Cognitive decline in the behavioral variant of frontotemporal dementia

Janine Diehl-Schmid, Susanne Bornschein, Corina Pohl, Hans Förstl, Alexander Kurz and Thomas Jahn

Department of Psychiatry of Technische Universität München, Munich, Germany

ABSTRACT

Background: Only a small number of studies on the natural disease course in behavioral variant frontotemporal dementia (bvFTD) have been conducted. This is surprising because knowledge about the progression of symptoms is a precondition for the design of clinical drug trials.

Methods: The aim of the present study was to examine the cognitive decline of 20 patients with mild bvFTD over one year using the Consortium to Establish a Registry for Alzheimer’s Disease – Neuropsychological Assessment Battery (CERAD-NAB).

Results: Within an average follow-up interval of 13 months, patient scores declined significantly in the Mini-mental-State-Examination (MMSE) and the CERAD-NAB subtests of naming, verbal and nonverbal memory. No significant changes were found in the CERAD-NAB subtests of category fluency, recognition, and visuoconstruction. The average annualized decline on the MMSE was 4.0 ± 4.9 points. Ceiling effects were detected in Figures Copy, Word List Recognition and Modified Boston Naming Test. Though the included patient group was rather homogeneous regarding severity of dementia, the cognitive changes were very heterogeneous.

Conclusion: Given the heterogeneity of cognitive decline, the design of a test battery for clinical trials in FTD will be challenging. A cognitive battery should definitely include the MMSE, Word List Learning and Word List Delayed Recall.

Key words: frontotemporal dementia (FTD), cognitive decline, CERAD-NAB, MMSE, neuropsychology, annual rate of change

Introduction

The behavioral variant of frontotemporal dementia (bvFTD), sometimes referred to as frontal variant FTD, fvFTD (Hodges and Miller, 2001) or just FTD (Neary et al., 1998), is a disorder caused by degeneration predominantly of the frontal cortex. The prevalence is low with about 18 cases per 100,000 persons (Borroni et al., 2009).

BvFTD is a progressive disorder and, according to the revised Lund-Manchester criteria from 1998, is characterized by insidious onset and gradual progression, early decline in personal and social behavior, early emotional blunting and early loss of insight into the condition (Neary et al., 1998).

One of the first symptoms of bvFTD is the dissolution of social attachment. Patients are indifferent and unconcerned about their spouses, relatives or friends. Social behavior often becomes superficial, tactless, rude and sometimes foolish. Disinhibition may be observed. Patients are inflexible, rigid, and unable to adjust to their environment. Affect is typically shallow and monotonous. Some patients show euphoric or even hypomanic mood. Patients may be overactive and show motor restlessness or pressure of speech. Apathetic behavior, however, is more frequent (Neary et al., 1998; Diehl and Kurz, 2002). As the disease progresses, cognitive symptoms always emerge (Kramer et al., 2003).

The pattern of cognitive impairment in bvFTD has been investigated in numerous cross-sectional studies, mainly in comparison with cognitively healthy individuals or patients with Alzheimer’s disease (AD). In patients with bvFTD, episodic memory and semantic memory are relatively...
preserved compared to patients with AD (Kramer et al., 2003). Also, spatial abilities tend to be intact (Diehl and Kurz, 2002). In contrast, patients with bvFTD are significantly impaired in tests of verbal (letter and semantic category) fluency (Diehl et al., 2005).

While the progression of AD has been thoroughly characterized, the natural course of bvFTD is less well known. However, with the advent of disease-modifying drugs it is important to understand accurately the evolution of symptoms and the progression of cognitive decline. Initial follow-up studies regarding behavioral symptoms have been published (Marczinski et al., 2004; Diehl-Schmid et al., 2006).

Follow-up studies focusing on cognitive symptoms are now beginning to characterize the longitudinal change in FTLD spectrum disorders. The rate of cognitive decline in FTLD has been mostly investigated in small studies using the Mini-mental State Examination (MMSE) (Pasquier et al., 2006; Chow et al., 2006). Other studies have used more comprehensive neuropsychological test batteries but have failed to differentiate between the language and behavioral variants of FTLD since they included patients who had FTLD-pathology at autopsy (Rascovsky et al., 2008; Grossman et al., 2008). To date, only a few studies have focused on the behavioral variant of FTD, comparing the rate of cognitive decline with other FTLD-variants or AD, respectively (Blair et al., 2007; Wicklund et al., 2007; Knopman et al., 2008; Libon et al., 2009).

It is still unclear in which cognitive domains changes can be measured within a one-year observational period and of a magnitude that would allow the detection of modest drug effects in the form of a delay or arrest of progression. All follow-up studies mentioned above included patients who were enrolled at different clinical stages of the disease. However, the annual rate of cognitive change in FTD is supposed to be strongly associated with the severity of dementia at baseline as has been shown in AD (Stern et al., 1994). With regard to the probable non-linearity of symptom progression the present study aimed to investigate the cognitive decline in patients who were relatively homogeneous with regard to severity of dementia at enrolment. All patients included in the study were diagnosed with bvFTD and were at a mild stage of the disease. Such patients would typically be included in pilot drug trials. The scheduled follow-up period was one year, which is a time period often chosen for drug trials. For the assessment of cognitive ability we used the Consortium to Establish a Registry for Alzheimer’s Disease – Neuropsychological Assessment Battery (CERAD-NAB; Welsh et al., 1994), an instrument that is widely administered in memory clinics worldwide (Chandler et al., 2005) and which has proven to be useful in the diagnosis and differential diagnosis of bvFTD, SD and AD previously (Diehl et al., 2005).

**Methods**

Over one year, 44 patients with FTD (bvFTD) were diagnosed in our memory clinic. The diagnosis was established according to the revised Lund-Manchester criteria (Neary et al., 1998) by consensus of two psychiatrists with profound experience in the field of FTLD (J.D-S. and A.K.). Diagnoses were based on information gathered from psychiatric assessment including the Geriatric Depression Scale (GDS; Sheikh and Yesavage, 1986) and a detailed informant interview including standardized questionnaires. Behavioral disturbances were assessed using the Frontal Behavioural Inventory (FBI; Kertesz et al., 1997) or the Neuropsychiatric Inventory (NPI; Cummings, 1996). An extensive neuropsychological workup (see below) was performed. All patients had a neurological examination, showing that none of the patients suffered from corticobasal degeneration, progressive supranuclear palsy or amyotrophic lateral sclerosis (ALS) at baseline. A routine laboratory work-up was performed. Structural neuroimaging (cranial computed tomography or magnetic resonance imaging) was performed to exclude significant vascular lesions, tumors and other structural changes except atrophy. 18F-fluorodeoxyglucose positron emission tomography (18F-FDG PET) was performed on all patients. All included patients showed frontal or frontotemporal hypometabolism in PET.

For this study we aimed to include patients at a mild stage of disease. Out of the 44 patients diagnosed with bvFTD within one year, 31 had a Clinical Dementia Rating (CDR) score of 0.5 or 1 and were therefore available for study inclusion. However, 12 patients were lost to follow-up, since they died (n = 1), could not be contacted (n = 2), or their proxy refused participation (n = 1). Two patients were excluded as they had developed ALS. Mostly due to administrative reasons we were not able to perform the follow-up visit for all patients exactly 12 months after baseline, so we allowed a time frame from 10 to 17 months. Six patients were excluded from the study because the follow-up visit could not be performed within this time frame. Patients were not excluded from the analyses for any other reason, so that finally 20 patients with bvFTD were included in the study, all of them in a mild stage of dementia at baseline (CDR = 0.5 (N = 7) or 1 (N = 13)). Informed consent according to the Declaration of Helsinki was obtained for every
patient. The study protocol was approved by the local ethics committee.

Neuropsychological assessments at the first visit (baseline) and the follow-up visit included psychometric tests which focused on memory, language, and executive functions (i.e. California Verbal Learning Test, Aachener Aphasie Test, Color-Word-Test, Trail Making Test A and B). All patients underwent the CERAD-NAB. This instrument comprises the following subtests: (1) Category Fluency (animals), (2) modified Boston Naming Test, (3) MMSE, (4) Word List Learning trial 1 to 3, (5) Word List Delayed Recall, (6) Word List Recognition, (7) Figures Copy, and (8) Figures Delayed Recall. From the Word List Recognition subtest, a discriminability measure is calculated as the number of correct “yes” plus number of correct “no” responses divided by 20 and then multiplied by 100, which results in the percentage of correct responses. Two additional measures were calculated as follows: Word List Savings is calculated as Word List Delayed Recall divided by word list trial 3 multiplied by 100. Figures Savings is calculated as Figures Delayed Recall divided by Figures Copy multiplied by 100. Furthermore, a CERAD Global Score was obtained by summing scores from the individual subtests, excluding the MMSE, into a total composite – a method used by other groups previously (Chandler et al., 2005).

A follow-up examination was performed an average of 13.3 ± 2.3 months later (10 to 17 months). Severity of dementia (as measured by CDR) worsened during this time: at follow-up two patients had a CDR global score = 0.5, seven patients were rated CDR = 1 and 11 patients were rated CDR = 2. The follow-up assessment included psychiatric assessment, neurological examination and informant interview including standardized questionnaires. A follow-up 18F-FDG PET was performed in the majority of patients. Neuropsychological assessments again included psychometric tests that focused on memory, language, and executive functions as well as the CERAD-NAB.

Non-parametric Wilcoxon-tests were used to compare raw scores and normative (age, sex and education corrected) z-scores at baseline and follow-up ($\alpha = 0.05$; two-tailed). Normative z-scores were calculated in order to compare the pre/post differences of all CERAD-NAB subtests on a common metric. In addition, a global z-score was determined by averaging the z-scores from all subtests. The estimated population means and standard deviations for the z-transformations of the raw scores were taken from the CERAD-NAB (German version) norms provided by the Basel Memory Clinic (www.memory-clinic.ch), which include data from 1100 healthy older subjects. Since the majority of the CERAD-NAB variables are moderately to extremely skewed in the normative sample, the resulting z-scores cannot be interpreted within a standard normal distribution model, but nevertheless serve as a common metric for all CERAD-NAB subtest variables and the summarizing global score, as intended.

Furthermore the annual rate of change (ARC) of each CERAD-NAB subscore was calculated as follows: raw score at follow-up visit minus raw score at baseline divided by the time in months multiplied by 12. To detect potential floor or ceiling effects, the number of patients who scored at the highest level (ceiling) and the worst level (floor) was counted for each CERAD subtest score.

Results

The demographic characteristics and MMSE scores at baseline are shown in Table 1. The mean CERAD-NAB raw scores at baseline and follow-up and the corresponding annualized rates of change (ARC) are shown in Table 2. At baseline, patients performed significantly worse compared to cognitively healthy persons on all CERAD subtests except Figures Copy (see Table 1). Within an average follow-up interval of 13 months patients declined significantly in the following CERAD-NAB subtests: modified Boston Naming Test, MMSE, Word List Learning Total, Word List Delayed Recall, Figures Delayed Recall, and Figures Savings. No significant changes were found on the CERAD-NAB subtests Category Fluency, Word List Savings, Word List Recognition, and Figures Copy. The average annualized decline on the MMSE was 4.0 ± 4.9 points.

Figure 1 illustrates the pre/post differences in the normative (age, sex and education corrected) z-scores of the CERAD-NAB. In addition to the CERAD global z-score, which demonstrates a highly significant overall cognitive decline from pre to post testing, six out of ten more specific test variables also indicate significant ($0.05 \leq p \leq 0.001$) deterioration from pre to post testing.

Significant floor effects were not found in any subtests. There were several instances of ceiling effects, particularly at baseline (naming, recognition, visuoconstruction) (see Table 3).

In 15 out of 20 patients the FBI was administered at baseline and at follow-up. Mean FBI score at baseline was 24.3 ± 9.8 (6–37). At follow-up the mean FBI score was 29.7 ± 9.8 (6–51). The increase on the FBI was statistically significant ($p < 0.05$), reflecting a worsening of behavioral abnormalities between baseline and follow-up.
The results of a small number of clinical trials that have been mostly single-centered and open label. These studies identified trazodone and selective serotonergic reuptake inhibitors (Huys et al., 2006) as useful in the symptomatic treatment of bvFTD, whereas treatment with memantine does not positively influence cognition or behavior (Diehl-Schmid et al., 2008; Boxer et al., 2009).

Table 1. Education, age, age at onset of first symptoms, duration of disease and MMSE score at baseline (mean values and standard deviations)

<table>
<thead>
<tr>
<th>EDUCATION (YEARS)</th>
<th>AGE (YEARS)</th>
<th>AGE AT ONSET (YEARS)</th>
<th>DURATION OF DISEASE (YEARS)</th>
<th>MMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>13.4 (4.2)</td>
<td>63.0 (7.9)</td>
<td>59.1 (7.6)</td>
<td>3.9 (2.4)</td>
</tr>
<tr>
<td></td>
<td>24.7 (2.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Mean (SD) CERAD-NAB subtest raw scores at baseline and follow-up together with corresponding mean (SD) annualized rates of change (ARC)

<table>
<thead>
<tr>
<th></th>
<th>MAXIMAL POSSIBLE SCORE</th>
<th>BASELINE RAW SCORES</th>
<th>FOLLOW-UP RAW SCORES</th>
<th>WILCOXON’S Z</th>
<th>ARC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category Fluency</td>
<td>(about 40)</td>
<td>9.2 (4.2)</td>
<td>8.1 (5.4)</td>
<td>−0.856</td>
<td>−1.0 (3.7)</td>
</tr>
<tr>
<td>Naming Test</td>
<td>15</td>
<td>12.5 (3.0)</td>
<td>11.0 (4.2)</td>
<td>−2.172*</td>
<td>−1.3 (2.6)</td>
</tr>
<tr>
<td>MMSE</td>
<td>30</td>
<td>24.7 (2.7)</td>
<td>20.4 (5.8)</td>
<td>−2.848**</td>
<td>−4.0 (4.9)</td>
</tr>
<tr>
<td>Word List Learning</td>
<td>30</td>
<td>14.5 (4.0)</td>
<td>10.4 (5.3)</td>
<td>−2.861**</td>
<td>−3.8 (4.9)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word List Delayed Recall</td>
<td>10</td>
<td>3.4 (2.4)</td>
<td>2.3 (2.2)</td>
<td>−2.047*</td>
<td>−1.1 (2.1)</td>
</tr>
<tr>
<td>Word List Savings</td>
<td>100</td>
<td>54.6 (34.3)</td>
<td>43.8 (39.4)</td>
<td>−1.107</td>
<td>−9.9 (39.0)</td>
</tr>
<tr>
<td>Word List recognition (discriminability)</td>
<td>100</td>
<td>84.5 (19.0)</td>
<td>76.5 (21.0)</td>
<td>−1.270</td>
<td>−7.7 (25.9)</td>
</tr>
<tr>
<td>Figures Copy</td>
<td>11</td>
<td>10.1 (1.1)</td>
<td>9.8 (1.7)</td>
<td>−0.743</td>
<td>−0.3 (1.6)</td>
</tr>
<tr>
<td>Figures Delayed Recall</td>
<td>11</td>
<td>6.2 (3.2)</td>
<td>4.2 (3.7)</td>
<td>−2.306*</td>
<td>−1.8 (3.3)</td>
</tr>
<tr>
<td>Figures Savings</td>
<td>100</td>
<td>61.3 (35.1)</td>
<td>40.6 (36.0)</td>
<td>−2.396*</td>
<td>−19.2 (32.4)</td>
</tr>
</tbody>
</table>

*p ≤ 0.05; **p ≤ 0.01; SD: standard deviation; MMSE = Mini-mental State Examination.

Table 3. CERAD-NAB subtests: floor and ceiling values in %; N = 20

<table>
<thead>
<tr>
<th></th>
<th>FLOOR BASELINE</th>
<th>FLOOR FOLLOW-UP</th>
<th>CEILING BASELINE</th>
<th>CEILING FOLLOW-UP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category Fluency</td>
<td>0</td>
<td>5</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>Modified Boston Naming Test</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Mini-mental State Examination</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Word List Learning Total</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Word List Delayed Recall</td>
<td>10</td>
<td>30</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Word List Savings</td>
<td>10</td>
<td>30</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>Word List recognition (discriminability)</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>Figures Copy</td>
<td>0</td>
<td>0</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>Figures Delayed Recall</td>
<td>10</td>
<td>25</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Figures Savings</td>
<td>10</td>
<td>25</td>
<td>15</td>
<td>10</td>
</tr>
</tbody>
</table>

n.d. = not defined.

Discussion

To date there are no FDA-approved drugs for the treatment of bvFTD – neither for symptom relief nor for modification of the course of the disease. A couple of treatment options are supported by the results of a small number of clinical trials that
A general problem of clinical trials in FTLD is the small patient samples. Furthermore, the available data are clearly limited by the different clinical criteria for subject enrolment and by the use of heterogeneous outcome measures that are in part insensitive to change. For this reason, it is crucial to identify suitable longitudinal outcome measures and develop assessment instruments that are best to detect disease progression (Knopman, 2007).

In our study, patients with mild bvFTD performed significantly worse at baseline on tests of language and memory than healthy controls (Figure 1). There were no differences between patients and controls in a test of visuoconstructional abilities. These findings were expected and are consistent with the results of other groups (Rogers et al., 2006; Wicklund et al., 2007). In the majority of cases cognitive deterioration accompanies behavioral disturbances at mild stages of bvFTD.

Over the course of about one year a significant decline of cognitive function was observed in patients with mild bvFTD, as is most clearly demonstrated by the CERAD-NAB global z-score. The follow-up exam revealed that within approximately one year the patients significantly worsened on the MMSE, on tests of naming (modified Boston Naming Test) and on tests of verbal and non-verbal memory (Word List Learning Total, Word List Delayed Recall, Figures Delayed Recall, and Figures Savings). This result corresponds to findings of pathology and neuroimaging studies which showed that with the progression of FTD the pathological changes surpass the borders of the frontal lobe and involve the medial/posterior temporal cortex, a region that is important for verbal and nonverbal memory function and object recognition (Whitwell et al., 2004; Diehl et al., 2006).

In contrast, performance in Category Fluency and Figures Copy remained almost stable between baseline and follow-up. The below average result in Category Fluency at baseline did not deteriorate significantly. A possible explanation for this might be that the patients’ semantic word fluency performance was already low at baseline (number of animals given: $M = 9.2; \text{SD} = 4.2; \text{range: 2–16}$) so that it could not deteriorate much further or was even slightly improved in some cases at follow-up ($M = 8.1; \text{SD} = 5.4; \text{range: 0–21}$). Patients’ performance in Figures Copy remained intact throughout the study. This finding is in keeping with the results of other studies which show that the visuospatial function in bvFTD is unimpaired until the later stages (Hodges and Patterson, 2007) consistent with the finding that – except in patients

---

**Figure 1.** Mean z-scores (circles) with standard deviations (vertical lines) of the CERAD global score and ten single test variables at baseline and follow-up.

- **CF** = Category Fluency; **BNT** = modified Boston Naming Test; **MMSE** = Mini-mental-State Examination; **WL** = Word List Learning; **FIG** = Figures

The gray horizontal stripe around $z = 0$ marks the conventional normative range of each test variable ($-1 \leq z \leq +1$).
with progranuline gene mutations (Rohrer et al., 2008) – the posterior parietal cortex is scarcely affected from neurodegeneration until advanced disease stages (Ferrr, 1999).

We did not detect definite floor effects in any of the subtests. However, there were ceiling effects: Figures Copy, Word List Recognition and modified Boston Naming test were too easy, resulting in many patients scoring at the highest level even at follow-up.

An interesting result of the study was the heterogeneity of changes not only in cognition but also in behavior as measured with the FBI, reflected by large standard deviations in many variables. This might in part be due to the small sample size, but it also points to the fact that bvFTD is very heterogeneous regarding symptom profile and disease course – a major challenge for the design of clinical trials.

Wicklund et al. (2007) recently examined the rates of cognitive decline in patients with mild to moderate bvFTD by also using subtests of the CERAD-NAB. Between 20 and 28 patients had follow-up examinations in the distinct neuropsychological subtests. In this study a best linear unbiased predictor (BLUP) analysis was used, which represents an ARC for each subject using all available longitudinal data from that subject. The ARC in the memory tests is lower (Word List Learning Total: −0.7; Word List Delayed Recall: −0.5) compared to our study, whereas the ARC of animal fluency was slightly higher (−1.7). These differences can most probably be explained by significantly different methodologies (i.e. patient inclusion, statistics). The ARC of the MMSE in the Wicklund study (−4.5) is almost identical with our finding of −4.0, and it is higher than 3.3, which was the ARC of patients with AD found in a large meta-analysis of approximately 3500 patients with AD (Han et al., 2008). More marked ARCs of the MMSE have been found in some other studies: Rascovsky et al. (2005) described an ARC of the MMSE of −6.7 points; Chow et al. (2006) found a change of −4.7 points. However, the latter studies also include language variants of FTLD, which have been demonstrated to decline faster in the MMSE – which is heavily dependent on language skills – than the behavioral variant (Osher et al., 2008).

A unique multi-center study by Knopman and colleagues (2008) prospectively followed 78 patients with frontotemporal lobar degeneration (including 36 patients with bvFTD) over one year using an extensive battery of cognitive and behavioral measures. The results were somewhat different from our study: The ARC on the MMSE of the bvFTD patients was lower (M = −2.45; SD = 4.32) than in the present investigation. In contrast, there was a much higher ARC in category fluency (M = −6.40; SD = 5.95), reflecting a fast/marked decline. The authors highlighted that word fluency tasks (letter fluency and category fluency) showed the largest ratios of mean change to standard deviation, “suggesting that, among individual tests, measures of fluency might be the most efficient for detecting change in a clinical trial” (Knopman et al., 2008, p. 2966). In our study, however, the Category Fluency task was the CERAD subtest that showed the smallest annual rate of change within one year (ARC mean = −1.0, SD = 3.7). While such inconsistent findings point to the necessity for more systematic research into neuropsychological change patterns in patients with bvFTD and other frontotemporal lobar degenerations, the results of the study by Knopman et al. are hardly comparable with our findings. Their multi-center study included patients with bvFTD who were less severely ill at baseline (as is evident from higher MMSE and better results on tests of memory and language at baseline). Furthermore, the variation in patients’ performance in the multi-center study was less homogeneous at baseline than in our study as is shown by higher standard deviations in the cognitive tests of the multi-center study.

A limitation of the present study is that no pathological verification of clinical diagnosis is available. It is well known that bvFTD is pathologically heterogeneous (Hodges et al., 2004). Differences in neuropsychological performance between tau-positive and tau-negative patients (Grossman et al., 2007) as well as different longitudinal patterns (Grossman et al., 2008) have been described. The follow-up results in groups with different histopathologies would be of great interest. However, given that we do not have reliable in vivo biomarkers for diagnosis of pathologic subtypes, inclusion into clinical trials will follow clinical diagnosis rather than pathological subtypes.

Another limitation of our study is the relatively small number of patients. However, it was a particular aim of the study to include a patient sample that was as homogeneous as possible regarding stage of the disease and disease entity. Therefore, we excluded patients with ALS from this study. Even if there are common overlaps between ALS and bvFTD, ALS is to be considered as a distinct entity. Investigating a heterogeneous group of patients can lead to spurious findings that cannot be generalized. Furthermore, due to the low prevalence of the disease it is unlikely that many more than 20 patients with bvFTD can be enrolled in a single-center study. The way forward might be to expand and extend international collaboration.
between international centers with expertise in this field.

For the z-transformations of the raw scores of the CERAD-NAB we used the CERAD-NAB norms provided by the Basel Memory Clinic, which include data from N = 1100 healthy older subjects. To facilitate comparisons across published studies we could instead have compared the CERAD-NAB results of the patients with the results of cognitively healthy control persons tested in our clinic at baseline and follow-up.

Taken together, using the CERAD-NAB we were able to detect a decline of cognitive functioning in a sample of patients with mild bvFTD. In most of the subtests a significant change within one year could be detected. However, Figures Copy, Word List Recognition and modified Boston Naming Test showed considerable ceiling effects and therefore seem not suitable for detecting disease progression at the early stage of bvFTD. The performance in Category Fluency and Figures Copy remained stable within one year. Therefore we conclude that a cognitive battery designed for clinical drug trials should include the MMSE, Word List Learning and Word List Delayed Recall. While the MMSE, which focuses on language and memory and fails to assess frontal functions, might not be a suitable instrument to detect early changes in bvFTD, it has proven to be sensitive to cognitive deterioration. For a patient population with mild bvFTD, tests of object naming and visuconstruction have to be more difficult than the tests used in our study. Furthermore, a standardized test battery for patients with frontotemporal lobar degeneration must include a sufficient number of tests of executive function. The recent study by Knopman and colleagues suggests an executive/working memory composite consisting of a simplified Trailmaking ratio, Backward Digit Span, Digit Symbol, Stroop Test, Number Cancellation, and a specially constructed sum-of-errors measure (Knopman et al., 2008).

**Conflict of interest**

None.

**Description of authors’ roles**

J. Diehl-Schmid designed the study and wrote the paper; S Bornschein did the neuropsychological assessments; C Pohl collected the data; H. Foerstl assisted in writing the paper; A Kurz supervised the data collection and assisted in writing the paper; and T Jahn was responsible for the statistical design and performed the statistical analyses.

**References**


