

UCSF

UC San Francisco Previously Published Works

Title

Interrater reliability of the new criteria for behavioral variant frontotemporal dementia

Permalink

<https://escholarship.org/uc/item/3c79414m>

Journal

Neurology, 80(21)

ISSN

0028-3878

Authors

LaMarre, Amanda K

Rascovsky, Katya

Bostrom, Alan

et al.

Publication Date

2013-05-21

DOI

10.1212/wnl.0b013e318293e368

Peer reviewed

# Interrater reliability of the new criteria for behavioral variant frontotemporal dementia

Amanda K. LaMarre, PhD  
Katya Rascovsky, PhD  
Alan Bostrom, PhD  
Parnian Toofanian, JD, PsyD  
Sarah Wilkins, BA  
Sharon J. Sha, MD  
David C. Perry, MD  
Zachary A. Miller, MD  
Georges Naasan, MD  
Robert Jr Laforce, MD, PhD  
Jayne Hagen, PhD  
Leonel T. Takada, MD  
Maria Carmela Tartaglia, MD  
Gail Kang, MD  
Douglas Galasko, MD  
David P. Salmon, PhD  
Sarah Tomaszewski Farias, PhD  
Berneet Kaur, MD  
John M. Olichney, MD  
Lovingly Quitania Park, PhD  
Mario F. Mendez, MD, PhD  
Po-Heng Tsai, MD  
Edmond Teng, MD, PhD  
Bradford Clark Dickerson, MD  
Kimiko Domoto-Reilly, MD  
Scott McGinnis, MD  
Bruce L. Miller, MD  
Joel H. Kramer, PsyD

Correspondence to  
Dr. LaMarre:  
LaMarreA@radonc.ucsf.edu

Supplemental data at  
[www.neurology.org](http://www.neurology.org)

## ABSTRACT

**Objective:** To evaluate the interrater reliability of the new International Behavioural Variant FTD Criteria Consortium (FTDC) criteria for behavioral variant frontotemporal dementia (bvFTD).

**Methods:** Twenty standardized clinical case modules were developed for patients with a range of neurodegenerative diagnoses, including bvFTD, primary progressive aphasia (nonfluent, semantic, and logopenic variant), Alzheimer disease, and Lewy body dementia. Eighteen blinded raters reviewed the modules and 1) rated the presence or absence of core diagnostic features for the FTDC criteria, and 2) provided an overall diagnostic rating. Interrater reliability was determined by  $\kappa$  statistics for multiple raters with categorical ratings.

**Results:** The mean  $\kappa$  value for diagnostic agreement was 0.81 for possible bvFTD and 0.82 for probable bvFTD (“almost perfect agreement”). Interrater reliability for 4 of the 6 core features had “substantial” agreement (behavioral disinhibition, perseverative/compulsive, sympathy/empathy, hyperorality;  $\kappa = 0.61$ – $0.80$ ), whereas 2 had “moderate” agreement (apathy/inertia, neuropsychological;  $\kappa = 0.41$ – $0.6$ ). Clinician years of experience did not significantly influence rater accuracy.

**Conclusions:** The FTDC criteria show promise for improving the diagnostic accuracy and reliability of clinicians and researchers. As disease-altering therapies are developed, accurate differential diagnosis between bvFTD and other neurodegenerative diseases will become increasingly important. *Neurology*® 2013;80:1973–1977

## GLOSSARY

**bvFTD** = behavioral variant frontotemporal dementia; **DLB** = dementia with Lewy bodies; **FTDC** = International Behavioural Variant FTD Criteria Consortium; **lvPPA** = logopenic variant primary progressive aphasia; **nfPPA** = nonfluent variant primary progressive aphasia; **SE** = standard error; **svPPA** = semantic variant primary progressive aphasia; **UC** = University of California; **UCSF** = University of California San Francisco.

Behavioral variant frontotemporal dementia (bvFTD) is a clinical syndrome characterized by profound changes in personality and behavior. Initially believed to be rare, it has now been determined to be a common cause of early-onset dementia, equal in prevalence to Alzheimer disease (AD) in individuals younger than 65 years.<sup>1,2</sup> In the absence of definitive antemortem biomarkers for the disease, diagnosis is based on the presence or absence of symptoms. Unfortunately, clinical diagnosis of this syndrome remains challenging, especially within the primary care community.

To address this issue, an international group of experts (International Behavioural Variant FTD Criteria Consortium) recently developed empirically based criteria for bvFTD (FTDC criteria<sup>3</sup>) that reflect major advances in our understanding of the disease. A recent validation study using pathology-confirmed cases substantiates its sensitivity in detecting bvFTD.<sup>3</sup>

The FTDC criteria were designed for broad-based use, but their sensitivity and specificity depend to some degree on how reliably they are applied. As such, the primary purpose of this study was to assess the interrater reliability of the new FTDC criteria for bvFTD within a heterogeneous group of patients with neurodegenerative disease. A secondary purpose was to determine

From the Memory and Aging Center (A.K.L., A.B., P.T., S.W., S.J.S., D.C.P., Z.A.M., G.N., J.H., L.T.T., M.C.T., G.K., B.L.M., J.H.K.), Department of Neurology, University of California San Francisco; Department of Neurology (K.R.), University of Pennsylvania, Philadelphia; Clinique de Mémoire (R.J.L.), Département des Sciences Neurologiques, Centre Hospitalier Affilié Universitaire de Québec, Canada; Department of Neuroscience (D.G., D.P.S.), University of California San Diego; Department of Neurology (S.T.F., B.K., J.M.O., L.Q.P.), University of California Davis; Department of Neurology (M.F.M., P.-H.T., E.T.), University of California Los Angeles; and Department of Neurology (B.C.D., K.D.-R., S.M.), Harvard Medical School, Boston, MA.

Go to [Neurology.org](http://Neurology.org) for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

whether length of experience with diagnosis and management of neurodegenerative disease influences interrater reliability.

**METHODS** **Participants. Raters.** Fifteen neurologists and 3 neuropsychologists were recruited to participate in the study. Their experience with diagnosing dementia syndromes ranged from 1 to 33 years (median = 7.0 years). At the time of their participation, all raters were working at the following academic medical centers: University of California (UC) San Francisco (UCSF), UC Davis, UC Los Angeles, UC San Diego, and Harvard University/Massachusetts General Hospital.

**Video module participants.** Twenty patients and their caregivers were recruited through our research program at the UCSF Memory and Aging Center. Patients were diagnosed via comprehensive multidisciplinary patient evaluation and consensus case conference. Our evaluation consists of a history and neurologic examination, neuropsychological testing, caregiver interviews/questionnaires, and diagnostic imaging. A consensus diagnosis is then made after presentation from each of the professionals involved in the patient's evaluation. Given that none of our patients is deceased, for the purposes of this study, we consider our diagnosis the "gold standard." Patient diagnoses included AD (n = 4), dementia with Lewy bodies (DLB) (n = 1), bvFTD (n = 5), nonfluent variant primary progressive aphasia (nfPPA) (n = 3), semantic variant primary progressive aphasia (svPPA) (n = 4), and logopenic variant primary progressive aphasia (lvPPA) (n = 3). Demographic information for these participants is listed in the table.

**Standard protocol approvals, registrations, and patient consents.** This research was approved by the UCSF Human Research Protection Program Independent Review Board and the Human Research Institutional Review Boards at UC Davis, UC Los Angeles, UC San Diego, and Harvard University/Massachusetts General Hospital. Written informed consent was obtained from all participants (or their guardians).

**Materials.** Raters were asked to evaluate each patient module using the FTDC criteria,<sup>3</sup> which are detailed in appendix e-1 on the *Neurology*<sup>®</sup> Web site at [www.neurology.org](http://www.neurology.org). Modules were created for a total of 20 patients with 1 of 6 diagnoses (AD, DLB, bvFTD, nfPPA, svPPA, lvPPA). Each module was standardized and contained 3 types of media:

1. Written history and summarized test results:
  - a. History of presenting illness, medical, social, and family history, summary of the physical and neurologic examination, current medications, neuropsychological testing

results (UCSF Bedside Cognitive Screening Battery<sup>4</sup> and Benson Figure Test<sup>5</sup>), and scores from neuropsychiatric and functional measures (Clinical Dementia Rating Scale,<sup>6</sup> Neuropsychiatric Inventory,<sup>7</sup> and Functional Activities Questionnaire<sup>8</sup>).

2. Videotaped interactions with the patient and/or caregiver (see appendix e-1 for additional information):
  - a. Neurologic examination: spontaneous speech, motor speech, eye movements, and motor and gait evaluation.
  - b. Language testing: examination of spontaneous speech, object naming, irregular word reading, single word and sentence comprehension, single word and sentence repetition, and knowledge regarding famous faces.
  - c. Caregiver interview: questions relating to time frame and nature of first symptoms, changes in specific domains of cognition, motor function, sleep, behavior, and activities of daily living, new-onset psychiatric symptoms, and questions specific to bvFTD (e.g., disinhibition, apathy, decreased empathy, compulsivity, hyperorality).
3. Magnetic resonance imaging:
  - a. Structural MRI scans showing T1- and T2-weighted images.

**Procedure.** The written portion of the patient modules was posted on a secure Web site, which raters accessed remotely. In addition to the modules, the Web site included separate written descriptions of the specific questions asked during the neurologic examination and caregiver interview and a description of each of the neuropsychological tasks. The FTDC criteria for bvFTD were also posted (alongside a downloadable PDF with check boxes that the raters were asked to complete while reviewing the module). The videos and MRI scans were mailed directly to the raters. Each rater was blinded to the patient's diagnosis. The raters were not made aware of the total proportion of bvFTD cases within the sample. The raters' instructions were posted on the Web site and listed as follows: "Each rater will carefully review the patient information presented on this site (e.g., Module 1), review the imaging data and watch the 3 videos for each module (neurologic exam, language testing and caregiver interview). The rater will then use the information presented to rate the patient on each symptom listed on the FTDC criteria form which you can download as a PDF." After reviewing the materials, the raters 1) rated each patient on each core diagnostic feature for the FTDC criteria (yes, no, don't know), and 2) provided an overall diagnostic rating (e.g., yes, no, don't know) for both possible bvFTD and probable bvFTD. Completed modules were mailed back to the study coordinator, along with the DVDs and MRI disc.

**Statistical analysis.** Rater agreement by core symptoms (behavioral disinhibition, apathy/inertia, loss of sympathy/empathy,

Table	Demographics for video module participants <sup>a</sup>				
	AD/DLB (n = 5)	bvFTD (n = 5)	nfPPA (n = 3)	svPPA (n = 4)	lvPPA (n = 3)
Age, y	63 (57-85)	68 (57-70)	75 (62-78)	63 (60-71)	55 (54-80)
Sex, M:F	3:2	4:1	2:1	2:2	3:0
Education, y	16 (14-20)	15 (13-19)	18 (16-20)	15 (13-18)	12 (12-16)
MMSE score	24.0 (23-28)	25.0 (22-29)	27.0 (23-30)	24.5 (21-28)	22.0 (18-26)
CDR Scale score	1.0 (0.5-2)	2.0 (1-2)	0.0 (0-1)	0.75 (0.5-1)	0.5 (0.5-2)

Abbreviations: AD = Alzheimer disease; bvFTD = behavioral variant frontotemporal dementia; CDR = Clinical Dementia Rating; DLB = dementia with Lewy bodies; lvPPA = logopenic variant primary progressive aphasia; MMSE = Mini-Mental State Examination; nfPPA = nonfluent variant primary progressive aphasia; svPPA = semantic variant primary progressive aphasia.  
<sup>a</sup> Values are presented as median (range) unless otherwise noted.

perseverative or compulsive behaviors, hyperorality, neuropsychological profile), diagnostic imaging, and diagnosis (possible or probable bvFTD) were determined via the  $\kappa$  statistic for multiple raters, with categorical ratings. The  $\kappa$  statistic is the rate of observed agreement between all possible pairs of ratings adjusted for the proportion of agreement expected to occur by chance.<sup>9</sup> Guidelines regarding magnitude of agreement suggest the following: values 0–0.20 = slight agreement, 0.21–0.40 = fair agreement, 0.41–0.60 = moderate agreement, 0.61–0.80 = substantial agreement, and 0.81–1 = almost perfect agreement.<sup>10</sup> To determine whether any center was an outlier for rater agreement, the analyses were repeated by removing each center systematically (by removing all of the individual raters from the center) and then rerunning the analysis. Logistic regression was applied to determine whether number of years of experience was statistically related to the rater's accuracy regarding diagnosis. All missing items (39/3,240 = 1.2% of observations) and items marked "Don't Know" (194/3,240 = 6% of observations) were excluded from the analysis.

**RESULTS Diagnostic agreement among raters per core feature/diagnostic imaging.** Interrater reliabilities for the core diagnostic features of the FTDC criteria are shown in the figure. Four of the core features (behavioral disinhibition, loss of sympathy/empathy, perseverative/compulsive behaviors, and hyperorality) demonstrated "substantial" agreement and the remaining 2 core features (apathy/inertia and neuropsychological profile) displayed "moderate" agreement. Interrater reliability for diagnostic imaging (significant atrophy in frontal/temporal lobes) showed "substantial" agreement (mean  $\kappa$  value: 0.66; standard error [SE]: 0.016).

**Diagnostic agreement among raters/sites for diagnosis.** The mean  $\kappa$  value for diagnostic agreement was 0.81

(SE: 0.12) for possible bvFTD, and 0.82 (SE: 0.08) for probable bvFTD, indicating that among the 18 raters, "almost perfect agreement" was achieved. When this analysis was repeated by systematically removing all of the raters from individual sites (thereby leaving 4 of 5 sites in the analysis), mean  $\kappa$  values for both possible and probable bvFTD remained essentially unchanged, suggesting that ratings remained consistent across sites without significant outliers.

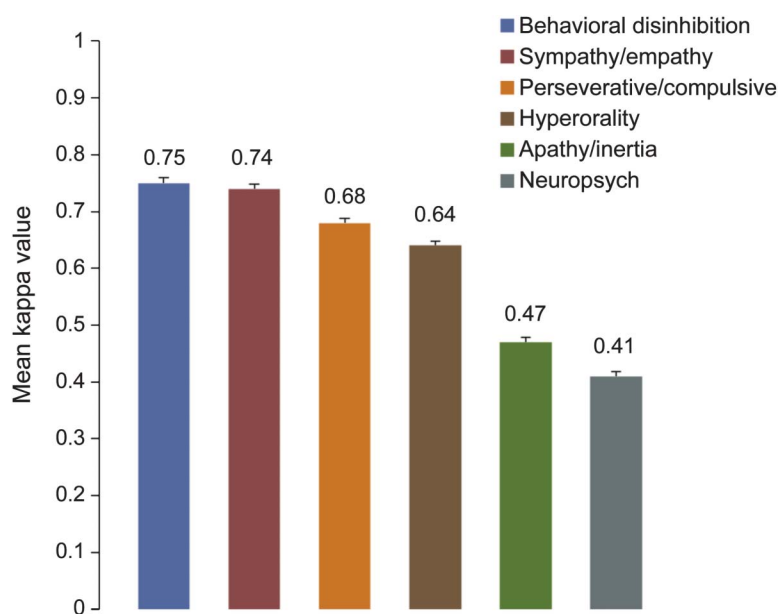
**Rater accuracy vs years of experience diagnosing dementia.** Logistic regression was applied to determine whether number of years of experience in diagnosing dementia influenced rater accuracy for diagnosis of possible and probable bvFTD. Using our center's multidisciplinary diagnosis as the "gold standard diagnosis," we coded each rater's yes/no diagnosis of probable and possible bvFTD as either correct (1) or incorrect (0). These values were then entered into the logistic regression, using number of years experience as the continuous predictor variable. Based on all 20 patient modules, the odds ratio was 1.03 (95% confidence interval: 0.98–1.08;  $p = 0.26$ ), indicating that more years of experience did not result in greater accuracy.

**Which patients were misclassified as bvFTD?** In addition to determining rater accuracy for patients who carried a diagnosis of bvFTD, we were interested in determining which patients might be susceptible to misclassification. Review of rater responses for each patient regarding diagnosis (e.g., yes/no for possible and probable bvFTD) indicated that, of all 5 non-bvFTD patient groups (AD/DLB, nfPPA, svPPA, and lvPPA) only patients with svPPA ( $n = 4$ ) were misclassified as having bvFTD. Raters misclassified the 4 patients with svPPA as meeting criteria for both possible and probable bvFTD on average of 32% of responses.

**DISCUSSION** The results of this study demonstrate the reliability of the new FTDC criteria for bvFTD. Furthermore, they indicate that clinicians with varying years of experience and professional backgrounds can accurately and reliably apply the FTDC criteria to different neurodegenerative disease presentations.

In 1999, Lopez et al.<sup>11</sup> evaluated the sensitivity, specificity, and reliability of diagnostic criteria for several neurodegenerative syndromes, including bvFTD. Using the Lund-Manchester criteria,<sup>12</sup> they found a  $\kappa$  value of 0.75 based on a total of 4 expert raters. In the present study, we found reliability coefficients above 0.8 for both possible and probable bvFTD, demonstrating improved reliability compared with the Lund-Manchester criteria. Moreover, whereas Lopez et al. achieved their results through the use of only expert raters, we were able to demonstrate excellent

**Figure** Interrater reliability of each core clinical symptom of the FTDC criteria



Mean  $\kappa$  values (standard error) per core symptom; total number of raters = 18. FTDC = International Behavioural Variant FTD Criteria Consortium.

interrater reliability for the FTDC criteria using raters from multiple centers with varying levels of expertise (e.g., residents, fellows, and experienced clinicians).

Importantly, there were few cases in which misclassification occurred. Of those that did arise, they were confined exclusively to cases with svPPA. Although the diagnostic criteria for svPPA are based on changes in language,<sup>13</sup> it is well known that patients with svPPA often demonstrate significant alterations in emotion and behavior. Indeed, several researchers have investigated behavioral symptomatology in patients with svPPA and have found significant overlap between svPPA and bvFTD regarding both particular behaviors (e.g., disinhibition, reduced empathy, compulsions, altered food preferences) and pattern of neuroanatomical degeneration.<sup>14–18</sup> As such, it seems possible that certain cases of svPPA may actually meet diagnostic criteria for both syndromes (both bvFTD and svPPA).

In addition to evaluating diagnostic accuracy, demonstrating the reliability of a set of diagnostic criteria is important because it helps to clarify which items in the criteria lead to diagnostic disagreement. For example, 2 “core” clinical symptoms in the FTDC criteria only achieved moderate agreement among raters (early apathy and neuropsychological profile). Given the ubiquity of apathy in most neurodegenerative diseases, this finding was surprising. Additionally, in the validity study by Rascovsky et al.,<sup>3</sup> it was the second most frequent feature endorsed in pathology-confirmed cases of bvFTD. One potential reason for this issue may be attributable to clinician differences in the operational definition of apathy or the interpretation of “early” apathy.<sup>19</sup> Another potential reason may relate to the idiosyncratic pattern of frontal-lobe degeneration unique to each patient. For example, patients with predominant degeneration of the medial frontal lobes tend to display significant levels of apathy, whereas those with predominant ventromedial prefrontal cortex degeneration tend to show greater levels of disinhibition.<sup>20</sup> A number of issues may have contributed to rater disagreement for the neuropsychological profile. For example, test scores can be variable, even within a cognitive domain (e.g., 2 impaired scores on memory testing, whereas one is average), leading to difficulty in making a judgment about whether a domain is impaired or not. In addition, it can be difficult to interpret neuropsychological test scores without the benefit of observing a patient’s behavior during testing. A patient can perform poorly on a test for more than one reason (e.g., perform poorly on a figure copy task because of inattention, rather than a true visuospatial deficit).<sup>21</sup> Regardless, these findings make it clear that these symptom definitions will require clarification in future revisions of the criteria.

We chose to use standardized vignette-based modules that each clinician reviewed and rated. This method is arguably very different from a “live” interaction between a clinician and the patient/caregiver, whereby the clinician pursues his or her own line of questioning based on the answers provided by the patient and informant. An alternative approach that may have better simulated a “real-life” clinical interaction would have been to have individual clinicians interview the same series of patients in person, such that each clinician could gather the information required for diagnosis using his or her own clinical judgment. It is possible that this method may have generated even larger  $\kappa$  values, as clinicians could clarify responses and gather more data than the set quantity of information provided by the vignettes. However, the alternate is also possible, should the amount and clarity of information provided in our vignettes be much greater than that gathered by the typical clinician.

For many clinicians, the most significant diagnostic challenge regarding bvFTD occurs when patients present with atypical early-onset AD<sup>22</sup> or overlapping psychiatric symptoms such as depression, compulsivity, or mania.<sup>23</sup> Unfortunately, our sample size for each diagnostic group was small (5 maximum) and weighted toward cases of bvFTD or primary progressive aphasia. Although our study represents a significant first step toward determining the reliability of the FTDC criteria, future studies should consider using a larger and more diverse set of patients to examine the criteria’s reliability.

The findings of this study support the use of the new FTDC criteria. It is important that major neurodegenerative disease research centers incorporate these criteria into their clinics and research programs so that we may evaluate their utility and continue to improve on them. As the population ages and disease-modifying therapies for neurodegenerative disease are developed, rapid, accurate diagnosis of bvFTD will be of increasing importance.

#### AUTHOR CONTRIBUTIONS

Dr. LaMarre: study design and concept, data acquisition, analysis, interpretation, drafting and revision of manuscript for intellectual content. Dr. Rascovsky: study design and concept, data acquisition, revision of manuscript for intellectual content. Dr. Bostrom: analysis and interpretation of data, revision of manuscript for intellectual content. Dr. Toofanian, Ms. Wilkins, Dr. Sha, Dr. Perry, Dr. Z. Miller, Dr. Naasan, Dr. Laforce, Dr. Hagen, Dr. Takada, Dr. Tartaglia, and Dr. Kang: data acquisition, revision of manuscript for intellectual content. Dr. Galasko, Dr. Salmon, and Dr. Farias: study design, data acquisition, revision of manuscript for intellectual content. Dr. Kaur: data acquisition, revision of manuscript for intellectual content. Dr. Olichney: study design, data acquisition, revision of manuscript for intellectual content. Dr. Quitania Park, Dr. Mendez, Dr. Tsai, Dr. Teng, Dr. Dickerson, Dr. Domoto-Reilly, and Dr. McGinnis: data acquisition, revision of manuscript for intellectual content. Dr. B. Miller: study concept and design, revision of manuscript for intellectual content. Dr. Kramer: study concept and design, obtaining funding, data acquisition, supervision of project, revision of manuscript for intellectual content.



## STUDY FUNDING

Funding for this project was provided by the California Department of Public Health and the NIH National Institute on Aging (grants P50 AG023501 and P01 AG019724). Dr. Rascovsky is supported by grants AG17586, AG32953, and NS44266.

## DISCLOSURE

A. LaMarre, K. Rascovsky, A. Bostrom, P. Toofanian, and S. Wilkins report no disclosures. S. Sha receives partial funding from the University of California San Francisco Alzheimer's Disease and Research Center. D. Perry, Z. Miller, G. Naasan, R. Laforce, J. Hagen, L. Takada, M. Tartaglia, and G. Kang report no disclosures. D. Galasko has served on an advisory board for Elan Pharmaceuticals, on DSMBs for Elan, Janssen, and Balance Therapeutics, and has received research support from the NIH (NIA 5P50AG005131, NIA 5P01AG020206) and the Michael J. Fox Foundation (20112536). D. Salmon serves as a consultant to Bristol-Myer Squibb, and receives research support from the NIH (NIA P50-AG05131 [Co-I and Project Leader], NIA U01-AG10483 [Co-I], R01 AG012674 [Co-I], R01 DC011492 [Co-I]). S. Farias receives grant funding from the NIH/NIA (grants AG031252, AG10129, AG031563, AG10220, and AG021028) although none of these grants supported the present study. She performs neuropsychological evaluations in her clinical practice and bills for these services. B. Kaur reports no disclosures. J. Olichney receives funding from NIH grant P30AG10129, and received funding from grants RO1AG018442 and RL1AG032115 and the State of California Alzheimer's Disease Program. He also receives support from Genentech for a clinical drug trial and has served as a consultant for Lundbeck Pharmaceuticals. L. Quitania Park reports no disclosures. M. Mendez receives support from NIA grant R01AG034499-03 (10%). He has received compensation for individual speaking engagements, as an editor for *UpToDate*, and for book royalties. He has also participated in a clinical drug trial for Forest Pharmaceuticals. P. Tsai serves as principal investigator for a research study (12-AVR-131) sponsored by Avanir Pharmaceuticals. E. Teng has served as a consultant for NeuroVision Imaging. Mrs. Teng owns stock in GE Healthcare and Cerner Corporation. B. Dickerson is a consultant for Pfizer and En Vivo Inc. He receives funding from the Alzheimer's Association and NIH grants MH097094, NS077059, AG030311, and AG005134. K. Domoto-Reilly received travel compensation from TauRx, and has received travel expenses and honoraria for lectures and educational activities not funded by industry. Dr. Domoto-Reilly receives research support from NIA grant AG0323065 and has received research support from NIH grant DA023427. S. McGinnis reports no disclosures. B. Miller receives grant support from the NIH/NIA and has nothing to disclose related to this article. Dr. Miller serves as a consultant for TauRx and Allon Therapeutics. He has also received a research grant from Novartis. He is on the Board of Directors for the John Douglas French Foundation for Alzheimer's Research and for The Larry L. Hillblom Foundation. J. Kramer receives grant support from the NIH/NIA and has received support from the California Department of Public Health. Go to [Neurology.org](http://Neurology.org) for full disclosures.

Received December 11, 2012. Accepted in final form February 19, 2013.

## REFERENCES

1. Knopman DS, Roberts RO. Estimating the number of persons with frontotemporal lobar degeneration in the US population. *J Mol Neurosci* 2011;45:330–335.
2. Ratnavalli E, Brayne C, Dawson K, Hodges JR. The prevalence of frontotemporal dementia. *Neurology* 2002;58:1615–1621.
3. Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 2011;134:2456–2477.
4. Kramer JH, Jurik J, Sha SJ, et al. Distinctive neuropsychological patterns in frontotemporal dementia, semantic

dementia, and Alzheimer disease. *Cogn Behav Neurol* 2003;16:211–218.

5. Possin KL, Laluz VR, Alcantar OZ, Miller BL, Kramer JH. Distinct neuroanatomical substrates and cognitive mechanisms of figure copy performance in Alzheimer's disease and behavioral variant frontotemporal dementia. *Neuropsychologia* 2011;49:43–48.
6. Morris JC. The clinical dementia rating (CDR): current version and scoring rules. *Neurology* 1993;43:2412–2414.
7. Cummings JL. The neuropsychiatric inventory: assessing psychopathology in dementia patients. *Neurology* 1997;48:S10–S16.
8. Pfeffer RI, Kurosaki TT, Harrah CH Jr, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol* 1982;37:323–329.
9. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
11. Lopez OL, Litvan I, Catt KE, et al. Accuracy of four clinical diagnostic criteria for the diagnosis of neurodegenerative dementias. *Neurology* 1999;53:1292–1299.
12. The Lund and Manchester Groups. Clinical and neuropathological criteria for frontotemporal dementia. *J Neurol Neurosurg Psychiatry* 1994;57:416–418.
13. Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. *Neurology* 2011;76:1006–1014.
14. Seeley WW, Bauer AM, Miller BL, et al. The natural history of temporal variant frontotemporal dementia. *Neurology* 2005;64:1384–1390.
15. Josephs KA, Whitwell JL, Knopman DS, et al. Two distinct subtypes of right temporal variant frontotemporal dementia. *Neurology* 2009;73:1443–1450.
16. Rosen HJ, Gorno-Tempini ML, Goldman WP, et al. Patterns of brain atrophy in frontotemporal dementia and semantic dementia. *Neurology* 2002;58:198–208.
17. Rosen HJ, Allison SC, Schauer GF, Gorno-Tempini ML, Weiner MW, Miller BL. Neuroanatomical correlates of behavioural disorders in dementia. *Brain* 2005;128:2612–2625.
18. Whitwell JL, Przybelski SA, Weigand SD, et al. Distinct anatomical subtypes of the behavioural variant of frontotemporal dementia: a cluster analysis study. *Brain* 2009;132:2932–2946.
19. Clarke DE, Ko JY, Kuhl EA, van Reekum R, Salvador R, Marin RS. Are the available apathy measures reliable and valid? A review of the psychometric evidence. *J Psychosom Res* 2011;70:73–97.
20. Liu W, Miller BL, Kramer JH, et al. Behavioral disorders in the frontal and temporal variants of frontotemporal dementia. *Neurology* 2004;62:742–748.
21. Thompson JC, Stopford CL, Snowden JS, Neary D. Qualitative neuropsychological performance characteristics in frontotemporal dementia and Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2005;76:920–927.
22. Kelley BJ, Boeve BF, Josephs KA. Cognitive and noncognitive neurological features of young-onset dementia. *Dement Geriatr Cogn Disord* 2009;27:564–571.
23. Woolley JD, Khan BK, Murthy NK, Miller BL, Rankin KP. The diagnostic challenge of psychiatric symptoms in neurodegenerative disease: rates of and risk factors for prior psychiatric diagnosis in patients with early neurodegenerative disease. *J Clin Psychiatry* 2011;72:126–133.