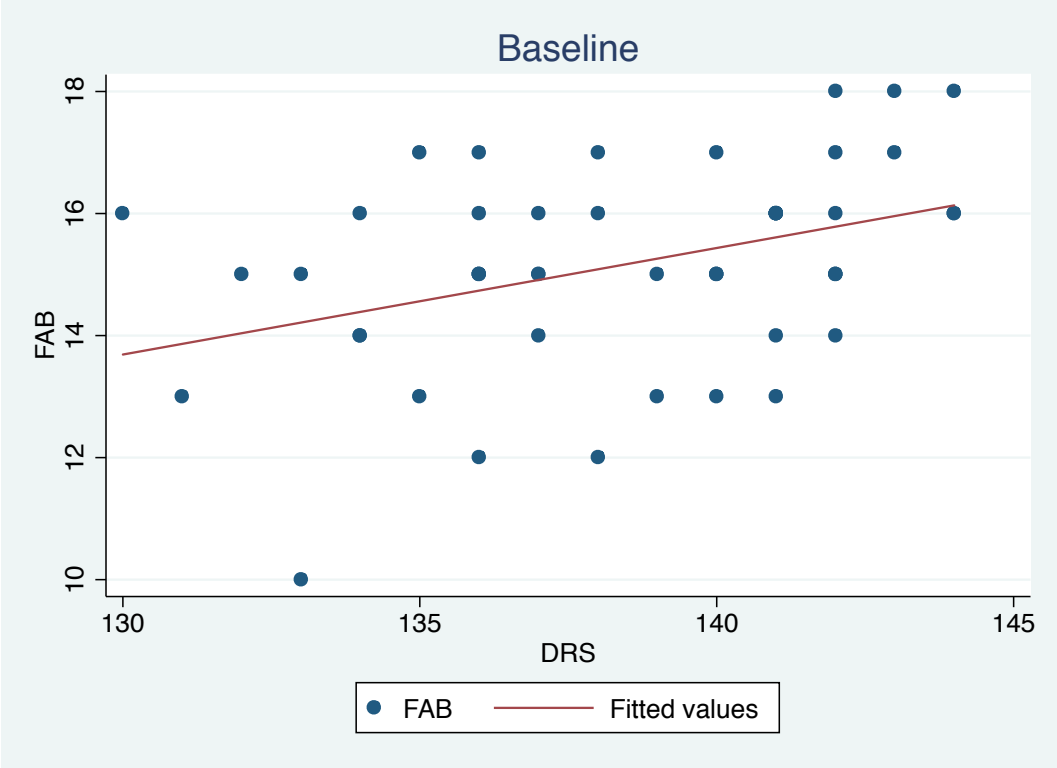
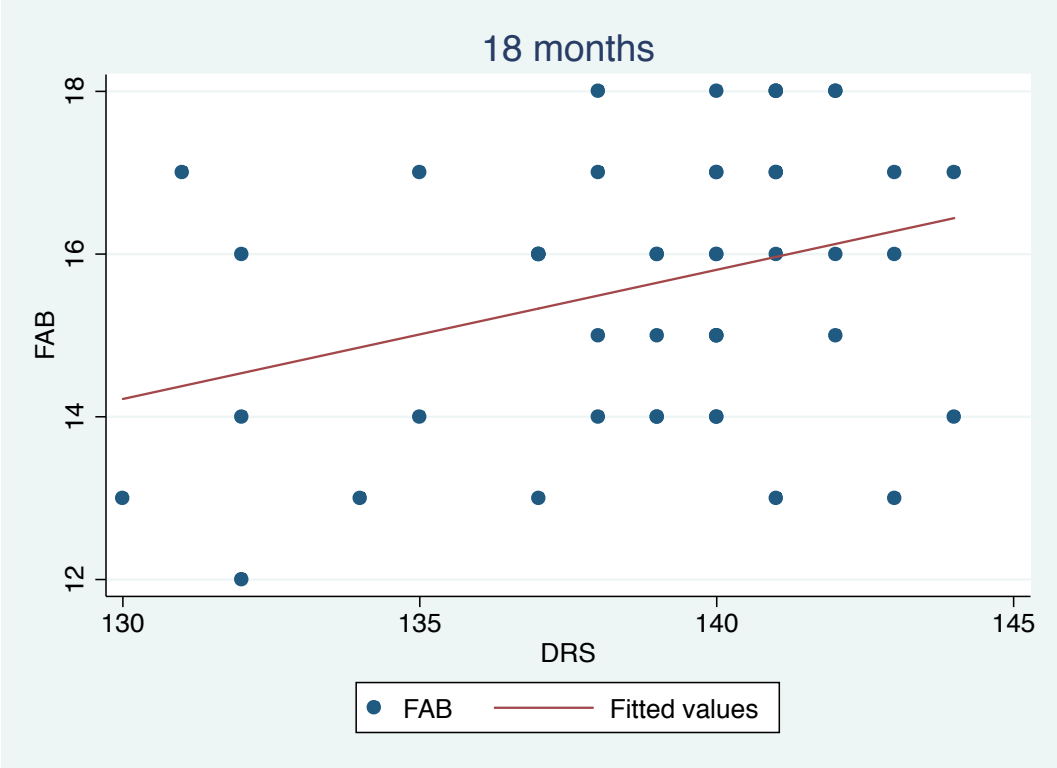

36 months
FAB
5
10
15
20
100
110
120
130
140
DRS
FAB
Fitted values

Figure 5. Control group: Correlation of FAB with DRS at A) 0 months, B) 18 months, and C) 36 months assessments.

A)



B)


18 months
FAB
12
14
16
18
130
135
140
145
DRS
FAB
Fitted values

C)

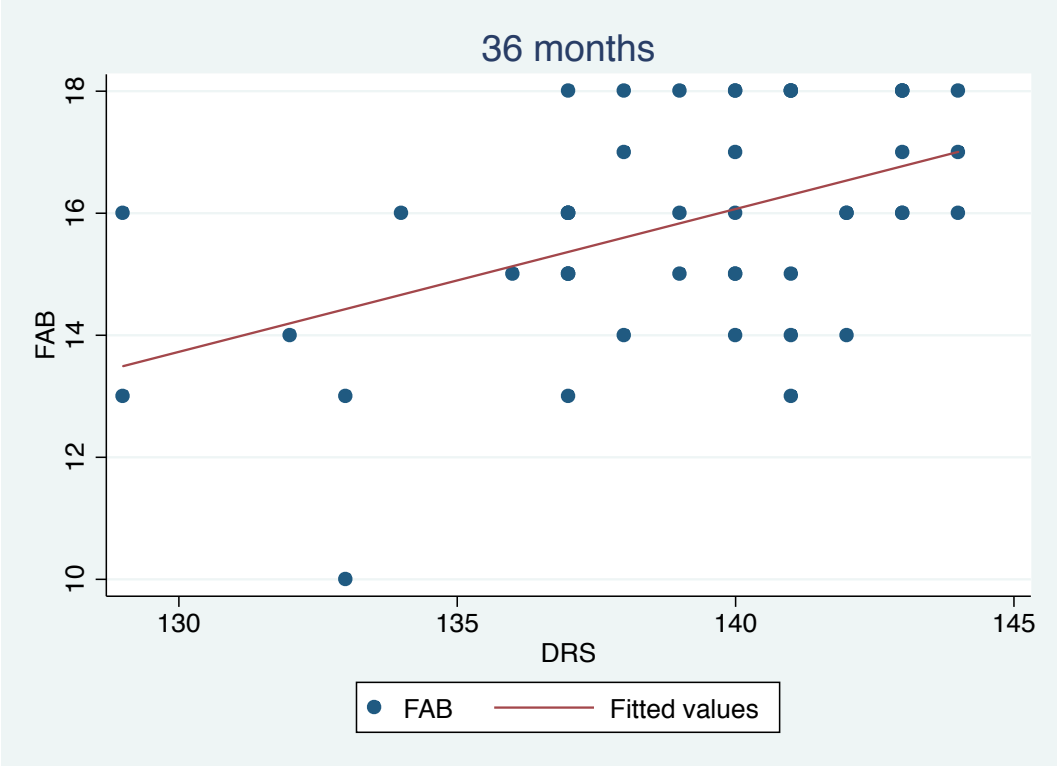
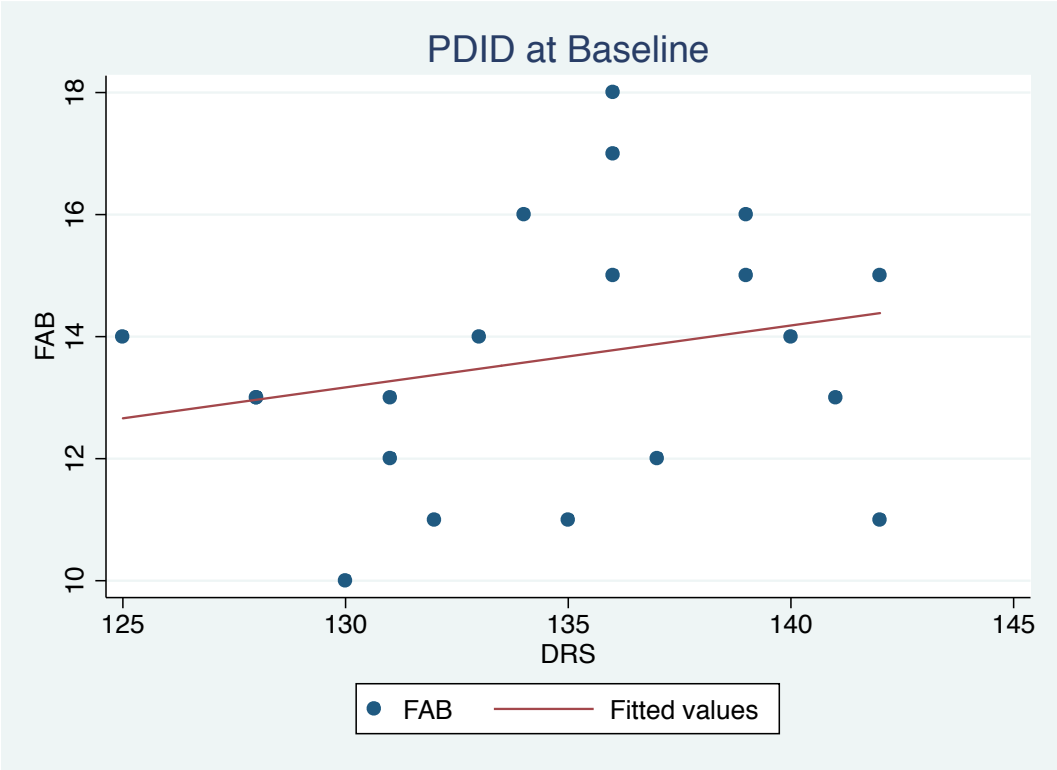
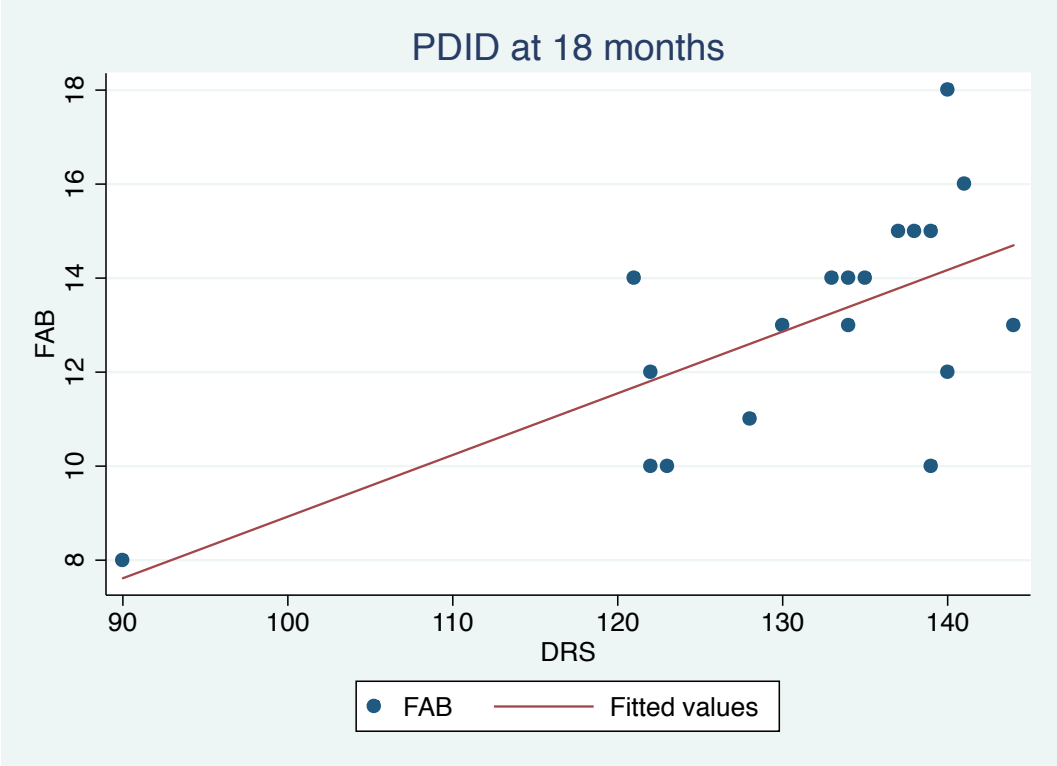

36 months
FAB
10
12
14
16
18
130
135
140
145
DRS
FAB
Fitted values

Figure 6. PDID group: Correlation of FAB with DRS at A) 0 months, B) 18 months, B1) 18 months excluding outliers, and C) 36 months assessments.

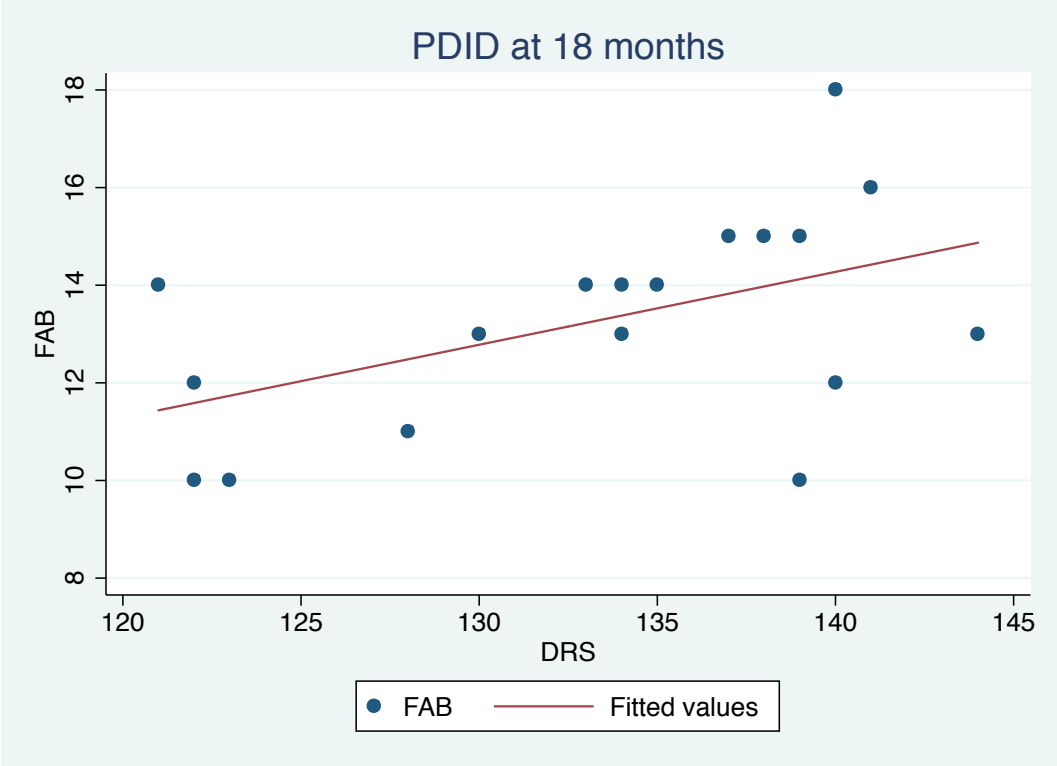
A)



B)


PDID at 18 months
FAB
8
10
12
14
16
18
90
100
110
120
130
140
DRS
FAB
Fitted values

B1)


PDID at 18 months
FAB
8
10
12
14
16
18
120
125
130
135
140
145
DRS
FAB
Fitted values

C)

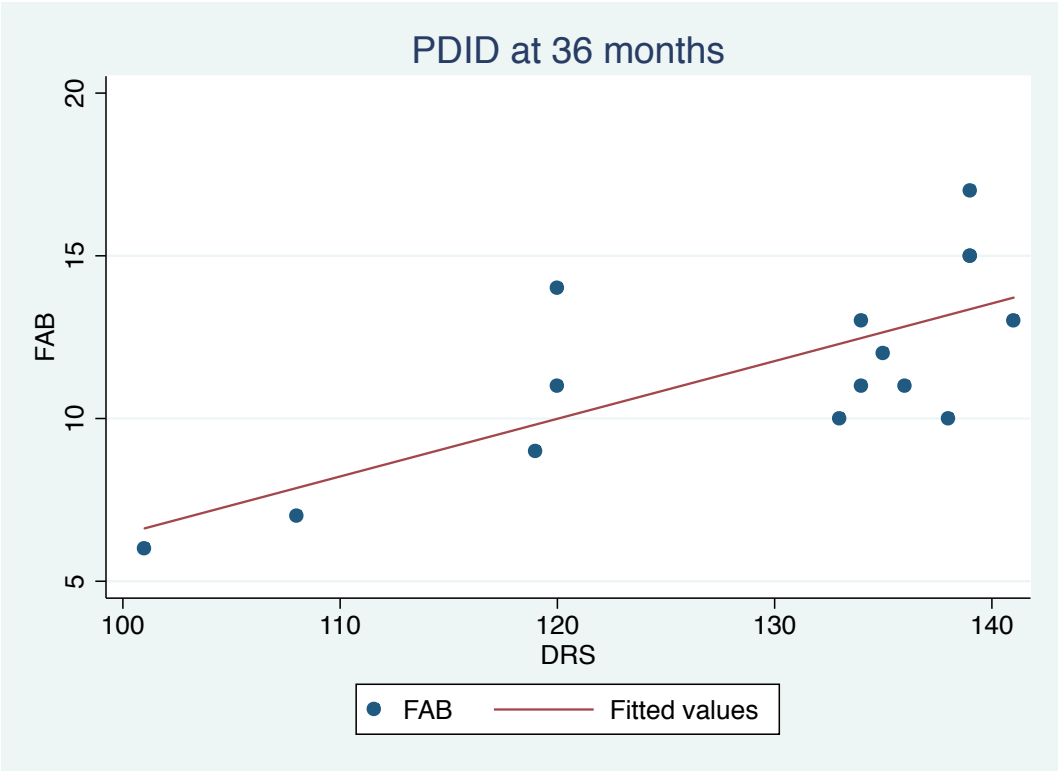
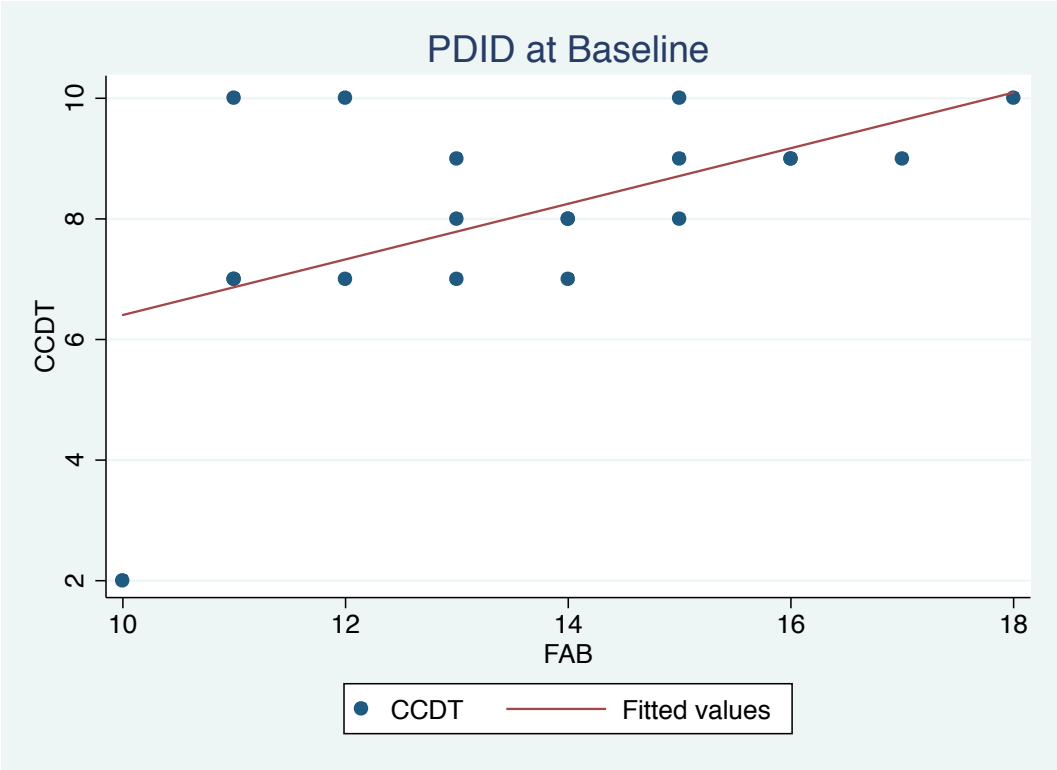
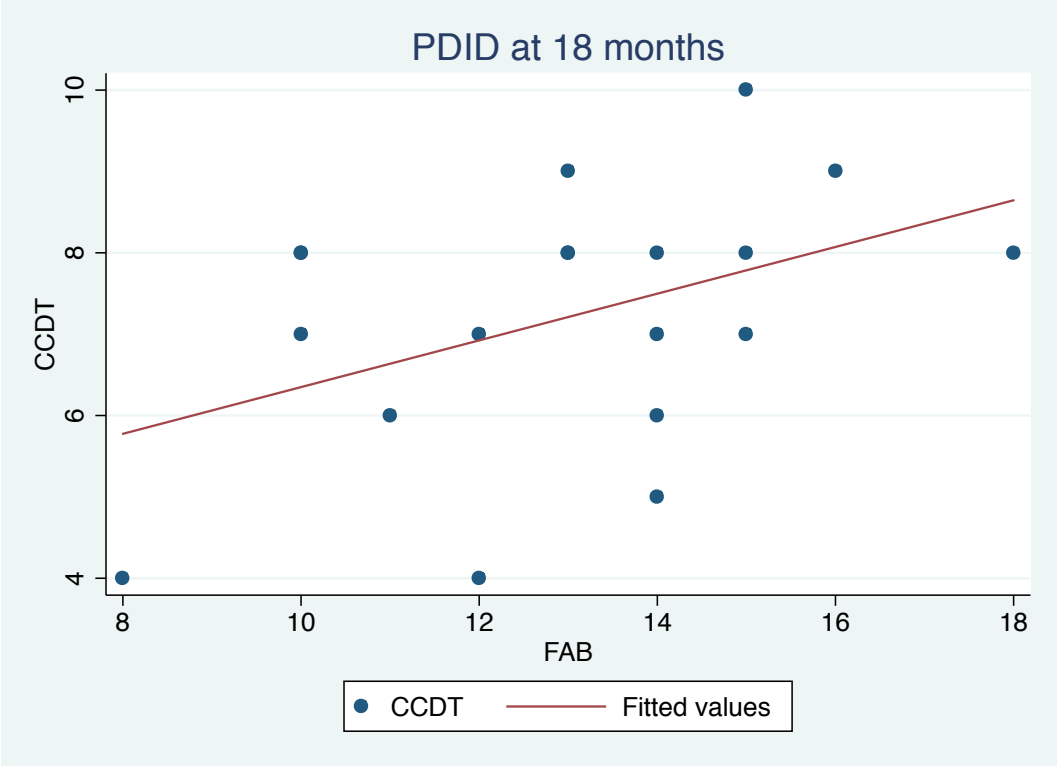

PDID at 36 months
FAB
5
10
15
20
100
110
120
130
140
DRS
FAB
Fitted values

Figure 7. PDID group: Correlation of CCDT with FAB at A) 0 months, B) 18 months, B1) 18 months excluding outliers, and C) 36 months assessments.

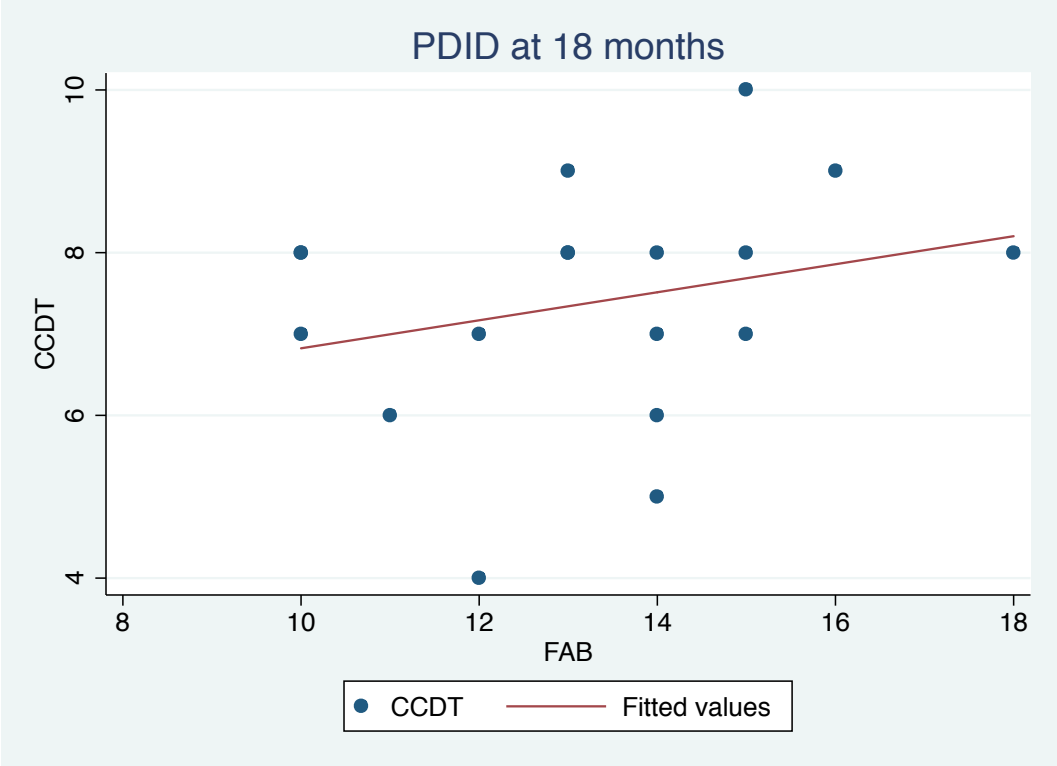
A)



B)


PDID at 18 months
CCDT
FAB
4
6
8
10
8
10
12
14
16
18
CCDT
Fitted values

B1)


PDID at 18 months
CCDT
FAB
4
6
8
10
8
10
12
14
16
18
CCDT
Fitted values

C)

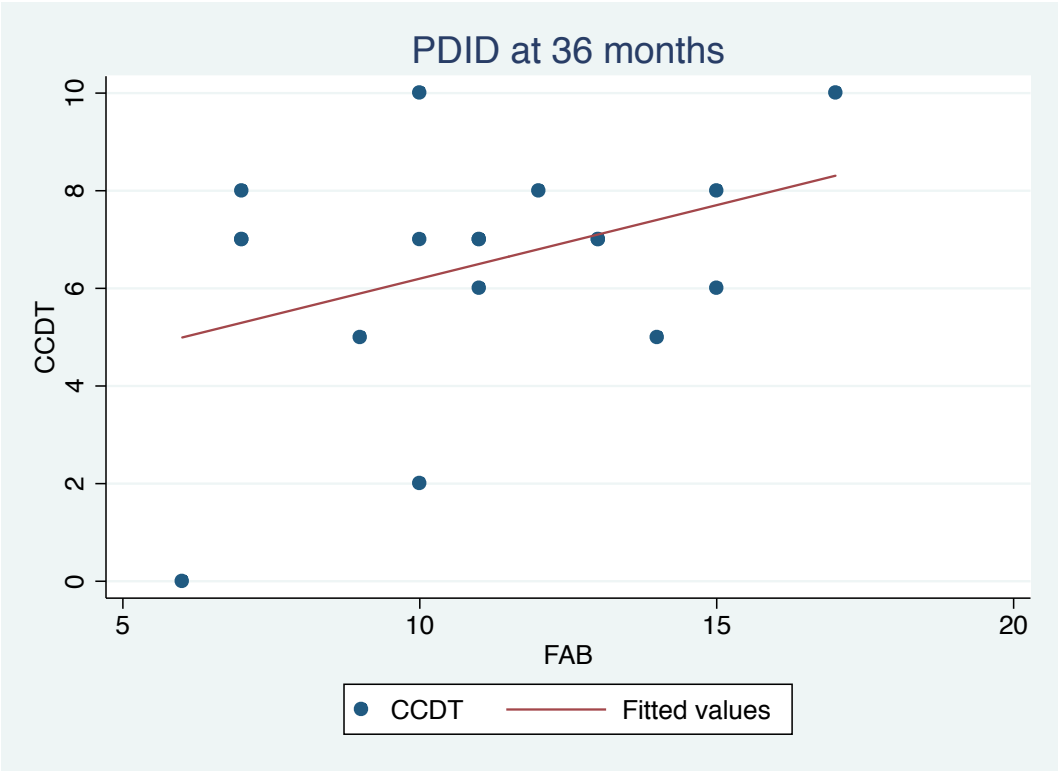
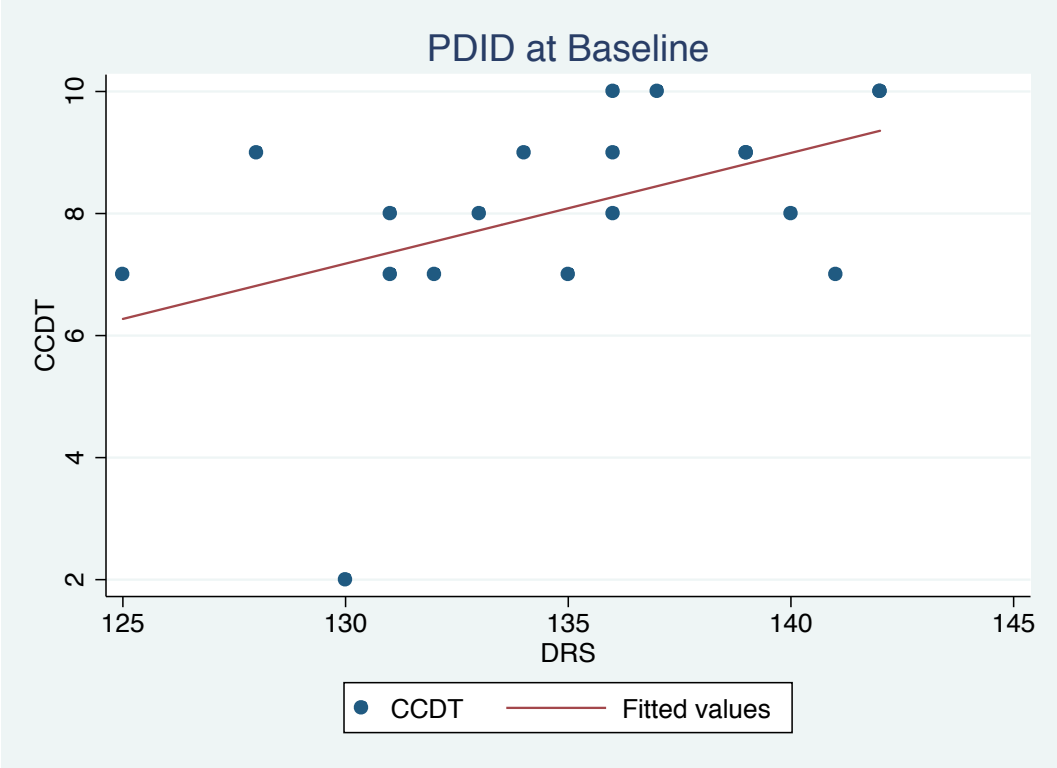
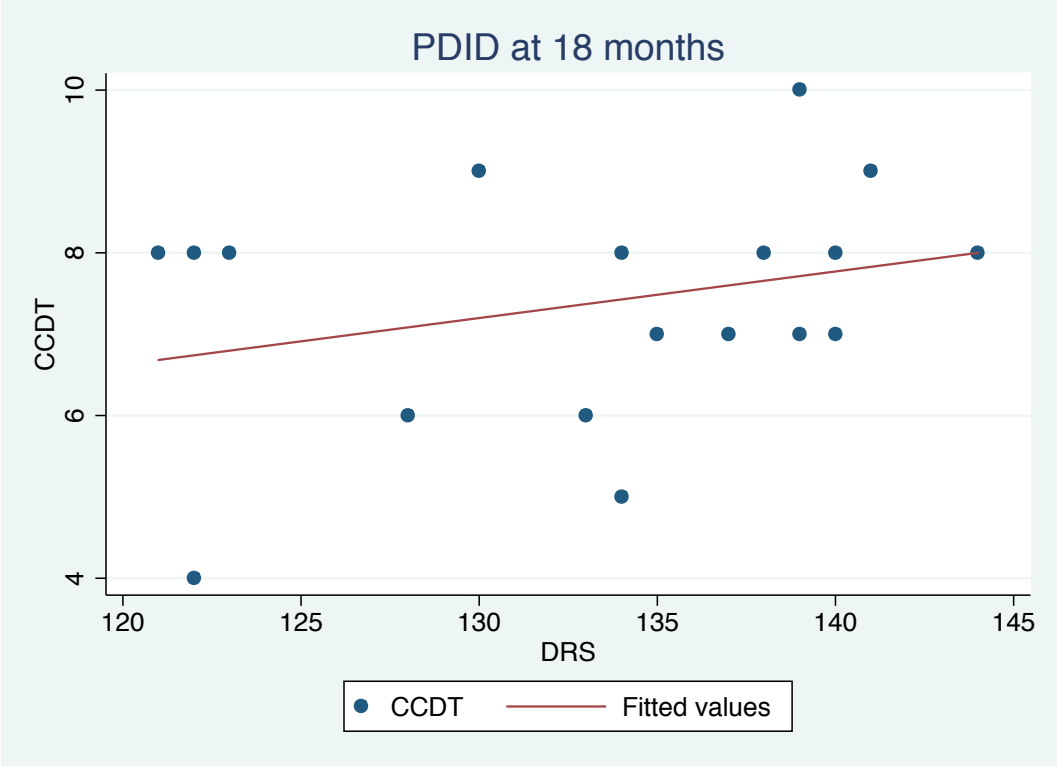

PDID at 36 months
CCDT
0
2
4
6
8
10
5
10
15
20
FAB
CCDT
Fitted values

Figure 8. PDID group: Correlation of CCDT with DRS at A) 0 months, B) 18 months, B1) 18 months excluding outliers, and C) 36 months assessments.

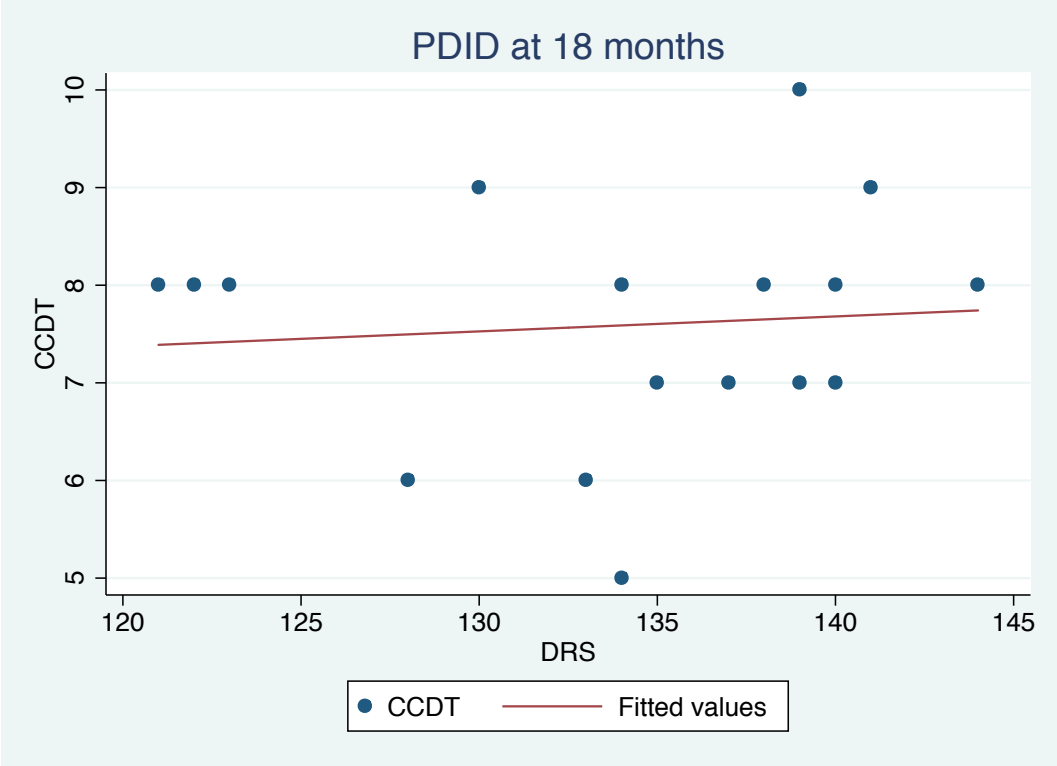
A)



B)


PDID at 18 months
CCDT
DRS
4
6
8
10
120
125
130
135
140
145
CCDT
Fitted values

B1)


PDID at 18 months
CCDT
DRS
5
6
7
8
9
10
120
125
130
135
140
145
CCDT
Fitted values

C)

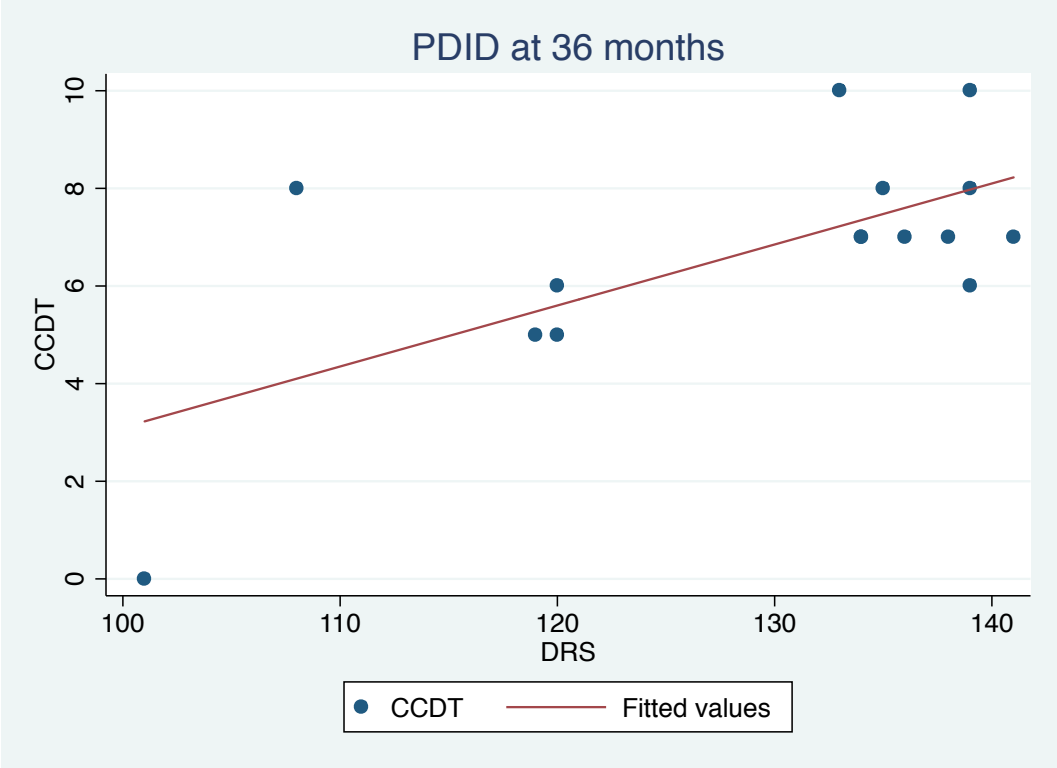
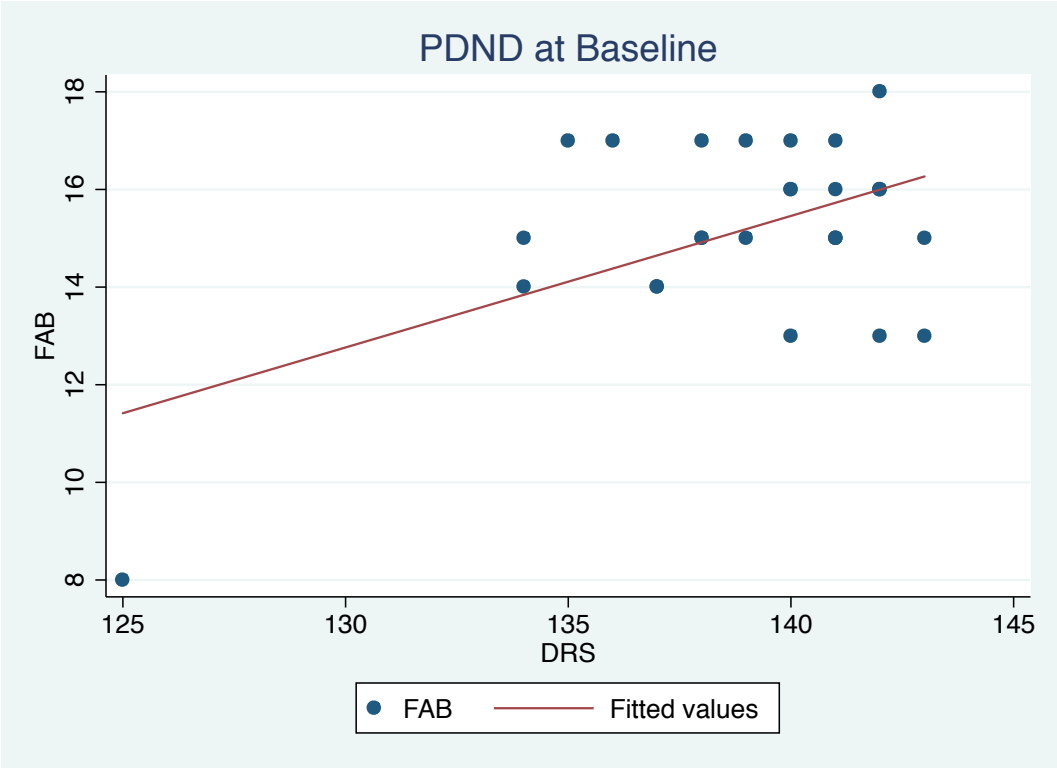
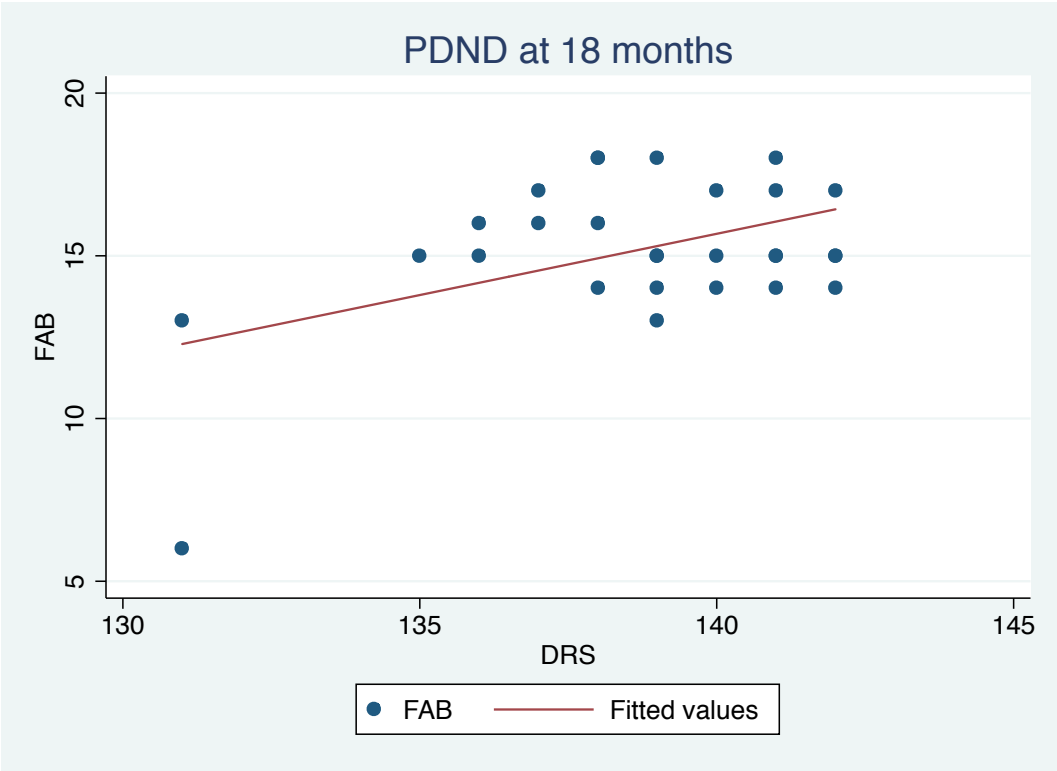

PDID at 36 months
CCDT
0
2
4
6
8
10
100
110
120
130
140
DRS
CCDT
Fitted values

Figure 9. PDND group: Correlation of FAB with DRS at A) 0 months, B) 18 months, B1) 18 months excluding outliers, and C) 36 months assessments.

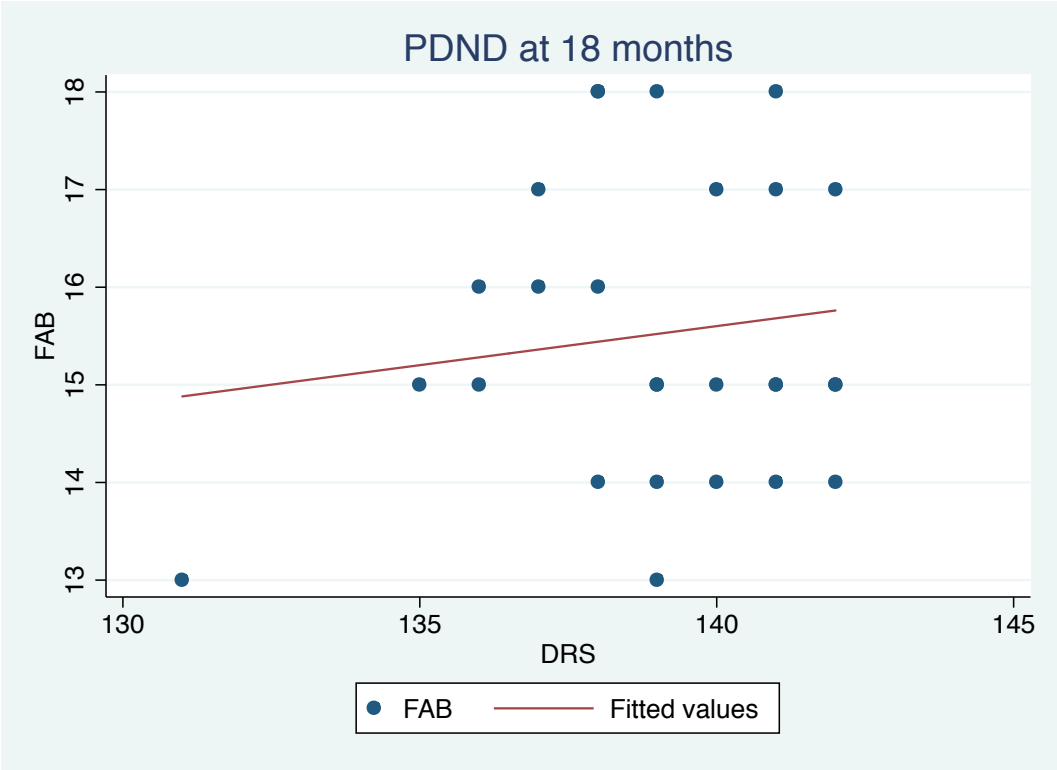
A)



B)


PDND at 18 months
FAB
20
15
10
5
130
135
140
145
DRS
FAB
Fitted values

B1)


PDND at 18 months
FAB
18
17
16
15
14
13
130
135
140
145
DRS
FAB
Fitted values

C)

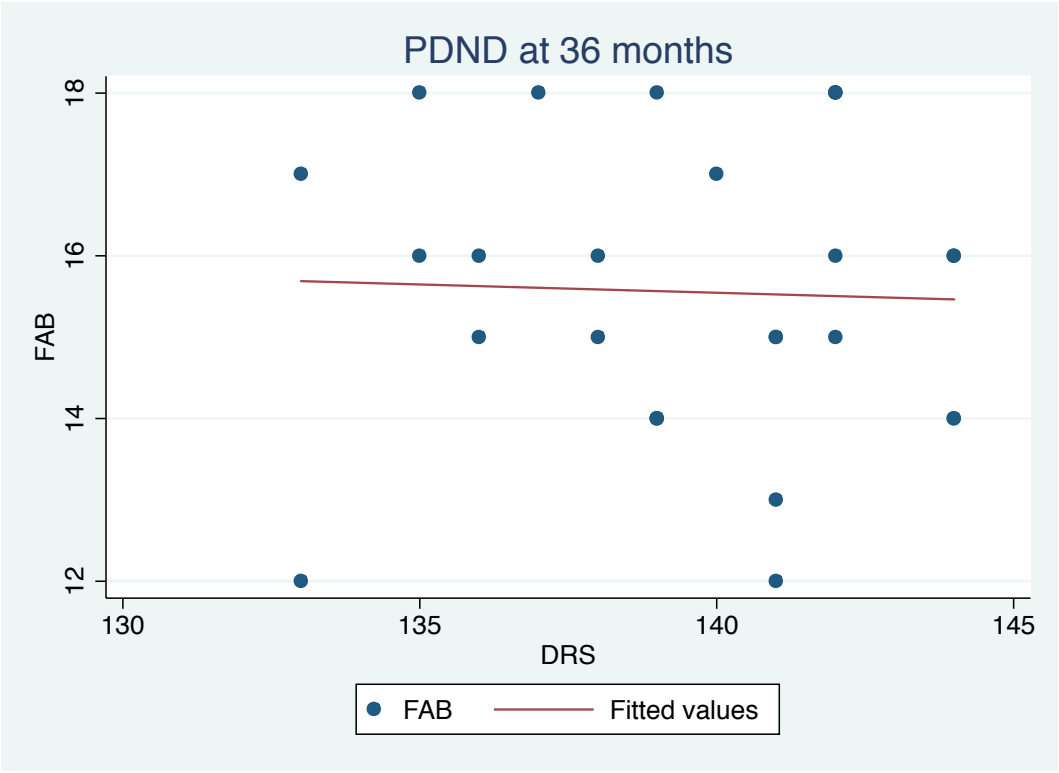
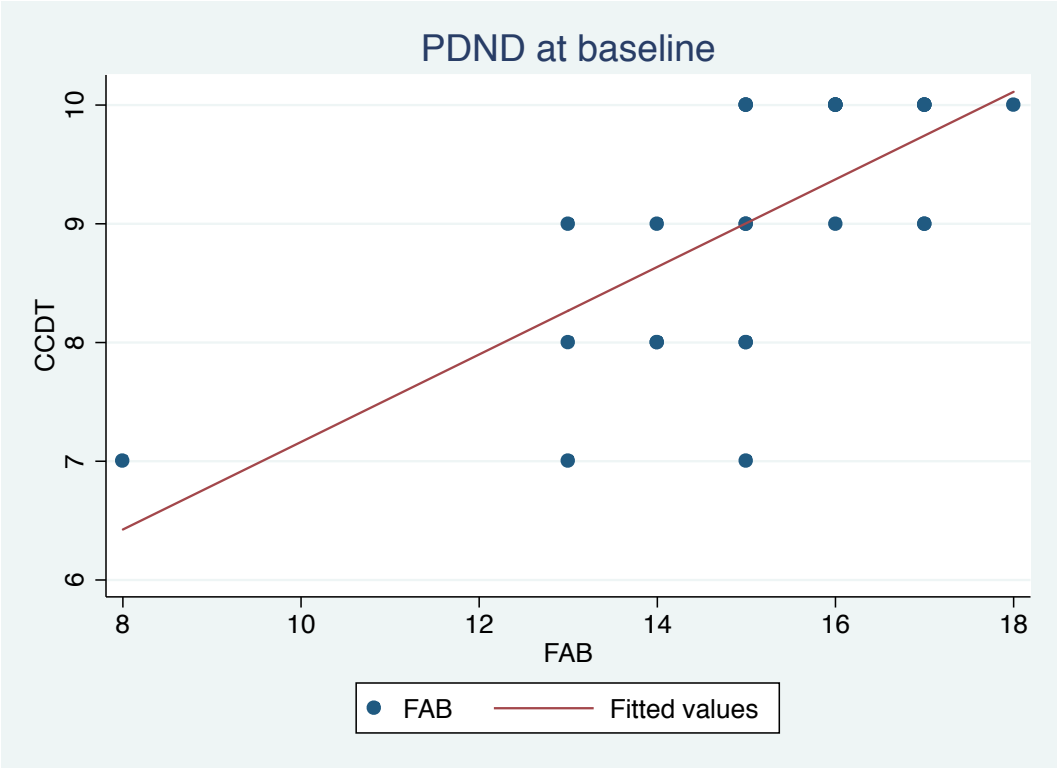
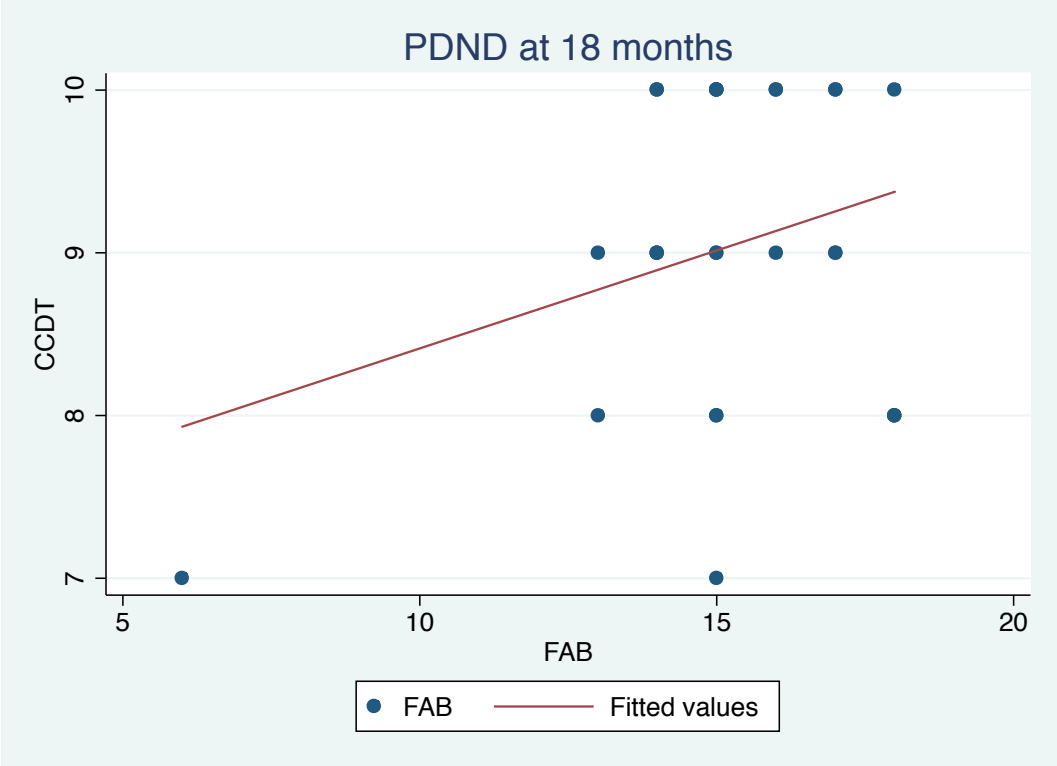

PDND at 36 months
FAB
18
16
14
12
130
135
140
145
DRS
FAB
Fitted values

Figure 10. PDND group: Correlation of CCDT with FAB at A) 0 months, B) 18 months, B1) 18 months excluding outliers, and C) 36 months assessments.

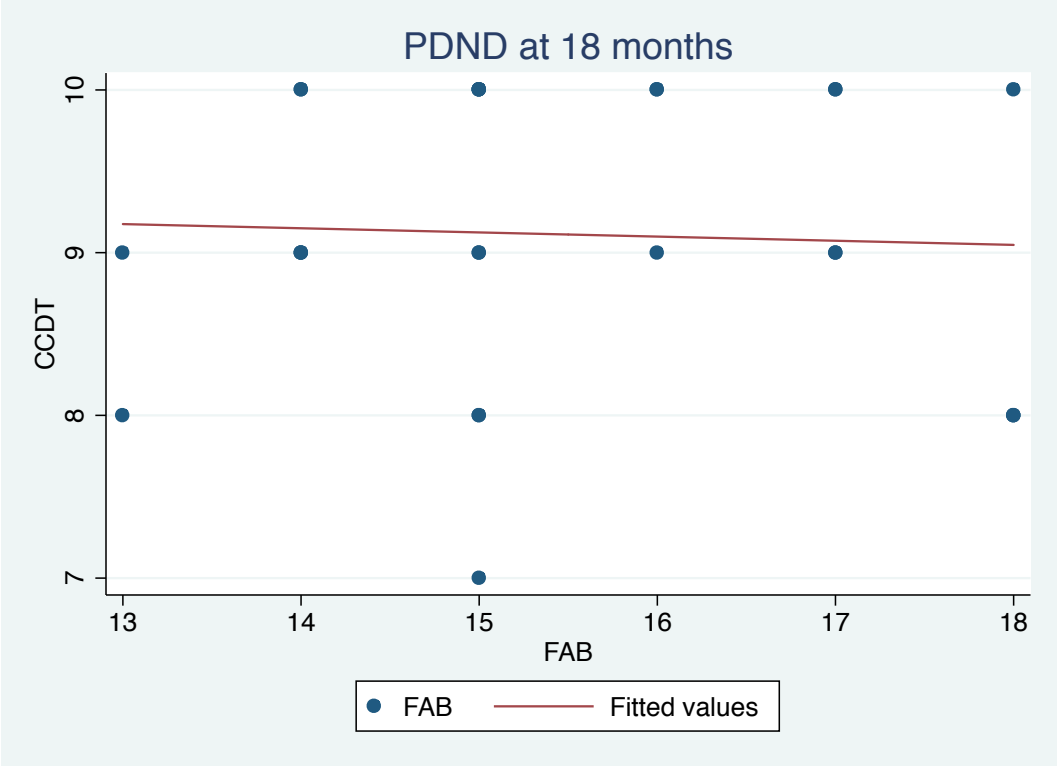
A)



B)


PDND at 18 months
CCDT
10
9
8
7
5
10
15
20
FAB
FAB
Fitted values

B1)


PDND at 18 months
CCDT
10
9
8
7
13
14
15
16
17
18
FAB
FAB
Fitted values

C)

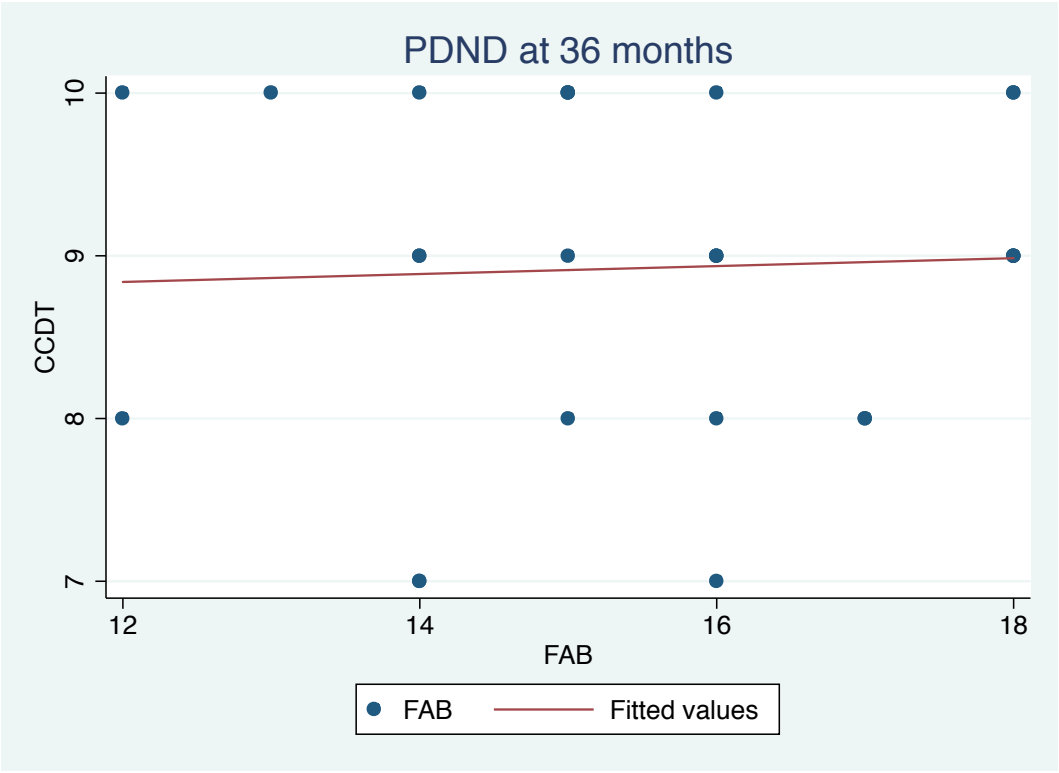
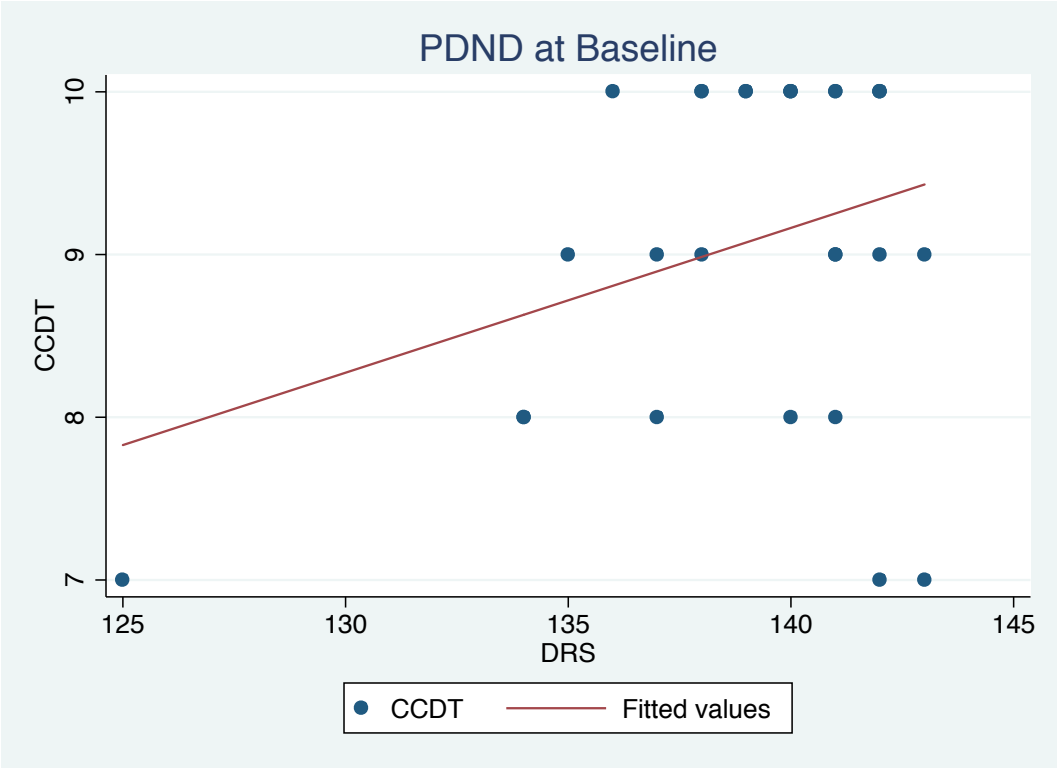
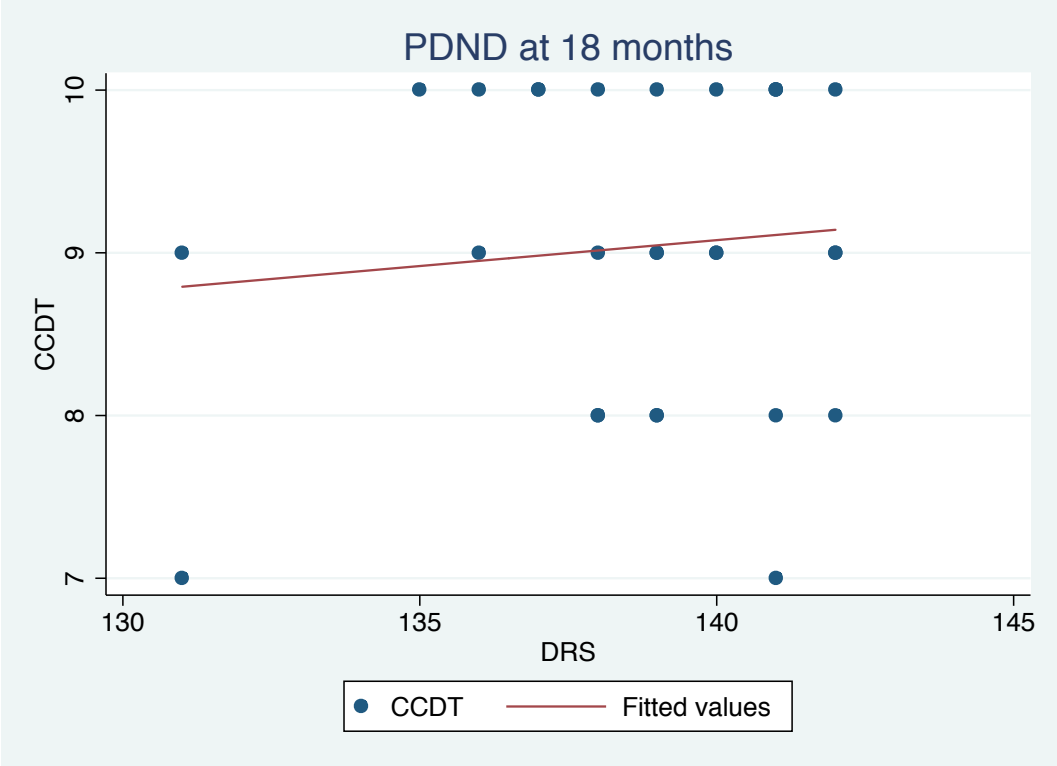


Figure 11. PDND group: Correlation of CCDT with DRS at A) 0 months, B) 18 months, B1) 18 months excluding outliers, and C) 36 months assessments.

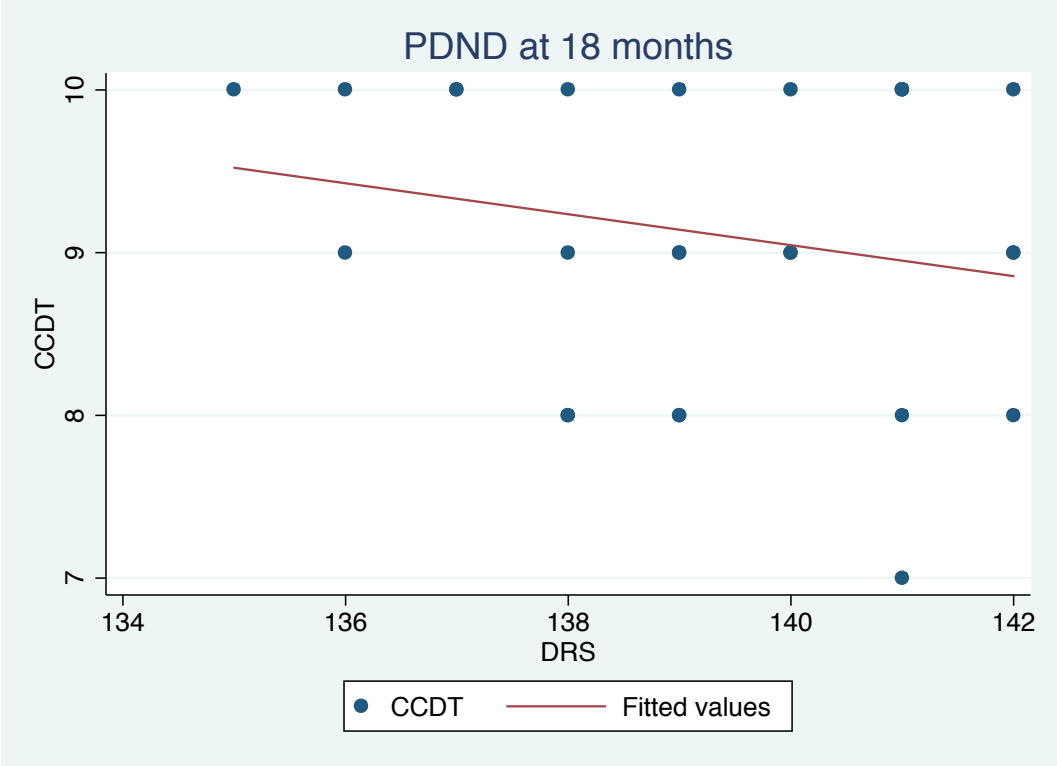
A)



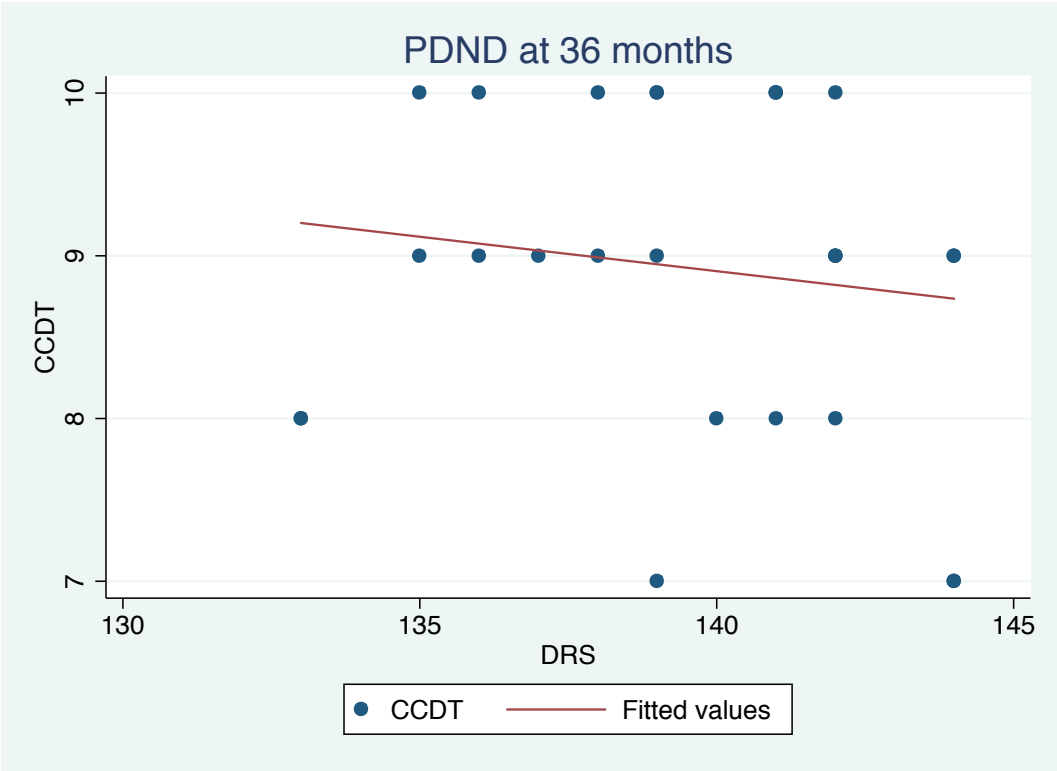
B)


PDND at 18 months
CCDT
10
9
8
7
130
135
140
145
DRS
CCDT
Fitted values

B1)


PDND at 18 months
CCDT
10
9
8
7
134
136
138
140
142
DRS
CCDT
Fitted values

C)


PDND at 36 months
CCDT
10
9
8
7
130
135
140
145
DRS
CCDT
Fitted values

6. Discussion

This study, in addition to estimating the incidence of PDD, highlights the utility of several neuropsychological assessment tools to track change in PD patients over time. Our PD patients had an incidence of PDD that was similar to previously published rates. [73, 74] The survival analysis clearly demonstrates the higher incidence of cognitive decline in PD patients when compared to controls. Three tools, the CDT, FAB and DRS, appear to be of benefit in monitoring for progression to dementia in patients with PD. The prospectively collected data allowed for assessing changes that are usually not easily identified in cross-sectional studies of PD.

In multiple analyses in this study older age was repeatedly associated with development of dementia, even in the subgroup analysis. In addition to age, higher H&Y scores were also associated with decline on the assessment scales. In this cohort where dementia developed in a quarter, and in 41.67% of PD patients, cognitive brain defects were seen to decline over time in tests that evaluated frontal and posterior cognitive brain functions, more so in patients developing dementia.

One of the main objectives of this study was to detect change in the measurements used. In addition to correcting for important confounders, we used random effect models for analyzing our longitudinal data set, considering that it is usually not possible to create the exact experimental situation each time a measurement is obtained from a participant. Another strength for this technique is that it takes into account the scores at baseline. General multivariate analyses would not be suitable for data sets amenable to such variation. This method allowed us to discriminate the changes occurring over time from the existing differences at baseline. [75]

PD patients had lower scores on the quantitative clock drawing tests in comparison to controls at each time of assessment, most likely highlighting the cognitive dysfunction among these patients even before dementia can be diagnosed. This was seen in both spontaneous and copy conditions, but the differences were larger in the spontaneous condition. The SCDT may thus be a better tool to determine the presence of executive dysfunction early on in the disease course, but as shown in the longitudinal analysis, may not be as helpful in assessing progression. Additionally, the SCDT could not consistently

differentiate between the two PD subgroups, which is probably because of the development of executive dysfunction (which is considered a component of spontaneous clock drawing) in most PD patients, even without progressing to cortical dementia, where the clock drawing copy is expected to become impaired. These findings are consistent with an early prospective study. [76, 77]

The CCDT was able to consistently discriminate within the PD subgroup, and change over time, suggesting these patients may be on their way to develop dementia. The SCDT and CCDT should be used in combination in the clinical setting, using the SCDT to detect impairment, and if abnormalities do not improve on the CCDT, then the individual is more likely to be progressing to dementia (Figure 12). Interestingly, in the comparison between PD and controls, the CCDT showed significant reduction in the number of stimulus bound, perseverative, and conceptual errors, these are primarily frontal functions. This was seen both the PD and control participants, but more appreciated in the PD group as they were infrequent among controls. The spatial/planning deficits remained significantly different between the two groups, which is most likely explained by the development of visual-spatial dysfunction in PD patients due to dementia. This suggests that the presence of errors on the CDT early on may be a sign of progression to dementia, particularly if these errors do not improve in the copy condition.

It has been previously demonstrated that difficulties in clock drawing that improve to a copy stimulus suggest a problem with executive abilities; where as difficulties with both situations may suggest visual spatial impairment. [78, 79] Similar to our study, Cahn-Weiner et al. looked at the CDT in AD, PDD and LBD patients using the Rouleau method. They noted among the qualitative errors that planning and conceptualization difficulties were the most prominent, but they were not able to discriminate between the three groups by their qualitative errors.[31] Among the important predictors that have been identified in a prospective study for the development of a more rapid cognitive decline in PD were low scores on the pentagon copy at baseline. [76] This is consistent with our findings about the CCDT. In an earlier report, Cormack et al. 2004 studied pentagon copying among AD, PD, LBD, and PDD patients, and found that there was similar performance between DLB and PDD patients. [80]

We did not examine qualitative subcomponents of the CDT longitudinally as this would result in too many variables to analyze in relation to the size of the cohort. Also, the interpretation of graphic difficulties in PD patients is difficult, since it may correlate with motor symptoms rather than represent cognitive defects. Additionally, the correlation between raters was not optimal in some of the subcomponents of the qualitative scale. Conversely, the presence or absence of spatial errors correlated well between raters and occurred most frequently among the qualitative measures. We recommend it should still be used in assessing the CDT. In addition to the Rouleau scale, multiple other scales provide useful information.[28, 81] The utility of the CDT in differentiating MCI from dementia in non-PD patients is still a topic of debate. In some studies it has not been found useful in cognitive impairment cases but helpful in identifying those with dementia. [32, 33]

As previous studies have shown, [54, 59] the MMSE scale's inadequacy to distinguish cognitively impaired patients with PD from controls is demonstrated here. This is particularly true early in the course of PD. The difference in the MMSE in the entire cohort was not significant until the third assessment, mostly from the effect of patients converting to dementia over that time. Although the MMSE scores were consistently and significantly lower in the PDID, the decline over time was not significant enough to distinguish the group from PDND, while the differences early on in the disease are small.

The FAB was useful in monitoring progression in patients, in both the primary and subgroup analysis. This is in keeping with its utility demonstrated in other studies [40, 45] and the prominent role that is played by executive dysfunction in PDD. The battery's utility is likely attributed to its subcomponents that have been found abnormal in frontal defects, such as phonemic fluency. [82] More recently however, studies have found semantic fluency to be more predictive of PDD. [76] The FAB is a useful bedside tool that, unlike the DRS, can be easily incorporated into a clinical assessment due to its brevity.

Not surprisingly the DRS showed progressive change over time and was consistently lower in the PD group during the primary analysis, and in the PDID group in the subgroup analysis. The DRS is accurate in detecting deficits in those with mild impairments [12] and sensitive in PDD. [41] It has been previously proposed that a cut-off on

the DRS of 132 had good accuracy in detecting dementia in PD. [52] In this study, mean values of the PDID subgroup reached 132 at the 18-month assessment, which is in keeping with this recommended cut-off. When considering the whole cohort, the mean DRS scores at any time point were always above 132. The DRS utility in clinical practice is limited by the long duration required for administration. [54] A score of 123 has been recommended to discriminate between PDD and PDND, none of the mean DRS scores in any of the groups were that low. The lowest mean score (129.1) obtained by the PDID group was at least 10 points lower than that of the PDND group, at 36-month assessment.

There does not seem to be much numerical decline in the scales despite their statistical significance. Slow decline has also been demonstrated in other cohorts. Williams-Grey et al. did not find much change in the proportion of those with cognitive impairment during a 3.5-year follow up of PD. [76] In the follow up study at five years, deficits based on abnormal fronto-striatal tasks did not relate to dementia while tasks related to posterior regions did. [77]

The FAB and DRS correlated significantly in the cohort, and in the subgroup the CCDT, FAB, and DRS each correlated significantly well with one another. The FAB was previously found to be correlated with the DRS in a mixed group of dementia patients, [40] and to be significantly lower in PD patients compared to controls. [44] It is interesting that the CCDT correlated well with the FAB and the DRS, this is supported by a study which suggested parietal lobe dysfunction being responsible for some impairment on the FAB. [46] The CDT thus may be a useful screen for detecting PDD. The DRS assesses multiple cognitive domains, explaining why it correlated well with the other scales.

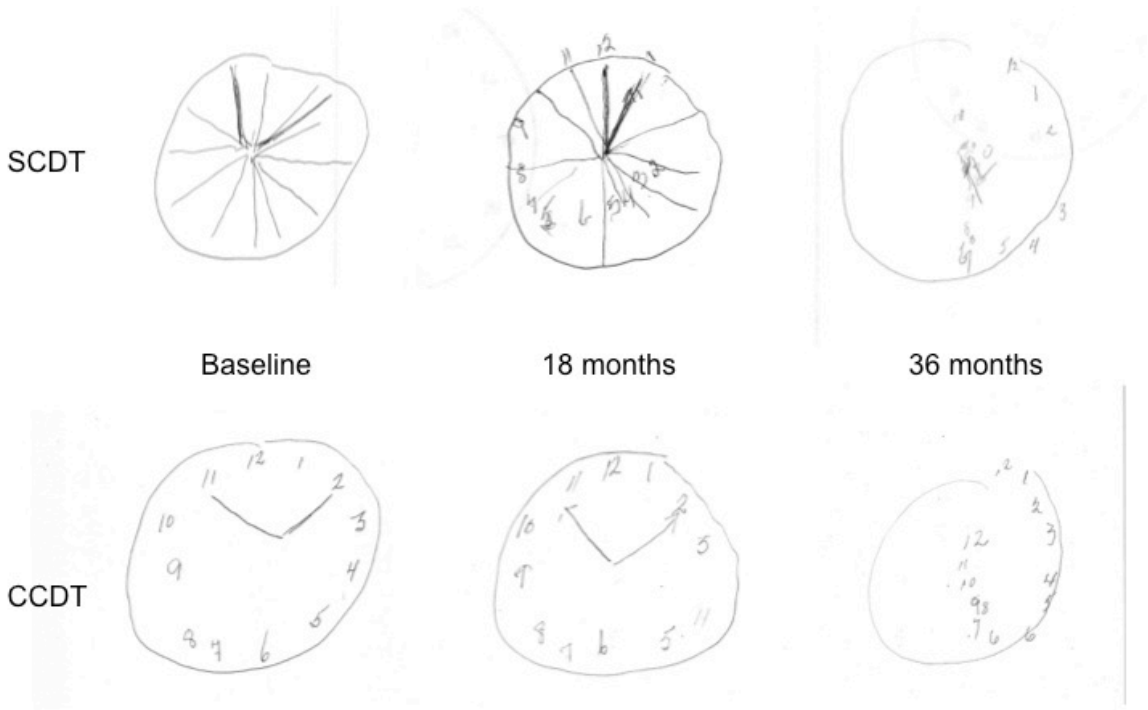
Impairments on neuropsychological tools identify impairments in different cognitive domains. The FAB is primarily designed to identify cognitive dysfunction localized to the frontal lobe. The MMSE and DRS assess multiple domains that correspond to different brain regions. Perhaps counter intuitively, the CDT can be applied in a fashion similar to the latter two scales in that it assesses multiple domains when both a spontaneous and copy format are used. Functional MRI studies have also suggested the requirement of the posterior parietal regions among the different stages of the CDT. [83, 84]

With regard to study limitations, our cohort was only followed for a period of three years; therefore, it is difficult to determine how the results would have varied had the period been longer. One study observed that the mean duration of PD at the time criteria for dementia first became fulfilled was 13.8 ± 6.3 years. [4] Similarly, not much change was found in the proportion of patients with dementia in another cohort study of PD patients followed for 3.5 years. [76] Attrition may be of concern as dementia was only determined at the 36-month assessment and prior to that there were six deaths and three dropouts. Another potential source of bias is that some PD patients and controls were from the same social circuit and could possibly have shared factors that influence the risk of developing a particular outcome.

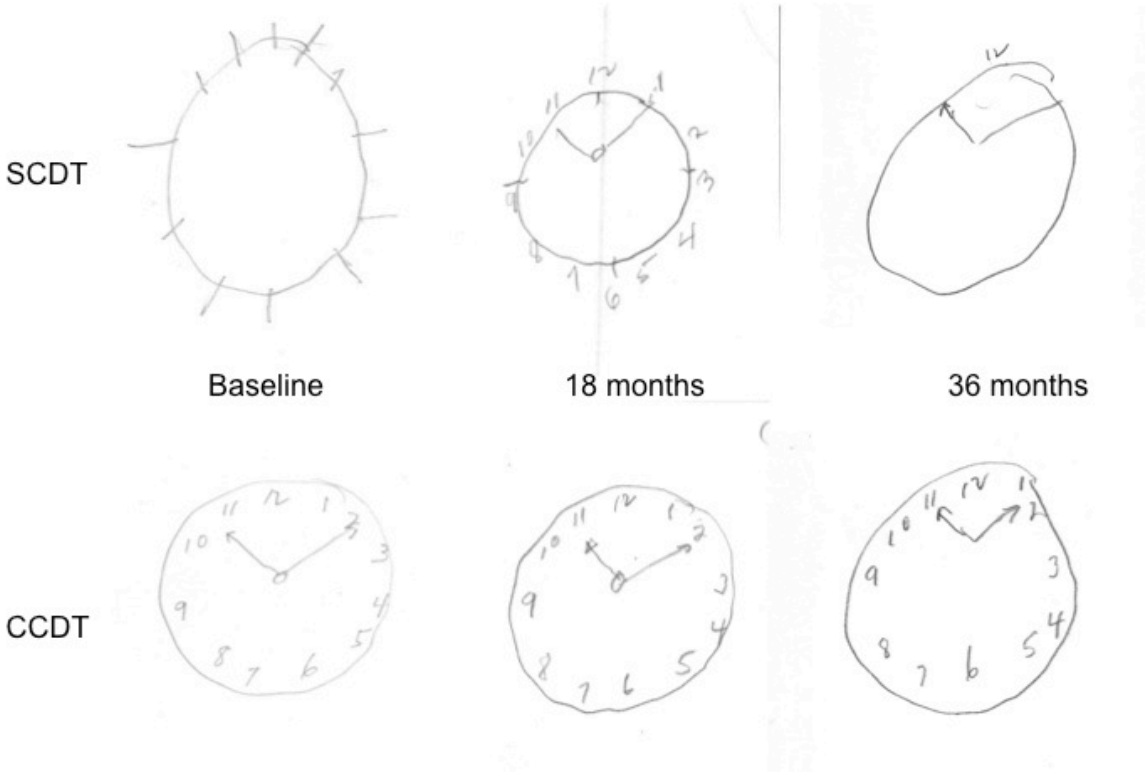
In summary, tests that assess executive dysfunction such as the CDT, FAB and DRS are useful in discriminating non-demented PD and control patients with early cognitive impairments. The CCDT, FAB, and DRS are able to discriminate PDID from PDND patients over time and suggest which patient may be progressing to dementia. Thus they are also likely to be useful in following the cognitive status of individual patients over time. The spontaneous and copy stages of the CDT should be used together, as it can help differentiate between those with PD executive dysfunction only, from PD patients developing dementia.

Figure 12. Examples of performance on the Clock Drawing Test: A) A patient with PDID exhibits impaired performance on the SCDT starting at baseline with relatively preserved performance on the CCDT until 36 months with similar impairment in both situations. B) There is impaired performance on the SCDT in a PDND patient at each time point while the performance on the CCDT remained intact. C) A PDID patient with impairments in both CDT situations from baseline with more pronounced difficulties in the SCDT. Performance declines over time in both situations.

A)



B)



C)

SCDT

Baseline 18 months 36 months

CCDT

References

1. Parkinson J. An essay on the shaking palsy. 1817. J Neuropsychiatry Clin Neurosci 2002;14:223-36; discussion 222.

2. Frank S, Jankovic J. Parkinson disease.

3. Aarsland D, Zaccai J, Brayne C. A systematic review of prevalence studies of dementia in parkinson's disease. Mov Disord 2005;20:1255-1263.

4. Aarsland D, Andersen K, Larsen JP, Lolk A, Kragh-Sorensen P. Prevalence and characteristics of dementia in parkinson disease: An 8-year prospective study. Arch Neurol 2003;60:387-392.

5. Marder K, Leung D, Tang M, et al. Are demented patients with parkinson's disease accurately reflected in prevalence surveys? A survival analysis. Neurology 1991;41:1240-1243.

6. Aarsland D, Bronnick K, Williams-Gray C, et al. Mild cognitive impairment in parkinson disease: A multicenter pooled analysis. Neurology 2010;75:1062-1069.

7. The Alzheimer Society. Rising tide: The impact of dementia on canadian society. 2009.

8. Wimo A, Winblad B, Jonsson L. The worldwide societal costs of dementia: Estimates for 2009. Alzheimers Dement 2010;6:98-103.

9. Rodriguez-Ferreiro J, Cuetos F, Herrera E, Menendez M, Ribacoba R. Cognitive impairment in parkinson's disease without dementia. Mov Disord 2010;25:2136-2141.

10. Muslimovic D, Post B, Speelman JD, Schmand B. Cognitive profile of patients with newly diagnosed parkinson disease. Neurology 2005;65:1239-1245.

11. Muslimovic D, Post B, Speelman JD, De Haan RJ, Schmand B. Cognitive decline in parkinson's disease: A prospective longitudinal study. J Int Neuropsychol Soc 2009;15:426-437.

12. Rosenthal E, Brennan L, Xie S, et al. Association between cognition and function in patients with parkinson disease with and without dementia. Mov Disord 2010;25:1170-1176.

13. Galasko D, Bennett D, Sano M, et al. An inventory to assess activities of daily living for clinical trials in alzheimer's disease. the alzheimer's disease cooperative study. Alzheimer Dis Assoc Disord 1997;11 Suppl 2:S33-9.

14. Ardila A, Surloff C, Mark V. Executive dysfunction. 2011.

15. Leh SE, Petrides M, Strafella AP. The neural circuitry of executive functions in healthy subjects and parkinson's disease. Neuropsychopharmacology 2010;35:70-85.

16. Ibarretxe-Bilbao N, Junque C, Marti MJ, Tolosa E. Brain structural MRI correlates of cognitive dysfunctions in parkinson's disease. J Neurol Sci 2011;310:70-74.

17. Camicioli R, Gee M, Bouchard TP, et al. Voxel-based morphometry reveals extra-nigral atrophy patterns associated with dopamine refractory cognitive and motor impairment in parkinsonism. Parkinsonism Relat Disord 2009;15:187-195.

18. Camicioli R, Moore MM, Kinney A, Corbridge E, Glassberg K, Kaye JA. Parkinson's disease is associated with hippocampal atrophy. Mov Disord 2003;18:784-790.

19. Aarsland D, Perry R, Brown A, Larsen JP, Ballard C. Neuropathology of dementia in parkinson's disease: A prospective, community-based study. Ann Neurol 2005;58:773-776.

20. Janvin CC, Larsen JP, Salmon DP, Galasko D, Hugdahl K, Aarsland D. Cognitive profiles of individual patients with parkinson's disease and dementia: Comparison with dementia with lewy bodies and alzheimer's disease. Mov Disord 2006;21:337-342.

21. Noe E, Marder K, Bell KL, Jacobs DM, Manly JJ, Stern Y. Comparison of dementia with lewy bodies to alzheimer's disease and parkinson's disease with dementia. Mov Disord 2004;19:60-67.

22. Aarsland D, Litvan I, Salmon D, Galasko D, Wentzel-Larsen T, Larsen JP. Performance on the dementia rating scale in parkinson's disease with dementia and dementia with lewy bodies: Comparison with progressive supranuclear palsy and alzheimer's disease. J Neurol Neurosurg Psychiatry 2003;74:1215-1220.

23. CRITCHLEY M, RUSSELL WR, ZANGWILL OL. Discussion on parietal lobe syndromes. Proc R Soc Med 1951;44:337-346.

24. Freund B, Gravenstein S, Ferris R, Burke BL, Shaheen E. Drawing clocks and driving cars. J Gen Intern Med 2005;20:240-244.

25. Rouleau I, Salmon DP, Butters N, Kennedy C, McGuire K. Quantitative and qualitative analyses of clock drawings in alzheimer's and huntington's disease. Brain Cogn 1992;18:70-87.

26. Zhou A, Jia J. The value of the clock drawing test and the mini-mental state examination for identifying vascular cognitive impairment no dementia. Int J Geriatr Psychiatry 2008;23:422-426.

27. Riedel O, Klotsche J, Spottke A, et al. Cognitive impairment in 873 patients with idiopathic parkinson's disease. results from the german study on epidemiology of parkinson's disease with dementia (GEPAD). J Neurol 2008;255:255-264.

28. Freedman M, Leach L, Kaplan E, Winocur G, Shulman K, Delis DC. Clock Drawing: A Neuropsychological Analysis. New York: Oxford University Press, 1994.

29. Libon DJ, Swenson RA, Barnoski EJ, Sands LP. Clock drawing as an assessment tool for dementia. Arch Clin Neuropsychol 1993;8:405-415.

30. Gnanalingham KK, Byrne EJ, Thornton A. Clock-face drawing to differentiate lewy body and alzheimer type dementia syndromes. Lancet 1996;347:696-697.

31. Cahn-Weiner DA, Williams K, Grace J, Tremont G, Westervelt H, Stern RA. Discrimination of dementia with lewy bodies from alzheimer disease and parkinson disease using the clock drawing test. Cogn Behav Neurol 2003;16:85-92.

32. Powlishta KK, Von Dras DD, Stanford A, et al. The clock drawing test is a poor screen for very mild dementia. Neurology 2002;59:898-903.

33. Nishiwaki Y, Breeze E, Smeeth L, Bulpitt CJ, Peters R, Fletcher AE. Validity of the clock-drawing test as a screening tool for cognitive impairment in the elderly. Am J Epidemiol 2004;160:797-807.

34. Babins L, Slater ME, Whitehead V, Chertkow H. Can an 18-point clock-drawing scoring system predict dementia in elderly individuals with mild cognitive impairment? J Clin Exp Neuropsychol 2008;30:173-186.

35. Kaplan E. A process approach to neuropsychological assessment. In: Clinical Neuropsychology and Brain Function: Research, Measurement, and Practice. Washington, DC: American Psychological Association, 1988, .

36. Philpot M. The clock-drawing test: A critique. Int Psychogeriatr 2004;16:251-256.

37. Shulman KI. Clock-drawing: Is it the ideal cognitive screening test? Int J Geriatr Psychiatry 2000;15:548-561.

38. Piovezan MR, Teive HA, Piovesan EJ, Mader MJ, Werneck LC. Cognitive function assessment in idiopathic parkinson's disease. Arq Neuropsiquiatr 2007;65:942-946.

39. Stella F, Gobbi LT, Gobbi S, Oliani MM, Tanaka K, Pieruccini-Faria F. Early impairment of cognitive functions in parkinson's disease. Arq Neuropsiquiatr 2007;65:406-410.

40. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: A frontal assessment battery at bedside. Neurology 2000;55:1621-1626.

41. Kaszas B, Kovacs N, Balas I, et al. Sensitivity and specificity of addenbrooke's cognitive examination, mattis dementia rating scale, frontal assessment battery and mini mental state examination for diagnosing dementia in parkinson's disease. Parkinsonism Relat Disord 2012;18:553-556.

42. Ben Sassi S, Nabli F, Hentati E, et al. Cognitive dysfunction in tunisian LRRK2 associated parkinson's disease. Parkinsonism Relat Disord 2012;18:243-246.

43. Beato R, Amaral-Carvalho V, Guimaraes HC, et al. Frontal assessment battery in a brazilian sample of healthy controls: Normative data. Arq Neuropsiquiatr 2012;70:278-280.

44. Bugalho P, Vale J. Brief cognitive assessment in the early stages of parkinson disease. Cogn Behav Neurol 2011;24:169-173.

45. Lima CF, Meireles LP, Fonseca R, Castro SL, Garrett C. The frontal assessment battery (FAB) in parkinson's disease and correlations with formal measures of executive functioning. J Neurol 2008;255:1756-1761.

46. Matsui H, Udaka F, Miyoshi T, et al. Frontal assessment battery and brain perfusion image in parkinson's disease. J Geriatr Psychiatry Neurol 2006;19:41-45.

47. Mattis S. Mental status examination for organic mental syndrome in the elderly patient. In: Geriatric Psychiatry: A Handbook for Psychiatrists and Primary Care Physicians. New York: Grune & Stratton, 1976, .

48. Monsch AU, Bondi MW, Salmon DP, et al. Clinical validity of the mattis dementia rating scale in detecting dementia of the alzheimer type. A double cross-validation and application to a community-dwelling sample. Arch Neurol 1995;52:899-904.

49. Smith GE, Ivnik RJ, Malec JF, Kokmen E, Tangalos E, Petersen RC. Psychometric properties of the mattis dementia rating scale. Assessment 1994;1:123-132.

50. Paolo AM, Troster AI, Glatt SL, Hubble JP, Koller WC. Differentiation of the dementias of alzheimer's and parkinson's disease with the dementia rating scale. J Geriatr Psychiatry Neurol 1995;8:184-188.

51. Brown GG, Rahill AA, Gorell JM, et al. Validity of the dementia rating scale in assessing cognitive function in parkinson's disease. J Geriatr Psychiatry Neurol 1999;12:180-188.

52. Matteau E, Dupre N, Langlois M, et al. Mattis dementia rating scale 2: Screening for MCI and dementia. Am J Alzheimers Dis Other Demen 2011;26:389-398.

53. Villeneuve S, Rodrigues-Brazete J, Joncas S, Postuma RB, Latreille V, Gagnon JF. Validity of the mattis dementia rating scale to detect mild cognitive impairment in parkinson's disease and REM sleep behavior disorder. Dement Geriatr Cogn Disord 2011;31:210-217.

54. Kulisevsky J, Pagonabarraga J. Cognitive impairment in parkinson's disease: Tools for diagnosis and assessment. Mov Disord 2009;24:1103-1110.

55. Llebaria G, Pagonabarraga J, Kulisevsky J, et al. Cut-off score of the mattis dementia rating scale for screening dementia in parkinson's disease. Mov Disord 2008;23:1546-1550.

56. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189-198.

57. Camicioli R, Wild K. Assessment of the elderly with dementia. In: Handbook of Neurological Rating Scales. Second ed. New York, New York: Demos medical Publising, 2006, 195.

58. Zadikoff C, Fox SH, Tang-Wai DF, et al. A comparison of the mini mental state exam to the montreal cognitive assessment in identifying cognitive deficits in parkinson's disease. Mov Disord 2008;23:297-299.

59. Nazem S, Siderowf AD, Duda JE, et al. Montreal cognitive assessment performance in patients with parkinson's disease with "normal" global cognition according to mini-mental state examination score. J Am Geriatr Soc 2009;57:304-308.

60. Litvan I, Bhatia KP, Burn DJ, et al. Movement disorders society scientific issues committee report: SIC task force appraisal of clinical diagnostic criteria for parkinsonian disorders. Mov Disord 2003;18:467-486.

61. Camicioli R, Sabino J, Gee M, et al. Ventricular dilatation and brain atrophy in patients with parkinson's disease with incipient dementia. Mov Disord 2011;26:1443-1450.

62. Camicioli R, Majumdar SR. Relationship between mild cognitive impairment and falls in older people with and without parkinson's disease: 1-year prospective cohort study. Gait Posture 2010;32:87-91.

63. de Frias CM, Dixon RA, Fisher N, Camicioli R. Intraindividual variability in neurocognitive speed: A comparison of parkinson's disease and normal older adults. Neuropsychologia 2007;45:2499-2507.

64. Rouleau I, Salmon DP, Butters N. Longitudinal analysis of clock drawing in alzheimer's disease patients. Brain Cogn 1996;31:17-34.

65. Fahn S, Elton RL, UPDRS program members. Unified Parkinson's disease rating scale. In: *Recent Developments in Parkinson's Disease*. Macmillan Health Care Information ed. Florham Park, NJ: 1987, 153.

66. Hoehn MM, Yahr MD. Parkinsonism: Onset, progression and mortality. Neurology 1967;17:427-442.

67. van Marwijk HW, Wallace P, de Bock GH, Hermans J, Kaptein AA, Mulder JD. Evaluation of the feasibility, reliability and diagnostic value of shortened versions of the geriatric depression scale. Br J Gen Pract 1995;45:195-199.

68. Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. J Am Geriatr Soc 1968;16:622-626.

69. Hachinski VC, Iliff LD, Zilhka E, et al. Cerebral blood flow in dementia. Arch Neurol 1975;32:632-637.

70. Morris JC. The clinical dementia rating (CDR): Current version and scoring rules. Neurology 1993;43:2412-2414.

71. American Psychiatric Association.
Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR Fourth Edition (Text Revision). Washington, DC: American Psychiatric Association, 2000.

72. Kawas C, Gray S, Brookmeyer R, Fozard J, Zonderman A. Age-specific incidence rates of alzheimer's disease: The baltimore longitudinal study of aging. Neurology 2000;54:2072-2077.

73. Aarsland D, Andersen K, Larsen JP, Lolk A, Nielsen H, Kragh-Sorensen P. Risk of dementia in parkinson's disease: A community-based, prospective study. Neurology 2001;56:730-736.

74. Hobson P, Meara J. Risk and incidence of dementia in a cohort of older subjects with parkinson's disease in the united kingdom. Mov Disord 2004;19:1043-1049.

75. Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982;38:963-974.

76. Williams-Gray CH, Foltynie T, Brayne CE, Robbins TW, Barker RA. Evolution of cognitive dysfunction in an incident parkinson's disease cohort. Brain 2007;130:1787-1798.

77. Williams-Gray CH, Evans JR, Goris A, et al. The distinct cognitive syndromes of parkinson's disease: 5 year follow-up of the CamPaIGN cohort. Brain 2009;132:2958-2969.

78. Gnanalingham KK, Byrne EJ, Thornton A. Clock-face drawing to differentiate lewy body and alzheimer type dementia syndromes. Lancet 1996;347:696-697.

79. Royall DR, Polk M. Dementias that present with and without posterior cortical features: An important clinical distinction. J Am Geriatr Soc 1998;46:98-105.

80. Cormack F, Aarsland D, Ballard C, Tovee MJ. Pentagon drawing and neuropsychological performance in dementia with lewy bodies, alzheimer's disease,

parkinson's disease and parkinson's disease with dementia. Int J Geriatr Psychiatry 2004;19:371-377.

81. Suhr J, Grace J, Allen J, Nadler J, McKenna M. Quantitative and qualitative performance of stroke versus normal elderly on six clock drawing systems. Arch Clin Neuropsychol 1998;13:495-502.

82. Jacobs DM, Marder K, Cote LJ, Sano M, Stern Y, Mayeux R. Neuropsychological characteristics of preclinical dementia in parkinson's disease. Neurology 1995;45:1691-1696.

83. Ino T, Asada T, Ito J, Kimura T, Fukuyama H. Parieto-frontal networks for clock drawing revealed with fMRI. Neurosci Res 2003;45:71-77.

84. Trojano L, Fragassi NA, Chiacchio L, et al. Relationships between constructional and visuospatial abilities in normal subjects and in focal brain-damaged patients. J Clin Exp Neuropsychol 2004;26:1103-1112.

8. Appendix

Appendix 1-Table of results of mixed effect analysis of PD versus controls.

SCDT

	Coefficient (change on the SCDT score)	Z score, P value
PD group (vs controls)	-1.30	-3.57, <0.0001
Time	0.01	-1.09, 0.277
Interaction (Group x Time)	-0.002	-0.19, 0.85
Education (years)	-0.002	-0.04, 0.97
Age	-0.16	-4.81, <0.0001
Sex (males)	0.19	0.62, 0.534

CCDT

	Coefficients (change on the CCDT score)	Z score, P value
PD group (vs controls)	-0.5	-1.91, 0.056
Time	-0.001	-1.32, 0.188
Interaction (Group x Time)	-0.011	-1.35, 0.176
Males	0.22	1.04, 0.297
Age	-0.085	-3.81, < 0.0001
Education	0.02	0.65, 0.517

MMSE

	Coefficients (change on the MMSE score)	Z score, P value
PD group (vs controls)	-0.348	-0.79, 0.431
Time	-0.0199	-2.12, 0.034
Interaction (Group x Time)	-0.025	-1.89, 0.059
Males	-0.324	-0.85, 0.398
Age	-0.148	-3.67, < 0.0001
Education	0.613	1.05, 0.294

FAB

	Coefficients (change on the FAB score)	Z score, P value
PD group (vs controls)	-0.707	-1.84, 0.066
Time	0.018	2.08, 0.038
Interaction (Group x Time)	-0.039	-3.22, 0.001
Males	0.708	2.16, 0.031
Age	-0.206	-5.96, < 0.0001
Education	0.073	1.46, 0.145

DRS

	Coefficients (change on the DRS score)	Z score, P value
PD group (vs controls)	-1.793	-1.59, 0.111
Time	0.01	0.43, 0.664
Interaction (Group x Time)	-0.066	-1.99, 0.046
Males	0.327	0.33, 0.741
Age	-0.408	-3.89, < 0.0001
Education	0.145	0.96, 0.337

Appendix 2-Table of results of mixed effect analysis of PD subgroup, comparing PDID with PDND.

CCDT

	Coefficients (change on the CCDT score)	Z score, P value
PDID group (vs PDND)	-0.513	-1.27, 0.203
Time	-0.002	-0.22, 0.824
Interaction (Group x Time)	-0.03	-2.05, 0.040
Age	-0.057	-1.58, 0.115
Males	0.515	1.82, 0.069
Education	-0.006	-0.12, 0.907
H & Y	-0.674	-3.42, 0.001

MMSE

	Coefficients (change on the MMSE score)	Z score, P value
PDID group (vs PDND)	-0.866	-1.10, 0.270
Time	-0.022	-1.66, 0.098
Interaction (Group x Time)	-0.031	-1.46, 0.145
Age	0.011	0.14, 0.889
Males	-0.191	-0.30, 0.760
Education	0.044	0.40, 0.688
H & Y	-1.514	-4.58, < 0.0001

FAB

	Coefficients (change on the FAB score)	Z score, P value
PDID group (vs PDND)	-0.141	-0.23, 0.821
Time	0.013	1.07, 0.285
Interaction (Group x Time)	-0.076	-3.91, < 0.0001
Age	-0.192	-3.25, 0.001
Males	0.829	1.76, 0.078
Education	0.007	0.09, 0.930
H & Y	-0.693	-2.46, 0.014

DRS

	Coefficients (change on the DRS score)	Z score, P value
PDID group (vs PDND)	-1.227	-0.60, 0.545
Time	0.025	0.71, 0.476
Interaction (Group x Time)	-0.207	-3.59, < 0.0001
Age	-0.378	-1.91, 0.057
Males	0.662	0.42, 0.676
Education	-0.125	-0.45, 0.655
H & Y	-2.61	-2.63, 0.009