Supplementary Online Content


eMethods. Assessments and Analysis

eFigure. Cluster Solutions

eTable. Patient Characteristics

This supplementary material has been provided by the authors to give readers additional information about their work.
eMethods. Assessments and Analysis

Neuropsychological assessment

All patients completed a battery of neuropsychological tests designed to assess major domains of cognition.

Memory: Verbal episodic memory was evaluated with the California Verbal Learning Test–Short Form (CVLT), which includes a list of 9-item words, presented over 4 learning trials. Performance on verbal learning was scored by combining results from the four trials, for maximum of 36 points. Immediate (short delay) and delayed CVLT were assessed by free recall of the list at 30-second and 10-minute intervals respectively. Visual memory (Benson delayed recall) was assessed by asking the patients to draw the Benson figure from memory after a 10-minute delay, and scored on a 17 point scale. Visual recognition was tested by asking the patient to identify the correct figure from four options after a 10-minute delay.

Language: Confrontation naming was assessed with a 15-item short form of the Boston Naming Test. Repetition was assessed by having patients repeat 5 phonemically complex sentences. Verbal agility was evaluated by having patients rapidly articulate a multi-syllabic word and was measured as the number of repetitions completed correctly within 5 seconds. Surface dyslexia was tested by having patients read 6 irregular words and measured as the number correct out of 6. Syntax comprehension was measured using a subset of 5 items from the Boston Diagnostic Aphasia Evaluation for which the examiner read a sentence aloud, and the patient had to select from among 4 options the picture that best matched the sentence. Lexical fluency was assessed using ‘D words’ and category fluency (semantic fluency) with ‘animals’.

Visuospatial: Patients were asked to copy a complex figure (Benson figure) as the object of visual construction and the accuracy was scored on a 17 point scale. The Number Location subtest of the Visual Object Space Perception (VOSP) test required the patient to precisely locate a stimulus on a two-dimensional plane, requiring dorsal-stream (“where”) visual processing and scored out of 10. The face recognition subtest of the Comprehensive Affect Testing System (CATS) is a ventral-stream task involving 12 trials where the patient determined whether two faces are the same or different.

Emotion naming: The affect matching subtest of the CATS contained 16 trials where the patient was shown a photo of an emotional face and required to select the correct label from a list (i.e. ‘happy’, ‘sad’, ‘angry’, ‘frightened’, ‘surprised’, ‘disgusted’ or ‘neutral’).

Executive: Simple auditory attention was assessed by digit span forward and working memory by digit span backward. The modified trail-making test required the patient to serially alternate between numbers and days of the week, evaluating set-shifting and sequencing. To adjust for the fact that some patients do not complete the task within the required time window of 2-minutes, the dependent measure was calculated as the number of correct connections made per minute. The number of errors made during trial making test measured the error insensitivity. Non-verbal generation ability was measured using the Filled Dots condition from the design fluency subscale of the Delis-Kaplan Executive Function Scale (DKEFS) and was scored as the number of correct designs generated within a minute. The Stroop tests required the patient to rapidly state the color of the ink in which the blocks of letters are printed, and scored as the number of correct responses within a minute. For the Stroop color naming condition the text and the ink color are matched, thus it is a simple measure of information processing speed. On the Stroop interference trial, the text and the ink color are different, thus it is used as a measure of cognitive control. Calculation ability was tested by five paper-and-pencil arithmetic problems of varying degree of difficulty.

Socioemotional assessment:

We used both face-to-face tests and informant questionnaires to assess different aspects of socioemotional function of bvFTD patients. The face-to-face tests included all three subtests of The Awareness of Social Inference Test (TASIT) and the UCSF Cognitive Theory of Mind (cTOM) test. Emotion naming was evaluated using the TASIT-EET (TASIT-Emotion Evaluation Test), which consists of 14 items. The patients were asked to identify the emotion, depicted in brief (~20s) video vignettes in which professional actors show one of the 7 basic emotions (multiple choice options: happy, surprised, sad, anxious, angry, disgusted, and neutral). To measure sensitivity to the paralinguistic elements of sarcasm, we used the TASIT-SIM (TASIT-Social Inference–Minimal), which consists of short video vignettes in which actors use voice, facial expression, and body posture/gestures to express either
sincere or sarcastic sentiments with semantically neutral scripts. We used five video vignettes (Simple Sarcasm (SSR) items), each with four questions to characterize the patient’s comprehension of the exchange. To measure complex social cognition, we used the TASIT SIE (Social Inference–Enriched) which consists of 16 video vignettes and assesses the ability to comprehend complex social exchanges in order to detect lies versus sarcasm. The voice, facial expression, and prosody of the actors indicated the intended meaning of the exchange, together with other contextual clues that revealed the speakers’ intentions. To evaluate patients’ capacity for non-emotional perspective-taking/theory of mind, we used the UCSF cognitive TOM test, which consists of eight video vignettes of actors with limited perspective on the other’s actions, followed by questions arranged in ascending order of executive demand (i.e., control, first-order ToM, and second-order ToM questions). Informant questionnaires included the Interpersonal Reactivity Index (IRI) and Interpersonal Adjective Scale (IAS). Empathic Concern (EC) and Perspective Taking (PT) subscales of the IRI evaluated the patients’ emotional and cognitive empathy, respectively. The 7-item EC scale measures the capacity to attend to and appropriately respond to another’s needs and emotional expressions. The 7-item PT subscale measures the patients’ tendency to spontaneously employ empathic perspective taking in their typical social interactions, using cognitive resources to put themselves in others’ shoes. The IAS is a well-validated questionnaire based on the circumplex model of personality, which aims to measure individual differences in interpersonal traits along the orthogonal axes of power and affiliation. We used the interpersonal warmth (LM) and interpersonal assertiveness (PA) subscales of the IAS. Socioemotional function was considered impaired at a Z-score<1.35 based on published normative samples of age-matched healthy controls.

Neuropathology

Twenty four patients were subjected to an extensive dementia-oriented postmortem assessment at UCSF, following a standard protocol described previously. The brains were procured within 18 hours postmortem. The majority of the brains was cut into 8-10-mm thick coronal slabs and alternately fixed, in 10% neutral buffered formalin for 72 hours, or rapidly frozen; the remaining were hemi-sectioned sagittally with 1 hemisphere fixed in 10% neutral-buffered formalin and the other rapidly frozen. Tissue blocks covering dementia-related regions of interest were dissected from the fixed slabs, and basic and immunohistochemical stains were applied following standard diagnostic procedures developed for patients with dementia. All immunohistochemical runs included positive control sections to exclude technical factors as a cause of absent immunoreactivity.

Voxel Based Morphometry (VBM) analysis

Gray matter volumes were assessed in a subset of patients for whom a unified structural Magnetic Resonance Image (MRI) acquisition protocol had been performed (1.5, 3, or 4 Tesla). Analysis was performed using the VBM8 toolbox of statistical parametric mapping version 8 (SPM8) (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). All images were obtained within 12 months of the MEG evaluation. Atrophy patterns were determined by comparison to images from 44 age-matched control participants who were evaluated at the UCSF Memory and Aging Center. The structural images were corrected for bias field, segmented into grey matter, white matter, and CSF, and initially normalized to Montreal Neurological Institute (MNI) space using the Unified Segmentation procedure implemented in the SPM8 running on Matlab (MathWorks, Natick, MA). More anatomically precise inter-subject registration was then performed with the Diffeomorphic Anatomical Registration through Exponentiated Lie algebra (DARTEL) toolbox by warping each patient’s image to a template created from elderly normal control patients. The modulated images were smoothed with an 8-mm full-width-at-half-maximum isotropic Gaussian kernel for group analyses; differences between patient and control groups were analyzed with t-tests. Age, gender and scan-strength were included into the model as covariates, and total intracranial volume was included as a nuisance covariate. All group analyses were corrected for multiple comparisons with a family wise error of P<0.05.

Cluster analysis and principal component analysis

We estimated the mean voxel volume of specific mutually exclusive ROIs defining the salience network (SN) and semantic appraisal network (SAN), bilaterally, as identified by previous functional imaging studies. Left and right temporal poles, left and right gurus recti, left and right amygdala, and left and right subcallosal areas comprised...
the bilateral SAN. Left and right anterior insula, left and right anterior cingulate, left and right basal forebrain, left and right frontal operculum, left and right posterior orbital gyri, and left and right and amygdala comprised the bilateral SN. We used the evalclusters function of the MatLab statistical toolbox to identify distinct patient subgroups based on the distinct volumetric patterns of the 18 ROIs. Specifically, we used the kmeans clustering algorithm, which partitions data into mutually exclusive clusters, and returns the index of the cluster to which it has assigned each observation. In our analysis we specified the clustering evaluation criterion as ‘gap’, and used the distance metric of squared Euclidean distance. Comparison of clustering solutions for three, four and five clusters clearly demonstrated that four clusters give the optimum separation without overlap (eFIGURE 1). We used the same 18 ROIs in a principal component analysis (PCA), using the pca function of MatLab statistical toolbox. Next we plotted each patient’s mean SN and SAN volumes onto the first two components. The mean SN and SAN volumes were calculated as the average of the ROIs listed under each network as defined above.

In our initial analysis we started with a list of 28 ROIs, each one uniquely contributing to 4 potentially bvFTD-affected networks, including SN,23,25 SAN,25,26 default-mode (DMN),27,28 and frontoparietal (FPN) networks.29,30 With this relatively extensive list of ROIs from four networks, the PCA/cluster analysis showed that DMN was non-contributory to the 1st and 2nd dimensions of PCA, demonstrating identical results from the cluster analysis either with or without DMN included into the model. Inclusion of FPN ROIs reduced the sensitivity of the cluster analysis, limiting the model’s ability to separate patients into distinct clusters. The model using SN and SAN alone performed optimally in producing distinct clusters of patients. Based on these initial data reduction steps we retained only the SN and SAN networks in the cluster/PCA analysis.

We ran the clustering algorithm both with and without including the scanner strength as an additional variable into the model. For 97% of the cohort the cluster assignment was identical in both models demonstrating that scanner strength had minimal influence on dimensions generated by volumetric assessment. This finding is consistent with previous reports of automated volumetric analysis of patients with neurodegenerative diseases demonstrating only minimal effect of scanner strength when scans are preprocessed using SPM-type voxel-wise methods.31 Based on these results we present the data from the model without including the scanner-strength.

Polytomous logistic regression

A polytomous logistic regression model was used to assess the stability of the clusters derived from cluster analysis.32 We used the MatLab statistical toolbox function mnrval in polytomous regression analysis, which models a multinomial regression for nominal outcomes and estimate the category probabilities. The output of this model predicts the probability that an observation belongs to each category of a categorical response variable. For each patient, the model used the volume estimates of the 18 ROIs, and the cluster assignment, and predicted the probability of falling into each of the four clusters.

Regression diagnostics to identify outliers in the rate of disease progression

We used regression diagnostics of leverage and Cook’s distance measures to determine the outliers in the distribution of rate of disease progression among the subset of bvFTD patients who had more than one evaluation. We applied these methods to the overall sample of bvFTD patients who have more than one evaluation (n=59), and to each subgroup of bvFTD patients (SN-FT, n=16 ; SN-F, n=14; SAN, n=7; subcortical, n=22). In both approaches one patient who belonged to the subcortical group was identified as an outlier. The Cooks distance values ranged from 0.00002 to 0.2031, with a mean of 0.0159 ± (0.03). The leverage statistic values ranged from 0.01 to 0.279, with a mean of 0.029 ± (0.03). The patient who was identified as an outlier reported 0.095 and 0.203 as Cook’s distance value and the leverage statistic respectively.
eReferences


7. Spreen O, Benton AL. Neuropsychology Center Comprehensive Examination for Aphasia. Victoria, BC: Neuropsychology Laboratory, University of Victoria; 1977.


Silhouette plots demonstrating the cluster indices output from kmeans. From left to right the subplots show cluster solutions derived for 3 clusters, 4 clusters and 5 clusters. Each cluster is represented by a silhouette, which is based on the comparison of its tightness and separation. The silhouette value displays a measure of how close each point in one cluster to points in the neighboring clusters. The silhouette value ranges from 0 to 1, where 0 indicates points that are not distinctly in one cluster or another and 1 indicates points that are clearly distant from neighboring clusters. Negative values indicate points that are overlapped between clusters and potentially assigned with a larger error. The average silhouette values are a quantitative way to compare between solutions where higher numbers indicate better separation.
**eTable. Patient Characteristics**

<table>
<thead>
<tr>
<th></th>
<th>bvFTD patients (n=90)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at disease onset – yr</strong></td>
<td>55.1 ± 9.7</td>
</tr>
<tr>
<td><strong>Age at first evaluation – yr</strong></td>
<td>61.1 ± 8.2</td>
</tr>
<tr>
<td><strong>Male sex – no. (%)</strong></td>
<td>54 (60.0)</td>
</tr>
<tr>
<td><strong>White – no. (%)</strong></td>
<td>82 (91.1)</td>
</tr>
<tr>
<td><strong>Right handedness – no. (%)</strong></td>
<td>82 (91.1)</td>
</tr>
<tr>
<td><strong>Education – yr</strong></td>
<td>16.9 ± 9.6</td>
</tr>
<tr>
<td><strong>MMSE at initial evaluation</strong></td>
<td>23.0 ± 6.3</td>
</tr>
<tr>
<td><strong>CDR at initial evaluation</strong></td>
<td>1.3 ± 0.7</td>
</tr>
<tr>
<td><strong>CDRSOB at initial evaluation</strong></td>
<td>6.9 ± 3.3</td>
</tr>
<tr>
<td><strong>FTLD-CDRSOB at initial evaluation</strong></td>
<td>9.2 ± 3.7</td>
</tr>
<tr>
<td><strong>Mutation carriers – no. (%)</strong></td>
<td>24 (28.6)</td>
</tr>
<tr>
<td>C9orf72 – no. (%)</td>
<td>14 (58.3)</td>
</tr>
<tr>
<td>GRN – no. (%)</td>
<td>6 (25.0)</td>
</tr>
<tr>
<td>MAPT – no. (%)</td>
<td>4 (16.7)</td>
</tr>
<tr>
<td>GRN T allele frequency – no. (%)</td>
<td>41 (33.1)$^c$</td>
</tr>
<tr>
<td>Tau H1 allele frequency – no. (%)</td>
<td>101 (82.8)</td>
</tr>
</tbody>
</table>

Abbreviations: bvFTD = behavioral variant frontotemporal dementia; CDR = Clinical Dementia Rating; CDRSOB = CDR Sum of Boxes; C9ORF72 = chromosome 9 open reading frame 72 hexanucleotide expansions; FTLD-CDRSOB = Frontotemporal Lobar Degeneration - modified CDRSOB; GRN = progranulin; MAPT = microtubule-associated protein tau; MMSE = Mini Mental State. Age, MMSE, CDR, CDRSOB, and FTLD-CDRSOB, at initial evaluation represent patient characteristics measured within one year of first presentation. Values for age, education, MMSE, CDR, CDRSOB, and CDR-FTLD, are mean ± standard deviations. Race or ethnic group was self-reported. Scores on the MMSE range from 0 to 30, with higher scores denoting better cognitive function. Scores on the CDR, CDRSOB, and FTLD-CDRSOB range from 0 to 3, 0 to 18 and 0 to 24, respectively, with higher scores denoting more disability.

$a$ n=89 bvFTD patients had FTLD-CDRSOB recorded at their first evaluation.

$b$ percentages are calculated for the patients who underwent genetic testing (n=84).

$c$ P=.017, compared to allele frequency of normal population using Chi-square test.