Single-subject classification of presymptomatic frontotemporal dementia mutation carriers using multimodal MRI

Rogier A. Feis\textsuperscript{a,b,*}, Mark J.R.J. Bouts\textsuperscript{a,b,c}, Jessica L. Panman\textsuperscript{a,d}, Lize C. Jiskoot\textsuperscript{a,d}, Elise G.P. Dopper\textsuperscript{a,d,e}, Tijn M. Schouten\textsuperscript{a,b,c}, Frank de Vos\textsuperscript{a,b,c}, Jeroen van der Grond\textsuperscript{a}, John C. van Swieten\textsuperscript{d,i}, Serge A.R.B. Rombouts\textsuperscript{b,h}

\textsuperscript{a} Department of Radiology, Leiden University Medical Centre, Leiden, the Netherlands
\textsuperscript{b} Leiden Institute for Brain and Cognition, Leiden University, Leiden, the Netherlands
\textsuperscript{c} Institute of Psychology, Leiden University, Leiden, the Netherlands
\textsuperscript{d} Department of Neurology, Erasmus Medical Centre, Rotterdam, the Netherlands
\textsuperscript{e} Alzheimer Centre & Department of Neurology, Neuroscience Campus Amsterdam, VU University Medical Centre, Amsterdam, the Netherlands
\textsuperscript{f} Institute of Psychology, Leiden University, Leiden, the Netherlands
\textsuperscript{h} Department of Clinical Genetics, Neuroscience Campus Amsterdam, VU University Medical Centre, Amsterdam, the Netherlands

\textbf{A B S T R A C T}

\textbf{Background:} Classification models based on magnetic resonance imaging (MRI) may aid early diagnosis of frontotemporal dementia (FTD) but have only been applied in established FTD cases. Detection of FTD patients in earlier disease stages, such as presymptomatic mutation carriers, may further advance early diagnosis and treatment. In this study, we aim to distinguish presymptomatic FTD mutation carriers from controls on an individual level using multimodal MRI-based classification.

\textbf{Methods:} Anatomical MRI, diffusion tensor imaging (DTI) and resting-state functional MRI data were collected in 55 presymptomatic FTD mutation carriers (8 microtubule-associated protein Tau, 35 progranulin, and 12 chromosome 9 open reading frame 72) and 48 familial controls. We calculated grey and white matter density features from anatomical MRI scans, diffusivity features from DTI, and functional connectivity features from resting-state functional MRI. These features were applied in a recently introduced multimodal behavioural variant FTD (bvFTD) classification model, and were subsequently used to train and test unimodal and multimodal carrier-control models.

\textbf{Classification performance was quantified using area under the receiver operator characteristic curves (AUC).}

\textbf{Results:} The bvFTD model was not able to separate presymptomatic carriers from controls beyond chance level (AUC = 0.570, p = 0.11). In contrast, one unimodal and several multimodal carrier-control models performed significantly better than chance level. The unimodal model included the radial diffusivity feature and had an AUC of 0.646 (p = 0.021). The best multimodal model combined radial diffusivity and white matter density features (AUC = 0.680, p = 0.005).

\textbf{Conclusions:} FTD mutation carriers can be separated from controls with a modest AUC even before symptom-onset, using a newly created carrier-control classification model, while this was not possible using a recent bvFTD classification model. A multimodal MRI-based classification score may therefore be a useful biomarker to aid earlier FTD diagnosis. The exclusive selection of white matter features in the best performing model suggests that the earliest FTD-related pathological processes occur in white matter.

\textbf{Abbreviations:} 3DT1w, 3-dimensional T1-weighted; AUC, Area under the receiver operating characteristics curve; AxD, Axial diffusivity; C9orf72, Chromosome 9 open reading frame 72; DTI, Diffusion tensor imaging; DWI, Diffusion-weighted imaging; FA, Fractional anisotropy; FCor, Full correlations; (bv)FTD, (behavioural variant) Frontotemporal dementia; GM, Grey matter; GMD, Grey matter density; GRN, Progranulin; ICA, Independent component analysis; MAPT, Microtubule-associated protein Tau; MD, Mean diffusivity; MMSE, Mini-mental state examination; (rs-f)MRI, (resting-state functional) Magnetic resonance imaging; Poor, Sparse L1-regularised partial correlations; RD, Radial diffusivity; ROC, Receiver operating characteristics; TBSS, Tract-based spatial statistics; WM, White matter; WMD, White matter density

\* Corresponding author at: Department of Radiology, Leiden University Medical Centre, Leiden, the Netherlands.

\textbf{E-mail addresses:} r.a.feis@lumc.nl (R.A. Feis), m.j.r.j.bouts@fsw.leidenuniv.nl (M.J.R.J. Bouts), j.panman@erasmusmc.nl (J.L. Panman), l.c.jiskoot@erasmusmc.nl (L.C. Jiskoot), e.dopper@erasmusmc.nl (E.G.P. Dopper), t.m.schouten@fsw.leidenuniv.nl (T.M. Schouten), f.de.vos@fsw.leidenuniv.nl (F. de Vos), j.van_der_grond@lumc.nl (J. van der Grond), j.c.vanswieten@erasmusmc.nl (J.C. van Swieten), s.a.r.rombouts@lumc.nl (S.A.R.B. Rombouts).

https://doi.org/10.1016/j.nicl.2018.07.014
Received 4 April 2018; Received in revised form 29 June 2018; Accepted 15 July 2018
Available online 17 July 2018
2213-1582/ © 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).
1. Introduction

Frontotemporal lobar degeneration is a common cause of early-onset dementia with a similar prevalence to Alzheimer's disease in the presenile population (Ratnavalli et al., 2002; Harvey et al., 2003; Rabinovici and Miller, 2010; Seeaara et al., 2011; Rascovsky et al., 2011). Although there are clinical disease criteria for the different clinical variants of frontotemporal dementia (FTD) (Rascovsky et al., 2011; Gorno-Tempini et al., 2011), diagnosis is often complicated and delayed by clinical heterogeneity. This hinders clinicians in providing accurate prognosis, effective disease management and developing new treatments (Mohs et al., 2001; Mendez et al., 2007; Mendez, 2009; Pressman and Miller, 2014).

Multimodal magnetic resonance imaging (MRI) has been suggested as a promising biomarker to improve on diagnostic standards in FTD. In FTD patients, MRI revealed specific patterns of neurodegeneration, involving grey matter (GM) and white matter (WM) atrophy (Whitwell et al., 2011) as a promising biomarker to improve on diagnostic standards in FTD. In order to distinguish presymptomatic FTD mutation carriers from controls, we applied two models. First, we applied a recent behavioural variant FTD (bvFTD)-control classification model (Bouts et al., 2018) to our MRI data to investigate whether the model separates presymptomatic mutation carriers from controls. We shall refer to this model as the “bvFTD model”. In a second analysis, we trained a new classification model on the presymptomatic mutation carriers and controls’ data, which we evaluated using cross-validation. We shall refer to this model as the “carrier-control model”. MRI pre-processing, feature selection and classification were performed identically to previous work (Bouts et al., 2018).

2. Participants

This retrospective study partially included previously published (Dopper et al., 2014; Papsma et al., 2017; Jiskoot et al., 2016) and newly acquired data from the Erasmus Medical Centre and Leiden University Medical Centre. Participants and clinical investigators were blinded to the participants’ DNA status. The study was conducted in accordance with regional regulations and the Declaration of Helsinki. The Erasmus Medical Centre and Leiden University Medical Centre local medical ethics committees approved the study, and every participant provided written informed consent.

For the current study, we included 55 presymptomatic FTD mutation carriers (8 MAPT, 35 GRN, 12 C9orf72) and 48 healthy familial controls (6 MAPT family, 31 GRN family and 11 C9orf72 family) between May 2010 and March 2016. These subjects were recruited from a cohort of healthy first-degree relatives of FTD patients with either a MAPT, GRN or C9orf72 mutation (FTD-Risk Cohort; FTD-RisC) and visited the Erasmus Medical Centre for a one-day assessment in order to ascertain asymptomatic status, collect clinical data, and determine DNA status as described before (Dopper et al., 2014; Papsma et al., 2017; Jiskoot et al., 2016). Participants were considered asymptomatic in the absence of (1) behavioural, cognitive, or neuropsychiatric change reported by the participant or knowledgeable informant, (2) cognitive disorders on neuropsychiatric tests, (3) motor neuron disease signs on neurologic examination, and (4) other FTD (Rascovsky et al., 2011; Gorno-Tempini et al., 2011) or amyotrophic lateral sclerosis (Ludolph et al., 2015) criteria. Healthy controls were assumed to have equal FTD risk as the general population. For a more detailed description of the recruitment protocol, see earlier work (Dopper et al., 2014; Papsma et al., 2017; Jiskoot et al., 2016). Inclusion criteria for the current study were: age between 40 and 70 years, and availability of a T1-weighted 3-dimensional MRI (3DTWI) scan, a diffusion-weighted imaging (DWI) dataset, and a resting-state fMRI T2*-weighted (rs-fMRI) scan. Exclusion criteria were: current or past neurologic or psychiatric disorders, history of drug abuse, large image artefacts, and gross brain pathology other than atrophy.

For details on the sample on which the bvFTD model was trained, please refer to Bouts et al. (2018) (Bouts et al., 2018). In short, 23 bvFTD patients and 35 controls between 40 and 80 years old were included to undergo a clinical assessment and MRI between November 2009 and November 2012. The MRI acquisition protocol was similar to the protocol applied in the current sample of carriers and controls. Image processing steps were identical to processing steps in the current sample.
2.3. MRI data acquisition

All subjects were scanned at the Leiden University Medical Centre using a 3 T MRI scanner (Achieva, Philips Medical Systems, Best, The Netherlands) with an 8-channel SENSE head coil. The imaging protocol included a whole-brain near-isotropic 3DT1w sequence for cortical and subcortical tissue-type segmentation, a DWI sequence for assessments of white matter integrity, and a rs-fMRI for the calculation of functional connectivity measures. Participants were instructed to lie still with their eyes closed and not to fall asleep during rs-fMRI. For scan parameters, see Table 1.

2.4. Image pre-processing

For 3DT1w images, the following pre-processing steps were performed: bias field correction (N4ITK (Tustison et al., 2010)), brain extraction (FSL BET (Smith, 2002)), non-linear registration to the MNI152 2 × 2 × 2 mm T1 template (FNIRT (Anderson et al., 2007)), tissue-type segmentation (SPM12 (Friston et al., 2007)) and segmentation of deep grey matter structures, including the bilateral thalamus, caudate nucleus, putamen, globus pallidum, nucleus accumbens, amygdala and hippocampus (FIRST (Patenneau et al., 2011)).

Pre-processing for DTI datasets included correction of motion and eddy-current induced distortion (eddy correct (Leemans and Jones, 2009)), calculation of voxel-wise measures of fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AxD, largest eigenvalue), and radial diffusivity (RD, average of the two remaining eigenvalues, DTIFIT (Smith et al., 2004)). A global mean FA image was created by non-linearly registering FA maps to the FMRIB58_FA template, and tract-based spatial statistics (FSL TBSS (Smith et al., 2006)) was used to extract FA, MD, AxD and RD values using the standard FSL TBSS skeleton. The skeleton was thresholded at 0.2 to ensure skeleton extracted values originate from WM.

For rs-fMRI data, pre-processing included motion correction (Jenkinson et al., 2002), brain extraction, spatial smoothing using a Gaussian kernel with a full width at half maximum of 3 mm, grand mean intensity normalisation, motion artefact removal, and high-pass temporal filtering (cut-off frequency = 0.01 Hz). Motion artefacts were removed using a single-session independent component analysis (ICA) to decompose the rs-fMRI data into distinct statistically independent components. Subsequently, motion-related components were automatically identified and removed using the ICA-based automatic removal of motion artefacts (ICA-AROMA, version 0.3-beta) procedure (Pruijm et al., 2015). Registration to standard space was performed in two steps. First, a temporal mean image calculated from the 4D rs-fMRI volume was registered to the 3DT1w image using Boundary-Based Registration (Greve and Fischl, 2009). Next, resulting registration parameters were concatenated to the 3DT1w-to-MNI152 template registration parameters to obtain the final registration parameters.

All registration and segmentation steps were critically reviewed and errors were corrected accordingly.

Table 1
MRI sequence parameter settings.

|                | Slices | TR (ms) | TE (ms) | Flip angle (°) | Matrix (mm) | Voxel size (mm) | Duration (min) |
|----------------|--------|---------|---------|----------------|-------------|----------------|---------------|
| 3DT1w          | 140    | 9.8     | 4.6     | 8              | 256 × 256   | 0.88 × 0.88 × 1.20 | 4.57           |
| DWI            | 70     | 8250    | 80      | 90             | 128 × 128   | 2.00 × 2.00 × 2.00 | 8.48           |
| rs-fMRI        | 58     | 2200    | 30      | 80             | 80 × 80     | 2.75 × 2.75 × 2.99* | 7.28           |

Scan protocol of whole-brain near-isotropic 3DT1-weighted (3DT1w), diffusion-weighted imaging (DWI), and resting-state functional MRI T2*+weighted MRI (rs-fMRI) on a 3T scanner at the Leiden University Medical Centre. Abbreviations: TR: repetition time; TE: echo time.

2.5. Feature selection

Cortical GM density (GMD) and WM density (WMD) were calculated as a weighted average of their respective regional WM or GM probability (SPM segmentation) weighted by the probability of a voxel being part of that specific tract or region. The latter probabilities were derived from the 48 Harvard-Oxford probabilistic anatomical brain atlas cortical regions (split into left and right) and from the Johns-Hopkins University white-matter tractography atlas for 20 WM tract regions. Voxels with region probability values < 25% were excluded. This provided a measure of brain atrophy of a specific GM region or WM tract. For deep GM regions, GMD values were calculated as the regions' volume (FIRST segmentations) divided by total intracranial volume. This resulted in a feature vector of 110 average GMD values (48 left cortical, 48 right cortical and 14 deep GM regions) and a feature vector of 20 average WMD values per subject.

DTI-based features were calculated by projecting each subject’s FA, MD, AxD and RD values onto the TBSS group skeleton on a voxel-wise basis. Like the WMD features, the 20 WM tracts of the probabilistic JHU white-matter tractography atlas were then used to calculate a weighted mean value per tract per subject. This resulted in 4*20 feature vectors of mean FA, MD, AxD and RD values per subject.

In order to calculate the functional connectivity features, all processed rs-fMRI images were combined in a temporally concatenated independent component analysis (ICA (Beckmann and Smith, 2004)), with dimensionality fixed at 70 components and an ICA threshold of 0.99 (Smith et al., 2013). This meant that each voxel included in the ICA map was 99 times more likely to be part of that component than to be caused by Gaussian background noise. For each subject, we calculated the mean time course for each component, weighted by the ICA weight map and GM probability of that component’s region. These mean time courses were subsequently used to determine the functional connectivity of a component with the 69 other components. Functional connectivity was either expressed as full correlations (FCor) or as sparse L1-regularised partial correlations (PCor) between the components’ time courses. Partial correlations were calculated using the graphical lasso algorithm (Friedman et al., 2008). The functional connectivity measures resulted in two feature vectors of each (70 * 69)/2 = 2415 (partial) correlations per subject. Finally, we concatenated all feature vectors into one vector per subject.

2.6. bvFTD model

For our first analysis, a bvFTD patient-control classification model (Bouts et al., 2018) was applied to each subject’s extracted feature vector. We applied the best performing, multimodal model that discriminated bvFTD patients from controls, which included the features FA, GMD, and FCor (Bouts et al., 2018), as well as age and gender. Each subject’s feature vector was fed into the model, resulting in a probability score from 0 to 1, where 0 represents a control subject and 1 represents a bvFTD patient. Extrapolated to our subjects, these scores showed how alike our presymptomatic FTD mutation carriers and healthy controls are to bvFTD patients.
2.7. Carrier-control model

For the second analysis, feature vectors were used to train a logistic elastic net regression algorithm (Bouts et al., 2018; Zou and Hastie, 2005; Friedman et al., 2010; Schouten et al., 2016). The elastic net regression procedure estimates a sparse regression model that includes only a subset of the provided features by imposing a penalty for including features and for the weight of each feature. This way, elastic net provides a solution for the imbalance between the large number of features and the small number of subjects. Age and gender were included into the model without penalty to ensure that estimated feature regression coefficients were conditional on subject age and gender. Here, a probability score of 0 represented a control subject and 1 represented a presymptomatic FTD mutation carrier.

2.8. Cross-validation

Similarly to previous work (Bouts et al., 2018; Schouten et al., 2016), we trained our carrier-control model in a nested 10-fold cross-validation scheme to reduce classification bias. One part of the data (e.g. 10%) was set apart as a test set and served to test the generalised classification performance of the elastic net regression model. The remaining parts (90%) were used to train the model. However, in addition to the classification performance, we also wanted to determine the optimal penalty size without overestimating classification performance (Varma and Simon, 2006; Kriegeskorte et al., 2009). To this end, we used a second, nested 10-fold cross-validation loop on the training set over a grid of hyperparameters to determine the optimal penalty. In the nested loop, we estimated the model’s hyperparameters that corresponded with the lowest binomial deviance, a goodness-of-fit measure that evaluates the difference between the predicted and actual observations. Next, these hyperparameters and corresponding penalties were used to train a model using the training set of the outer loop. Finally, the classification performance was tested on the test set of the outer loop. This process was repeated ten times to make sure that each subject was part of the test set at least once. Since the test set of the outer loop was neither used for model training, nor for parameter optimisation, potential prediction bias was reduced as much as possible (Kriegeskorte et al., 2009). The entire classification procedure was repeated 50 times to average classification outcome variability resulting from random partitioning in training and test folds. All classification analyses and evaluations were implemented in R version 3.3.2 (R core 2010, GLMnet package (Friedman et al., 2010)).

2.9. Classification performance

For both analyses, we quantified classification performances using receiver operating characteristic (ROC) curves. ROC curves were calculated by shifting the threshold for classifying an individual as patient (bvFTD model analysis) or carrier (carrier-control model analysis) from 0 to 1, and plotting the true positive rate (sensitivity) versus the false positivity (1 − specificity) for each intermediate point. The area under this ROC curve (AUC) is a measure of classification performance insensitive to the distribution between the groups (Fawcett, 2006). Additionally, we calculated the optimal operating point on the curve to calculate the model’s sensitivity, specificity and classification accuracy, given equal class distribution and equal penalty for false positive and false negative predictions. For the carrier-control model analysis, we averaged AUC, accuracy, sensitivity and specificity values from the 50 times repeated nested cross-validations.

2.10. Multimodal classification

To obtain the best multimodal carrier-control model using several feature vectors, we performed step-wise feature concatenation as previously described (Bouts et al., 2018; Schouten et al., 2016). First, we assessed classification performance for each feature separately. Subsequently, we added a new feature to the best performing feature combination (i.e. highest AUC) of the previous step until all features were included in the model. The best performing feature combination will be referred to as the multimodal carrier-control model.

2.11. Statistical analysis

Statistical analyses of non-imaging data were performed using R (R Core 2016, Vienna, Austria). We tested for carrier-control differences using unpaired t-tests (age and education), the Mann-Whitney U test (mini-mental state examination (MMSE) scores (0–30)) and the χ² test (gender distribution). Probability scores were compared using Mann-Whitney U tests for overall carrier-control contrasts, and Kruskal Wallis H tests and Dunn post-hoc tests for comparisons between all four groups (MAPT, GRN, C9orf72 and controls). To compare models’ AUC values against chance level, we used permutation tests (N = 5,000) (Noirhomme et al., 2014). In order to correct for multiple comparisons, we took the maximum AUC difference of the family of tests for each permutation. Then we compared the observed AUC difference to the new distribution of maximum AUC differences to get a family-wise error rate corrected p-value. The alpha level required for statistical significance was set at p < 0.05.

3. Results

3.1. Demographics

In total, 103 subjects met the inclusion criteria (Table 2). Mean age was similar for mutation carriers (52.0 ± 8.6 years) and healthy controls (54.2 ± 7.5 years). The proportion of female participants between mutation carriers (67%) and healthy controls (58%) was not different (p = 0.3). Education level was similar between groups (mutation carriers, 13.6 ± 2.9 years; healthy controls, 13.2 ± 2.4 years). MMSE was similarly distributed between groups (median [min-max], mutation carriers: 30 (Agosta et al., 2012; McMillan et al., 2012; McMillan et al., 2014; Mahoney et al., 2014; Daianu et al., 2016; Zhou et al., 2010; Zhou and Seeley, 2014), healthy controls: 29 (Agosta et al., 2012; McMillan et al., 2012; McMillan et al., 2014; Mahoney et al., 2014; Mahoney et al., 2014; Daianu et al., 2016; Zhou et al., 2010; Zhou and Seeley, 2014)).

3.2. bvFTD model

Application of the bvFTD model resulted in low bvFTD probability scores for most subjects (Fig. 1A), and the bvFTD probability scores were not significantly different in presymptomatic carriers (median = 0.015) than controls (median = 0.005, p = 0.22). ROC analysis of the bvFTD probabilities resulted in an AUC of 0.570, which was not significantly better than chance level (p = 0.11). Separated by gene (Fig. 1B), there were no differences between the four groups’ bvFTD probability scores (p = 0.60). BvFTD probability scores of the original patients and controls used for cross-validation of the bvFTD model.

Table 2

Demographics.

|                | Carrier (n = 55)” | Control (n = 48) | P-value |
|----------------|-------------------|-----------------|---------|
| Age            | 52.0 (8.6)        | 54.2 (7.5)      | 0.2     |
| Gender, γ (%)  | 37 (67%)          | 28 (58%)        | 0.3     |
| Education, γ<bb> | 13.6 (2.9)       | 13.2 (2.4)      | 0.5     |
| MMSE           | 30 (24-30)        | 29 (24-30)      | 0.5     |

Abbreviations: MMSE: mini-mental state examination.

* 8 MAPT, 35 GRN, 12 C9orf72.
** Values denote mean (standard deviation).
*** Values denote median (range).
**** Education values were missing for four carriers and two controls.
Fig. 1. Classification results bvFTD model.

Box and scatter plot of each subject's bvFTD probability score on a scale from 0 (representing control) to 1 (representing bvFTD patient) after application of the bvFTD model. Groups are defined by carrier status (Fig. 1A) and genetic status (Fig. 1B). Probability scores were not significantly different for carriers and controls (p = 0.22), and did not differ between the four genetic groups (p = 0.60). Probability score results of the bvFTD patients and controls on which the bvFTD model was cross-validated were added for reference (Fig. 1C, data courtesy of Bouts et al. (2018) (Bouts et al., 2018)). Abbreviations: C9orf72: chromosome 9 open reading frame 72; GRN: progranulin; MAPT: microtubule-associated protein tau.
Bouts et al., 2018)). Abbreviations: AxD: axial diffusivity; FA: fractional anisotropy; FCor: full correlations between ICA components; WMD: white matter density; AUC: area under the ROC curve. Multimodal models result from step-wise addition of measures to the best performing classiﬁcation model used RD in combination with the WMD feature. Most models that outperformed chance included RD and WMD, and outperformed chance with an AUC of 0.646, 0.616, and 0.608, respectively. Of additional features were added to this model (Table 4). Interestingly, all models that outperformed chance level included RD, and most included several white matter features, such as WMD and the diffusivity features (i.e. FA, MD, AxD and/or RD).

Application of the best performing multimodal carrier-control model resulted in carrier probability scores (Fig. 2), which were different between carriers and controls (median = 0.629) and controls (median = 0.453, p = 0.001, Fig. 2A). Furthermore, there was a difference between the four groups’ carrier probability scores (p = 0.006) when separated by gene (Fig. 2B). Post-hoc tests revealed that GRN carriers had higher carrier probability scores than controls (Bonferroni family-wise error rate corrected p = 0.009). The other groups did not differ from each other.

4. Discussion

This study investigated whether presymptomatic FTD mutation carriers with MAPT, GRN and C9orf72 mutations can be individually distinguished from healthy controls using MRI. Using a recently introduced MRI-based classiﬁcation model trained on established bvFTD patients and controls, nearly all FTD mutation carriers and controls had low probability scores. The bvFTD model was therefore not able to separate carriers from controls better than chance level. However, MRI-based classiﬁcation models that were trained on our own sample were able to separate carriers from controls better than chance level. In our carrier-control model, the RD feature proved sufﬁcient to separate carriers from controls better than chance, but the best performing model used RD in combination with the WMD feature. Most models that outperformed chance used diffusivity features in combination with the WMD feature, supporting the hypothesis that WM alterations are the ﬁrst to appear in preclinical FTD pathology.

In an effort to improve on the FTD diagnostic criteria, single-subject classiﬁcation using MRI measures has recently received signiﬁcant attention (Möller et al., 2015b; McMillan et al., 2014; Davatzikos et al., 2008; Raamana et al., 2014; Koikkalainen et al., 2016; Wang et al., 2016; Meyer et al., 2017; Bron et al., 2017; Bouts et al., 2018). A recent multimodal classiﬁcation study incorporated structural, DTI and arterial spin labelling data to classify FTD (behavioural and language variants) from cognitively normal controls, and achieved an AUC of 0.96 (Bron et al., 2017). Another classiﬁcation study included tissue density, DTI and rs-fMRI measures, and achieved an AUC of 0.924 for bvFTD versus cognitively normal controls (Bouts et al., 2018). These high classiﬁcation performances are promising, but they are based on established FTD cases. It is unclear how FTD patient models generalise to earlier FTD stages, where brain alterations are less distinct. To test this, we applied a bvFTD model (Bouts et al., 2018) on FTD mutation carriers in a presymptomatic stage. We hypothesised that if the bvFTD model would be able to recognise early-stage FTD pathology, our

### Table 3
**ROC characteristics.**

| Modality | AUC | Min–max | Sensitivity | Specificity | Accuracy | FWER Corr P-value (AUC > chance) |
|----------|-----|---------|-------------|-------------|----------|----------------------------------|
| GM D     | 0.487 | (0.429–0.537) | 0.481 | 0.573 | 0.524 | 0.937 |
| WMD      | 0.616 | (0.546–0.670) | 0.583 | 0.646 | 0.612 | 0.076 |
| FA       | 0.511 | (0.450–0.558) | 0.516 | 0.559 | 0.536 | 0.760 |
| MD       | 0.608 | (0.563–0.652) | 0.561 | 0.668 | 0.611 | 0.124 |
| AxD      | 0.565 | (0.516–0.609) | 0.548 | 0.613 | 0.579 | 0.365 |
| RD       | 0.646 | (0.603–0.689) | 0.610 | 0.688 | 0.646 | 0.021 |
| FCor     | 0.534 | (0.459–0.592) | 0.545 | 0.568 | 0.555 | 0.650 |
| PCor     | 0.510 | (0.469–0.545) | 0.508 | 0.578 | 0.541 | 0.619 |
| Multimodal | 0.680 | (0.615–0.725) | 0.602 | 0.739 | 0.666 | 0.005 |

Presymptomatic FTD mutation carriers versus controls classiﬁcation. Multimodal represents the best combination from our step-wise multimodal procedure (i.e. WMD & AxD). Bold: best-performing model. Italic: mean AUC signiﬁcantly higher than chance level after family-wise error rate correction.

### Table 4
**Multimodal classiﬁcation performance.**

| Step: combined with: | RD | WMD | AxD | MD | GMD | FA | FCor | PCor |
|---------------------|----|-----|-----|----|-----|----|------|------|
| 1: –                |    | 0.646 | 0.616 | 0.565 | 0.608 | 0.487 | 0.511 | 0.534 | 0.510 |
| 2: RD               | –  | –  | 0.680 | 0.636 | 0.639 | 0.640 | 0.615 | 0.582 | 0.549 |
| 3: RD + WMD         | –  | –  | –  | 0.679 | 0.677 | 0.659 | 0.660 | 0.600 | 0.579 |
| 4: RD + WMD + AxD   | –  | –  | –  | –  | 0.645 | 0.636 | 0.641 | 0.612 | 0.595 |
| 5: RD + WMD + AxD + MD | –  | –  | –  | –  | –  | 0.626 | 0.615 | 0.613 | 0.607 |
| 6: RD + WMD + AxD + MD + GMD | –  | –  | –  | –  | –  | –  | 0.629 | 0.623 | 0.628 |
| 7: RD + WMD + AxD + MD + GMD + FA | –  | –  | –  | –  | –  | –  | –  | 0.635 | 0.631 |
| 8: RD + WMD + AxD + MD + GMD + FA + FCor | –  | –  | –  | –  | –  | –  | –  | –  | 0.622 |

Mean AUC values from 50 repetitions. Multimodal models result from step-wise addition of measures to the best performing classiﬁcation model of the previous step, starting with the best performing single MRI measure (i.e. RD). Bold: best performing model. Italic: mean AUC signiﬁcantly higher than chance level after family-wise error rate correction.

Abbreviations: AxD: axial diffusivity; FA: fractional anisotropy; FCor: full correlations between ICA components; WMD: white matter density; AUC: area under the ROC curve.
presymptomatic FTD mutation carriers would have higher probability scores than controls. We found that it was not possible to separate carriers from controls significantly better than chance using this model, as most carriers and controls had very low bvFTD probability scores. This could indicate that presymptomatic differences present in FTD mutation carriers are too subtle to be picked up by a classification model that was trained established bvFTD patients. However, it could also mean that most of our mutation carriers were still too far from conversion to have significant FTD-related changes. Since the bvFTD model was trained on patients, it stands to reason that classification of carriers and controls becomes more accurate as mutation carriers approach conversion. Vice versa, one might expect the carriers with high probability scores to be closer to symptom-onset than carriers with lower probabilities. Although it was not statistically significant, there was a trend towards older age in carriers with a bvFTD probability score higher than 0.25 than in the rest of the carrier group (data not shown). It can therefore not be entirely ruled out that age is partly associated with a higher bvFTD score. Longitudinal research is warranted to formally test whether this model captures presymptomatic FTD-related changes as mutation carriers approach conversion.

By training classifiers on presymptomatic FTD mutation carriers and controls, we obtained a unimodal carrier-control model based on the RD feature and several multimodal carrier-control models that significantly outperformed chance level. This suggests that classification models should be trained using early-stage FTD patients or presymptomatic FTD mutation carriers instead of advanced FTD cases, in order to be sensitive to early-stage FTD pathology. Furthermore, our carrier-control models demonstrate that MRI-based machine learning is powerful enough to detect subtle pathological changes associated with FTD even before symptom-onset and on a single-subject level. Although classification performance beyond chance level is an important finding, it must be noted that AUCs of 0.646 and 0.680 are modest and far from
sufficient for diagnostic use in the clinic. This is at least partly explained by our heterogeneous sample, as we included carriers of several genes in order to obtain sufficient sample size for robust cross-validation. Heterogeneity further arose from the uncertain time to onset in our sample. Investigating a uniform population a few years before symptom-onset might lead to higher classification performance, but these data were not available to us.

On a pathological level, it has been argued that neurodegeneration in FTD starts in the WM (Rohrer et al., 2013; Möller et al., 2016; Suri et al., 2014; Canu et al., 2017). Our results support this hypothesis, as the only unimodal model that outperformed chance was based on the RD feature, which was furthermore included in all multimodal models that significantly outperformed chance. Additionally, the majority of models that outperformed chance were based on WMD and DTI features. This means that our carrier-control model was able to combine subtle WM differences from the diffusion-weighted scans and the structural 3DT1w scan to classify a subject as mutation carrier or control.

In addition to the uncertain time-to-onset, there were several other limitations. Firstly, the bvFTD model was trained on a relatively small sample of 23 bvFTD patients and 35 controls. A model based on a larger sample might capture the heterogeneity of bvFTD pathology more completely, which could benefit generalisation to our presymptomatic sample. Furthermore, the model was trained on sporadic bvFTD patients, while it was applied to carriers of MAPT, GRN and C9orf72 genes. Since correlations between genetics, pathology and phenotype are not fully elucidated (Mann and Snowden, 2017), care must be taken not to over interpret our results. Specifically, pathological changes associated with non-behavioural variants (Seeaer et al., 2011) may be insufficiently recognised by the bvFTD model. Lastly, we used nested cross-validation to estimate out-of-sample performance for the carrier-control model, which minimises prediction bias (Kriegeskorte et al., 2011). But these data were not available to us.

5. Conclusion

Our data show that presymptomatic FTD mutation carriers can be distinguished from healthy controls on an individual level using a new multimodal MRI-based carrier-control classification model, while this was not possible using a recent bvFTD classification model. A multimodal MRI-based classification score may therefore be a useful biomarker to aid earlier FTD diagnosis. Successful single-subject recognition of early-stage or presymptomatic FTD may facilitate more precise subject recruitment into clinical trials. Furthermore, our multimodal MRI-based carrier-control classification model supports the hypothesis that FTD-related neurodegeneration starts in WM.

Acknowledgements

The authors of this work were supported by the Leiden University Medical Centre MD/PhD Scholarship (to RAF), ZonMw programme Memorabel project 73305013, JPND PreFrontAlz consortium project 733050142 (to JCS), and NWO VICI grant 016-130-667 (to SARBR). The views expressed are those of the authors and not necessarily those of the funding sources. The funding sources had no involvement in the design of the study; in the collection, analysis and interpretation of data; in the preparation of the manuscript; and in the decision to submit the article for publication. The authors report no conflict of interest.

References

Agosta, F., Scola, E., Canu, E., Marcone, A., Magnani, G., Sarro, L., et al., 2012. White matter damage in frontotemporal lobar degeneration spectrum. Cereb. Cortex 22 (12), 2705-2714.
Anderson, J.L.R., Jenkinson, M., Smith, S., 2007. Non-Linear Registration Aka Spatial Normalization. FMRIB Technical Report TR07JA2. FMRIB Centre, Oxford, United Kingdom.
Beckmann, C.F., Smith, S.M., 2004 Feb. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imaging 23 (2), 137–152.
Berndt, W., Wehr, J., Rinaldi, B., Houtz, M., Sayah, S., Camuamat, A., et al., 2018 Feb 1. Early cognitive, structural, and microstructural changes in presymptomatic C9orf72 carriers younger than 40 years. JAMA Neurol. 75 (2), 236-245.
Borroni, B., Alberici, A., Premi, E., Archetti, S., Garibotto, V., Agostì, C., et al., 2008 Jun. Brain magnetic resonance imaging structural changes in a pedigree of asymptomatic progranulin mutation carriers. Rejuvenation Res. 11 (3), 585-595.
Borroni, B., Alberici, A., Cercignani, M., Premi, E., Serra, L., Cerini, C., et al., 2012. Granulin mutation drives brain damage and reorganization from preclinical to symptomatic FTD. Neurobiol. Aging 33 (10), 2506-2520.
Bouts, M.J.R.J., Müller, C., Haflemeijer, A., van Swieten, J.C., Dopper, E., van der Flier, W.M., et al., 2018 Mar 27. Single subject classification of Alzheimer’s disease and behavioral variant frontotemporal dementia using anatomical, diffusion tensor, and resting-state functional magnetic resonance imaging. Zhang Y, editor. J. Alzheimers Dis. 62 (4), 1827-1839.
Bron, E.E., Smits, M., Papma, J.M., Steketee, R.M.E., Meijboom, B., de Groot, M., et al., 2017 Aug 16. Multiparametric computer-aided differential diagnosis of Alzheimer’s disease and frontotemporal dementia using structural and advanced MRI. Eur. Radiol. 27 (8), 3372-3382.
Canu, E., Agosta, F., Mandic-Stojmenovic, G., Stožeković, T., Stefanova, E., Imagi, A., et al., 2017. Multimodal MRI to distinguish early onset Alzheimer’s disease and behavioural variant of frontotemporal dementia. NeuroImage Clin. 15, 428-438.
Cash, D.M., Bocchetta, M., Thomas, D.L., Dick, K.M., van Swieten, J.C., Borroni, B., et al., 2018 Feb 1. Patterns of grey matter atrophy in genetic frontotemporal dementia: Results from the GENFI study. Neurobiol. Aging 62, 191-196.
Chao, L.L., Schiff, N., Cvejic, E.M., Mueller, S.G., Rosen, H.J., Goremo-Tempini, M.L., et al., 2007 Nov 1. Patterns of white matter atrophy in frontotemporal lobar degeneration. Arch. Neurol. 64 (11), 1619.
Daiu M., Mendez, M.F., Jim, Y., Melrose, R.J., Jimenez, E.E., et al., 2016 Oct 29. An advanced white matter tract analysis in frontotemporal dementia and early-onset Alzheimer’s disease. Brain Imaging Behav. 10 (4), 1038-1053.
Davatzikos, C., Resnick, S.M., Wu, X., Parmpi, P., Clark, C.M., 2008 Jul. Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. NeuroImage 41 (4), 1220-1227.
Dopper, E.G.P., Bormbouts, S.A.R.B., Jiskoot, L.C., den Heijer, T., de Graaf, J.R.A., de Koning, I., et al., 2014 Jul 8. Structural and functional brain connectivity in pre-symptomatic familial frontotemporal dementia. Neurology 83 (2), e18-e26.
Farb, N.A.S., Grady, C.L., Strother, S., Tang-Wai, D.F., Masellis, M., Black, S., et al., 2013 Jul. Abnormal network connectivity in frontotemporal dementia: evidence for pre-frontal isolation. Cortex 49 (7), 1856-1873.
Fawcett, T., 2006 Jun 1. An introduction to ROC analysis. Pattern Recogn. Lett. 27 (8), 861-874.
Friedman, J., Hastie, T., Tibshirani, R., 2008 Jul. Sparse inverse covariance estimation with the graphical lasso. Biostatistics 9 (3), 432-441.
Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1-22.
Frings, L., Yew, B., Flanagan, E., Lam, B.Y.K., Huil, M., Huppert, H. et al., 2014 Mar 3. Quantitative grey and white matter changes in frontotemporal dementia and Alzheimer’s disease. Esteban FJ, editor. PLoS One 9 (3), e98014.
Friston, K.J., Ashburner, J., Kiebel, S., Nichols, T., Penny, W.D., 2007. Statistical Parametric Mapping: The Analysis of Functional Brain Images. Elsevier/Academic Press (87pp.).
Gorno-Tempini, M.L., EL, Hillel A., Weintraub, S., Kertesz, A., Mendez, M., SFF, Cappa, et al., 2011. Classification of primary progressive aphasia and its variants. Neurology 76 (11), 1006-1014.
Greve, D.N., Fischl, B., 2009 Oct. Accurate and robust brain image alignment using boundary-based registration. NeuroImage 48 (1), 63–72.
Haflemeijer, A., Möller, C., Dopper, E.G.P., Jiskoot, L.C., Schouten, T.M., van Swieten, J.C., et al., 2015. Resting state functional connectivity differences between behavioral variant frontotemporal dementia and Alzheimer’s disease. Front. Hum. Neurosci. 9, 474.
Haflemeijer, A., Möller, C., Dopper, E.G.P., Jiskoot, L.C., van den Berg-Huysmans, A.A., van Swieten, J.C., et al., 2016. Differences in structural covariance brain networks between behavioral variant frontotemporal dementia and Alzheimer’s disease. Hum. Brain Mapp. 37 (3), 978-988.
Harvey, R.J., Skelton-Robinson, M., Rosser, M.N., 2003 Sep. The prevalence and causes of dementia in people under the age of 65 years. J. Neurol. Neurosurg. Psychiatry 74 (9), 1206-1209.
Huyé, E.D., Armstrong, N., Momeni, P., Grafman, J., 2008 Nov 14. Challenges and new opportunities in the investigation of new drug therapies to treat frontotemporal dementia. Expert Opin. Ther. Targets 12 (11), 1307-1316.
Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the boundary-based registration. NeuroImage 48 (1), 63–72.
Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I., 2009 May. Circular analysis in systems neuroscience: the dangers of double dipping. Nat. Neurosci. 12 (5), 535-540.
