RESEARCH ARTICLE

Mild Cognitive Impairment as a Risk Factor for Parkinson's Disease Dementia

Jeroen Hoogland, MD,[1]* Judith A. Boel, MSc,[1,2] Rob M.A. de Bie, MD, PhD,[1] Ronald B. Geskus, PhD,[3] Ben A. Schmand, PhD,[2,4] John C. Dalrymple-Alford, PhD,[5] Connie Marras, MD, PhD,[6] Charles H. Adler, MD, PhD,[7] Jennifer G. Goldman, MD, MS,[8] Alexander I. Tröster, PhD,[9] David J. Burn, MD, PhD,[10] Irene Litvan, MD, PhD,[11] and Gert J. Geurtsen, PhD,[4] on behalf of the MDS Study Group "Validation of Mild Cognitive Impairment in Parkinson Disease"

[1]*Department of Neurology, Academic Medical Center Amsterdam, Amsterdam, The Netherlands*
[2]*Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands*
[3]*Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center Amsterdam, Amsterdam, The Netherlands*
[4]*Department of Medical Psychology, Academic Medical Center Amsterdam, Amsterdam, The Netherlands*
[5]*New Zealand Brain Research Institute, Brain Research New Zealand, Centre of Research Excellence, Christchurch, New Zealand*
[6]*Morton and Gloria Shulman Movement Disorders Centre and the Edmond J. Safra Program in Parkinson's Disease, Toronto Western Hospital, University of Toronto, Toronto, Ontario, Canada*
[7]*Arizona Parkinson's Disease Consortium, Mayo Clinic Arizona, Scottsdale, Arizona, USA, and Banner Sun Health Research Institute, Sun City, Arizona, USA*
[8]*Department of Neurological Sciences, Section of Parkinson Disease and Movement Disorders, Rush University Medical Center, Chicago, Illinois, USA*
[9]*Department of Clinical Neuropsychology and Center for Neuromodulation, Barrow Neurological Institute, Phoenix, Arizona, USA*
[10]*Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, United Kingdom*
[11]*Department of Neurosciences University of California San Diego, Movement Disorder Center, San Diego, California, USA*

ABSTRACT: **Background**: The International Parkinson and Movement Disorder Society criteria for mild cognitive impairment in PD were recently formulated.
Objectives: The aim of this international study was to evaluate the predictive validity of the comprehensive (level II) version of these criteria by assessment of their contribution to the hazard of PD dementia.
Methods: Individual patient data were selected from four separate studies on cognition in PD that provided information on demographics, motor examination, depression, neuropsychological examination suitable for application of level II criteria, and longitudinal follow-up for conversion to dementia. Survival analysis evaluated the predictive value of level II criteria for cognitive decline toward dementia as expressed by the relative hazard of dementia.
Results: A total of 467 patients were included. The analyses showed a clear contribution of impairment according to level II mild cognitive impairment criteria, age, and severity of PD motor symptoms to the hazard of dementia. There was a trend of increasing hazard of dementia with declining neuropsychological performance.
Conclusions: This is the first large international study evaluating the predictive validity of level II mild cognitive impairment criteria for PD. The results showed a clear and unique contribution of classification according to level II criteria to the hazard of PD dementia. This finding supports their predictive validity and shows that they contribute important new information on the hazard of dementia, beyond known demographic and PD-specific factors of influence. © 2017 International Parkinson and Movement Disorder Society

Key Words: Parkinson's disease; mild cognitive impairment; dementia; neuropsychological tests; survival analyses

*Correspondence to: Jeroen Hoogland, Department of Neurology, Academic Medical Center Amsterdam, Meibergdreef 9, 1105AZ, Amsterdam, the Netherlands; E-mail: j.hoogland@amc.nl

Funding agencies: Funding sources for study were the Michael J. Fox Foundation and Dutch Parkinson Foundation.

Relevant conflicts of interest/financial disclosures: John C. Dalrymple-Alford reports grants from Neurological Foundation of New Zealand, Canterbury Medical Research Foundation (NZ), Health Research Council of New Zealand, and Freemasons of New Zealand. Connie Marras reports funding from the Canadian Institutes of Health Research and Michael J. Fox Foundation. Charles A. Adler reports grants from the NINDS, Arizona Biomedical Research Commission, and Michael J. Fox Foundation. Jennifer G. Goldman reports grants from NIH and the Michael J. Fox Foundation. Alexander I. Tröster reports grants from the Michael J. Fox Foundation during the conduct of the study. Irene Litvan reports grants from the Michael J. Fox Foundation during the conduct of the study. Gert J. Geurtsen reports grants from the Michael J. Fox Foundation and Dutch Parkinson Foundation during the conduct of the study.

Full financial disclosures and author roles may be found in the online version of this article.

Received: 30 September 2016; **Revised:** 9 March 2017; **Accepted:** 10 March 2017

Published online 00 Month 2017 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.27002

Cognitive deficits have been increasingly recognized as important manifestations of Parkinson's disease (PD). Two important landmarks in this process were the formulation of clinical criteria for Parkinson's disease dementia (PDD)[1] and, more recently, clinical criteria for Parkinson's Disease with Mild Cognitive Impairment (PD-MCI).[2] The International Parkinson and Movement Disorder Society (MDS) PD-MCI Validation Study Group was initiated to validate the PD-MCI criteria, starting with the evaluation of their prognostic value for the development of PDD.[3] The underlying rationale is that mild cognitive impairment in PD may be regarded as a stage between normal cognition and PDD, that is, as "cognitive decline that is not normal for age, but with essentially normal functional activities."[4] Previous research shows that cognitive decline in PD is frequent,[5,6] can start early in the disease,[7-9] and has a heterogeneous presentation.[10] Some patients rapidly decline toward PDD, and others have extended periods of cognitive health, or only minimal impairment. Whereas the ability to identify patients with a high risk of rapid cognitive decline is of distinct importance for both clinical care and intervention trials, there is a pressing lack of validated markers. PD-MCI is a possible clinical marker and can be assessed in an abbreviated (level I) or comprehensive (level II) manner according to the MDS PD-MCI diagnostic criteria.[2] We conducted a large international study of longitudinal individual patient data to evaluate whether level II PD-MCI criteria are a prognostic indicator of cognitive decline to dementia.

Materials and Methods

Data Inclusion

The formation of the MDS PD-MCI Validation Group has been described earlier.[3] Members contributed individual patient data from either ongoing or completed studies in PD. Studies were eligible if they met the following inclusion criteria: (1) longitudinal data with at least 75 patients at first measurement and at least 67% participation on at least one subsequent visit (as a quality check); (2) known PDD status at follow-up; and (3) neuropsychological and disease data to which the level II PD-MCI criteria could be applied. The latter required data from studies with standardized neuropsychological testing by qualified personnel that included at least two tests per each of five cognitive domains (i.e., attention/working memory, executive function, language, memory, and visuospatial function) and a measure of gradual cognitive decline (a subjective measure provided by either the patient, informant, or clinician).

All available demographic and clinical data were retrieved, including information on age, sex, years of education, PD duration, global cognitive measures, neuropsychological test scores, either UPDRS part III[11] or MDS UPDRS-III,[12] Hoehn and Yahr scores,[13] Mini–Mental State Exam (MMSE) scores,[14] and depression. Within included studies, patients with disease duration of more than 25 years since PD symptom onset at first measurement were excluded to enhance uniformity of the data. Patients were also excluded when they had PDD at first measurement. The method used to diagnose PDD was allowed to differ across sites and is described below and in Table 1. The same holds for indicators of depression.

Application of the PD-MCI Criteria

Level II PD-MCI criteria have been described by Litvan and colleagues.[2] The two major requirements are impairment on comprehensive formal neuropsychological testing and gradual cognitive decline, while not fulfilling PDD criteria.

Impairment on Formal Neuropsychological Testing

Two tests per cognitive domain were selected in each study database to enhance comparability between studies in application of the level II PD-MCI criteria. The selection was based on the expert consensus of experienced neuropsychologists from all participating centers in the MDS PD-MCI study group. Neuropsychological performance on the resulting 10 tests per subject was interpreted against published norms where available. Otherwise, normative scores were derived from a local control sample using multiple regression techniques correcting for the effects of age, sex, and education. Impairment was rated crossing cutoffs of –1 standard deviation (SD), –1.5 SD, and –2 SD from the mean for at least two tests. Patients that did not cross the –1 SD level for at least two tests were labeled as "no neuropsychological impairment." Note that patients were classified according to their lowest pair of test performances and could only belong to one of the four groups.

Cognitive Decline and Functional Independence

Cognitive decline reported by the patient, caregiver, and/or clinician is required for PD-MCI. Because the criteria do not specify a method to determine this, there were no restrictions on the methods used to assess gradual cognitive decline. Although functional independence is included in the PD-MCI criteria to rule out PDD, because we had already excluded PDD patients, we did not include additional measures of functional independence in the application of the PD-MCI criteria.

TABLE 1. Cohort details of the included studies

Cohort		AZSAND (n = 101)	CARPA (n = 112)	NZBRI (n = 136)	Toronto (n = 118)
Cohort type		Open community volunteers cohort	Closed incident clinic cohort	Closed prevalent clinic cohort	Closed prevalent clinic cohort
Follow-up (range in years)		Yearly (range, 0.5-6.0)	0, 3, 5, and 8 years (range, 1.4-9.0)	baseline+approximately 2-yearly up to 6 years (range, 0.8-6.2)	0, 1, and 2 years (range, 0.5-3.3)
PD criteria		UKPDS Brain Bank	Gelb	UKPDS Brain Bank	UKPDS Brain Bank
PDD criteria		MDS PDD and DSM-IV	Based on MMSE and FIM[a]	MDS PDD	MDS PDD
Normative scores		Control group	Published norms	Published norms	Published norms
Subjective cognitive decline[b]	Patient	UPDRS-I item I	PDQL[27] items 31 and 34	PDQ-39[28] item 32 and CDR[29] memory items	Abbreviated NBI patient version[25]
	Significant other			CDR memory items	Abbreviated NBI caregiver version
	Clinician	UPDRS-I item I			
Indicator of depression (absent/present)		use of antidepressants	HADS depression subscore ≥ 11[30]	NPI depression subscale total score frequency $\times$ severity) ≥ 4[31]	GDS 15 score ≥ 5[30]

The table shows the cohort types, diagnostic criteria, reference used for evaluation of neuropsychological performance, and the measures of subjective cognitive decline for each of the studies.
[a]Refer to the main text for more details.
[b]References are provided for scales not further mentioned in the text.
CDR, Clinical Dementia Rating; GDS15, Geriatric Depression Scale 15; HADS, Hospital Anxiety and Depression Scale; NBI, Neurobehavioral Signs and Symptoms Abbreviated Inventory; NPI, Neuropsychiatric Inventory; PDQ-39, Parkinson's Disease Questionnaire 39; PDQL, Parkinson's Disease Quality of Life Questionnaire.

PD-MCI: Levels of Impairment and Its Subtypes

When patients had signs of cognitive decline and impairment on two neuropsychological tests, they were categorized as level II PD-MCI and classified according to the severity of their impairment as based on different cut-off scores for determining impairment. This resulted in four cognitive status groups: one without cognitive impairment and three PD-MCI groups with increasing levels of impairment (PD-MCI according to –1, –1.5, and –2 SD). Furthermore, the PD-MCI patients were classified according to their cognitive domain of impairment when only one domain was affected (single-domain PD-MCI) and classified as multidomain PD-MCI if the impaired tests covered multiple domains.

Statistical Analysis

Imputation

Because 10 neuropsychological tests and a subjective measure of gradual cognitive decline were all needed to apply the PD-MCI scoring, we anticipated the need for an imputation method. Multiple imputation (MI) is the method of choice for complex incomplete data problems.[15] Advantages are that MI can use the relations between the observed measures and preserve them in the imputations, and that it takes uncertainty with regard to the missing data into account.[16] The R statistical software[17] provides a flexible approach to multiple imputation in the *mice* package.[15] Twenty imputations were created within the original studies using predictive mean matching, and all variables to be evaluated in further analyses were included in the imputation model; an exception was made for derived variables, which were computed based on the (imputed) underlying variables (i.e., individual neuropsychological tests were imputed, not the cognitive status group classification), and for the time to event, which was replaced by the Nelson Aalen estimate of the cumulative baseline hazard as is best practice.[18] All analyses were performed on the imputed data and were pooled using Rubin's rules,[19] unless stated otherwise. In short, Rubin's rules are used to derive one overall estimate and variance from the multiple imputations while accounting for both the within and between imputation variability.

Survival Analysis

We used a Cox proportional hazards model with counting process formulation as implemented in the R *survival* package[20,21] and *rms* package.[22] The event of interest was the development of PDD. The time of PD symptom onset was used as the start time; the time from PD symptom onset to PDD or censoring as the follow-up time. Onset of PDD was estimated to be halfway between the actual observation of PDD and the observation preceding that moment, given that the exact time of onset is unknown. Individuals not developing PDD were censored at their last visit. Use of duration since symptom onset as the principal time axis

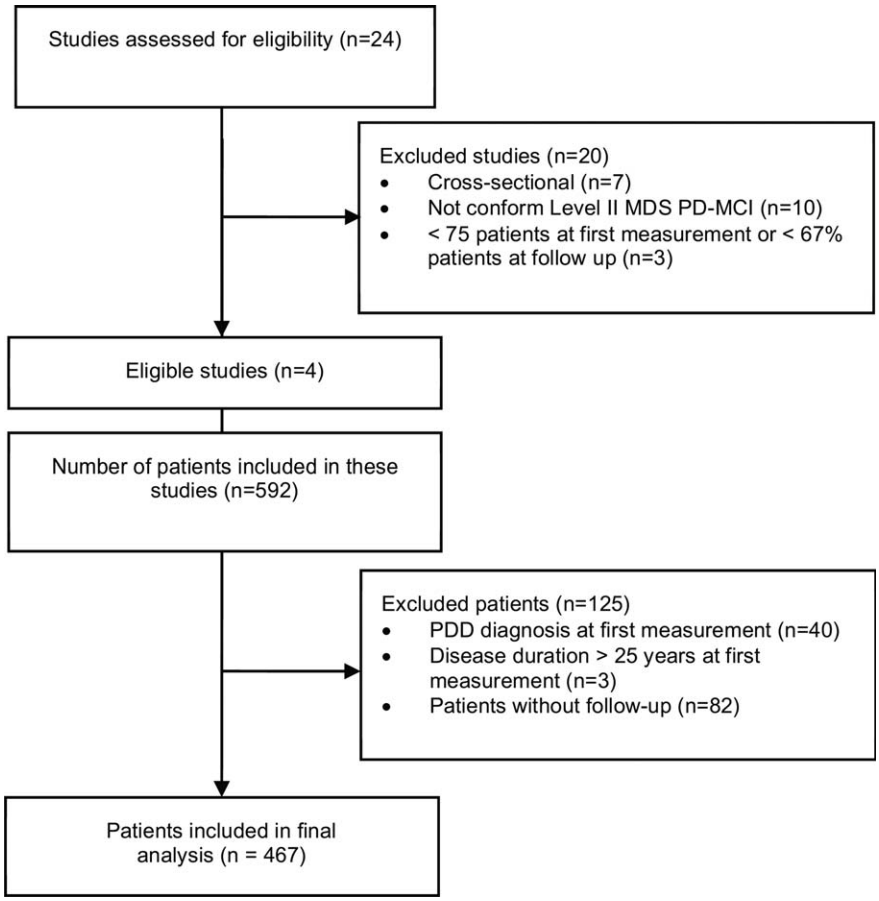

FIG. 1. Flow chart showing the data inclusion process.

allowed for correction for left truncation, which refers to the situation where patients were already at risk of PDD before they were included in the study (as in prevalence cohorts). While this was necessary because of the structure of the data, it precluded estimation of the effect of disease duration. Included predictors were age, sex, years of education, UPDRS-III, an indicator of depression, and the four categories of cognitive status at first measurement. No cognitive impairment was used as the reference group for the latter. The rate of PDD was allowed to differ between the original studies by use of study site as a stratum variable. Possible nonlinearity was examined using restricted cubic splines with four knots for all continuous predictors. Proportionality was assessed globally and, if necessary, per covariate by testing for a difference from zero of the correlation coefficient between Kaplan–Meier transformed survival time and the scaled Schoenfeld residuals. The scaled Schoenfeld residuals were inspected visually to detect nonlinear patterns possibly invalidating this test. The C statistic was used to quantify predictive value and is defined as the proportion of all possible patient pairs for whom the ordering of observed and predicted survival times is concordant. A bootstrapping approach using 200 samples was used to estimate slope shrinkage and corrected Nagelkerke R^2 as suggested in Harrell and colleagues.[23] For the C statistics, slope shrinkage, and Nagelkerke R^2 estimates, their median over imputations was obtained because Rubin's rules do not apply to their distributions.

Results

Data Inclusion

Twenty-three validation study group sites contributed individual patient data from 24 studies. Figure 1 schematically displays the inclusion process. A total of 467 patients from four large longitudinal cohort studies fulfilled the inclusion criteria. These studies will be referred to as the AZSAND cohort, the CARPA cohort,[7] the NZBRI cohort,[24] and the Toronto cohort.[25] The AZSAND cohort is part of the Arizona Study of Aging and Neurodegenerative Disease.[26] Cohort details are summarized in Table 1. Both open and closed cohorts and both incident and prevalent cohorts were followed. Follow-up length, frequency, and intervals differed between the studies. PD was diagnosed according to standard criteria.[32,33] Neuropsychological scores in the CARPA, NZBRI, and Toronto cohorts were adjusted for age, education, and/or sex, where applicable, based on published norms. In the AZSAND cohort, normative scores were

TABLE 2. Tests that were used for Level II MDS PD-MCI analyses

Studies	Language	Attention	Executive Functioning	Memory	Visuospatial Functioning
CARPA	WAIS-III Similarities Category fluency	Trail Making Test A Stroop interference	MWCST perseverative errors Tower of London total moves score	RAVLT delayed recall RBMT delayed recall	JOLO GIT Legkaarten[a]
NZBRI	Boston Naming Test Category fluency (DKEFS)	Digit Ordering Test WAIS-III Digit Span total	Stroop interference (DKEFS) Trail Making Test B	CVLT II long delayed recall RCF delayed recall	JOLO RCF copy
AZSAND	Boston Naming Test Category fluency	WMS-R Digit Span backward WMS-R Digit Symbol	Stroop interference Letter fluency	RAVLT delayed recall WMS-R Logical Memory delayed recall	JOLO Clock Drawing Test
Toronto	Boston Naming Test Category fluency (DKEFS)	WAIS-III Letter Number Sequencing WAIS-III Digit Span total	Category switching (DKEFS) Stroop interference (DKEFS)	CVLT II long delayed recall RCF delayed recall	JOLO RCF copy

All obtained scores were normative scores. For tests with more than one main outcome available, such as the Tower of London, the specific score used is mentioned. Note that, based on expert consensus, the same test can appear in multiple domains and more general tests can appear outside of their primary domain. This reflects difference in availability between individual studies. Therefore, semantic fluency can be the best available language test and the Stroop interference can appear in both the attention and executive domain.
[a]Dutch; this is a tangram-like visuospatial subtest of the Dutch Groningen Intelligence Test.
CVLT, California Verbal Learning Test; DKEFS = Delis–Kaplan Executive Function System; JOLO = Judgment of Line Orientation Test; MWCST = Modified Wisconsin Card Sorting Test; RAVLT = Rey Auditory Verbal Learning Test; RBMT = Rivermead Behavioral Memory Test; RCF = Rey Complex Figure; WAIS-R/III= Wechsler Adult Intelligence Scale version Revised/III; WMS-R = Wechsler Memory Scale-Revised.

derived from a sample of 708 non-PD community volunteers. All anticipated demographic and clinical information was available in all four centers except for the H & Y scores. Either UPDRS-III[11] or MDS UPDRS-III[12] was used to assess motor function, and comparable scores on the scale of the UPDRS-III were derived using conversion guidelines.[34] These will further be referred to as UPDRS-III* to reflect the mixed nature. Patients in the ASZAND cohort were mostly assessed in practically defined *off* state, whereas the others were mostly assessed in *on* state. Furthermore, PDD classification differed across centers. MDS PDD criteria[1] were used in the NZBRI, Toronto, and AZSAND cohorts, with additional use of Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria[35] in the latter. The CARPA study used an MMSE[14] cutoff as used in the Dubois screening criteria[36] (<26) in combination with the cognitive items of the Functional Independence Measures (FIM)[37] ($\geq$1 item with a score $\leq$5), or an MMSE score below 21 regardless of functional impairment. The available information on depression differed across centers as well, as indicated in Table 1. These measures were all summarized to either indicate presence of absence of depression to enable shared analyses using all available data. Detailed information on the neuropsychological examinations is available Table 2.

Missing Values

The percentage of missing values in age, sex, years of education, and PD symptom duration ranged from 0% to 1%. Gradual cognitive decline had 2% missing values, the UPDRS-III* 6%, and indicators of depression were missing in 9%. One percent had a missing attention test, versus 2% for visuospatial and executive function respectively, 10% for memory, and 16% for language. PD-MCI depends on multiple measures and was missing in 22%. Because only a small portion of the neuropsychological tests is missing and their mutual relations are strong, the conditions are such that multiple imputation is expected to perform well.

Individual Patient Data Descriptives

Descriptive statistics are shown in Table 3. Included patients (n = 467) had a mean age of 69 years, male predominance (63%), a median duration since PD symptom onset of 4 years, and a median UPDRS-III* score of 20. Fifteen percent had a positive indicator for depression. Forty percent of the patients had neuropsychological impairment, roughly equally divided over the PD-MCI groups. Regarding PD-MCI subtypes, the majority of patients (92%) had multidomain impairment. The other 8% were scattered over the single domains. To provide an overview of the available data over time, Supporting Figure 1 shows the distribution of observation periods.

Table 3 shows that 69 patients developed PDD during follow-up (14.3%). Only 6.4% of patients without any cognitive impairment at first measurement developed PDD, progressively increasing to 50.0% of the PD-MCI group fulfilling the –2 SD cutoff on first neuropsychological evaluation. Further examination of the involvement of individual cognitive domains in PD-MCI classification revealed that each of the five domains was commonly impaired and that their association with conversion to PDD was heterogeneous. Details are provided in Supporting Table 1.

Note that in this section on descriptive statistics, important characteristics of the data, among which

TABLE 3. Descriptive statistics at first measurement and conversion to PDD

		AZSAND (n = 101)	CARPA (n = 112)	NZBRI (n = 136)	Toronto (n = 118)	**Overall (n = 467)**
Age, years (mean, SD; range)		72.8 (8.5; 46-86)	65.7 (10.4; 32-84)	66.0 (8.2; 42-80)	71.1 (5.4; 60-84)	**68.7 (8.8; 32-86)**
Sex, male (frequency, %)		66 (65.3)	59 (52.7)	87 (64.0)	81 (68.6)	**293 (62.7)**
Education, years (mean, SD; range)		15.5 (2.6; 8-23)	11.6 (2.5; 7-18)	13.0 (2.9; 8-20)	15.9 (2.4; 8-20)	**14.0 (3.1; 7-23)**
MMSE (median, IQR; range)		29 (27-30; 16-30)	28 (27-29; 22-30)	28 (26-29; 21-30)	29 (27-30; 22-30)	**28 (27-29; 16-30)**
PD symptom duration, years (median, IQR; range)		8.0 (4.0-12.0; 0-23)	1.3 (1.0-1.8; 0-7)	4.0 (2.3-7.0; 1-20)	5.0 (3.0-9.8; 1-21)	**4.0 (2.0-8.0; 0-23)**
UPDRS III* (median, IQR; range)		20 (11-31; 2-52)	15 (11-21; 5-39)	24 (17-32; 3-69)	20 (15-26; 1-48)	**20 (13-28; 1-69)**
Positive indicator of depression (frequency, %),		29 (28.7)	10 (8.9)	28 (20.6)	2 (1.7)	**69 (14.8)**
PD-MCI count (frequency, %)	No impairment	75 (74.2)	47 (42.0)	86 (63.3)	72 (61.0)	**280 (60.0)**
	−1 to −1.5 SD	10 (9.9)	24 (21.4)	15 (11.0)	18 (15.3)	**67 (14.3)**
	−1.5 to −2 SD	8 (7.9)	29 (25.9)	14 (10.3)	7 (5.9)	**58 (12.4)**
	below −2 SD	8 (7.9)	12 (10.7)	21 (15.4)	21 (17.8)	**62 (13.3)**
Conversion to PDD by cognitive classification (frequency/n, %)	No impairment	5 (41.7)	5 (23.8)	7 (28.0)	1 (9.1)	**18/280 (6.4)**
	−1 to −1.5 SD	1 (8.3)	4 (19.0)	3 (12.0)	0 (0.0)	**8/67 (11.9)**
	−1.5 to −2 SD	1 (8.3)	5 (23.8)	4 (16.0)	2 (18.2)	**12/58 (20.7)**
	below −2 SD	5 (41.7)	7 (33.3)	11 (44.0)	8 (72.7)	**31/62 (50.0)**
Person years of follow−up by cognitive classification (years, years per event)	No impairment	208 (78.8)	299 (45.0)	295 (71.6)	138 (65.7)	**940 (60.6)**
	−1 to −1.5 SD	28 (10.6)	153 (23.0)	40 (9.7)	34 (16.1)	**255 (16.5)**
	−1.5 to −2 SD	18 (6.8)	164 (24.7)	25 (6.1)	10 (4.8)	**217 (14.0)**
	below −2 SD	10 (3.8)	48 (7.2)	52 (12.6)	28 (13.3)	**138 (8.9)**

IQR, interquartile range.

the inter-relation between covariates, left truncation, and censoring, cannot be taken into account. However, the survival analyses reported in the following section take these aspects into account.

Survival Analyses

A Cox model including all covariates and stratified by the study sites met the proportionality assumption (chi-square, 11.55; df, 8; *P* value, 0.17). Therefore, the covariate effects did not significantly change over time since PD symptom onset. Furthermore, all non-linear terms were not significant (chi-square, 10.90; df, 6; *P* value, 0.09) and left out of the model. The results for the final model are shown in Table 4 and Supporting Figure 2. The hazard of PDD was higher with increasing age and UPDRS-III* score. Sex, years of education, and the indicator of depression had no significant effect. PD-MCI below the –1.5 SD cutoff clearly raised the hazard of PDD.

As a measure of the predictive value of the survival model, the median observed C statistic was 0.85, which means that in 85% of the possible paired patient comparisons, prediction and outcome are concordant. (i.e., the predicted time to event was shorter for the one first developing PDD). The median bootstrap corrected Nagelkerke R^2 was 0.28. The median slope shrinkage estimate was 0.86, indicating that an estimated 14% of the model fit was due to overfitting.

To increase insight into the relative contribution of the individual predictors to the hazard of PDD, bootstrap corrected R^2 values were derived for different submodels. A model stratified on study site and including only age explained 10.3% of the variance. Adding sex, years of education, UPDRS*, and an indicator of depression increased this to 12.4%. PD-MCI bridges the gap to 28.0%. Age and PD-MCI therefore clearly have the largest contribution.

Discussion

We found level II MDS PD-MCI to be clearly related to the hazard of PDD after controlling for demographic characteristics, severity of PD, and indicators of depression. This constitutes a new contribution to the PD-MCI literature, adds to the validation process of the level II MDS PD-MCI criteria, and thereby supports their application. In more detail, the analyses showed an increase in hazard with decreasing performance on neuropsychological examination, while correcting for possible confounders. This can be interpreted as a relative increase in the rate of

TABLE 4. Multivariate Cox proportional hazards model evaluating the hazard of PDD

	β	SE	95% CI	HR (e^{β})	z-Statistic	*P* Value
Age (per year)	0.09	0.02	(0.05; 0.13)	1.09	4.36	<0.005
Sex (male)	0.26	0.29	(-0.32; 0.83)	1.29	0.88	0.38
Years of education	0.00	0.05	(–0.11; 0.10)	1.00	–0.06	0.95
PD-MCI –1 to –1.5 SD	0.71	0.45	(–0.18; 1.59)	2.02	1.56	0.12
PD-MCI –1.5 to –2 SD	1.24	0.45	(0.37; 2.11)	3.46	2.78	<0.01
PD-MCI below –2 SD	2.42	0.35	(1.73; 3.11)	11.25	6.89	<0.005
UPDRS-III*	0.03	0.01	(0.01; 0.06)	1.03	2.49	0.01
Depression indicator	–0.22	0.47	(–1.14; 0.70)	0.80	–0.47	0.64

The overall model chi-square (df, 8) was 97.5 ($P < 0.005$). The reference categories were female and no cognitive impairment. For continuous variables, hazard ratios are expressed per unit difference on their scale of measurement (years and UPDRS-III* points respectively).
HR, hazard ratio; SE, standard error of β; 95% CI, 95% confidence interval for β.

conversion to PDD. The results indicated that this relative difference was constant over time, that is, the relative difference between cognitively healthy patients and PD-MCI patients at each consecutive time point since symptom onset was the same. As an example, the relative hazard to develop PDD for patients in the –1.5 to –2 SD group was, at any time since PD symptom onset, estimated to be approximately 3.5 times higher than for the cognitively healthy patients. This effect is comparable to the effect of an age difference of approximately 14 years or an increase in UPDRS-III* score of 37 points. The increase in hazard with increasing age and increasing PD severity is consistent with literature reviews by Aarsland and Kurtz[38] and Litvan and colleagues.[4]

Furthermore, the pattern of increase in hazard with each successive degree of neuropsychological impairment gives new insight into the use of cutoffs for the level II criteria when predicting PDD. Namely, a selection of one cutoff loses important information, due to grouping of patients with a different hazard of PDD. While the ease of use of one cutoff is evident, the current study was able to show that progressive impairment keeps adding to the hazard of PDD. This is in line with the view of mild cognitive impairment as a stage on the continuum between normal cognition and PDD. Regarding clinical relevance, the view of the MDS PD-MCI validation study group is that impairment beyond the –1.5 SD cutoff represents clinically meaningful decline. The –1.5 to –2 SD group, with an estimated 3.5 times increase in the hazard of PDD when compared to the normal cognition group, is also deemed to be sufficiently far from the development of PDD to be a meaningful subgroup. These patients could, for instance, be an interesting subpopulation for medication trials aiming to halt progression of cognitive decline.

Concerning the PD-MCI subclassification in single- and multidomain impairment, the groups of single-domain impairment were too small to provide useful statistical inference. Only 8% had single-domain impairment, without any meaningful pattern over the five cognitive domains. This is in agreement with an earlier study on the level II PD-MCI criteria in an individual PD cohort.[39] These findings may reflect widespread cognitive deficits or lack of sufficient specificity among current cognitive test measures, or could result from a bias toward multiple domain impairment in the MDS PD-MCI criteria.

There are several strengths in the current study. First, the predictive effect of level II PD-MCI was assessed over an extensive follow-up period. Second, it is the first study to uniformly apply the level II PD-MCI criteria in a varied and large international sample of PD patients. As specifically allowed in the MDS PD-MCI criteria, patients were examined with a variety of instruments reflecting the variability that exists across different international centers. Furthermore, a broad spectrum of disease duration was available with first assessments on patients ranging from 0 to 23 years since PD symptom onset. Under these heterogeneous circumstances, level II PD-MCI strongly contributed to the hazard of PDD. A downside to the approach favoring external validity is the limited comparability of the used measurement instruments. Each study followed their own local procedures, and this resulted in a variety of applied measures, with little overlap. This precluded evaluation of the prognostic value of individual measures in the aggregated data, and it necessitated expert consensus-based selection from the available measures to rate PD-MCI. The difference in study designs also impeded easy interpretation of the PD-MCI prevalence values and resulted in different operationalizations of conversion to PDD. Future prospective, large-scale studies could be designed with these issues in mind.

A limitation of the study is the small number of conversions to PDD and the limited number of studies, which impeded analyses of possibly important interaction effects, interstudy variability, and the derivation of time to PDD conversion estimates for possible clinically meaningful subgroups. Consequently, it precludes direct generalization of the results to the individual patient level, given that this should take the between center variability into account. Furthermore, the effect of duration since symptom onset could not be estimated, because it

was needed as the time axis to correct for the left-truncated structure of the data relating to the inclusion of prevalence cohorts. A further limitation is the possibility of informative censoring, also known as attrition bias. In any longitudinal cohort study design, patients returning for follow-up assessments may differ from those who do not (e.g., depending on patient health and its relation to the incentive to participate). Informative censoring denotes the situation where leaving the study is not independent of the study outcome. In the current setting, lacking information on mortality as a potentially important competing risk renders the analyses prone to violation of the assumption of noninformative censoring. In other words, when mortality and PDD share a biological cause, censoring because of mortality is informative of the hazard of PDD. Unfortunately, the current data cannot be used to estimate this possible influence and there is no research on this relation in PD. Informative censoring in general illustrates a challenge of longitudinal studies in an advancing neurodegenerative disorder and cautions that study samples may be less representative over time, leading to biased estimates.*

Our results represent a basic validation of level II MDS PD-MCI as a risk factor for PDD, but are not exhaustive, given the multiple ways that level II MDS PD-MCI criteria can be applied. The available data directed the focus to detection of cognitive impairment by means of normative neuropsychological test scores. However, the criteria can also be fulfilled by decline on serial cognitive testing or decline from premorbid level. While these options are specifically mentioned in the criteria, their operationalization is not yet clearly defined. In general, the multitude of available options in application of the PD-MCI criteria achieves a greater flexibility for their use, but that can also be a potential limitation. Differences in allowed measures, cut-off scores, and definitions of impairments should lead to caution when comparing different applications of the criteria. We recommend that future research further operationalizes the PD-MCI criteria across diverse populations. ■

Acknowledgments: This study was conducted on behalf of the International Parkinson and Movement Disorders Society PD-MCI Validation Study Group, consisting of the byline authors and the members listed in the Appendix.

Appendix

North America

VA San Diego Healthcare System, Departments of Psychiatry and Neurosciences, University of California San Diego, Movement Disorder Center, San Diego, CA (J. Vincent Filoteo, PhD);
Departments of Psychiatry and Neurology, University of Pennsylvania School of Medicine, and Parkinson's Disease and Mental Illness Research, Philadelphia Veterans Affairs Medical Center, Philadelphia, PA (Daniel Weintraub, MD);
Arizona Parkinson's Disease Consortium, Mayo Clinic Arizona, Scottsdale, AZ (John N. Caviness, MD) and Banner Sun Health Research Institute, Sun City, AZ (Christine Belden, PhD);
Veterans Affairs Puget Sound Health Care System and Department of Neurology, University of Washington School of Medicine, Seattle, WA (Cyrus P. Zabetian, MD, MSc, and Brenna A. Cholerton, PhD);
Lou Ruvo Center for Brain Health, Neurological Institute, Cleveland Clinic (James B. Leverenz, MD);
Department of Neurology, Hershey Medical Center (Xuemei Huang, MD, PhD, and Paul J. Eslinger, PhD);
Morton and Gloria Shulman Movement Disorders Centre and the Edmond J. Safra Program in Parkinson's Disease, Toronto Western Hospital, University of Toronto, Toronto, Ontario, Canada (Sarah Duff-Canning, PhD);
Department of Neurological Sciences, Section of Parkinson Disease and Movement Disorders, Rush University Medical Center, Chicago, IL (Bryan Bernard, PhD, and Glenn Stebbins, PhD)
The Centre for Applied Neurogenetics, University of British Columbia, Vancouver, British Columbia (Matthew Farrer, PhD);

Australia and New Zealand

New Zealand Brain Research Institute, Brain Research New Zealand, Centre of Research Excellence, Christchurch, New Zealand (Tim J. Anderson, FRACP, and Daniel J. Myall, PhD);
Brain & Mind Research Institute, The University of Sydney, Sydney, Australia (Sharon L. Naismith, PhD, Simon J.G. Lewis, PhD) and Neuroscience Research Australia, University of New South Wales, Randwick, Australia (Glenda M. Halliday, PhD);

Asia

Department of Neurology, National Taiwan University Hospital, College of Medicine, National Taiwan University, Taipei, Taiwan (R.M. Wu, MD, PhD);
Institute of Behavioral Medicine, College of Medicine, National Cheng Kung University, Tainan, Taiwan (R.L. Yu, PhD);

Europe

John Van Geest Centre for Brain Repair, University of Cambridge, Cambridge, United Kingdom (Caroline H. Williams-Gray, MRCP, PhD, David P. Breen, MRCP, PhD, and Roger A. Barker, MRCP, PhD);

*Unless the informative censoring is a competing risk that can be analyzed as such.

Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, United Kingdom (Alison J. Yarnall, PhD);
Department of Medical Psychology, section Medical Neuropsychology, VU Medical Center, Amsterdam, The Netherlands (Ted Koene, MSc, and Martin Klein, PhD);
Paracelsus-Elena-Klinik, Kassel, Germany, and University Medical Center Goettingen, Department of Neurosurgery and Institute of Neuropathology, Goettingen, Germany (Brit Mollenhauer, MD), Paracelsus-Elena-Klinik, Kassel, Germany, and University Medical Center Goettingen, Department of Neurosurgery, Goettingen, Germany, (Claudia Trenkwalder, MD);
Movement Disorders Unit, Neurology Department, Hospital and Institute of Biomedical Research Sant Pau, "CIBERNED," Barcelona, Spain (Jaime Kulisevsky, MD, PhD, and Javier Pagonabarraga, MD, PhD) and "Universitat Oberta de Catalunya" (Jaime Kulisevsky, MD, PhD);
Department of Neurology, Hospital Donostia, Donostia, San Sebastian and Ikerbasque, Basque Foundation for Science, Bilbao, Spain (Maria C. Rodriguez-Oroz, MD, PhD, and Carmen Gasca-Salas, MD, PhD)
Department of Psychiatry and Clinical Psychobiology, Faculty of Medicine, University of Barcelona, Spain (Carme Junque, PhD, and Barbara Segura, PhD);
Neurogenerative disease center, University of Salerno, Salerno, Italy (Paolo Barone, PhD) and Department of Psychology, Second University of Naples, Italy (Gabriella Santangelo, PhD);
Fifth Local Sanitary Unit, Hospital Psychology, Pisa, Italy, and Department of Surgical, Medical, Molecular, and Critical Area Pathology, Pisa University School of Medicine, Italy (Marco Timpano Sportiello PhD and Davide M. Cammisuli, PhD);
Parkinson Unit, Fondazione Ospedale San Camillo IRCCS, Venice, Italy (Roberta Biundo, PhD, Angelo Antonini, PhD, and Luca Weis, PhD);
The Norwegian Centre for Movement Disorders, Department of Neurology, and Memory Clinic, Stavanger University Hospital, Stavanger, Norway (Kenn Freddy Pedersen, PhD, and Guido Alves, PhD).

References

1. Emre M, Aarsland D, Brown R, et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. Mov Disord 2007;22:1689-1707.
2. Litvan I, Goldman JG, Tröster AI, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. Mov Disord 2012;27:349-356.
3. Geurtsen GJ, Hoogland J, Goldman JG, et al. Parkinson's disease mild cognitive impairment: application and validation of the criteria. J Parkinsons Dis 2014;4:131-137.
4. Litvan I, Aarsland D, Adler C. MDS task force on mild cognitive impairment in Parkinson's disease: critical review of PD-MCI. Mov Disord 2011;26:1814-1824.
5. Hely M, Reid W. The Sydney multicenter study of Parkinson's disease: the inevitability of dementia at 20 years. Mov Disord 2008; 23:837-844.
6. Buter T, Hout A Van den, Matthews F. Dementia and survival in Parkinson disease a 12-year population study. Neurology 2008;70: 1017-1022.
7. Muslimović D, Post B, Speelman J, Schmand B. Cognitive profile of patients with newly diagnosed Parkinson disease. Neurology 2005;65:1239-1245.
8. Foltynie T, Brayne CE, Robbins TW, Barker RA. The cognitive ability of an incident cohort of Parkinson's patients in the UK. The CamPaIGN study. Brain 2004;127:550-560.
9. Aarsland D, Brønnick K, Larsen JP, Tysnes OB, Alves G. Cognitive impairment in incident, untreated Parkinson disease: the Norwegian ParkWest study. Neurology 2009;72:1121-1126.
10. Kehagia AA, Barker RA, Robbins TW. Neuropsychological and clinical heterogeneity of cognitive impairment and dementia in patients with Parkinson's disease. Lancet Neurol 2010;9:1200-1213.
11. Fahn S, Elton R, Members of the UPDRS Development Committee. The Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden C, Calne D, Goldstein M, eds. Recent Developments in Parkinson's Disease, Vol. 2. Florham Park, NJ: Macmillan Healthcare Information; 1987:153-163, 293-304.
12. Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord 2008;23:2129-2170.
13. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. Neurology 1967;17:427-442.
14. Folstein M, Folstein S, McHugh P. 'Mini-mental state': a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189-198.
15. Buuren S, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. J Stat Softw 2011;45:1-67.
16. Buuren, S. Flexible Imputation of Missing Data. Boca Raton, FL: Chapman & Hall/CRC; 2012.
17. R Development Core Team. R: a language and environment for statistical computing. 2014. http://www.r-project.org. Access date: 9 March 2017.
18. White IR, Royston P. Imputing missing covariate values for the Cox model. Stat Med 2009;28:1982-1998.
19. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation:current practice and guidelines. BMC Med Res Methodol 2009;9:57.
20. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York: Springer; 2000.
21. Therneau TM. A Package for Survival Analysis in S. R package version 2.37-7. 2014. http://cran.r-project.org/package=survival. Access date: 9 March 2017.
22. Harrell FE, Jr. rms: Regression Modeling Strategies. R package version 4.3-0. 2015. http://cran.r-project.org/package=rms. Access date: 9 March 2017.
23. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-387.
24. Dalrymple-Alford JC, Livingston L, MacAskill MR, et al. Characterizing mild cognitive impairment in Parkinson's disease. Mov Disord 2011;26:629-636.
25. Marras C, Armstrong M, Meaney C. Measuring mild cognitive impairment in patients with Parkinson's disease. Mov Disord 2013;28:626-633.
26. Beach T, Adler C, Sue L. Arizona Study of Aging and Neurodegenerative Disorders and Brain and Body Donation Program. Neuropathology 2015;35:354-389.
27. de Boer AG, Wijker W, Speelman JD, de Haes JC. Quality of life in patients with Parkinson's disease: development of a questionnaire. J Neurol Neurosurg Psychiatry 1996;61:70-74.
28. Peto V, Jenkinson C, Fitzpatrick R. PDQ-39: a review of the development, validation and application of a Parkinson's disease quality of life questionnaire and its associated measures. J Neurol 1998; 245(Suppl 1):S10-S14.

29. Morris J. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993;43:2412-2414.

30. Schrag A, Barone P, Brown RG, et al. Depression rating scales in Parkinson's disease: critique and recommendations. Mov Disord 2007;22:1077-1092.

31. Caputo M, Monastero R, Mariani E, et al. Neuropsychiatric symptoms in 921 elderly subjects with dementia: a comparison between vascular and neurodegenerative types. Acta Psychiatr Scand 2008; 117:455-464.

32. Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. Arch Neurol 1999;56:33-39.

33. Hughes A, Daniel S, Kilford L, Lees A. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry 1992;55:181-184.

34. Goetz CG, Stebbins GT, Tilley BC. Calibration of Unified Parkinson's Disease Rating Scale scores to Movement Disorder Society-unified Parkinson's disease rating scale scores. Mov Disord 2012; 27:1239-1242.

35. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. Washington, DC: American Psychiatric Association; 2000.

36. Dubois B, Burn D, Goetz C, et al. Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. Mov Disord 2007;22:2314-2324.

37. Ottenbacher K, Hsu Y. The reliability of the functional independence measure: a quantitative review. Arch Phys Med Rehabil 1996;77:1226-1232.

38. Aarsland D, Kurz MW. The epidemiology of dementia associated with Parkinson disease. J Neurol Sci 2010;289:18-22.

39. Goldman J, Holden S, Bernard B, Ouyang B, Goetz CG, Stebbins GT. Defining optimal cutoff scores for cognitive impairment using Movement Disorder Society Task Force criteria for mild cognitive impairment in Parkinson's disease. Mov Disord 2013;28:1972-1979.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.