Accuracy of the Clinical Evaluation for Frontotemporal Dementia

Mario F. Mendez, MD, PhD; Jill S. Shapira, RN, PhD; Aaron McMurtray, MD; Eliot Licht, MD; Bruce L. Miller, MD

Background: Without a definitive clinical test, the early diagnosis of frontotemporal dementia (FTD) can be difficult.

Objective: To evaluate the accuracy of the clinical evaluation for FTD.

Design: Retrospective assessment of consensus criteria for FTD, neuropsychological measures, magnetic resonance images, and single-photon emission computed tomography/positron emission tomography (SPECT/PET) scans at baseline compared with a standard of subsequent clinical diagnosis after follow-up and reevaluation to year 2.

Setting: University hospital.

Patients: A total of 134 patients referred for clinical evaluation of suspected FTD. These patients had 1 or more core or supportive features of FTD in the absence of another etiology on initial assessment.

Main Outcome Measures: Sensitivities, specificities, and predictive values of consensus criteria for FTD, magnetic resonance images, and SPECT/PET scans at initial assessment.

Results: The sensitivities and specificities for the diagnosis of FTD were 36.5% and 100.0% for consensus criteria, 63.5% and 70.4% for magnetic resonance images, and 90.5% and 74.6% for SPECT/PET scans, respectively. With a previous prevalence of nearly 50% for FTD, the positive predictive value was greatest for consensus criteria (100.0%), and the negative predictive value was greatest for SPECT/PET (89.8%). The initial neuropsychological results did not distinguish FTD, but the pattern of progression (worse naming and executive functions and preserved constructional ability) helped establish the diagnosis at year 2.

Conclusions: Consensus criteria for FTD and neuropsychological measures lacked sensitivity for FTD; however, neuroimaging, particularly functional brain studies, greatly increased the sensitivity of detecting FTD. The clinical diagnosis of FTD needs to combine neuropsychiatric features with SPECT or PET findings while following the changes on neuropsychological tests.

Arch Neurol. 2007;64:830-835

Frontotemporal Dementia (FTD) is a devastating neurodegenerative disease that commonly afflicts people in middle age, when they are in the prime of life. The early diagnosis of FTD is critical for developing management strategies and interventions for these patients. Most individuals who develop FTD manifest with new-onset behavioral symptoms, and, in the absence of a biomarker, the clinical diagnosis depends on recognizing all the core or necessary, neuropsychiatric features of FTD.

Practicing physicians continue to have difficulty diagnosing early FTD. These patients often lack all the necessary core features for the clinical diagnosis of FTD and fail to meet diagnostic criteria on initial assessment. Many of the initial symptoms of this disorder are compatible with a range of neurologic and psychiatric disorders. Conversely, physicians misdiagnose FTD in patients with Alzheimer disease (AD), other neurologic disorders, and primary psychiatric conditions. Consequently, patients with FTD can go from physician to physician delaying diagnosis and risking inappropriate therapy. Despite this diagnostic confusion, there are few data on the accuracy of a clinical evaluation for FTD.

This study aims to clarify the usefulness of the clinical evaluation in diagnosing FTD among neurologic and psychiatric patients suspected of having that disorder. This population of patients with "possible FTD" is the most fitting population for assessing the accuracy of the clinical evaluation for FTD. Among these patients, we evaluated the initial sensitivity and specificity of consensus criteria, neuropsychological screening tests, and clinical neuroimaging compared with a standard of subsequent clinical diagnosis at year 2.
This study evaluated and followed up 134 patients with suspected FTD. They were evaluated when first seen and again after 2 years of clinical follow-up. In this population, we determined the sensitivities, specificities, and predictive values of the initial consensus criteria for FTD, neuropsychological testing, and clinical structural and functional neuroimaging in predicting a clinical diagnostic standard of FTD at year 2. The study further determined the diagnoses of patients who did not meet the criteria for FTD at year 2.

**METHODS**

**OVERVIEW**

We reviewed all referrals to the University of California at Los Angeles FTD and Neurobehavior Program for consideration of the behavioral variant of FTD during a 4-year period. All potential participants with FTD in this study were initially evaluated in the neurologic clinics of the University of California at Los Angeles, including the Focal-type Dementias Clinic and the Frontotemporal Dementia and Neurobehavior Clinic. The participants were community-based patients referred by other physicians for further opinion. This study excluded patients with language-predominant variants (primary progressive aphasia or semantic dementia) and frontotemporal lobar degeneration. This study included only patients who had a second assessment at year 2 or who died.

The referred patients had seen psychiatrists, neurologists, and primary care physicians for new-onset behavioral symptoms in middle to late life resulting in suspected FTD. After the assessment failed to reveal an etiology, the physicians referred their patients to our FTD program for consideration of that diagnosis. These patients had a greater likelihood of FTD than other, unselected patients with similar neuropsychiatric symptoms because they had already been screened for obvious neurologic or psychiatric disorders by their physicians. Hence, these patients constitute the base population of interest in the consideration of the sensitivity and specificity of the diagnosis of FTD.

**INITIAL CLINICAL ASSESSMENT**

The patients underwent a baseline workup. All of the patients were evaluated by 1 of us (M.F.M.). In addition to interview, neurologic examination, and laboratory tests, the patients underwent dementia scales, neuropsychological screening measures, and review of their neuroimaging findings. The dementia scales included the Mini-Mental State Examination and the Clinical Dementia Rating Scale, a good measure of functional impairment in FTD. If not already completed by their physicians and available for review, the patients also underwent magnetic resonance imaging (MRI) and either single-photon emission computed tomography (SPECT) or positron emission tomography (PET).

The neuropsychological screening measures included the Consortium to Establish a Registry in Alzheimer’s Disease (CERAD) and the Frontal Assessment Battery. The neurologist administers these tests as part of the evaluation during the clinic visits. The Consortium to Establish a Registry in Alzheimer’s Disease includes verbal fluency, the Boston Naming Test (short 15-item version), constructions (copy of a circle, a rhombus, overlapping rectangles, and a cube), and memory tests (word list memory for 10 words on Trials I-III, word list recall, and a true-false memory recognition test). The Consortium to Establish a Registry in Alzheimer’s Disease memory measures were abbreviated to 2 scores: the Savings Score (Delayed Recall Learning/Learning Trial III as a proportion of 10 words) and the accurate “yes” answers on Recognition. The Frontal Assessment Battery consists of 6 items rated on scales from 0 to 3 and encompassing similarities, 5 word fluency, alternate programs, alternate tapping, the go/no-go test, and grasp reflex testing.

The diagnostic assessment included clinical interpretations of MRI and SPECT or PET. The studies were obtained at different medical centers using different scanners; therefore, we could not directly combine MRI or SPECT/PET information from the different sites. Moreover, some patients had SPECT scans and others had PET scans; earlier referrals were more likely to have had SPECT than PET scans. Nevertheless, the examiners reread all outside studies, brought by hard copy or by CDROM for review, and those obtained at the University of California at Los Angeles. We interpreted the imaging based on the absence or presence of frontotemporal changes in the absence of corresponding changes in more posterior areas of the brain. The technique used for rereading clinical images from different sources has been previously reported. Blinded visual repeated inspections involve a quadrant approach, with 2 raters grading the studies for atrophy, hypometabolism, or hypoperfusion on a 4-point scale (0=absent, 1=mild, 2=moderate, and 3=severe) for each of the left frontal, right frontal, left anterior temporal, and right anterior temporal regions. Previous interrater reliability for 2 of the raters was high ($r=0.71$ for 296 ratings; $P<.001$). This study analyzed these changes and their contribution to the diagnosis separately from core consensus criteria.

On initial interview the physicians determined the presence of the diagnostic features of FTD. Clinical FTD was defined as meeting the 5 core criteria of the clinical consensus criteria for FTD (Table 1). The decision about the presence of a core symptom was guided by the presence of a change in behavior that deviated from the patient’s “normal” baseline. For this study, possible FTD was defined as follows: (1) does not meet clinical consensus criteria for FTD (does not have all 5 necessary core diagnostic features (Table 1), (2) has an insidious onset and gradual progression of either 1 or more of the 5 core behavioral diagnostic features or 1 or more of the 6 supportive behavioral diagnostic features (Table 1), and (3) absence of a neurologic or psychiatric illness or condition that could entirely account for and explain the patient’s core or supportive neuropsychiatric behavior(s).

After initial assessment the patients were followed up for 2 years or until death. The surviving patients underwent a second diagnostic assessment that included clinical examinations, neuropsychological testing, and either SPECT or PET (MRI).
was not routinely repeated. The follow-up SPECT/PET scans were not necessarily consistent with the original SPECT/PET scans. Sensitivities, specificities, and predictive values of the consensus criteria, MRI, and SPECT/PET were assessed by comparing the FTD diagnoses at baseline (initial assessment) with the final diagnosis (follow-up) used as the clinical diagnostic standard.

**DATA ANALYSIS**

In addition to calculations of sensitivities, specificities, and predictive values, simple group comparisons involved t tests (age and educational level) and χ² tests (sex). All other comparisons involved factorial analysis of variance with 2 independent variables: eventual FTD vs non-FTD group membership at year 2 and initial vs year 2 results. The Tukey B test was used for post hoc analysis.

**RESULTS**

We identified 134 patients with possible FTD who were followed up to year 2 (Table 2). By year 2, 63 were diagnosed as having FTD (FTD patients) and 71 as having other conditions (non-FTD patients). There were no clear differences between FTD patients and non-FTD patients on demographic variables; however, the 2 groups differed on Mini-Mental State Examination and Clinical Dementia Rating Scale scores, with FTD patients having worse scores at year 2 compared with the non-FTD group, which included many psychiatric patients who actually improved on follow-up.

**CLINICAL DIAGNOSES**

On initial assessment, only 23 patients (17.2%) met all of the core consensus criteria for FTD. By year 2, 40 (29.9% of the total; 36.0% of the remaining 111 patients) additional patients developed other manifestations of FTD sufficient to meet the consensus criteria for that disorder. This represented a conversion rate of 20% per year; thus, 80% remained not converted in 1 year, and 64% (80% × 80%) remained not converted at year 2. Another 36 patients (26.9%) eventually had a diagnosis of a primary psychiatric disorder, 17 (12.7%) had AD (with some frontal features), 9 (6.7%) had another neurologic disorder, and 9 remained undiagnosed (Table 3). The psychiatric diagnoses were based on current psychiatric criteria and were corroborated by the team psychiatrists. Patients with AD met the research criteria for AD and were supported by deteriorations in memory and visuospatial skills. The remaining undiagnosed patients did not clearly meet any diagnostic criteria at year 2.

**CONSENSUS CRITERIA**

Sensitivities, specificities, and predictive values were calculated for the initial consensus criteria using the final diagnosis by year 2 as the definitive result. For the consensus criteria, sensitivity was 36.5% and specificity was near 100%, since all 23 patients initially diagnosed as having FTD progressed to other manifestations of the disorder. Using the pretest prevalence of 47% for FTD patients in this population, the positive predictive value of an initial consensus criteria diagnosis was 100% because there were no false positives, but the negative predictive value was only 64% owing to many false-negatives.
The frequent misdiagnosis of FTD is often difficult to diagnose early in its course. Patients with FTD initially manifest subtle personality or behavioral changes that could be caused by a range of other diagnostic conditions. In this study, the largest series to date of patients with suspected FTD, the consensus criteria for FTD lacked sensitivity for detecting FTD but had specificity in excluding other neurologic or psychiatric disorders. Neuropsychological screening measures also lacked sensitivity for detecting FTD but were helpful in confirming FTD across time. In contrast, neuroimaging, particularly functional imaging, was most sensitive and helped greatly in the initial detection of patients with FTD.

The accurate diagnosis of FTD is increasingly important for the development of new therapeutic options or disease-modifying therapies for this disease. However, physicians often misdiagnose many if not most patients with early FTD or are significantly delayed in arriving at the correct diagnosis. The few clinicopathologic studies available suggest a wide range of clinical accuracy for diagnosing FTD. Still other studies indicate a significant delay in diagnosis of 3 to 4 years or more. Even at tertiary care centers, the diagnostic rate for frontotemporal lobar degeneration reaches only approximately 85% by the time of death. The frequent misdiagnosis of FTD occurs because of the variable behavioral presentations of this disease.

### Table 4. Characteristics of Patients With Possible Frontotemporal Dementia: Neuropsychological Measures*

<table>
<thead>
<tr>
<th></th>
<th>FTD Patients</th>
<th>Non-FTD Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial</td>
<td>Year 2</td>
</tr>
<tr>
<td>Verbal digit span</td>
<td>6.48 ± 1.12</td>
<td>6.17 ± 1.37</td>
</tr>
<tr>
<td>Fluency: animals</td>
<td>13.71 ± 7.80</td>
<td>12.57 ± 6.71</td>
</tr>
<tr>
<td>Boston Naming Test</td>
<td>11.90 ± 2.41</td>
<td>10.30 ± 2.72</td>
</tr>
<tr>
<td>CERAD: Savings</td>
<td>6.22 ± 2.67</td>
<td>5.14 ± 2.46</td>
</tr>
<tr>
<td>CERAD: Recognition</td>
<td>7.95 ± 1.30</td>
<td>7.46 ± 1.20</td>
</tr>
<tr>
<td>Constructions score</td>
<td>9.59 ± 1.44</td>
<td>9.11 ± 1.70</td>
</tr>
<tr>
<td>Total FAB score</td>
<td>10.87 ± 3.39</td>
<td>9.44 ± 3.22</td>
</tr>
<tr>
<td>Similarities</td>
<td>1.96 ± 0.58</td>
<td>1.73 ± 0.55</td>
</tr>
<tr>
<td>Fluency: S word</td>
<td>2.14 ± 1.15</td>
<td>1.76 ± 0.93</td>
</tr>
<tr>
<td>Alternate programs</td>
<td>1.86 ± 1.00</td>
<td>1.71 ± 0.89</td>
</tr>
<tr>
<td>Alternate tapping</td>
<td>1.22 ± 0.73</td>
<td>0.91 ± 0.76</td>
</tr>
<tr>
<td>Go/no-go test</td>
<td>1.70 ± 0.56</td>
<td>1.40 ± 0.49</td>
</tr>
<tr>
<td>Grasp reflex testing</td>
<td>2.44 ± 0.62</td>
<td>2.32 ± 0.69</td>
</tr>
</tbody>
</table>

Abbreviations: CERAD, Consortium to Establish a Registry in Alzheimer’s Disease; FAB, Frontal Assessment Battery; Recognition, accurate yes answers on Recognition; Savings, Savings Memory Score (Delayed Recall Learning/Learning Trial III as a proportion of 10 words).

*Data are presented as mean ± SD unless otherwise indicated.
†Significant differences between the initial evaluation and year 2 on post hoc analysis.
‡Significant differences between the FTD and non-FTD groups on post hoc analysis.

### Comment

Frontotemporal dementia is often difficult to diagnose early in its course. Patients with FTD initially manifest subtle personality or behavioral changes that could be caused by a range of other diagnostic conditions. In this study, the largest series to date of patients with suspected FTD, the consensus criteria for FTD lacked sensitivity for detecting FTD but had specificity in excluding other neurologic or psychiatric disorders. Neuropsychological screening measures also lacked sensitivity for detecting FTD but were helpful in confirming FTD across time. In contrast, neuroimaging, particularly functional imaging, was most sensitive and helped greatly in the initial detection of patients with FTD.

The accurate diagnosis of FTD is increasingly important for the development of new therapeutic options or disease-modifying therapies for this disease. However, physicians often misdiagnose many if not most patients with early FTD or are significantly delayed in arriving at the correct diagnosis. The few clinicopathologic studies available suggest a wide range of clinical accuracy for diagnosing FTD. Still other studies indicate a significant delay in diagnosis of 3 to 4 years or more. Even at tertiary care centers, the diagnostic rate for frontotemporal lobar degeneration reaches only approximately 85% by the time of death.

The frequent misdiagnosis of FTD occurs because of the variable behavioral presentations of this disease.
The diagnosis of FTD is not facilitated by any available, objective biomarkers, and the accurate diagnosis of FTD depends on the recognition of clinical behavioral features. The earliest manifestations of FTD are usually changes in social and emotional behavior. Within the first few years after onset, neuropsychiatric symptoms usually precede or overshadow any cognitive disabilities. As demonstrated in this study, the profile of neuropsychological abnormalities in executive functions and language, with less impaired memory and visuospatial skills than AD, may emerge only as the disease progresses and lacks sensitivity in the beginning stages of FTD. Furthermore, the earliest behavioral manifestations of this disorder include a great deal of variability associated with variability in the earliest localization of the disease and possibly in the neuropathologic features.

In this study, patients with possible FTD had a range of diagnoses. Despite the fact that behavioral abnormalities are common in early FTD and that memory loss is a key feature in early AD, the single most common misdiagnosis for FTD seems to be AD. Both dementias have an insidious onset and progression of behavioral changes that cause substantial impairment in social or occupational functioning. Patients with possible FTD can have many other dementias or encephalopathies. Physicians may also mistake FTD for psychiatric conditions. One study reports that 7 of 12 FTD patients first saw psychiatrists for presumed primary psychiatric disorders. The apathy and disengagement of early FTD can mimic atypical depression, and the disinhibition and impulsivity of FTD may resemble anxiety or personality disorders. Finally, physicians may even confuse FTD patients with having schizophrenia or atypical psychosynthesis of late onset.

Although other criteria have been proposed, the most commonly used diagnostic criteria, the consensus criteria for FTD, focus on progressive declines in neuropsychiatric areas (Table 1). To be diagnosed as having FTD, patients must have all of the core diagnostic features. This study confirms that the consensus criteria for FTD are insensitive for the early diagnosis of FTD. In consensus criteria, neuropsychological and neuroimaging studies are supportive features only that increase specificity and the certainty of the diagnosis of FTD. This study shows that core diagnostic features are already sufficiently specific in distinguishing FTD from the range of neurologic and psychiatric disorders, and further specificity from supportive features is unnecessary. This reflects a deficiency in the consensus criteria because core diagnostic features and supportive features focus on specificity rather than on a combination of sensitive and specific measures for detecting and diagnosing FTD.

The supportive features can facilitate the early diagnosis of FTD in other ways. Although neuropsychological screening measures were not helpful in the initial detection of FTD, across time they showed worse naming, worsening executive functions, and better constructions among FTD patients compared with non-FTD patients. Use of MRI greatly increased the sensitivity of detecting FTD, and functional imaging with SPECT or PET was even better, achieving a sensitivity for detecting FTD of more than 90%. These findings suggest new diagnostic criteria that incorporate SPECT or PET as core diagnostic features rather than as supportive features and use neuropsychological tests for follow-up assessment and confirmation.

There are several additional considerations in interpreting the results of this study. First, it is a retrospective investigation that focuses on a potentially widely accessible clinical evaluation. This study benefits from the ecologic validity of this type of clinical evaluation and the potential generalizability of the findings. Second, this study involves a specific referral population of patients with possible FTD, including many patients with psychiatric problems. This reflects the possible FTD population of interest and the fact that FTD patients often first visit psychiatrists. Moreover, the pretest prevalence of FTD of approximately 50% is what might be generally expected among patients suspected of having FTD. Third, there was much variability in the neuroimaging. Again, this is the usual situation encountered in clinical settings, where there is available visual inspection of images from different scanners but little or no quantitative assessment of images. Fourth, the neuropsychological measures were limited. An extensive neuropsychological battery could have been more sensitive to the early changes of FTD, but extensive testing is not available or practical in many clinical settings. Finally, there was no neuropathologic diagnosis, only a clinical diagnostic standard. Ultimately, some of the diagnoses at year 2 may prove to be incorrect on autopsy. Consequently, we continue to follow up as many of these patients as possible to death and autopsy for clinicopathologic verification of their diagnoses.

Frontotemporal dementia is a devastating neurodegenerative disease, and the early diagnosis of this disorder is critical for developing treatment strategies and interventions for these patients. In the absence of a biomarker, however, FTD remains difficult to diagnose and may be confused with other neurologic or psychiatric disorders. Physicians need better and more sensitive diagnostic methods that are practical and easily available. Functional neuroimaging, in particular, can contribute to the neuropsychiatric criteria. Alternative diagnostic criteria to be tested in future studies should include several of the core or supportive behavioral features plus functional neuroimaging and serial neuropsychological measures.

Accepted for Publication: October 22, 2006.
Correspondence: Mario F. Mendez, MD, PhD, Neurobehavior Unit (116AF), VA Greater Los Angeles Healthcare, 11301 Wilshire Blvd, Los Angeles, CA 90073 (mmendez@ucla.edu).
Author Contributions: The corresponding author had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Mendez, McMurtray, and Miller. Acquisition of data: Mendez, Shapira, and Licht. Analysis and interpretation of data: Mendez. Drafting of the manuscript: Mendez, McMurtray, and Licht. Critical revision of the manuscript for important intellectual content: Mendez, Shapira, and Miller. Statistical analysis: Mendez. Obtained funding: Miller.
Administrative, technical, and material support: Shapira and Licht. Study supervision: Mendez.

Financial Disclosure: None reported.

Funding/Support: This research was supported in part by grant AG19724-01 from the National Institute on Aging, by the UCLA Alzheimer's Disease Center, and by the State of California.

REFERENCES