Journal of the International Neuropsychological Society (2018), **24**, 1084–1098

doi:10.1017/S1355617718000747

Evaluating Mild Cognitive Impairment in Essential Tremor: How Many and Which Neuropsychological Tests?

Tess E.K. Cersonsky,[1] Sarah Morgan,[1] Sarah Kellner,[1] Daphne Robakis,[1] Xinhua Liu,[2] Edward D. Huey,[3,4,5] Elan D. Louis,[1,6,7] AND Stephanie Cosentino[4,5]

[1]Division of Movement Disorders, Department of Neurology, Yale School of Medicine, Yale University, New Haven, Connecticut
[2]Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, New York
[3]Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York, New York
[4]Department of Neurology, College of Physicians and Surgeons, Columbia University, New York, New York
[5]Taub Institute for Research on Alzheimer's Disease and the Aging Brain, College of Physicians and Surgeons, Columbia University, New York, New York
[6]Department of Chronic Disease Epidemiology, Yale School of Public Health, Yale University, New Haven, Connecticut
[7]Center for Neuroepidemiology and Clinical Neurological Research, Yale School of Medicine, Yale University, New Haven, Connecticut

(RECEIVED February 14, 2018; FINAL REVISION June 28, 2018; ACCEPTED July 21, 2018; FIRST PUBLISHED ONLINE October 10, 2018)

Abstract

Objectives: Essential tremor (ET) confers an increased risk for developing both amnestic and non-amnestic mild cognitive impairment (MCI). Yet, the optimal measures for detecting mild cognitive changes in individuals with this movement disorder have not been established. We sought to identify the cognitive domains and specific motor-free neuropsychological tests that are most sensitive to mild deficits in cognition as defined by a Clinical Dementia Rating (CDR) of 0.5, which is generally associated with a clinical diagnosis of MCI. **Methods:** A total of 196 ET subjects enrolled in a prospective, longitudinal, clinical-pathological study underwent an extensive motor-free neuropsychological test battery and were assigned a CDR score. Logistic regression analyses were performed to identify the neuropsychological tests which best identified individuals with CDR of 0.5 (mild deficits in cognition) *versus* 0 (normal cognition). **Results:** In regression models, we identified five tests in the domains of Memory and Executive Function which best discriminated subjects with CDR of 0.5 *versus* 0 (86.9% model classification accuracy). These tests were the California Verbal Learning Test II Total Recall, Logical Memory II, Verbal-Paired Associates I, Category Switching Fluency, and Color-Word Inhibition. **Conclusions:** Mild cognitive difficulty among ET subjects is best predicted by combined performance on five measures of memory and executive function. These results inform the nature of cognitive dysfunction in ET and the creation of a brief cognitive battery to assess patients with ET for cognitively driven dysfunction in life that could indicate the presence of MCI. (*JINS*, 2018, *24*, 1084–1098)

Keywords: Movement disorders, Cognition, Mental status and dementia tests, Memory and learning tests, Memory, Executive function

INTRODUCTION

Essential tremor (ET) is among the most common movement disorders, with a prevalence of 4% among adults age ≥ 40 years (Dogu et al., 2003; Louis, Ottman, & Hauser, 1998). It is characterized primarily by kinetic tremor (Louis, 2009). Risk factors include older age and family history. The underlying pathophysiology is not completely understood, although the tremor is thought to involve an aberration in a cerebello-thalamo-cortical loop (Louis, 2014a). Recent literature indicates that ET is a complex syndrome with heterogeneous motor and non-motor features. Among the latter are cognitive impairments (Benito-Leon & Louis, 2013; Janicki, Cosentino, & Louis, 2013; Louis, 2010; Louis & Rao, 2014b; Mameli et al., 2013; Sinoff & Badarny, 2014), such as reductions in attention, executive function, visuospatial processing, and memory (Gasparini et al., 2001; Janicki et al., 2013; Kim et al., 2009; Lombardi, Woolston, Roberts, & Gross, 2001; Tröster et al., 2002); and increased risk of mild cognitive impairment (MCI) and dementia (Benito-León, Louis, Bermejo-Pareja, & NEDICES Study Group, 2006a, 2006b; Benito-León, Louis, Mitchell, & Bermejo-Pareja, 2011; Bermejo-Pareja, Louis, Benito-León, & NEDICES Study Group, 2007; Louis, Benito-León, Vega-Quiroga, Bermejo-Pareja, & NEDICES Study Group, 2010;

Correspondence and reprint requests to: Stephanie Cosentino; Columbia University Medical Center, 630 West 168th Street; P&S Mailbox 16; NY, NY 10032; E-mail: sc2460@cumc.columbia.edu

Thawani, Schupf, & Louis, 2009), deficits which are reflected in functional imaging studies, with abberant network connectivity observed to be associated with tremor features and cognitive dysfuncton (Benito-León et al., 2015). Cognitive dysfunction that occurs in ET, even when mild, is accompanied by a range of functional consequences (Louis et al., 2010, 2016, 2017; Rao, Gillman, & Louis, 2014; Rao, Uddin, Gillman, & Louis, 2013), thereby highlighting its importance.

MCI is defined as cognitive decline that is not explained by an individual's age or education but does not interfere with activities of daily life (Gauthier et al., 2006). Much of the work developing the concept and assessment of MCI has been performed in studies that assess conversion of MCI to Alzheimer's disease (AD) (Aretouli, Okonkwo, Samek, & Brandt, 2011; Clark et al., 2014; Ganguli et al., 2010; Josephs et al., 2011), yet MCI occurs more broadly in neurodegenerative conditions. An optimal battery for diagnosing MCI in Parkinson's disease (PD) was recently described by Goldman et al. (2015), but to our knowledge, no guidelines have been proposed for ET. The unique pathological substrates of ET, in conjunction with an increased risk of dementia and the challenge of assessing cognition independently of motor functioning, raise the question as to how MCI should be assessed in ET patients. Our study seeks to determine a short panel of neuropsychological tests that could be reliably used in a clinical setting to evaluate mild decline in cognitive functioning that would be consistent with the degree of impairment seen in MCI.

As a gold standard for mild changes in cognition, we use the Clinical Dementia Rating scale (CDR). This semi-structured interview with the participant and an informant (friend or family member) is used to evaluate changes in cognition across six domains of functioning (memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care) and generates a global score between 0 and 3 (0 = No Impairment, 0.5 = Questionable Impairment, 1 = Mild Dementia, 2 = Moderate Dementia, and 3 = Severe Dementia; some scales also include a score of 4 in terminal illness) (Morris, 1997).

In conjunction with impaired performance on objective neuropsychological testing, the CDR score informs clinical diagnoses of normal cognition, MCI, or dementia (Aretouli et al., 2011; Clark et al., 2014; Ganguli et al., 2010; Josephs et al., 2011). A score of 0.5 on Memory ("Consistent slight forgetfulness; partial recollection of events; 'benign' forgetfulness") automatically results in a global CDR of 0.5 or higher, but impairment in at least two other categories (such as Orientation, Judgment & Problem Solving, Community Affairs, Home & Hobbies, and Personal Care) can also lead to a CDR of 0.5 (Morris, 1993), which is generally associated with a clinical diagnosis of MCI (Abner et al., 2017; Peterson, 1999).

In this study, we administered an extensive, 4-hr motor-free protocol of 19 neuropsychological tests to a population of older adults with ET. Using regression analysis, we first identified the best two tests per domain for predicting MCI, per the Movement Disorders Society (MDS) Task Force Criteria on assessing PD-MCI (Litvan et al., 2012), which outline the need for neuropsychological testing that includes at least two tests in each of the five cognitive domains. To determine the optimal method for detecting mild changes in cognition, however, we ran additional regression models to identify the subset of tests and the most relevant cognitive domains (regardless of number of tests per domain) for best predicting mild impairment in cognition as defined by CDR 0.5. Based on our results, we discuss the nature of cognitive dysfunction in ET, contrast it with what is understood about cognitive dysfunction in PD, and provide guidance for development of a brief neuropsychological protocol that we believe would inform the assessment and detection of MCI in ET.

METHODS

Study Design, Assessments, and Assignment of Diagnoses

Subjects were enrolled in a prospective, longitudinal study of cognitive function in ET (Clinical-Pathological Study of Cognitive Impairment in Essential Tremor [COGNET], NINDS R01NS086736) beginning July 2014. The study aims to clinically characterize a cohort of ET subjects across three assessments (baseline, 18 months, and 36 months). For these analyses, only baseline data (collected July 2014 – July 2016) were included. Subjects were recruited through advertisements on a study website and other websites (International Essential Tremor Foundation) that listed the following eligibility criteria: (1) diagnosis of ET, (2) $\geq$ 55 years old, (3) no deep brain stimulation surgery for ET, (4) willingness to perform study measures and be a brain donor. Yale University and Columbia University Internal Review Boards approved study procedures, and signed informed consent was obtained upon enrollment. Demographic and clinical data on age, gender, ethnicity, and education were collected at baseline. Medications with cognition-enhancing, cognition-decreasing, and mood-modulating effects were noted.

The cognitive test battery, designed by a neuropsychologist (S.C.) specifically for this study, purposefully minimizes any tests involving motor abilities that could disadvantage ET subjects with moderate or severe tremors. In addition to the Mini-Mental State Examination [MMSE (Folstein, Folstein, & McHugh, 1975)], the Montreal Cognitive Assessment [MoCA (Nasreddine et al., 2005)], the Wechsler Test of Adult Reading [WTAR (Wechsler, 2001)], the NEO Personality Inventory (McCrae & Costa, 2010), and the Geriatric Depression Scale [GDS (Yesavage et al., 1986)], the test battery included assessments across five domains: *Attention* [Oral Symbol-Digit Modalities Test (SDMT) (Smith, 1982)], Digit Span Forward (Wechsler, 1997)]; *Executive Function* [Digit Span Backward (Wechsler, 1997), Delis-Kaplan Executive Function System (D-KEFS), Verbal Fluency Test (VFT), Color-Word Interference (CW), Sorting, and

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

20-Questions (Delis, Kaplan, & Kramer, 2001)]; *Visuospatial Abilities* [Benton Judgment of Line Orientation (JLO) (Benton, Sivan, des Hamsher, Varney, & Spreen, 1994), Benton Facial Recognition Test (BFRT) (Benton & Van Allen, 1968), Wechsler Adult Intelligence Scale IV (WAIS-IV), Visual Puzzles (Wechsler, 1997)]; *Language* [Multilingual Aphasia Examination (MAE), Token Test (Benton, des Hamsher, Rey, & Sivan, 1994), Boston Naming Test (BNT) (Kaplan, Goodglass, & Weintraub, 1983)]; and *Memory* [California Verbal Learning Test (CVLT-II) (Delis, Kramer, Kaplan, & Ober, 2000), Wechsler Memory Scale Revised (WMS-R), Logical Memory (LM) (Wechsler, 1987) and Verbal Paired Associates (VPA) (Wechsler, 2008)].

In-person assessments at subjects' homes were conducted by trained personnel and consisted of a clinical questionnaire; 19 neuropsychological tests; questionnaires evaluating mood, sleep, tremor experience, and physical activity; and a videotaped neurological examination. Using published normative data, participant raw scores were converted to *Z*-scores according to age, gender, and/or education.

The videotaped neurological examination was reviewed by a movement disorders neurologist (E.D.L., D.R.). Videotaped kinetic or postural tremor were rated (0–3) on 12 items, and total tremor score (range 0–36) was calculated. Additionally, ET diagnosis was confirmed using the Washington Heights-Inwood Genetic Study of ET (WHIGET) diagnostic criteria [moderate or greater amplitude kinetic tremor (tremor rating $\geq$ 2) during three or more tests or head tremor, in the absence of PD, dystonia, or another cause (Louis et al., 1997)] which are reliable (Louis, Ford, & Bismuth, 1998) and valid (Louis et al., 1999).

Designated informants were queried by means of telephone regarding the participant's level of functioning in the six CDR domains (Morris, 1997) and completed the Neuropsychiatric Inventory (Cummings et al., 1994), Frontal Behavioral Inventory (Kertesz, Davidson, & Fox, 1983), and Lawton Instrumental Activities of Daily Living Scale (Lawton & Brody, 1969). Considerations of tremor disability were not included in CDR calculations. If an informant was not available, CDR was determined through participant self-report and examiner impression. CDR was confirmed and a cognitive diagnosis was determined during diagnostic case conference with trained experts (E.D.H., S.C.).

Three primary cognitive diagnoses were assigned using clinical judgment and diagnostic specifications: (1) *Normal Cognition*; (2) *MCI* (CDR 0.5 and impairment on 2 MCI-designated tests); and (3) *Dementia* (CDR $\geq$ 1 and impairment in multiple domains). Impairment on a single test was defined as a *Z*-score $\leq$ -1.5. Normal cognition included: *No impairment* (CDR 0, no impairment on any test); *Impairment of unlikely clinical significance* (CDR 0, impairment on 1 test); *Impairment of possible clinical significance* (CDR 0 or 0.5, impairment in $\geq$ 1 test but not meeting operational criteria for MCI); *Questionable or Isolated Functional Impairment* (CDR 0.5, no impairment on any neuropsychological test).

Regarding diagnosis of MCI, individuals were classified as *single* or *multi-domain* and *amnestic* (a-MCI) or *non-amnestic* (na-MCI). As described by Collins et al. (2017), specific tests in each domain were *a priori* selected for diagnosis of MCI (Table 1) based on: (1) relative purity of measurement for the construct under evaluation (e.g., in the spatial domain, JLO, given its lesser demand on executive functioning than Visual Puzzles); (2) demonstrated utility of measures in previous studies; and (3) general availability of the measure to researchers who wish to replicate findings. Selecting specific tests in each domain also prevented over-sampling of domains with more sub-scores generated from a single test (e.g., immediate and delayed memory from a memory test as compared to a single score from a naming test).

Exclusion Criteria and Statistical Analysis

Exclusion criteria for these analyses included: (1) diagnosis of dementia (CDR $\geq$ 1; $n = 20$), (2) diagnosis of cognitive impairment related to substance use ($n = 3$), (3) diagnosis of PD or dystonia based upon videotaped neurological examination ($n = 13$). Demographics for remaining participants (CDR 0 or CDR 0.5; $n = 196$) were assessed for normality (Kolmogorov-Smirnov test) and analyzed with appropriate statistical tests (Chi-Square, Fisher exact test, Mann-Whitney *U* test, or *t* test). Speed-based tests with potential for voice tremor-interference (VFT, CW, SDMT) were compared. Analyses were completed using SPSS24, SAS 9.4, or R Software.

Predictive Model Analyses

Logistic models with single test predictors examined associations between group membership (CDR 0 and CDR 0.5) and each predictor individually, while logistic models with multiple predictors assessed simultaneous effect of predictors. To describe the goodness-of-fit of models, classification accuracy [area under receiver operating characteristics (ROC) curve, AUC] and generalized R-squared for predictability of individual outcome were calculated for all models. A schematic of this process is shown in Figure 1.

Cognitive domain selection

Neuropsychological tests were grouped by cognitive domain (Table 1). Aggregate scores for each domain were calculated by averaging individual test *Z*-scores within each domain, tested for normality, and compared using two-tailed *t* test or Mann-Whitney *U* Test.

A logistic forward step-wise regression procedure was conducted on a sample of participants with complete data on all test variables ($n = 151$). The procedure selected test variables based on a preset Wald test significance level (e.g., $\alpha = 0.15$) for a variable being added into or kept in the model. Domain mean scores ($n = 151$) were fit to a model (Model 1), from which significant domains were used as variables in an additional model (Model 2). Each domain mean was also isolated and fit to a logistic model to determine its strength as an independent predictor of CDR 0.5 (Model 3).

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

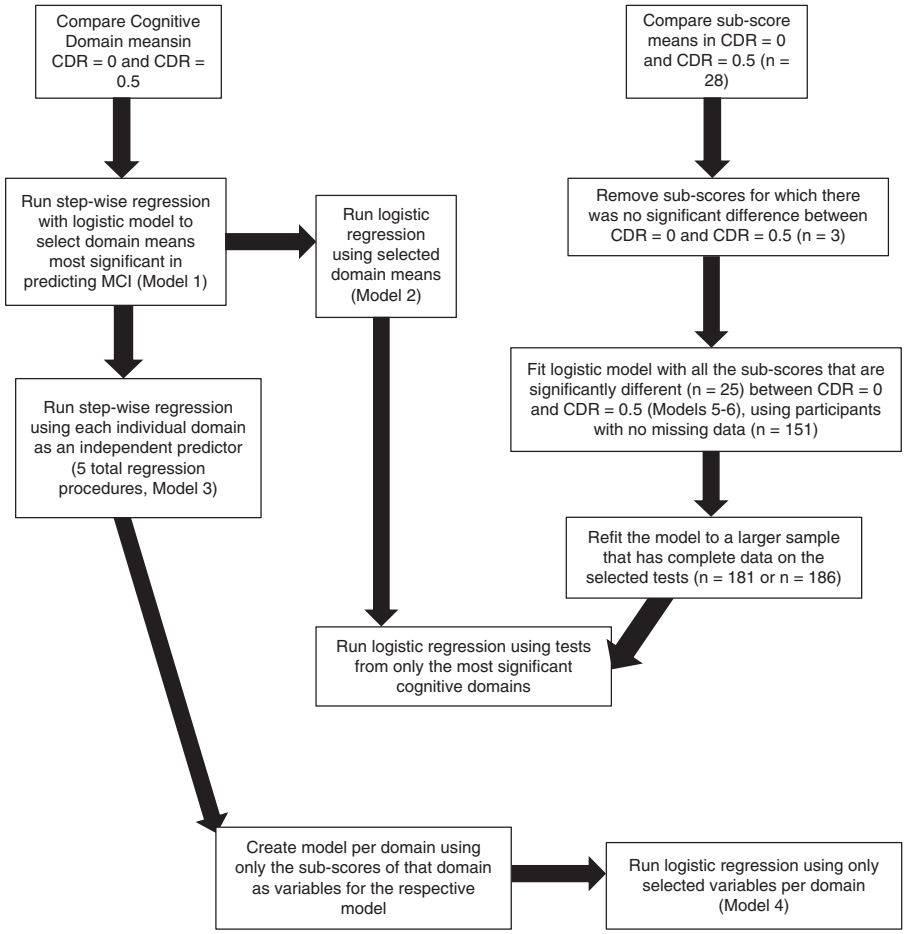

Fig 1. Predictive model analysis method. Procedure for model-based selection procedure to select five variables and assess significance of cognitive domains in the model.

Individual domain models

Logistic models were created for each domain individually with domain sub-scores as variables, in alignment with MDS Task Force Criteria. All sub-scores were included such that there were at least 2 tests per domain, including tests for which there was no significant difference between CDR groups. Importance of sub-scores was ranked by magnitude of estimated coefficient standardized by standard error. Significant sub-scores were fit as variables to domain-specific models (Model 4).

Test selection

Means for each of the 28 sub-scores (from different aspects of 19 tests) were calculated and compared using Mann-Whitney *U* Tests. Scores for which there was no significant difference ($p < .05$) between CDR groups were excluded as candidates from subsequent model-based variable selection. All significant sub-scores were fit to a model to predict CDR 0.5 (Models 5–6). The model of selected variables, fit using a sub-sample with complete data on the candidate tests ($n = 151$), was applied to a larger sample that had complete data on the selected variables.

Using Model to Predict Subtle Diagnostic Categories

Finally, the model of combined test scores with highest AUC was applied within each level of function (CDR 0 and 0.5) to evaluate the extent to which it could distinguish between subtle diagnostic classifications. Participant test-scores, weighted according to model coefficients, were fit to the logistic regression equation, yielding estimated probability of predicting impairment, which was compared using Spearman's correlation test (probability *vs*. diagnostic category severity) and *t* tests (mean probability).

RESULTS

Sample Characteristics

Table 2 demonstrates sample characteristics ($n = 196$). CDR 0 ($n = 148$) and CDR 0.5 ($n = 48$) groups did not significantly differ by gender, education, race, ethnicity, tremor score, ET duration, age of tremor onset, presence of voice tremor, number using cognition-decreasing or mood-modulating medications, or GDS ($p > .05$). The groups differed by age and number

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

Table 1. Cognitive tests currently in use for diagnosing MCI

Domain	Test	Sub-score	In use in current battery	Selected by Model 6
Memory	California Verbal Learning Test (CVLT-II)	Long-Delay Free Recall	*Yes*	No
		Total Recall	No	*Yes*
	Logical Memory	LM-I	No	No
		LM-II	*Yes*	*Yes*
	Verbal Paired Associates	VPA-I	No	*Yes*
		VPA-II	No	No
Executive Function	Verbal Fluency Test (VFT)	Letter Fluency	*Yes*	No
		Category Fluency	No	No
		Category Switching Fluency	No	*Yes*
		Category Switching Accuracy	No	No
	Color-Word Interference Test (CW)	Inhibition	*Yes*	*Yes*
		Naming	No	No
		Reading	No	No
		Switching	No	No
	Sorting	Confirmed Correct Sorts	*Yes*	No
		Free Sort Description Score	No	No
		Recognition Description Score	No	No
	20-Questions	Initial Abstraction Score	No	No
		Weighted Achievement Score	*Yes*	No
		Total Questions Asked	No	No
	Digit Span Backward		No	No
Attention	Symbol Digit Modalities Test (SDMT)		*Yes*	No
	Digit Span Forward		*Yes*	No
Visuospatial Abilities	Benton Facial Recognition Test (BFRT)		*Yes*	No
	Judgment of Line Orientation (JLO)		*Yes*	No
	Visual Puzzles		No	No
Language	Boston Naming Test (BNT)		*Yes*	No
	MAE Token Test		*Yes*	No

taking cognition-enhancing medications (Table 2); those taking cognition-enhancing medications (CDR 0.5) did not differ in domain means compared to those who do not.

Among the 28 neurological sub-scores (Table 3), 24 had incomplete data, with 181–195 participants per test, resulting in 48 subjects with missing data on at least one test. To reduce the impact of missing data, CW Reading, BFRT, and MAE Token Test were excluded from variable selection because they did not significantly differ between CDR groups ($p > .10$). The 25 remaining scores had complete data from 151 participants (123 NC, 28 MCI). There was no significant difference in scores on the CDR or speed based tasks as a function of voice tremor ($p > .05$, data not shown). An exception was the SDMT for which those with voice tremor appeared to perform better than those without. This effect is, therefore, unlikely to be due to presence or absence of voice tremor.

Cognitive Domain Selection

All cognitive domain scores were significantly different ($p < .05$) between CDR groups, except for the visuospatial domain ($p = .553$). Table 4 details results of the domain selection procedure; domain means for Memory and Executive Function were determined to be significantly different across groups (Model 1). These variables were then applied to an additional model (Model 2; AUC = 85.5%; $R^2 = 0.4569$). Only memory was seen to have a good fit as a domain independently predicting CDR 0.5 (Model 3; AUC = 86.2%; $R^2 = 0.4739$).

Individual Domain Models

Results for individual domain models with two significant tests per domain ($\alpha = 0.1$) are found in Table 5 (Model 4). Two tests were selected per domain for Attention, Visuospatial, and Language, and three were chosen per domain for Memory and Executive Function. High AUC was seen in only Memory (86.2%) and Executive Function (80.6%).

Test Selection

Applying the logistic model-based stepwise selection procedure to the 151 subject sub-set, four tests (Model 5) were selected at $\alpha = 0.10$: CVLT-II Total Recall, LM-II, VPA-I,

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

Table 2. Clinical and cognitive characteristics of sample

		CDR 0 $n = 148$	CDR 0.5 $n = 48$	p-Value
Demographics	Age, y	77.46 ± 9.69	82.38 ± 7.60	**.000**[c]
	Gender, *n* (% women)	95 (64.1)	27 (56.3)	.324[a]
	Education, y	15.79 ± 2.70	15.18 ± 2.99	.125[b]
	Race, *n* (% white)	141 (95.3)	46 (95.8)	.255[a]
	Ethnicity, *n* (% Non-Hispanic)	144 (97.3)	45 (93.8)	.430[a]
Tremor Features	Washington Heights-Inwood Total Tremor Score (range 0–36)	20.51 ± 5.85	20.86 ± 5.11	.586[b]
	ET duration, y	38.08 ± 21.97	40.65 ± 22.39	.474[b]
	Age of tremor onset, y	39.28 ± 21.77	41.59 ± 23.94	.600[b]
	Voice Tremor, *n* (%)	72 (48.6)	23 (47.9)	.930[a]
Medications	Cognition-Enhancing, *n* (%)	0 (0.0)	6 (12.5)	**.001**[a]
	Cognition-Decreasing, *n* (%)	76 (51.4)	28 (58.3)	.400[a]
	Mood-Modulating, *n* (%)	43 (29.1)	18 (37.5)	.272[a]
Cognitive and Neuropsychological Features	CDR Completed by informant, *n* (%)	102 (68.9)	39 (81.3)	.098[a]
	Lawton Instrumental Activities of Daily Living Scale (range 0–8)	7.64 ± 0.95	6.65 ± 1.70	**.000**[b]
	MMSE Sum (range 0–30)	28.99 ± 2.02	26.92 ± 3.79	**.000**[b]
	MoCA Sum (range 0–30)	25.80 ± 2.67	22.10 ± 2.69	**.000**[b]
	Memory Domain, Z-score	0.23 ± 0.65	−0.90 ± 0.77	**.000**[b]
	Attention Domain, Z-score	−0.10 ± 0.75	−0.58 ± 0.73	**.000**[d]
	Language Domain, Z-score	0.18 ± 0.46	−0.08 ± 0.63	**.000**[c]
	Visuospatial Domain, Z-score	0.59 ± 0.64	0.28 ± 0.54	.553[d]
	Executive Domain, Z-score	0.26 ± 0.51	−0.42 ± 0.61	**.000**[d]
	Geriatric Depression Scale (GDS, range 0-30)	6.13 ± 4.91	7.15 ± 5.76	.312[b]
Cognition-Enhancing Medications (CDR 0.5)	Domain	Taking	Not taking	p-Value
	Memory	−1.11 ± 0.66	−0.87 ± 0.78	.42[b]
	Executive Function	−0.74 ± 0.44	−0.38 ± 0.62	.13[b]
	Attention	−0.67 ± 0.78	−0.57 ± 0.73	.80[b]
	Visuospatial	−0.01 ± 0.71	0.32 ± 0.63	.42[b]
	Language	−0.26 ± 0.63	−0.05 ± 0.63	.44[b]

Note. All values are mean ± *SD* or number (percentage). Significant values are bold.
n = number, y = years.
[a]Chi square test.
[b]Mann-Whitney U test.
[c]Two-sample t- test with unequal variances (Levene's test for equality of variances $p < .05$).
[d]Two-sample t- test with equal variances (Levene's test for equality of variances $p > .05$).

VFT Switching (AUC = 86.7%; $R^2 = 0.487$). Five tests (Model 6) were selected at $\alpha = 0.15$: CVLT-II Total Recall, LM-II, VPA-I, VFT Switching, CW Inhibition (AUC = 86.9%; $R^2 = 0.504$) (Table 6; Figure 2).

When combining information from all domain analyses and individual test models, choosing only tests from the most significant domains (Memory and Executive Function) as variables for selection, the same five tests were chosen at $\alpha = 0.2$ as were chosen at $\alpha = 0.15$: CVLT-II Total Recall, LM-II, VPA-I, VFT Switching, and CW Inhibition (AUC = 86.8%; $R^2 = 0.502$; $n = 181$). At $\alpha = 0.1$, four tests were chosen: CVLT-II Total Recall, LM-II, VPA-I, and CW Inhibition (AUC = 86.7%; $R^2 = 0.487$; $n = 183$).

When applying the models to all participants with complete data on selected tests (Figure 3), the model with four selected tests (Model 5) remained a good fit (AUC = 86.7%; $R^2 = 0.495$; $n = 186$), as did the model with five selected tests (Model 6; AUC = 86.8%; $R^2 = 0.502$; $n = 181$).

In all models, increased odds of CDR 0.5 were associated with lower scores.

Using Model to Predict Cognitive Categories within CDR Level

Model 6 ($\alpha = 0.15$; five tests: CVLT-II Total Recall, LM-II, VPA-I, VFT Switching, CW Inhibition), with the highest AUC, was used for calculating probability of predicting cognitive categories within CDR 0 and CDR 0.5 (Table 7). Within CDR group, the model detected probability differences between cognitive categories, and within CDR 0, probability of impairment positively correlated with diagnostic category severity. Within CDR 0, the model distinguished *impairment*

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

Table 3. Neuropsychological tests by domain

Domain	Test/sub-score		CDR 0 Mean ± *SD* (*n*)	CDR 0.5 Mean ± *SD* (*n*)	*p*-Value
Memory	California Verbal Learning Test (CVLT-II)	Long-Delay Free Recall	−0.24 ± 0.74 (146)	−1.13 ± 1.09 (48)	**< .0001**[a]
		Total Recall	0.045 ± 1.03 (147)	−1.18 ± 1.08 (48)	**< .0001**[a]
	Logical Memory-Revised	LM-I	0.12 ± 0.95 (146)	−0.88 ± 1.09 (48)	**< .0001**[a]
		LM-II	0.17 ± 0.92 (148)	−0.97 ± 1.21 (47)	**< .0001**[a]
	Verbal Paired Associates	VPA-I	0.58 ± 0.91 (147)	−0.54 ± 1.00 (46)	**< .0001**[a]
		VPA-II	0.45 ± 0.99 (146)	−0.65 ± 1.05 (46)	**< .0001**[a]
Executive Function	Verbal Fluency Test (VFT)	Letter Fluency	0.31 ± 1.07 (148)	−0.35 ± 1.16 (48)	**.002**[a]
		Category Fluency	0.34 ± 1.03 (148)	−0.45 ± 1.15 (48)	**< .0001**[a]
		Category Switching Fluency	0.35 ± 1.13 (147)	−0.75 ± 1.19 (46)	**< .0001**[a]
		Category Switching Accuracy	0.20 ± 1.12 (147)	−0.73 ± 1.15 (46)	**< .0001**[a]
	Color-Word Interference Test (CW)	Inhibition	0.067 ± 0.89 (144)	−0.28 ± 1.06 (43)	**.002**[a]
		Naming	−0.28 ± 1.05 (146)	−0.82 ± 1.21 (47)	**.010**[a]
		Reading	−0.12 ± 0.97 (146)	−0.36 ± 1.04 (47)	.108[a]
		Switching	0.30 ± 0.97 (141)	−0.72 ± 1.42 (42)	**< .0001**[a]
	Sorting	Confirmed Correct Sorts	0.47 ± 0.79 (146)	−0.074 ± 0.73 (45)	**.0004**[a]
		Free Sort Description Score	0.29 ± 0.87 (146)	−0.21 ± 0.95 (45)	**.005**[a]
		Recognition Description Score	0.11 ± 0.99 (145)	−0.55 ± 0.83 (44)	**.0002**[a]
	20-Questions	Initial Abstraction Score	0.42 ± 1.00 (147)	−0.00015 ± 0.93 (46)	**.015**[a]
		Weighted Achievement Score	0.59 ± 1.05 (147)	−0.20 ± 1.40 (46)	**.0006**[a]
		Total Questions Asked	0.58 ± 0.89 (146)	−0.17 ± 1.31 (45)	**.0006**[a]
	Digit Span Backward		0.076 ± 0.97 (147)	−0.42 ± 0.88 (46)	**.003**[a]
Attention	Symbol Digit Modalities Test (SDMT)		0.26 ± 0.95 (148)	−0.34 ± 1.02 (48)	**< .0001**[b]
	Digit Span Forward		−0.47 ± 1.11 (147)	−0.84 ± 0.87 (46)	**.042**[a]
Visuo-spatial	Benton Facial Recognition Test (BFRT)		0.90 ± 1.04 (145)	0.88 ± 1.02 (47)	.746[a]
	Judgment of Line Orientation (JLO)		0.67 ± 0.88 (146)	0.21 ± 0.98 (47)	**.006**[a]
	Visual Puzzles		0.18 ± 0.95 (136)	−0.25 ± 0.77 (45)	**.012**[a]
Language	Boston Naming Test (BNT)		−0.24 ± 0.74 (148)	−0.69 ± 1.00 (47)	**.013**[a]
	MAE Token Test		0.61 ± 0.43 (146)	0.58 ± 0.64 (46)	.746[a]

Note. Values are mean ± *SD* (*n*). Significant values are bold.
[a]Mann-Whitney U-test.
[b]Two-tailed t-test.

of possible clinical significance, from *no impairment* and *impairment of unlikely clinical significance*. Within CDR 0.5, the model distinguished between *impairment of possible clinical significance* versus *questionable or isolated functional impairment* and a-MCI *versus* na-MCI.

DISCUSSION

Overview

Our study identifies five neuropsychological tests that are sensitive to mild cognitive problems in ET, defined by CDR 0.5. Results inform procedures for detecting cognitive impairment in ET and provide evidence for a targeted test battery. Although we identified two tests within each of the five cognitive domains in our individual domain models to predict CDR 0.5 (Model 4), which complies with the MDS Task Force Criteria for diagnosing PD-MCI (Litvan et al., 2012), our domain-based models (Models 1–3) were most accurate in identifying those with CDR 0.5 when using tests within Memory and Executive Function domains. In a separate test-based model (Model 6), the selection process identified five test scores, again within Memory and Executive Function domains, that were most accurate in identifying those with CDR 0.5: VPA-I (Immediate), CVLT-II Total

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

Table 4. Cognitive Domain selection

	Domain means[a]							
	All domains (Model 1)		Significant domains (Model 2)		Domain as a predictor[b] (Model 3)			
	Coefficient (b ± *SE*)	*p*-Value	Coefficient (b ± *SE*)	*p*-Value	Model	Coefficient (b ± *SE*)	AUC, *n*	R^2
Memory	−2.03 ± 0.48	**< .0001**	−1.93 ± 0.4342	**< .0001**	3a	−2.25 ± 0.35	86.2% (188)	0.4739
Executive Function	−1.19 ± 0.64	**.066**	−1.02 ± 0.52	**.048**	3b	−2.16 ± 0.47	77.1% (167)	0.2598
Attention	0.10 ± 0.41	.80	–	–	3c	−0.88 ± 0.25	68.9% (193)	0.1047
Visuospatial Abilities	0.24 ± 0.42	.57	–	–	3d	−0.83 ± 0.26	65.8% (179)	0.0928
Language	−0.17 ± 0.39	.66	–	–	3e	−0.64 ± 0.20	62.5% (195)	0.0775
AUC, *n*	86.5% (151)		85.5% (162)					
R^2	0.4770		0.4569					

Note. Coefficients (b ± *SE*) are standard parameter estimate ± Wald error. Significant values are in bold (based on α = 0.10).
AUC = area under ROC curve for classification accuracy.
[a]1 Domain model fit with all five domains as variables.
[b]Individual domain models fit with each domain as a variable for the respective Model 3a-e.

Table 5. Individual domain models (Model 4)

Domain	Test	Sub-score	Coefficient (b ± *SE*)	Ranking of Test	R^2	AUC
Memory (Model 4a)	California Verbal Learning Test (CVLT-II)	Long-Delay Free Recall	–	–	0.4670	86.2%
		Total Recall	−0.61 ± 0.18**	3		
	Logical Memory	LM-I	–	–		
		LM-II	−0.64 ± 0.23**	2		
	Verbal Paired Associates	VPA-I	−0.88 ± 0.27**	1		
		VPA-II	–	–		
Executive Function (Model 4b)	Verbal Fluency Test (VFT)	Letter Fluency	-0.52 ± 0.23*	2	0.3268	80.6%
		Category Fluency	–	–		
		Category Switching Fluency	−0.67 ± 0.19**	1		
		Category Switching Accuracy	–	–		
	Color-Word Interference Test (CW)	Inhibition	–	–		
		Naming	–	–		
		Reading	–	–		
		Switching	−0.50 ± 0.18**	3		
	Sorting	Confirmed Correct Sorts	–	–		
		Free Sort Description Score	–	–		
		Recognition Description Score	–	–		
	20-Questions	Initial Abstraction Score	–	–		
		Weighted Achievement Score	–	–		
		Total Questions Asked	–	–		
	Digit Span Backward		–	–		
Attention (Model 4c)	Symbol Digit Modalities Test (SDMT)		−0.58 ± 0.18**	1	0.1128	69.8%
	Digit Span Forward		−0.30 ± 0.18*	2		
Visuospatial Abilities (Model 4d)	Benton Facial Recognition Test (BFRT)		–	–	0.0934	65.9%
	Judgment of Line Orientation (JLO)		−0.37 ± 0.20	2		
	Visual Puzzles		−0.47 ± 0.23*	1		
Language (Model 4e)	Boston Naming Test (BNT)		−0.60 ± 0.20*	1	0.0689	61.7%
	MAE Token test		−0.05 ± 0.35	2		

Note. Selection criteria α = 0.10. Coefficients (b ± *SE*) are standard parameter estimate ± Wald error.
* $p < .05$, ** $p < .01$. Ranking is by |bl/Wald error.

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

Table 6. Test selection

	Selected from 25 sub-tests				Fit to larger sample			
	Model 5 ($n = 151$)		Model 6 ($n = 151$)		Model 5 ($n = 186$)		Model 6 ($n = 181$)	
Parameter	Coefficient (b ± *SE*)	*p*-Value	Coefficient (b ± *SE*)	*p*-Value	Coefficient (b ± *SE*)	*p*-Value	Coefficient (b ± *SE*)	*p*-Value
Intercept	−2.15 ± 0.38	< 0.0001	−2.06 ± 0.38	< 0.0001	−1.81 ± 0.30	< 0.0001	−1.77 ± 0.30	< 0.0001
CVLT-II Total Recall	−0.56 ± 0.22	0.0096	−0.50 ± 0.22	0.022	−0.60 ± 0.18	0.004	−0.48 ± 0.19	0.009
VPA-I	−0.88 ± 0.34	0.010	−0.86 ± 0.35	0.013	−0.74 ± 0.27	0.007	−0.78 ± 0.28	0.006
LM-II	−0.65 ± 0.30	0.033	−0.70 ± 0.31	0.025	−0.60 ± 0.25	0.016	−0.62 ± 0.23	0.015
VFT Switching	−0.46 ± 0.25	0.068	−0.40 ± 0.25	0.120	−0.59 ± 0.21	0.006	−0.50 ± 0.22	0.023
CW Inhibition	–	–	−0.48 ± 0.32	0.134	–	–	−0.57 ± 0.25	0.022
Classification Accuracy (AUC)	86.7%		86.9%		86.7%		86.8%	
R^2	0.4870		0.5042		0.4955		0.5017	

Note. Coefficients (b ± *SE*) are standard parameter estimate ± Wald Error
Model 1: α = 0.10, four parameters.
Model 2: α = 0.15, five parameters.

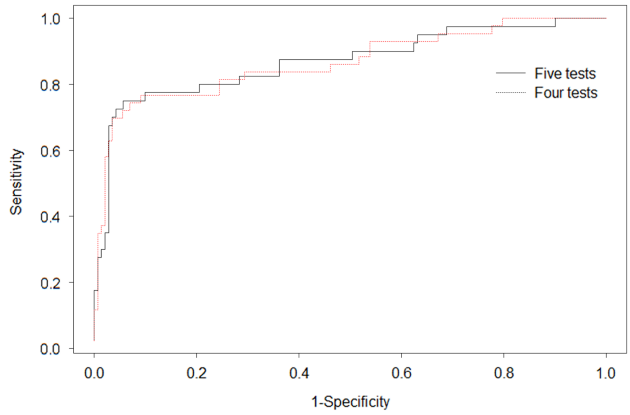

Fig 2. ROC curves based on the logistic models (Models 5–6) with predictors selected by different thresholds from all 25 sub-scores that were individually different between CDR = 0 and CDR = 0.5 ($n = 151$).

Recall, LM-II (Delayed), VFT Category Switching, and CW Inhibition.

Neuropsychological and Neuropathological Significance

The selection process identified performance in Memory and Executive Function domains as most sensitive to CDR 0.5, which is consistent with previous studies regarding cognitive dysfunction in ET (Benito-Leon & Louis, 2006; Gasparini et al., 2001; Higginson et al., 2008; Sahin et al., 2006). Executive Function has been identified as an area of impairment in ET (Frisina, Tse, Halbig, & Libow, 2009; Gasparini et al., 2001; Higginson et al., 2008; Kim et al., 2009; Lombardi et al., 2001; Passamonti et al., 2011; Sahin et al., 2006), although not necessarily the most common (Collins et al., 2017; Higginson et al., 2008; Sinoff & Badarny, 2014; Tröster et al., 2002). Executive deficit in ET is presumed to reflect the cerebello-thalamo-cortical basis of ET (Deuschl, Wenzelburger, Loffler, Raethjen, & Stolze, 2000; Middleton & Strick, 2000a, 2000b, 2001; Montgomery, Baker, Lyons, & Koller, 2000) as is seen in cerebellar cognitive affective syndrome (Janicki et al., 2013).

Test scores selected by this analysis are consistent with observations of impairment on both memory and executive measures in ET. Two of the selected memory scores were immediate recall measures from VPA (Wechsler, 2008) and CVLT-II (Delis et al., 2000). It has been suggested that CVLT-II can substitute for VPA, but Holster, Corsun-Ascher, Olivier, and Golden (2012) and others suggest that this substitution be made cautiously, as there is significant discrepancy between original VPA scores and CVLT-II converted VPA scores. Inclusion of both scores in this model may reflect the different processes tapped when recalling individual words (implicit free-recall) *versus* words stored as pairs (explicit associative learning) (Miller et al., 2012; Thiruselvam, Vogt, & Hoelzle, 2015).

The third memory measure selected, the LM subtest (Wechsler, 1987), measures narrative episodic memory and aligns with the Unified Data Set (UDS) 2.0, implemented nation-wide by the National Alzheimer's Coordinating Center in Alzheimer's Disease Research Centers (Chapman et al., 2016). Unlike CVLT-II and VPA-I measures selected by the model, the portion of LM was selected was delayed memory. As such, while inclusion of various verbal memory measures may appear to be redundant, the combination of these tests yields important information about clinical status likely reflecting relative differences in specific memory systems tapped by each test.

In the executive function domain, scores from VFT and CW tests (Delis et al., 2001) provide information regarding discrimination of individuals with CDR 0 *versus* 0.5. The specific VFT score chosen by the model, Category Switching, is a measure of the degree to which subjects can efficiently switch

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

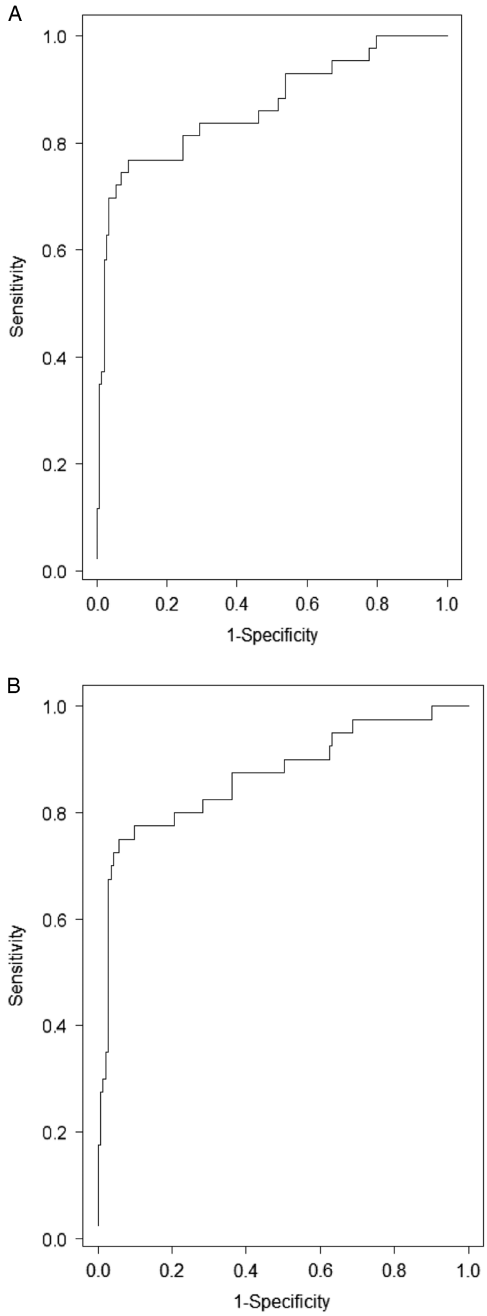

Fig 3. ROC curves based on the logistic models (Models 5–6) with predictors selected by different threshold, which fit to a larger sample with complete data on the selected tests: (a) four variables selected, $n = 168$; (b) five variables selected, $n = 181$.

back and forth between naming objects from different categories (e.g., "fruits" and "furniture"). Subjects with frontal lobe impairment or damage exhibit disproportionate impairment on Switching relative to Category Fluency (Delis et al., 2001), and because this task requires additional cognitive switching capabilities, discrepancies between this measure and standard measures of category fluency can indicate difficulties in cognitive flexibility.

The CW inhibition score reflects the subject's ability to inhibit the prepotent response of word reading in favor of the less automatic response, naming the ink color in which the word is printed. Poor performance on this measure is also considered to be an indicator of prefrontal (Vendrell et al., 1995) or frontostriatal (Koziol & Budding, 2009) dysfunction (Delis et al., 2001). Inclusion of this test, despite the small increment of significance it provides to models 5 and 6, makes the models more heterogenous as this is the only test which incorporates visual stimuli and does not require significant verbal ability.

It has been suggested that altered cerebellar-cortical pathways in ET are specifically involved in the executive control circuit that mediates focused attention in suppressing task irrelevant thoughts (Passamonti et al., 2011), which may underlie not only the executive tests selected, but selection of memory tests that have heavy attentional demands (Lombardi et al., 2001). Indeed, there is evidence of frontal-executive dysfunction in ET for memory tests that are not inherently organized (Lafo et al., 2015), and the immediate recall component of list learning tests such as CVLT-II and VPA appear to be particularly associated with executive dysfunction given their demands on organized and efficient retrieval of information (Tremont, Halpert, Javorsky, & Stern, 2000). However, formal tests of attention were not influential tests for detection of CDR 0.5, suggesting that deficits in attention alone cannot explain reduced performance on memory measures. Taken together, use of tests that combine both memory and executive function may be particularly vital for diagnosing cognitive impairment in ET.

Model as a Diagnostic Guideline

Of the five tests selected by the model, only two were used *a priori* to diagnose MCI: CW Inhibition and LM-II (Table 1). Model 6 was used to examine whether models could distinguish subtle gradations in diagnostic categories within CDR levels. Among those with CDR 0.5, the model differentiated between those who had no observable cognitive impairment despite their CDR rating and those whose cognitive profiles were considered to be *possibly clinically significant* (some signs of cognitive dysfunction but not meeting MCI criteria).

Additionally, the model distinguished between those whose cognitive profiles were considered to be *possibly clinically significant* and those diagnosed with MCI. This result, in conjunction with the fact that several of the tests selected by the model were not tests used to diagnose MCI, indicate that individuals with MCI performed worse on "non-MCI" tests as well. Although not tested directly, this finding lends support to the idea that tests selected *a priori* to diagnose MCI represent overall cognitive functioning. Lastly, among those with CDR 0.5, the model distinguished between a-MCI and na-MCI, showing that performance on the measures in this model is more sensitive to a-MCI cases. This result is unsurprising given that three of five tests in this model were memory tests and likely reflects the fact that most MCI cases (74.3%) were amnestic.

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

Table 7. Model as a predictor of cognitive impairment

CDR	Diagnostic category	Correlation[a] Coefficient (r_s)	Correlation[a] *p*-Value	Probability of impairment (%, mean ± *SD*)
0	Normal Cognition ($n = 65$)	0.462	**.000**	4.27 ± 5.00
	Impaired Performance with *Unlikely* Clinical Significance ($n = 55$)			11.2 ± 12.3
	Impaired Performance with *Possible* Clinical Significance ($n = 28$)			20.3 ± 21.4
0.5	Questionable or Isolated Functional Impairment ($n = 3$)	N/A		6.07 ± 4.81
	Impaired Performance with *Possible* Clinical Significance ($n = 10$)	− 0.571	**.000**	30.4 ± 26.8
	MCI (all, $n = 35$)			63.2 ± 22.9
	Amnestic MCI (single or multi-domain, $n = 26$)	N/A		71.0 ± 18.2
	Non-amnestic MCI (single or multi-domain, $n = 9$)			40.6 ±22.9

CDR	Comparison	*p*-Value[b]
0	Normal Cognition vs. Impaired Performance with *Unlikely* Clinical Significance	**.0026**
	Normal Cognition vs. Impaired Performance with *Possible* Clinical Significance	**.0006**
	Impaired Performance with *Unlikely* Clinical Significance vs. Impaired Performance with *Possible* Clinical Significance	**.048**
0.5	Impaired Performance with *Possible* Clinical Significance vs. Questionable or Isolated Functional Impairment	**.0275**
	Impaired Performance with *Possible* Clinical Significance vs. MCI (single or multi-domain)	**.00052**
	Amnestic MCI (single or multi-domain) vs. Non-amnestic MCI (single or multi-domain)	**.00025**

Note. Significant values ($p < .05$) are in bold.
SD = standard deviation.
[a]Spearman's correlation test between probability of impairment and diagnostic category severity for those with linear severity: all categories within CDR 0 and between impaired performance with *possible* clinical significance and MCI within CDR 0.5.
[b]Two-sample t-test comparing mean probability between the respective group.

Finally, the model was sensitive to subtle differences in diagnostic classifications within CDR 0. Specifically, the model was sensitive to *impairment of possible clinical significance* compared to those with strictly normal cognition and those with impairment considered to be of *unlikely clinical significance*. The model's ability to distinguish between subtle cognitive categories among individuals with no evidence of difficulty provides further evidence for the utility of these tests in characterizing cognitive functioning in individuals with ET. The model may, therefore, provide information to clinicians regarding the potential development of cognitive impairment in the future in the absence of daily functional difficulty.

ET-MCI *versus* PD-MCI

In PD-MCI, impairment has been noted in all domains, with emphasis on executive function, attention, and visuospatial function (Goldman & Litvan, 2012). Goldman et al. (2015) created a framework for an optimal PD-MCI battery based on impairment, defined by a cutoff score of 2 standard deviations (*SD*) below norms (Marras et al., 2013). They identified a neuropsychological test battery with 10 tests (2 per domain) that predicted PD-MCI with sensitivity of 81.3% and specificity of 85.7%. The tests chosen were (1) *Attention*: SDMT and Trail Making Test-A (TMT-A) (Reitan & Wolfson, 1993); (2) *Executive Function*: Clock Drawing (Goodglass & Kaplan, 1983) and Trail-Making Test-B (TMT-B) (Reitan & Wolfson, 1993)]; (3) *Language*: BNT and VFT Category Fluency, Animal Naming; (4) *Memory*: Free and Cued Selective Reminding Test (Grober & Buschke, 2009) and Figural Memory Learning and Delayed Recall (Wilson, Gilley, Bennett, Beckett, & Evans, 2000); (5) *Visuospatial Abilities*: JLO and MMSE Intersecting Pentagons (Bourke, Castleden, Stephen, & Dennis, 1995). While no tests directly overlap with those chosen by the ET-MCI model, identical batteries have yet to be given to both groups.

ET and PD have similar deficits in neuropsychological functioning relating to fronto-cerebellar circuits (basal ganglia circuit in PD and cerebello-thalamic-cortico loop in ET), which are implicated in attention, executive function, memory, and naming. However, there is disagreement regarding the extent to which the cognitive profiles of the two movement disorders differ (Puertas-Martin et al., 2016). Overall, PD groups show poorer performance in visuospatial tasks compared to ET groups (Gasparini et al., 2001; Higginson et al., 2008; Lombardi et al., 2001; Sanchez-Ferraro et al., 2017). Discrepancy in visuospatial impairment is reflected in high probability of detecting impairment using a single visuospatial test in PD (19.7–38.2%) in contrast to lack

https://doi.org/10.1017/S1355617718000747 Published online by Cambridge University Press

of significant difference in ET visuospatial domain scores between CDR 0 and CDR 0.5 ($p = .553$), as opposed to other domains. Visuospatial impairment is present in PD (Watson & Leverenz, 2010), likely caused by compromised functional basal ganglia loops that include the posterior parietal cortex (Middleton & Strick, 2000b).

These data suggest that the cognitive profiles of ET-MCI and PD-MCI indeed differ in the domains that are affected. Both manifest as impairment in executive function and attention domains, but in ET-MCI, impairment is emphasized in the memory domain, while in PD-MCI, impairment is emphasized in the visuospatial domain. Therefore, an optimal test battery for ET-MCI is likely different than one suited for PD-MCI.

Limitations

It is necessary to consider the speed of motor and verbal output required for neuropsychological tests when studying cognition in ET. While our battery was designed to reduce the reliance on rapid manual responses, there are several tests that required rapid verbal responses. However, it does not appear that performance on the latter tests was influenced by voice tremor.

Our analysis had several limitations inherent to diagnosis of cognitive impairment in ET. Although previous studies have identified CDR as stable over long periods of time (Williams, Roe, & Morris, 2009), CDR is less stable across intervals 1 year apart, especially for those with milder disease at baseline, older age, more underlying conditions contributing to cognitive decline, different informant, and different evaluators (Koepsell, Gill, & Chen, 2013). It is often the case that cognitive compromise is seen on formal neuropsychological testing before the CDR. However, the goal of the current study was to identify tests that are indicative of decrements in cognition, and the CDR is widely used to capture early changes that occur in the context of MCI.

Interview with the participant and the informant is necessary to understand the extent to which the participant is experiencing difficulty, and CDR is one of the few, if any, objective assessments of this. It is a subjective measure that may be susceptible to underestimations or overestimations by either informant or participant; such bias could be lessened by excluding self-CDR from analyses. It is worth noting that the model derived by comparing CDR 0 to CDR 0.5 was able to distinguish between cognitive classifications even among those with a CDR score of 0.

Participants in this study contacted us on their own volition. It is possible that participants who were concerned about developing cognitive impairment or had subjective cognitive complaints responded to advertisements, although in a study of unbiased ET cases ascertained directly from the population, cases were observed to perform more poorly compared to case-matched controls on neuropsychological evaluations (Benito-León, Louis, Bermejo-Pareja, & NEDICES Study Group, 2006). Since more individuals with CDR 0.5 were taking cognition-enhancing medications than those with CDR 0, it is possible that taking cognition-enhancing medication could confound our results; however, we observed that these individuals did not differ in neuropsychological performance compared to those with CDR 0.5 not taking cognition-enhancing medications.

There is disagreement regarding a distinction between ET-MCI and AD- (prodromal) MCI, but the tests chosen by the model suggest that there may be different aspects of memory affected. In ET, immediate memory measures (CVLT-II Total Recall, VPA-I) are implicated in MCI, whereas delayed memory measures are the best predictors of conversion from MCI to AD (Gainotti, Quaranta, & Vita, 2014). However, the profiles appear to share impairment in cognitive flexibility within executive function (Traykov, Rigaud, Cesaro, & Boller, 2007). Future analyses should include control, PD, and AD subjects as to further determine the extent to which cognitive profiles in ET are similar or dissimilar to those observed in these diseases.

Conclusion

We used a logistic regression selection procedure to select the best tests for predicting mild impairment in cognition, defined as CDR 0.5, with accuracy of 86.9%. This model is sensitive to mild changes in cognition and to subtle gradations in performance on neuropsychological testing within CDR level, as judged in the context of clinical consensus conference. Future analyses will include new models for predicting non-amnestic subtypes of ET-MCI (na-MCI) and models distinguishing ET-MCI and PD-MCI using the same neuropsychological battery. Once further data are collected for second and third intervals, we will be able to confirm this model with stable CDR. Future analyses of neuropathological data will address the neuropathological basis of cognitive heterogeneity in ET.

The neuropsychological assessment suggested by current analyses, in conjunction with independent CDR, can be performed by a neurologist, neuropsychologist, or trained associate. The inclusion of verbal-only tests raises the question of task-interference among these tests; however, evidence suggests that verbal tasks are not more susceptible to interference than non-verbal tasks (Williams, Sullivan, Morra, Williams, & Donovick, 2014; Williams & Donovick, 2008). Such an assessment lasts approximately 40 min (compared to 4 hr), but if time is limited, VPA-I, with the highest predictability, could be administered in 10 min. Optimally, administration would involve the following sequence: (1) immediate portions of CVLT-II and VPA, (2) LM immediate, (3) VFT and CW, and (4) LM delayed.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (grant number NINDSR01NS086736). This funding body played no role in the design of the study, the collection, analysis, and interpretation

of data, or the writing of the manuscript. No authors have conflicts of interest or competing financial interests.

REFERENCES

Abner, E.L., Kryscio, R.J., Schmitt, F.A., Fardo, D.W., Moga, D.C., Ighodaro, E.T., … Nelson, P.T. (2017). Outcomes after diagnosis of mild cognitive impairment in a large autopsy series. *Annals of Neurology*, *81*(4), 549–559.

Aretouli, E., Okonkwo, O.C., Samek, J., & Brandt, J. (2011). The fate of the 0.5s: Predictors of 2-year outcome in mild cognitive impairment. *Journal of the International Neuropsychological Society*, *17*(2), 277–288.

Benito-León, J., & Louis, E.D. (2013). Rate of cognitive decline during the premotor phase of ET: A prospective study. *Neurology*, *81*, 60–66.

Benito-León, J., Louis, E.D., Romero, J., Hernández-Tamames, J., Manzanedo, E., Álvarez-Linera, J., … Rocon, E. (2015). Altered functional connectivity in essential tremor: A resting-state fMRI study. *Medicine*, *94*(49), e1936.

Benito-León, J., Louis, E.D., Bermejo-Pareja, F., & NEDICES Study Group. (2006a). Elderly-onset essential tremor is associated with dementia. *Neurology*, *66*(10), 1500–1505.

Benito-León, J., Louis, E.D., Bermejo-Pareja, F., & NEDICES Study Group. (2006b). Population-based case-control study of cognitive function in essential tremor. *Neurology*, *66*(1), 69–74.

Benito-León, J., Louis, E.D., Mitchell, A.J., & Bermejo-Pareja, F. (2011). Elderly-onset essential tremor and mild cognitive impairment: A population-based study (NEDICES). *Journal of Alzheimer's Disease*, *23*(4), 727–735. doi:10.3233/JAD-2011-101572

Benton, A., & Van Allen, M. (1968). Impairment in facial recognition in patients with cerebral disease. *Cortex*, *4*, 344–358. doi:10.1016/S0010-9452(68)80018-8

Benton, A., des Hamsher, K., Rey, G., & Sivan, A. (1994). *Multilingual Aphasia Examination* (3rd ed.). Iowa City, IA: AJA Associates.

Benton, A., Sivan, A., des Hamsher, K., Varney, N., & Spreen, O. (1994). *Contributions to neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

Bermejo-Pareja, F., Louis, E.D., Benito-León, J., & NEDICES Study Group. (2007). Risk of incident dementia in essential tremor: A population-based study. *Movement Disorders*, *22*(11), 1573–1580.

Bourke, J., Castleden, C., Stephen, R., & Dennis, M. (1995). A comparison of clock and pentagon drawing in Alzheimer's disease. *Geriatric Psychiatry*, *10*(8), 703–705.

Chapman, K., Bing-Canar, H., Alosco, M., Steinberg, E., Martin, B., Chaisson, C., … Stern, R. (2016). Mini Mental State Examination and Logical Memory scores for entry into Alzheimer's disease trials. *Alzheimer's Research and Therapy*, *8*, 9. doi:10.1186/s13195-016-0176-z

Clark, D.G., Kapur, P., Geldmacher, D.S., Brockington, J.C., Harrell, L., DeRamus, T.P., … Marson, D.C. (2014). Latent information in fluency lists predicts functional decline in persons at risk for Alzheimer disease. *Cortex*, *55*, 202–218.

Collins, K., Rohl, B., Morgan, S., Huey, E., Louis, E., & Cosentino, S. (2017). Mild cognitive impairment subtypes in a cohort of elderly essential tremor cases. *Journal of the International Neuropsychological Society*, *23*, 1–10. doi:10.1017/S1355617717000170

Cummings, J.L., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D.A., & Gornbein, J. (1994). The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. *Neurology*, *44*, 2308–2314.

Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis Kaplan Executive Function System (DKEFS) Examiner's Manual*. San Antonio, TX: Pearson.

Delis, D., Kramer, J., Kaplan, E., & Ober, B. (2000). *California Verbal Learning Test - second edition (CVLT-II)*. San Antonio, TX: Psychological Corporation.

Deuschl, G., Wenzelburger, R., Loffler, K., Raethjen, J., & Stolze, H. (2000). Essential Tremor and cerebellar dysfunction clinical and kinematic analysis of intention tremor. *Brain*, *123*, 1568–1580.

Dogu, O., Sevim, S., Camdeviren, H., Sasmaz, T., Bugdayci, R., Aral, M., … Louis, E.D. (2003). Prevalence of essential tremor: Door-to-door neurologic exams in Mersin Province, Turkey. *Neurology*, *61*(12), 1804–1806.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Frisina, P., Tse, W., Halbig, T., & Libow, L. (2009). The pattern of cognitive-functional decline in elderly essential tremor patients: An exporatory-comparative study with Parkinson's and Alzheimer's disease patients. *Journal of American Medical Directors Association*, *10*, 238–242.

Gainotti, G., Quaranta, D., & Vita, M. (2014). Neuropsychological predictors of conversion from mild cognitive impairment to Alzheimer's disease. *Journal of Alzheimer's Disease*, *38*(3), 481–495. doi:10.3233/JAD-130881

Ganguli, M., Bilt, J.V., Lee, C.W., Snitz, B.E., Chang, C.C., Loewenstein, D.A., … Saxton, J.A. (2010). Cognitive test performance predicts change in functional status at the population level: The MYHAT Project. *Journal of the International Neuropsychological Society*, 761–770.

Gasparini, J., Bonifati, V., Fabrizio, E., Fabbrini, G., Brusa, L., Lenzi, G.L., … Meco, G. (2001). Frontal lobe dysfunction in essential tremor: A preliminary study. *Journal of Neurology*, *248*, 399–402.

Gauthier, S., Reisberg, B., Zaudig, M., Peterson, R.C., Ritchie, K., Broich, K., … Hampel, H. (2006). Mild cognitive impairment. *The Lancet*, *367*(9518), 1262–1270.

Goldman, J., & Litvan, I. (2012). Mild cognitive impairment in Parkinson's disease. *Minerva Medica*, *102*(6), 441–459.

Goldman, J., Holden, S., Ouyang, B., Bernard, B., Goetz, C.G., & Stebbins, G.T. (2015). Diagnosing PD-MCI by MDS Task Force criteria: How many and which neuropsychological tests? *Movement Disorders*, *30*, 402–406.

Goodglass, H., & Kaplan, E. (1983). *Clock Drawing Test: Assessment of aphasia and related disorders* (Vol. 2). Philadelphia: Lea and Febiger.

Grober, E., & Buschke, H. (2009). Genuine memory deficits in dementia. *Developmental Neuropsychology*, *3*, 13–36.

Higginson, C., Wheelock, V., Levine, D., King, D., Pappas, C., & Sigvardt, K. (2008). Cognitive deficits in essential tremor consistent with frontosubcortical dysfunction. *Journal of Clinical and Experimental Neuropsychology*, *30*, 760–765.

Holster, J., Corsun-Ascher, C., Olivier, T., & Golden, C. (2012). An evaluation of the CVLT-II substitution on the WMS-IV amongst a mixed clinical and nonclinical population. *Archives of Assessment Psychology*, *2*(1).

Janicki, S.C., Cosentino, S., & Louis, E.D. (2013). The cognitive side of essential tremor: What are the therapeutic implications? *Therapeutic Advances in Neurological Disorders*, *6*, 353–368.

Josephs, K.A., Whitwell, J.L., Weigand, S.D., Senjem, M.L., Boeve, B.F., Knopman, D.S., … Petersen, R.C. (2011). Predicting functional decline in behavioural variant frontotemporal dementia. *Brain*, *134*, 432–448.

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

Kertesz, A., Davidson, W., & Fox, H. (1983). Frontal behavioral inventory: Diagnostic criteria for frontal lobe dementia. *Canadian Journal of Neurological Sciences*, *24*(1), 29–36.

Kim, J.S., Song, I.U., Shim, Y.S., Park, J.W., Yoo, J.Y., Kim, Y.I., … Lee, K.S. (2009). Cognitive impairment in essential tremor without dementia. *Journal of Clinical Neurology*, *5*, 81–84.

Koepsell, T.D., Gill, D.P., & Chen, B. (2013). Stability of clinical etiologic diagnosis in dementia and mild cognitive impairment: Results from a multi-center longitudinal database. *American Journal of Alzheimer's Disease & Other Dementias*, *28*(8), 750–758.

Koziol, L.F., & Budding, D.E. (2009). Familiarity and novelty - Evaluating the frontostriatal system. In *Subcortical structures and cognition: Implications for neuropsychological assessment* (pp. 233–234). New York: Springer.

Lafo, J.A., Jones, J.D., Okun, M.S., Bauer, R.M., Price, C.C., & Bowers, D. (2015). Memory similarities between essential tremor and Parkinson's disease: A final common pathway? *Clinical Neuropsychology*, *29*(7), 985–1001. doi:10.1080/13854046.2015.1118553

Lawton, M., & Brody, E. (1969). Assessment of older people: Self-maintaining and isntrumental activities of daily living. *The Gerontologist*, *9*(3), 179–186.

Litvan, I., Goldman, J.G., Tröster, A.I., Schmand, B.A., Weintraub, D., Petersen, R.C., … Emre, M. (2012). Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force Guidelines. *Movement Disorders*, *27*(3), 349–356.

Lombardi, W., Woolston, D.J., Roberts, J.W., & Gross, R.E. (2001). Cognitive deficits in patients with essential tremor. *Neurology*, *57*, 785–790.

Louis, E.D. (2009). The essential tremors: A family of neurodegenerative disorders? *Archives of Neurology*, *66*, 1202–1208.

Louis, E.D. (2010). Functional correlates of lower cognitive test scores in essential tremor. *Movement Disorders*, *25*, 481–499.

Louis, E.D. (2014a). Re-thinking the biology of essential tremor: From models to morphology. *Parkinsonism & Related Disorders*, *20*(Suppl. 1), S88–S93.

Louis, E.D., & Rao, A.K. (2014b). Tandem gait performance in essential tremor patients correlates with cognitive function. *Cerebellum & Ataxias*, *1*, 19.

Louis, E.D., Benito-León, J., Vega-Quiroga, S., Bermejo-Pareja, F., & NEDICES Study Group. (2010). Cognitive and motor functional activity in non-demented community-dwelling essential tremor cases. *Journal of Neurology, Neurosurgery, and Psychiatry*, *81*(9), 997–1001.

Louis, E.D., Collins, K., Rohl, B., Morgan, S., Robakis, D., Huey, E.D., … Cosentino, S. (2016). Self-reported physical activity in essential tremor: Relationship with tremor, balance, and cognitive function. *Journal of the Neurological Sciences*, *336*, 240–245.

Louis, E.D., Ford, B., & Bismuth, B. (1998). Reliability between two observers using a protocol for diagnosign essential tremor. *Movement Disorders*, *13*, 287–293.

Louis, E.D., Kellner, S., Morgan, S., Collins, K., Rohl, B., Huey, E. D., … Cosentino, S. (2017). Cognitive dysfunction is associated with greater imbalance and falls in essential tremor. *Frontiers in Neurology*, *8*, 154.

Louis, E.D., Ottman, R., & Hauser, W.A. (1998). How common is the most common adult movement disorder? estimates of the prevalence of essential tremor throughout the world. *Movement Disorders*, *13*(1), 5–10.

Louis, E.D., Ottman, R., Ford, B., Pullman, S., Martinez, M., Fahn, S., … Hauser, W.A. (1997). The Washington Height-Inwood Genetic Study of Essential Tremor: Methodologic issues in essential-tremor research. *Neuroepidemiology*, *16*, 668–673.

Louis, E.D., Wendt, K.J., Albert, S.M., Pullman, S.L., Yu, Q., & Andrews, H. (1999). Validity of a performance-based test of function in essential tremor. *Archives of Neurology*, *56*, 841–846.

Mameli, F., Tomasini, E., Scelzo, E., Fumagalli, M., Ferrucci, R., Bertolasi, L., … Priori, A. (2013). Lies tell the truth about cognitive dysfunction in essential tremor: An experimental deception study with the guilty knowledge task. *Journal of Neurology, Neurosurgery, and Psychiatry*, *84*(9), 1008–1013.

Marras, C., Armstrong, M., Meaney, C., Fox, S., Rothberg, B., Reginold, W., … Duff-Canning, S. (2013). Measuring mild cognitive impairment in patients with Parkinson's disease. *Movement Disorders*, *28*(5), 626–633. doi:10.1002/mds.25426

McCrae, R., & Costa, P. (2010). *NEO inventories for the NEO Personality Inventory-3 (NEO-PI-3), NEO Five-Factor Inventory-3 (NEO-FFI-3), NEO Personality Inventory-Revised (NEO PI-R): Professional manual*. Lutz, FL: PAR.

Middleton, F., & Strick, P. (2000a). Basal ganglia and cerebellar loops: Motor and cognitive circuits. *Brain Research. Brain Research Reviews*, *31*, 236–250.

Middleton, F., & Strick, P. (2000b). Basal galgia output and cognition: Evidence from anatomical, behavioral, and clinical studies. *Brain Cognition*, *42*, 183–200.

Middleton, F., & Strick, P. (2001). Cerebellar projections to the prefrontal cortex of the primate. *Journal of Neuroscience*, *21*, 700–712.

Miller, J.B., Axelrod, B.N., Rapport, L.J., Hanks, R.A., Bashem, J. R., & Schutte, C. (2012). Substitution of California Verbal Learning Test, second edition for Verbal Paired Associates on the Wechsler Memory Scale, fourth edition. *The Clinical Neuropsychologist*, *26*(4), 599–608. doi:10.1080/13854046.2012.677478

Montgomery, E., Baker, K., Lyons, K., & Koller, W. (2000). Motor initiation and execution in essential tremor and Parkinson's disease. *Movement Disorders*, *15*, 511–515.

Morris, J. (1993). The clinical dementia rating (CDR): Current version and scoring rules. *Neurology*, *43*(11), 2412–2414.

Morris, J. (1997). Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *International Psychogeriatrics*, *9*(Suppl. 1), 173–176; discussion 177–178.

Nasreddine, Z., Phillips, N., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. doi:10.11/j.1532-5415.2005.5322.x

Passamonti, L., Novellino, F., Cerasa, A., Chiriaco, C., Rocca, F., Matina, M., … Quattrone, A. (2011). Altered cortical-cerebellar circuits during verbal working memory in essential tremor. *Brain*, *134*, 2274–2786.

Peterson, R. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*(3), 303–308.

Puertas-Martin, V., Villarejo-Galende, A., Fernandez-Guinea, S., Romero, J., Louis, E.D., & Benito-Leon, J. (2016). A comparison study of cognitive and neuropsychiatric features of essential

tremor and Parkinson's disease. *Tremor and Other Hyperkinetic Movements*, *6*, 431. doi:10.7916/D86H4HRN

Rao, A.K., Gillman, A., & Louis, E.D. (2014). Balance confidence and falls in nondemented essential tremor patients: The role of cognition. *Archives of Physical Medicine and Rehabilitation*, *95* (10), 1832–1837.

Rao, A.K., Uddin, J., Gillman, A., & Louis, E.D. (2013). Cognitive motor interference during dual-task gait in essential tremor. *Gait & Posture*, *38*(3), 403–409.

Reitan, R.M., & Wolfson, D. (1993). *The Trail-making Test* (Vol. 2). Tucson, AZ: Neuropsychology Press.

Sahin, H., Terzi, M., Ucak, S., Yapici, O., Basoglu, T., & Onar, M. (2006). Frontal functions in young patients with essential tremor: A case comparison study. *Journal of Neuropsychiatry and Clinical Neuroscience*, *18*, 64–72.

Sanchez-Ferraro, A., Benito-Leon, J., Louis, E., Cantador, I., Hernandez-Gallego, J., Puertas-Martin, V., … Bermejo-Pareja, F. (2017). Cognition in non-demented Parkinson's disease vs essential tremor: A population-based study. *Acta Neurologica Scandanavia*, *136*(5), 393–400. doi:10.1111/ane.12752

Sinoff, G., & Badarny, S. (2014). Mild cognitive impairment, dementia, and affective disorders in essential tremor: A prospective study. *Tremor and Other Hyperkinetic Movements (New York, N.Y.)*, *4*, 227.

Smith, A. (1982). *Symbol Digit Modalities Test*. Los Angeles: Western Psychological Services.

Thawani, S.P., Schupf, N., & Louis, E.D. (2009). Essential tremor is associated with dementia: Prospective population-based study in New York. *Neurology*, *73*(8), 621–625.

Thiruselvam, I., Vogt, E.M., & Hoelzle, J.B. (2015). The interchangeability of CVLT-II and WMS-IV Verbal Paired Associates Scores: A slightly different story. *Archives of Clinical Neuropsychology*, *30*(3), 248–255.

Traykov, L., Rigaud, A.S., Cesaro, P., & Boller, F. (2007). Le déficit neuropsychologique dans la maladie d'Alzheimer débutante. *L'Encéphale*, *33*(3), 310–316. doi:ENC-6-2007-33-3-0013-7006-101019-200730033

Tremont, G., Halpert, S., Javorsky, D., & Stern, R. (2000). Differential impact of executive dysfunction on verbal list learning and story recall. *Clinical Neuropsychology*, *14*(3), 295–302. doi:10.1076/1385-4046(200008)14:3;1-P;FT295

Tröster, A.I., Woods, S.P., Fields, J.A., Lyons, K.E., Pahwa, R., Higginson, C.I., … Koller, W.C. (2002). Neuropsychological deficits in essential tremor: An expression of cerebello-thalamo-cortical pathophysiology? *European Journal of Neurology*, *9*(2), 143–151.

Vendrell, P., Junqué, C., Pujol, J., Jurado, M., Molet, J., & Grafman, J. (1995). The role of prefrontal regions in the Stroop task. *Neuropsychologia*, *33*(3), 341–352. doi:10.1016/0028-3932(94)00116-7

Watson, G., & Leverenz, J. (2010). Profile of cognitive impairment in Parkinson disease. *Brain Pathology*, *20*(3), 640–645. doi:10.1111/j.1750-3639.2010.00373.x

Wechsler, D. (1987). *Wechsler Memory Scale - Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2001). *Wechsler Individual Achievement Test* (2nd ed.). San Antonio, TX: Harcourt Assessment.

Wechsler, D. (2008). *Wechsler Memory Scale* (4th ed.). San Antonio, TX: Pearson.

Williams, B.R., & Donovick, P.J. (2008). Questioning the rule of thumb: Can verbal tasks be administered during the CVLT-II delay interval? *The Clinical Neuropsychologist*, *22*(5), 807–812.

Williams, B.R., Sullivan, S.K., Morra, L.F., Williams, J.R., & Donovick, P.J. (2014). The similar effects of verbal and non-verbal intervening tasks on word recall in an elderly population. *The Clinical Neuropsychologist*, *28*(3), 505–513.

Williams, M.M., Roe, C.M., & Morris, J.C. (2009). Stability of the Clinical Dementia Rating: 1979–2007. *Achives of Neurology*, *66* (6), 773–777.

Wilson, R.S., Gilley, D.W., Bennett, D.A., Beckett, L.A., & Evans, D.A. (2000). Person-specific paths of cognitive decline in Alzheimer's disease and their relation to age. *Psychology and Aging*, *15*(1), 18–28.

Yesavage, J., Poon, L., Crook, T., Davis, K., Eisdorfer, C., Gurland, B., … Thompson, L. (1986). *Handbook of clinical memory assessment of older adults*. Washington, DC: American Psychiatric Press.