


# Assessing Fluctuating Cognition in Dementia Diagnosis: Interrater Reliability of the Clinician Assessment of Fluctuation

American Journal of Alzheimer's Disease & Other Dementias®  
2016, Vol. 31(2) 137-143  
© The Author(s) 2015  
Reprints and permission:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/1533317515603359  
aja.sagepub.com  


Kathleen Van Dyk, PhD<sup>1,2</sup>, Stephanie Towns, PsyD<sup>1,3</sup>,  
Oksana Tatarina, MA<sup>1</sup>, Philip Yeung, BA<sup>1</sup>, Jhedy Dorrejo, MA<sup>1</sup>,  
Laura B. Zahodne, PhD<sup>1</sup>, and Yaakov Stern, PhD<sup>1</sup>

## Abstract

Fluctuating cognition (FC) is a core feature of dementia with Lewy bodies (DLB) but is challenging to assess. This study assessed the reliability and validity of the Clinician Assessment of Fluctuation (CAF), which assesses FC in patients with dementia. Interrater agreement of CAF outcomes (FC present and FC severe) was evaluated between physicians and nonphysicians in 141 patients with Alzheimer's disease (AD) or DLB. Frequency of CAF outcomes by clinical and neuropathological diagnosis was examined. We found that interrater reliability was fair on FC present and almost perfect on FC severe, and both outcomes were higher in patients with clinical DLB than with clinical AD and were qualitatively more often endorsed in cases with neuropathological evidence of Lewy bodies. We conclude that the CAF is a reliable measure of FC and can be valuable in differential dementia diagnosis.

## Keywords

fluctuating cognition, dementia, dementia with Lewy Bodies, Alzheimer's Disease

## Background

Dementia with Lewy bodies (DLB) has been estimated to represent as high as 30.5% of cases with dementia, making it potentially the second most common cause of dementia.<sup>1</sup> One of the three core diagnostic features of DLB is the presence of fluctuating cognition (FC).<sup>2</sup> Fluctuating cognition is characterized by alternating periods of cognitive impairment and “normal or near-normal performance” and “pronounced variations in attention and alertness.”<sup>3</sup> Although FC can be present during the course of other neurodegenerative illnesses such as Alzheimer's disease (AD)<sup>4</sup> and Parkinson's disease dementia,<sup>5</sup> FC occurs with greater frequency and severity in DLB.<sup>4</sup> Thus, the presence and severity of FC is an important factor in the diagnosis of DLB.<sup>2,4,6</sup>

The clinical diagnosis of DLB has historically had low sensitivity, with poor interrater reliability.<sup>7-9</sup> A reliable method of identifying diagnostic criteria would have important clinical implications. The identification and evaluation of FC has been challenging and has relied primarily on subjective clinical judgment. In fact, FC has been identified as the most problematic of the DLB diagnostic criteria with regard to interrater agreement.<sup>4,7</sup> Additionally, FC may be related to other comorbid medical or psychiatric conditions, such as psychosis or delirium. As such, systematically assessing FC is essential for the provision of accurate diagnoses and effective clinical care. Indeed, including a formal measure of FC is recommended in

the current diagnostic guidelines of the DLB Consortium.<sup>2</sup> Within the guidelines, the following measures were suggested: the Clinician Assessment of Fluctuation (CAF),<sup>10</sup> One Day Fluctuation Assessment Scale,<sup>10</sup> and The Mayo Fluctuations Composite Scale.<sup>11</sup> The CAF rating scale was designed to provide experienced clinicians with a structured method for assessing FC associated with DLB,<sup>4</sup> and it uniquely queries both frequency and duration of endorsed symptoms. Measuring variations in cortical activity using electroencephalography (EEG) captures important objective data that can help determine the presence of FC in dementia<sup>12,13</sup>; however, interview-based clinical measures such as the CAF can be a more cost-effective and efficient method for assessing this symptom.

<sup>1</sup> Cognitive Neuroscience Division, Department of Neurology, Columbia University Medical Center, New York, NY, USA

<sup>2</sup> Department of Psychology, Queens College and The Graduate Center, City University of New York, New York, NY, USA

<sup>3</sup> Department of Clinical Psychology, Antioch University New England, Keene, NH, USA

This article was accepted under the editorship of the former Editor-in-Chief, Carol F. Lippa.

## Corresponding Author:

Yaakov Stern, PhD, 630 West 168th Street, P&S Box 16, New York, NY 10032.  
Email: ysl1@columbia.edu

Methods developed to assess delirium are also a consideration, as there is symptom overlap between delirium and DLB. For example, the Confusion Assessment Method Instrument (CAM)<sup>14</sup> provides an important means to assess the presence of delirium in older adults. Despite similarities between delirium and FC, it is nonetheless important to recognize that these syndromes are distinct, with different underlying etiologies. The assessment of each syndrome should involve specific targeted measures. The CAM, for instance, also captures additional symptoms relevant to the presence of delirium which are not typically relevant to the assessment of FC (eg, memory impairment, perceptual disturbances, psychomotor agitation). Thus, efficient targeted tools such as the CAF can play an important role in dementia assessment, and studies have supported its validity in patient populations. Clinician ratings on the CAF have been found to be significantly correlated with fluctuations in neural activity on EEG recordings.<sup>10</sup> Clinician ratings on the CAF have also been found to be significantly correlated with variability in neuropsychological measures of attention and processing speed, with strongest correlations found among patients with DLB compared to AD and vascular dementia.<sup>6,10</sup> Finally, strong correlations were also found between clinician ratings on the CAF and another FC measure designed to be completed by a nonphysician rater.<sup>10</sup> Despite evidence to support the validity of the CAF, there have been no prior studies of interrater reliability nor have there been any investigations of the utility of the CAF when used by nonphysician raters. Additionally, it has been proposed that determining the scale's relevance to established diagnostic criteria would also be valuable.<sup>6</sup>

The current study contributes to our ability to assess FC by examining the interrater reliability of the CAF and the relationship between CAF scores and subjectively assessed FC as part of applying the diagnostic criteria for DLB. We also explored the potential differential use of the CAF by physicians and nonphysician research assistants. Finally, we qualitatively examined CAF scores among patients with neuropathological evidence of AD, Lewy body disease, or both.

## Methods

### Participants

One hundred forty-one participants were identified from the Multicenter Study of Predictors of Disease Course in Alzheimer's Disease.<sup>15</sup> This multicenter longitudinal study was designed to collect clinical, neuropsychiatric, functional, and neuropsychological indices to characterize probable Alzheimer's disease (pAD) and predict outcomes. Specific methodology of the Predictors Study has been described previously.<sup>15,16</sup> In brief, patients were drawn from the second cohort of the Predictors Study. Patients with AD having mild dementia (as defined by a modified Mini-Mental Score Examination [MMSE]<sup>17</sup> above 29) were recruited from 3 centers: Columbia University, Johns Hopkins University School of Medicine, and Massachusetts General Hospital. Patients with a history of

certain psychotic or substance abuse disorders were excluded as were patients with a history of electroconvulsive treatment or any evidence of prior stroke.<sup>15</sup> In the second Predictors Study cohort, patients with DLB were recruited and followed using the same procedures as patients with pAD, except that an initial modified MMSE of 20 and above was allowed. For the purposes of this study, we selected only those participants who received both the physician and the nonphysician evaluation at the same visit interval and examined data collected from the first visit at which both took place. In approximately 80% of cases, the visit selected for analysis was the baseline visit or a follow-up visit within 1 year of baseline. The majority of the remaining visits (approximately 16%) took place within 3 years of baseline, and approximately 4% of the visits selected took place longer than 4 years from baseline. To our knowledge, there were no factors relevant to the proposed analyses that meaningfully and systematically varied by which visit both evaluations took place; rather, the reason for which a baseline visit was not sufficient appeared to be errors of omission or administration.

Patient demographics are summarized in Table 1. All patients had received a diagnosis of either pAD in accordance with 1984 National Institute of Neurological and Communicative Diseases and Stroke–Alzheimer's Disease and Related Disorders Association Workgroup criteria<sup>19</sup> or DLB in accordance with 1996 Consortium on DLB consensus diagnostic criteria.<sup>3</sup> Among patients with DLB, in 3 cases, FC and only one other diagnostic criterion met to achieve a diagnosis of DLB. Of the 141 patients, 98 died during the course of this study and 35 were brought to autopsy, and we examined these groups for possible demographic differences. We found no gender or race differences between those who came to autopsy and those who did not. Those who came to autopsy were found to have more years of education ( $\bar{x}$  = 15.63, standard deviation [SD] = 3.12) than those who did not ( $\bar{x}$  = 14.24, SD = 3.28;  $F_{(1,96)} = 4.20$ ,  $P = .04$ ).

Neuropathological criteria included the National Institute on Aging (NIA)-Reagan Institute criteria for AD and required intermediate likelihood<sup>20</sup> (1 case included as AD had only "senile change, Alzheimer's type" and not intermediate likelihood), and the by McKeith et al criteria were used to identify diffuse Lewy body disease,<sup>3</sup> requiring at least limbic Lewy bodies. Neuropathological findings of 3 cases did not suggest AD or diffuse Lewy body disease and were excluded from these analyses (Neuropathological diagnoses of these 3 cases included Parkinson's disease, stroke, progressive supranuclear palsy, corticobasal degeneration, atherosclerosis, and amyloid angiopathy.). The remaining subset of 32 participants were grouped by evidence of the following neuropathological findings: (1) pathological evidence of AD with no evidence of diffuse Lewy body disease (NAD), (2) pathological evidence of diffuse Lewy body disease with no evidence of AD (NLB), and (3) pathological evidence of both AD and diffuse Lewy body disease (NAD + NLB). Among these 32 participants, there was additional neuropathological evidence of stroke ( $n = 8$ ), multi-system atrophy ( $n = 4$ ), amyloid angiopathy ( $n = 12$ ), and/or atherosclerosis ( $n = 23$ ).

**Table 1.** Demographics by Diagnostic Group.

	pAD (n = 124)	DLB (n = 17)	
Age $\bar{x}$ (SD), range	75.90 ( $\pm 7.89$ ), 51-92	74.47 ( $\pm 6.51$ ), 57-84	$F_{(1,139)} = 0.51, P = .48$
No. of males	51	12	$\chi^2 = 5.25, P = .03$
MMSE $\bar{x}$ (SD)	20.58 ( $\pm 4.96$ ), n = 120	18.27 ( $\pm 6.26$ ), n = 15	$F_{(1,133)} = 2.73, P = .10$
mMMSE $\bar{x}$ (SD)	35.43 ( $\pm 9.12$ ), n = 120	34.43 ( $\pm 9.38$ ), n = 14	$F_{(1,132)} = 0.15, P = .70$
No. of autopsies/no. of deaths	26/83	9/15	$\chi^2 = 4.55, P = .04$

Abbreviations: pAD, probable Alzheimer's disease; DLB, dementia with Lewy bodies; MMSE, Mini-Mental State Examination<sup>18</sup>; mMMSE, modified Mini-Mental State Examination; SD, standard deviation.<sup>17</sup>

## Measures

In addition to the assessments described in previous reports,<sup>16</sup> all patients were evaluated with the CAF<sup>4</sup> and a DLB diagnostic criteria evaluation form adapted from the 1996 Consortium on DLB consensus diagnostic criteria.<sup>3</sup> Participants were asked to rate whether symptoms were present over the past few months prior to assessment. The CAF is a 2-item questionnaire that was developed to assess fluctuations in attention and alertness, a core symptom of DLB. The first item captures the presence of fluctuating level of consciousness and the second item captures the presence of fluctuating cognitive impairment. If either of the scale's 2 items is endorsed, frequency is assessed on a scale ranging from 1 to 4, with 4 being the most frequent, and duration of the symptom is assessed on scale ranging from 0 to 4, with 4 being the longest duration. These 2 values (frequency and duration) are multiplied to achieve a severity score, ranging from 0 to 16.

The DLB diagnostic criteria form is a 9-item questionnaire designed to document the symptoms of DLB. The first 3 items of the form are core criteria (ie, FC, visual hallucinations, and Parkinsonism), followed by 6 items related to supportive features. This study examined responses to the FC item. This consisted of 3 examples of FC (ie, episodes of going blank or switching off, periods of apparent spontaneous remission during which cognitive functions improve, and excessive daytime drowsiness with transient confusion on awakening) and 2 exclusionary conditions (ie, fluctuations typically occur in the late/afternoon/early evening and fluctuations are associated with a change in medication). If any 1 of the 3 examples was endorsed and both exclusionary conditions were denied, the participant met criteria for the FC item of the DLB diagnostic criteria form. See Table 2 for the specific CAF items and the FC diagnostic criteria item.

## Procedure

This study was approved by the institutional review boards (IRBs) at the participating institutions. All participants underwent an IRB-approved informed consent process prior to enrollment in the study. Participants were administered the CAF and DLB diagnostic criteria form by 2 types of independent raters: one of the study research assistants (RA) and one of the study physicians (MD). Evaluation of FC was completed prior to assessment of other features of DLB. Both raters'

assessments were typically completed within a 1-month window, and 86% were completed on the same day.

## Data Analysis

Two dichotomous outcome variables were derived from the CAF: (1) *FC present*: whether at least 1 of the CAF's 2 items was endorsed (yes/no) and (2) *FC severe*: whether the calculated severity score was  $>5$  (yes/no), based on published data indicating that scores above 5 have the best specificity and sensitivity to discriminate DLB from pAD.<sup>4</sup> Interrater reliability between the RAs and MDs on the CAF was examined for both variables using Cohen's kappa.

For each rater (RA and MD), we evaluated the relationship of pAD versus DLB clinical diagnosis to each CAF variable using  $\chi^2$  or Fisher's exact test as appropriate. There were 5 RA assessments and 58 MD assessments in which the CAF severity score was not calculated due to missing data.

We also examined intrarater agreement between both CAF variables and the FC item on the DLB diagnostic criteria form using Cohen's  $\kappa$ . Six cases were excluded due to missing item data.

We report the frequency of the following for each of the 3 neuropathology groups: the clinical diagnosis and the 2 CAF variables for each rater. All analysis were performed at the  $P = .05$  level using IBM SPSS Statistics software package v19.0 and SAS 9.2. Across agreement analyses, Cohen's  $\kappa$  values were interpreted according to the published standards of Landis and Koch.<sup>21</sup>

## Results

### Clinician Assessment of Fluctuation

Interrater reliability of the FC present variable was fair,  $\kappa = .22, P < .01$ . Interrater reliability of FC severe was almost perfect,  $\kappa = .89, P < .01$  (Table 3 and Figures 1 and 2).

### Clinical Diagnosis

When examining the relationship between CAF outcome variables and clinical diagnosis, FC present was endorsed for a greater proportion of cases with DLB compared to cases with pAD (RA:  $\chi^2 = 45.74, P < .01$ , MD:  $\chi^2 = 42.67, P < .01$ ; Table 4). Further, a greater proportion of participants

**Table 2.** Assessment Items.

DLB Diagnostic FC Criterion (Adapted From McKeith et al <sup>3</sup> )	CAF (Adapted From Walker et al <sup>4</sup> )
<p>I. FC with pronounced variations in attention and alertness</p> <p>Has the patient experienced:</p> <ul style="list-style-type: none"> <li>A. Episodes of going blank or switching off?</li> <li>B. Periods of apparent spontaneous remission during which cognitive functions improve?</li> <li>C. Excessive daytime drowsiness with transient confusion on awakening?</li> <li>D. If present, are fluctuations typically in the late afternoon/early evening?</li> <li>E. If present, were fluctuations associated with a change in medications?</li> </ul>	<ul style="list-style-type: none"> <li>A. Does the patient have spontaneous impaired alertness and concentration—ie, appears drowsy but awake, looks dazed, is not aware of what's going on? (This question refers to fluctuation in level of consciousness. Clear examples demonstrating impaired consciousness with variations in performance/cognition are required to receive a positive rating). This question refers to a severe impaired state in which the patient appears to be unconscious, but his/her eyes are open</li> <li>B. Has the level of confusion experienced by the patient tended to vary recently from day to day or week to week? For example, becoming worse, then perhaps improving for a while (ie, up and down)? (This question refers to fluctuation in level of cognitive impairment, not level of consciousness. Significant fluctuation is regarded as present if distinct examples of differences in performance/cognition can be given on at least 2 occasions within a 1 month period)</li> </ul>
<p>(FC is present if a, b, or c is 'Yes' and D and E are both 'No')</p>	<p>If fluctuating consciousness/cognition is present (ie, either A or B is 'yes'), a severity rating should be made:</p> <ul style="list-style-type: none"> <li>1. How often would you say that these periods of increased confusion occur?                             <ul style="list-style-type: none"> <li>1. About once per month</li> <li>2. More than once a month, but less than once per week</li> <li>3. At least once per week but not every day</li> <li>4. Every day</li> </ul> </li> <li>2. For how long do these periods of increased confusion/disturbed consciousness last?                             <ul style="list-style-type: none"> <li>1. Less than 10 second</li> <li>2. More than 10 seconds but less than 5 minutes</li> <li>3. More than 5 minutes but less than 1 hour</li> <li>4. More than 1 hour but less than a day</li> <li>5. More than a day</li> </ul> </li> </ul>

Abbreviation: CAF, Clinician Assessment of Fluctuation; FC, fluctuating cognition; DLB, dementia with Lewy bodies.

**Table 3.** CAF Interrater Agreement.

	MD Score		
	No	Yes	Unknown
FC present			
RA score			
No	69	0	46
Yes	4	11	10
Unknown	0	0	1
Observed agreement = 0.57			
Expected agreement = 0.44			
FC severe			
RA score			
No	70	0	
Yes	2	9	
Observed agreement = 0.98			
Expected agreement = 0.79			

Abbreviations: CAF, Clinician Assessment of Fluctuation; FC, fluctuating cognition; FC present, fluctuating cognition is endorsed (ie, "yes" to item A and/or B); FC severe, fluctuating cognition score  $\geq$  5; MD, study physician; RA, research assistant.

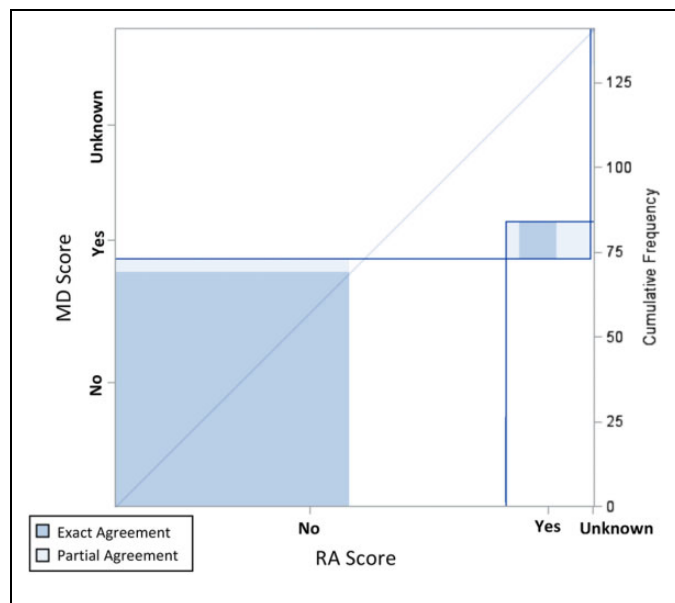
diagnosed with DLB had a positive FC severe score compared to pAD (RA:  $\chi^2 = 37.10, P < .01$ , MD:  $\chi^2 = 36.03, P < .01$ ). To further assess diagnostic utility, sensitivity and specificity analyses were conducted. For the RAs' ratings, sensitivity of the FC severe score was 57% and specificity was 95%. For the MDs' ratings, sensitivity of the FC severe score was 70% and specificity was 96%.

### Dementia With Lewy bodies Diagnostic Criteria

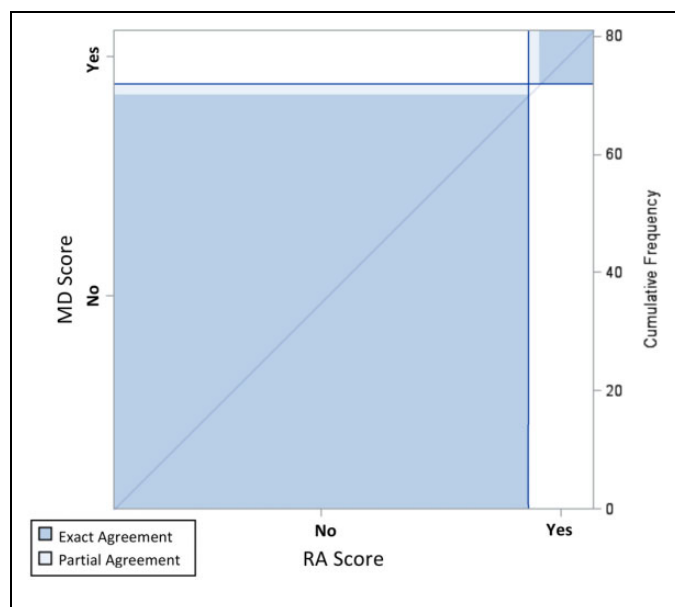
Moderate agreement was found between RAs' ratings between the CAF and the DLB diagnostic criteria: FC present,  $\kappa = .47, P < .01$ ; FC severe,  $\kappa = .42, P < .01$  (Table 5). For the MDs' ratings, only slight agreement was found between FC present and a positive response on the FC item of the DLB diagnostic criteria,  $\kappa = 0.15, P < .01$ ; however, there was substantial agreement between FC severe and a positive response on the FC item of the DLB diagnostic criteria,  $\kappa = .70, P < .01$ .

### Neuropathological Results

Thirty-two cases came to autopsy, and the number of years between their clinical evaluation and date of autopsy ranged



**Figure 1.** Inter-rater agreement plot for FC Present, 95% Confidence Interval.



**Figure 2.** Inter-rater agreement plot for FC Severe, 95% Confidence Interval.

from 0 to 7 ( $\bar{x}$  years = 2.94;  $\pm 1.95$ ). Of the 24 cases with a clinical diagnosis of pAD, 19 received a neuropathological diagnosis of AD, 1 received a neuropathological diagnosis of DLB, and 4 received a neuropathological diagnosis of both AD and DLB. Of the 8 cases with a clinical diagnosis of DLB, 3 received a neuropathological diagnosis of AD, 2 received a neuropathological diagnosis of Lewy body disease, and 3 received a neuropathological diagnosis of both AD and Lewy body disease. Each rater’s CAF outcome variables for each neuropathology group are presented in Table 6. In 22 cases, only AD pathology was found, in 3 cases, only diffuse Lewy

**Table 4.** CAF Variables by Clinical Diagnosis.

Sample	CAF Variable	RA Score, n	MD Score, n
Clinical pAD	FC present		
	Yes	12	3
	No	111	70
	Unkn/NA	1	51
	FC severe		
	Yes	6	3
Clinical DLB	FC present		
	Yes	13	8
	No	4	3
	Unkn/NA	0	6
	FC severe		
	Yes	8	7
	No	6	3

Abbreviations: CAF, Clinician Assessment of Fluctuation; DLB, dementia with Lewy bodies; FC, fluctuating cognition; FC present, fluctuating cognition is endorsed (ie, “yes” to item A and/or B); FC severe, fluctuating cognition score  $\geq 5$ ; MD, study physician; NA, not available; pAD, probable Alzheimer’s disease; RA, research assistant; UNKN, unknown.

**Table 5.** CAF Variables by FC Diagnostic Criterion.

Sample	CAF Variable	RA Score, n	MD Score, n
Positive FC on diagnostic criteria	FC present		
	Yes	12	7
	No	6	2
	Unkn/NA	1	5
	FC severe		
	Yes	7	7
Negative FC on diagnostic criteria	FC present		
	Yes	13	4
	No	107	71
	Unkn/NA	0	51
	FC severe		
	Yes	7	3
	No	112	71

Abbreviations: CAF, Clinician Assessment of Fluctuation; FC, fluctuating cognition; FC present, fluctuating cognition is endorsed (ie, “yes” to item A and/or B); FC severe, fluctuating cognition score  $\geq 5$ ; MD, study physician; NA, not available; RA, research assistant; UNKN, unknown.

body disease was found, and there was evidence of both diffuse Lewy body disease and AD pathology in 7 cases. Compared to cases with only AD pathology, a greater proportion of cases with neuropathological evidence of diffuse Lewy Body disease alone or with AD pathology had FC present by one or both raters.

## Discussion

Fluctuating cognition is a core symptom in the diagnosis of DLB, and clinicians have historically struggled to identify and quantify FC.<sup>4</sup> The current study is the first to examine interrater reliability of the CAF in a sample of adults with dementia. We found fair interrater reliability for the FC present scores and

**Table 6.** CAF Variables by Neuropathological Diagnosis.

Sample	CAF Variable	RA Score, n	MD Score, n
NAD	FC present		
	Yes	3	2
	No	18	12
	Unkn/NA	1	8
	FC severe		
	Yes	2	2
NLB	No	18	12
	FC present		
	Yes	1	1
	No	2	1
	Unkn/NA	0	1
	FC severe		
NAD + NLB	Yes	0	0
	No	2	1
	FC present		
	Yes	3	3
	No	4	4
	FC severe		
Yes	2	3	
No	4	4	

Abbreviations: CAF, Clinician Assessment of Fluctuation; FC, fluctuating cognition; FC present, fluctuating cognition is endorsed (ie, "yes" to item A and/or B); FC severe, fluctuating cognition score  $\geq 5$ ; MD, study physician; NA, not available; NAD, pathological evidence of AD and not diffuse Lewy Body disease; NAD + NLB, pathological evidence of both AD and diffuse Lewy Body disease; NLB, pathological evidence of diffuse Lewy body disease and not AD; RA, research assistant; UNKN, unknown.

almost perfect interrater reliability for FC severe scores. Our findings indicate that the CAF can provide an objective measure with good interrater reliability that can be utilized by either physicians or trained nonphysician providers. The severity score cutoff ( $>5$ )<sup>4</sup> produced very high agreement between raters and also proved to be better at differentiating patients diagnosed with DLB from patients with pAD among MD raters. Thus, the use of severity cutoff may result in a more reliable and specific measure of FC than merely endorsing the presence of FC on the CAF.

Although interrater agreement was sufficient on both CAF outcomes, we found higher agreement on FC severe compared to FC present. Examination of the data revealed that the CAF was applied somewhat differently by each rater group. The MD raters appeared less likely to endorse FC than the RAs and more commonly reported the presence of FC as "unknown," even among cases that ultimately received a clinical diagnoses of DLB. Examination of agreement on both measures showed that RA raters had comparable levels of agreement between DLB diagnostic criteria for FC and either CAF outcome measure. In contrast, MD ratings demonstrated only slight agreement between the FC diagnostic criterion and FC present scores, but substantial agreement was found with FC severe scores. These findings may reflect that MD raters are more stringent in their confident endorsement of the presence of FC, doing so only at higher levels of symptom severity. In light of this finding, employment of the CAF in research or clinical

settings may include recording the role of the assessor, so scores are considered accordingly.

Examination of neuropathological data revealed that, compared to those patients with only AD pathology, a greater proportion of patients with Lewy body pathology (either alone or in combination with Alzheimer's pathology) were rated as FC present or FC severe. Endorsement of FC present or FC severe was comparable between rater groups in this sample. These findings further support that assessing FC with the CAF can play an important role in differential dementia diagnosis. With a larger sample, future studies could investigate the relationship between the clinical outcomes of the CAF and neuropathological evidence of DLB and AD at varying stages of disease burden.

This study has limitations. First, the clinical DLB sample was relatively small, reflecting the difficulty in identifying patients with DLB in the clinic whose cognitive severity was mild enough to meet the entry criteria. The small sample size may also reflect referral biases and prevalence of dementia type. Additionally, although we did not find any significant differences between dementia groups on cognitive screening measures, the lower cutoff for DLB cases nonetheless raises the possibility for more advanced disease progression in the DLB group, which may have had an impact on FC ratings. In order to examine this possibility, we conducted follow-up analyses and found no relationship with either rater between MMSE and CAF scores, across the sample or within diagnostic group. It is also important to note that in most cases, the CAF was completed on the same day by both raters, which may have inflated reliability scores. We also note that CAF raters were not blinded to other clinical variables, including clinical diagnosis at study entry or other clinical symptoms, and it is possible that this knowledge had an impact on the obtained scores. Finally, it is not uncommon to discover Lewy body pathology at autopsy in cases with a clinical presentation of AD,<sup>22,23</sup> a potential limitation of our examination of FC in cases who came to autopsy.

Overall, our results support the use of the CAF rating scale as a reliable tool for capturing FC when used by either a physician or a nonphysician evaluator. Our findings indicate that CAF can play a valuable role in differential dementia diagnosis and also support the use of the  $>5$  cutoff score. Fluctuating cognition is a core symptom of DLB and one that has been historically difficult to capture. The CAF rating scale provides a means to assess the presence and severity of FC and can be a valuable tool in both clinical and research settings.

### Acknowledgments

The authors thank Prabha Siddarth for her help with the figures.

### Authors' Note

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The sponsors had no role in the study design or data interpretation.

## Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Predictors study is supported by NIA R01 AG007370 to Dr Stern, Dr Zahodne is supported by NIA T32 AG000261. This publication was also supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Number UL1 TR000040, formerly the National Center for Research Resources, Grant Number UL1 RR024156.

## References

- Zaccai J, McCracken C, Brayne C. A systematic review of prevalence and incidence studies of dementia with Lewy bodies. *Age and Ageing*. 2005;34(6):561-566.
- McKeith IG, Dickson DW, Lowe J, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. *Neurology*. 2005;65(12):1863-1872.
- McKeith IG, Galasko D, Kosaka K, et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. *Neurology*. 1996;47(5):1113-1124.
- Walker MP, Ayre GA, Cummings JL, et al. Quantifying fluctuation in dementia with Lewy bodies, Alzheimer's disease, and vascular dementia. *Neurology*. 2000;54(8):1616-1625.
- Varanese S, Perfetti B, Monaco D, et al. Fluctuating cognition and different cognitive and behavioural profiles in Parkinson's disease with dementia: comparison of dementia with Lewy bodies and Alzheimer's disease. *J Neurol*. 2010;257(6):1004-1011.
- Walker MP, Ayre GA, Perry EK, et al. Quantification and characterization of fluctuating cognition in dementia with Lewy bodies and Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2000;11(6):327-335.
- Mega MS, Masterman DL, Benson DF, et al. Dementia with Lewy bodies: reliability and validity of clinical and pathologic criteria. *Neurology*. 1996;47(6):1403-1409.
- Litvan I, MacIntyre A, Goetz CG, et al. Accuracy of the clinical diagnoses of Lewy body disease, Parkinson disease, and dementia with Lewy bodies: a clinicopathologic study. *Arch Neurol*. 1998;55(7):969-978.
- Nelson PT, Jicha GA, Kryscio RJ, et al. Low sensitivity in clinical diagnoses of dementia with Lewy bodies. *J Neurol*. 2010;257(3):359-366.
- Walker MP, Ayre GA, Cummings JL, et al. The clinician assessment of fluctuation and the one day fluctuation assessment scale. Two methods to assess fluctuating confusion in dementia. *Br J Psychiatry*. 2000;177:252-256.
- Ferman TJ, Smith GE, Boeve BF, et al. DLB fluctuations: specific features that reliably differentiate DLB from AD and normal aging. *Neurology*. 2004;62(2):181-187.
- Bonanni L, Thomas A, Tiraboschi P, Perfetti B, Varanese S, Onofri M. EEG comparisons in early Alzheimer's disease, dementia with Lewy bodies and Parkinson's disease with dementia patients with a 2-year follow-up. *Brain*. 2008;131(pt 3):690-705.
- Briel RCG, McKeith IG, Barker WA, Hewitt Y, Perry RH, Ince PG, Fairbairn AF. EEG findings in dementia with Lewy bodies and Alzheimer's disease. *J Neurol Neurosurg Psychiatry*. 1999;66(3):401-403.
- Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Ann Intern Med*. 1990;113(12):941-948.
- Stern Y, Folstein M, Albert M, et al. Multicenter study of predictors of disease course in Alzheimer disease (the "predictors study"). I. Study design, cohort description, and intersite comparisons. *Alzheimer Dis Assoc Disord*. 1993;7(1):3-21.
- Scarmeas N, Brandt J, Albert M, et al. Delusions and hallucinations are associated with worse outcome in Alzheimer disease. *Arch Neurol*. 2005;62(10):1601-1608.
- Stern Y, Sano M, Paulson J, Mayeaux R. Modified mini-mental state examination: validity and reliability. *Neurology*. 1987;37(suppl 1):179.
- Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatric Res*. 1975;12(3):189-198.
- McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34(7):939-944.
- The National Institute on Aging – Reagan Working Group. Consensus recommendations for the postmortem diagnosis of Alzheimer's disease. *Neurobiology of Aging*. 1997;18(suppl 4):S1-S2.
- Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33(2):363-374.
- Middleton L, Grinberg L, Miller B, Kawas C, Yaffe K. Neuropathologic features associated with Alzheimer disease diagnosis age matters. *Neurology*. 2011;77(19):1737-1744.
- Schneider JA, Arvanitakis Z, Leurgans SE, Bennett DA. The neuropathology of probable Alzheimer disease and mild cognitive impairment. *Ann Neurol*. 2009;66(2):200-208.