



BRAIN COMMUNICATIONS

Classification using fractional anisotropy predicts conversion in genetic frontotemporal dementia, a proof of concept

 Rogier A Feis,^{1,2,3} Jeroen van der Grond,¹ Mark J R J Bouts,^{1,2,3} Jessica L Panman,^{1,4} Jackie M Poos,^{1,4}  Tijn M Schouten,^{1,2,3} Frank de Vos,^{1,2,3} Lize C Jiskoot,^{1,4,5} Elise G P Dopper,^{1,4} Mark A van Buchem,^{1,2} John C van Swieten⁴ and Serge A R B Rombouts^{1,2,3}

Frontotemporal dementia is a highly heritable and devastating neurodegenerative disease. About 10–20% of all frontotemporal dementia is caused by known pathogenic mutations, but a reliable tool to predict clinical conversion in mutation carriers is lacking. In this retrospective proof-of-concept case-control study, we investigate whether MRI-based and cognition-based classifiers can predict which mutation carriers from genetic frontotemporal dementia families will develop symptoms ('convert') within 4 years. From genetic frontotemporal dementia families, we included 42 presymptomatic frontotemporal dementia mutation carriers. We acquired anatomical, diffusion-weighted imaging, and resting-state functional MRI, as well as neuropsychological data. After 4 years, seven mutation carriers had converted to frontotemporal dementia ('converters'), while 35 had not ('non-converters'). We trained regularized logistic regression models on baseline MRI and cognitive data to predict conversion to frontotemporal dementia within 4 years, and quantified prediction performance using area under the receiver operating characteristic curves. The prediction model based on fractional anisotropy, with highest contribution of the forceps minor, predicted conversion to frontotemporal dementia beyond chance level (0.81 area under the curve, family-wise error corrected $P=0.025$ versus chance level). Other MRI-based and cognitive features did not outperform chance level. Even in a small sample, fractional anisotropy predicted conversion in presymptomatic frontotemporal dementia mutation carriers beyond chance level. After validation in larger data sets, conversion prediction in genetic frontotemporal dementia may facilitate early recruitment into clinical trials.

- 1 Department of Radiology, Leiden University Medical Centre, 2333 ZA, Leiden, the Netherlands
- 2 Leiden Institute for Brain and Cognition, Leiden University, 2333 ZA, Leiden, the Netherlands
- 3 Institute of Psychology, Leiden University, 2333 AK, Leiden, the Netherlands
- 4 Department of Neurology, Erasmus Medical Centre, 3015 GD, Rotterdam, the Netherlands
- 5 Dementia Research Centre, University College London, London, WC1N 3AR, UK

Correspondence to: Rogier A. Feis, MD Department of Radiology, Leiden University Medical Centre, Albinusdreef 2, 2333 ZA, Leiden, the Netherlands
E-mail: r.a.feis@lumc.nl

Keywords: frontotemporal dementia; MAPT protein, human; GRN protein, human; multimodal MRI; classification

Abbreviations: 3DT_{1w} = T₁-weighted 3-dimensional MRI; AUC = area under the receiver operating characteristics curve; AxD = axial diffusivity; (bv)FTD = (behavioral variant) frontotemporal dementia; *C9orf72* = chromosome 9 open reading frame 72; DWI = diffusion-weighted imaging; DTI = diffusion tensor imaging; FA = fractional anisotropy; FCor = full correlations between ICA components; FTD-RisC = frontotemporal dementia risk cohort; GMD = gray matter density; GRN = granulin; GENFI =

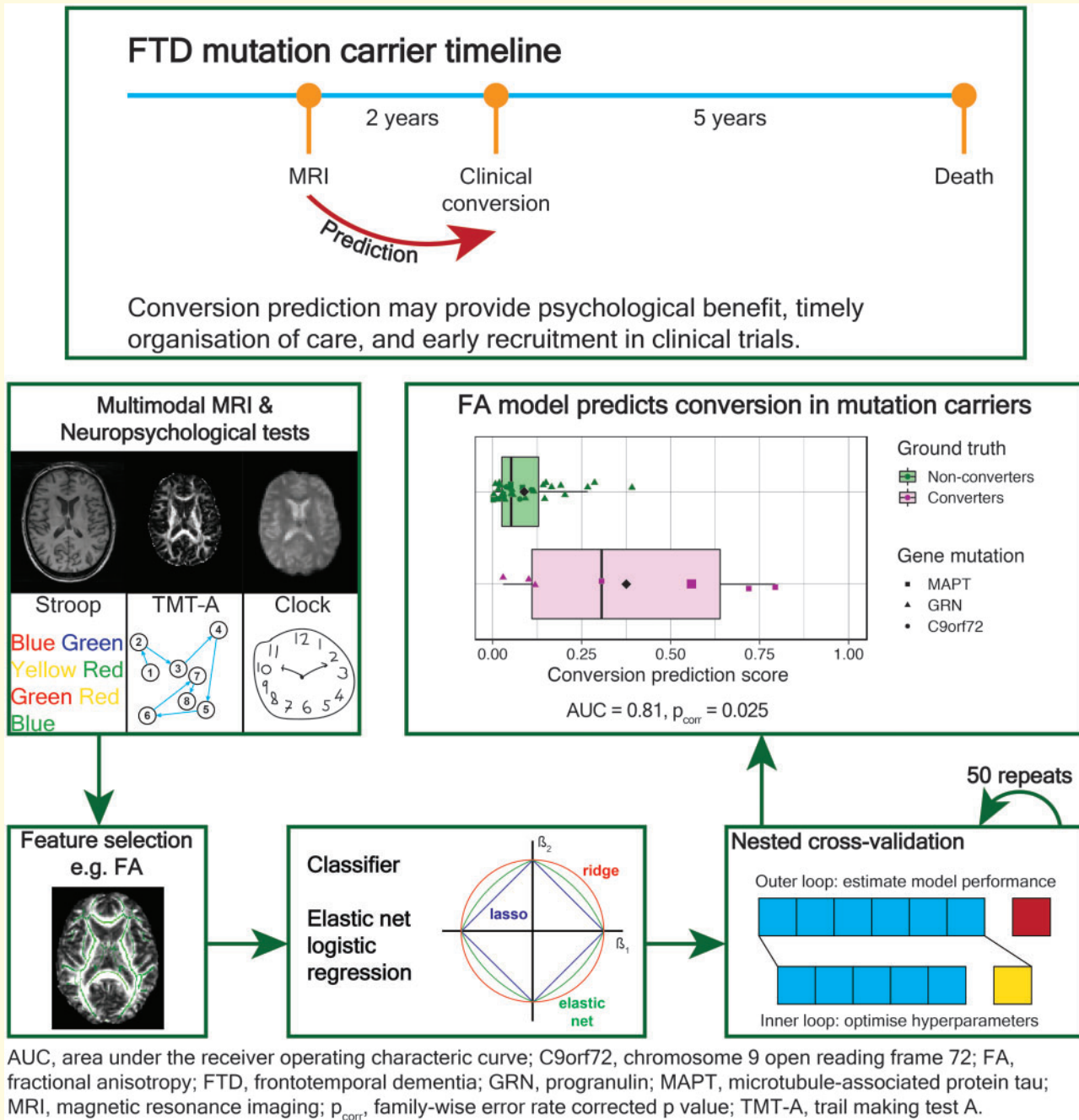
Received November 4, 2019. Revised April 29, 2020. Accepted May 11, 2020. Advance Access publication June 11, 2020

© The Author(s) (2020). Published by Oxford University Press on behalf of the Guarantors of Brain.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

genetic frontotemporal dementia initiative; ICA = independent component analysis; MAPT = microtubule-associated protein tau; MD = mean diffusivity; MMSE = mini-mental state examination; MRI = magnetic resonance imaging; PCor = sparse L1-regularized partial correlations between ICA components; (nfv)PPA = (non-fluent variant) primary progressive aphasia; RD = radial diffusivity; rs-fMRI = resting-state functional MRI; TBSS = tract-based spatial statistics; WM(D) = white matter (density)

Graphical Abstract



Introduction

Frontotemporal dementia (FTD) is a highly heritable and devastating neurodegenerative disease that often occurs at a presenile age (Ratnavalli *et al.*, 2002; Harvey *et al.*,

2003; Hogan *et al.*, 2016). Patients typically present with behavioral symptoms (behavioral variant FTD; bvFTD; Rascovsky *et al.*, 2011) or with language disorders (primary progressive aphasia; PPA; Gorno-Tempini *et al.*, 2011), but may also develop amyotrophic lateral sclerosis

(Lomen-Hoerth *et al.*, 2002) or Parkinsonian syndromes such as corticobasal syndrome and progressive supranuclear palsy (Josephs *et al.*, 2006). About 10–20% of all FTD is caused by three known pathogenic mutations: microtubule-associated protein tau (*MAPT*), progranulin (*GRN*) and chromosome 9 open reading frame 72 (*C9orf72*) repeat expansion (Seelaar *et al.*, 2008; Rohrer *et al.*, 2009; Benussi *et al.*, 2015). These mutations have a high penetrance and an autosomal dominant inheritance pattern. As such, roughly 50% of family members from a genetic FTD pedigree will develop an FTD variant.

Uncertainty about if and when a person will develop FTD is a major burden (Riedijk *et al.*, 2009; Cohn-Hokke *et al.*, 2018), and some subjects undergo genetic testing to eliminate this uncertainty. However, when a mutation is found, this raises the question when symptoms will emerge. A reliable prognosis may provide psychological benefit, and contributes to the timely organization of adequate care. Moreover, conversion prediction in genetic FTD may facilitate early recruitment into clinical trials for disease modifying treatments, which could lead to improved trial efficacy (Tsai and Boxer, 2016). However, accurate conversion prediction is currently not possible. Expected time to symptom onset, calculated using the ages of FTD onset of family members, has been used as proxy in research settings (Rohrer *et al.*, 2015; Jiskoot *et al.*, 2018a), but these estimates are inaccurate due to the large variation in age of onset between and within genetic FTD families (van Swieten and Heutink, 2008; Jiskoot *et al.*, 2018a). A more reliable clinical tool is needed to predict conversion in genetic FTD.

MRI-based classification is a promising diagnostic biomarker for FTD (Raamana *et al.*, 2014; Klöppel *et al.*, 2015; Koikkalainen *et al.*, 2016; Bron *et al.*, 2017; Meyer *et al.*, 2017; Bouts *et al.*, 2018), and can distinguish presymptomatic FTD mutation carriers from controls beyond chance (Feis *et al.*, 2019a). Classification with cognitive measures has also been advocated as diagnostic FTD biomarker (Wang *et al.*, 2016). Therefore, MRI- and cognition-based classification may be suitable candidates for conversion prediction in genetic FTD. In recent observational studies, FTD mutation carriers were followed up until conversion, and multimodal MRI data were analysed using mass univariate techniques such as voxel-based morphometry and diffusion tensor imaging (DTI; Jiskoot *et al.*, 2019), or using classification scores calculated from multimodal MRI and a classification model based on FTD patients (Feis *et al.*, 2019b). These studies suggest that brain changes as found on MRI appear relatively explosively in the final years before symptom onset. However, it is yet unknown whether conversion in such subjects could have been predicted based on pre-conversion MRI scans or neuropsychological data.

Here, we study whether MRI-based features and cognitive measures have predictive value for FTD conversion in genetic FTD. To this end, we train prediction models on MRI-based features and cognitive measures to predict which presymptomatic FTD mutation carriers develop symptoms within 4 years ('converters'), and which mutation carriers do not ('non-converters').

Materials and methods

Participants

This retrospective case-control study included FTD mutation carriers from the FTD-Risk Cohort (FTD-RisC; Dopper *et al.*, 2014; Jiskoot *et al.*, 2016, 2019; Meeter *et al.*, 2016; Papma *et al.*, 2017; Jiskoot *et al.*, 2018), a longitudinal study that follows healthy, 50% at-risk family members of genetic FTD patients on a 2-year basis. For the current study, we included the data of 43 FTD mutation carriers, who entered FTD-RisC between May 2010 and November 2014. Inclusion criteria were age at MRI scan between 40 and 70 years old, availability of a T₁-weighted 3-dimensional MRI (3DT_{1w}) scan, a diffusion-weighted imaging (DWI) dataset, a resting-state functional MRI T₂*-weighted (rs-fMRI) scan, and availability of clinical follow-up after 2 and 4 years. Genotyping, as described in previous work (Dopper *et al.*, 2014; Jiskoot *et al.*, 2016), revealed nine *MAPT* mutation carriers, 28 *GRN* mutation carriers, and six *C9orf72* repeat expansion carriers. For seven FTD mutation carriers who had converted to FTD (four *MAPT* and three *GRN* mutation carriers; 'converters', see 'Conversion' section below), we included the MRI data of their last visit before conversion. For one converter (*MAPT* mutation carrier), these MRI data were excluded due to the presence of artefacts, and we included MRI data of the previous visit. The time between MRI and conversion ranged between 11 and 41 months (mean = median = 22 months, SD = 10.5 months). More details on this group's demographics are provided in the results. The remaining mutation carriers were confirmed to be asymptomatic after 4 years follow-up in accordance with established diagnostic criteria for bvFTD (Rascovsky *et al.*, 2011), PPA (Gorno-Tempini *et al.*, 2011) and amyotrophic lateral sclerosis (Ludolph *et al.*, 2015), and were termed 'non-converters'. Follow-up after 6 years was not complete for this group: 25 non-converters were followed up after 6 years and all remained asymptomatic at that point. For non-converters, we included the FTD-RisC baseline MRI data. One non-converter with a *GRN* mutation was excluded due to incomplete neuropsychological data. Other exclusion criteria were current or past neurologic (other than dementia) or primary psychiatric disorders, history of drug abuse, large image artefacts and gross brain pathology other than atrophy.

Standard protocol approvals, registrations and patient consents

Participants and clinical investigators were blinded to the participants' genetic status, except for those that underwent predictive testing at their own request. For converters, genetic counselling was offered to the patient and family members, and genetic status was unblinded to confirm the presence of the pathogenic mutation. 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 (Feis *et al.*, 2019b).

Conversion

Conversion was determined in a multidisciplinary consensus meeting of the Erasmus Medical Centre FTD Expertise Centre, involving neurologists (J.C.v.S.), neuropsychologists (J.L.P., J.M.P., L.C.J.), neuroradiologists, a clinical geneticist and a care consultant (Feis *et al.*, 2019b). In the consensus meetings, information from the medical history, neuropsychological assessment and MRI of the brain were reviewed. The timing of symptom onset was estimated from heteroanamnesic information provided by knowledgeable informants (e.g. siblings, spouses). After the clinical diagnosis was made, and if the subject and family agreed, genetic status was unblinded for confirmation. Four *MAPT* mutation carriers and one *GRN* mutation carrier converted to bvFTD, while two *GRN* mutation carriers converted to non-fluent variant PPA (nfvPPA). BvFTD converters presented with progressive behavioral deterioration, functional decline and frontal and/or temporal lobe atrophy on structural MRI, fulfilling the international diagnostic consensus criteria for bvFTD with definite frontotemporal lobar degeneration pathology (Rascovsky *et al.*, 2011). The two nfvPPA converters presented with isolated language difficulties and no impairment in daily living activities. Both showed a non-fluent, halting speech with sound errors and agrammatism, fulfilling the diagnostic criteria for nfvPPA (Gorno-Tempini *et al.*, 2011). For more detailed information on conversion criteria, and for a full description of the converters' clinical profile (see Jiskoot *et al.*, 2018b, 2019).

MRI data acquisition

Subjects were scanned at the Leiden University Medical Centre using a 3T MRI scanner (Achieva, Philips Medical Systems, Best, the Netherlands) with an 8-channel SENSE head coil (Feis *et al.*, 2019a). The imaging protocol included a whole-brain near-isotropic 3D_{T1w} sequence for cortical and subcortical tissue-type segmentation, a DWI sequence for assessments of white matter

(WM) diffusivity, and rs-fMRI sequence 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. Scan parameters are provided in Table 1.

Image preprocessing

Preprocessing of MRI scans was performed similarly to previous work (Bouts *et al.*, 2018; Feis *et al.*, 2019a, b). All registration and segmentation steps were critically reviewed and errors were corrected accordingly.

For 3D_{T1w} images, we 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 gray matter structures, including the bilateral thalamus, caudate nucleus, putamen, globus pallidum, nucleus accumbens, amygdala and hippocampus (FIRST; Patenaude *et al.*, 2011). Our choice to use SPM segmentation was based on the segmentation tool comparison by Kazemi and Noorzadeh (2014).

Preprocessing of DWI data sets included correction of motion and eddy-current induced distortions (eddy correct; Leemans and Jones, 2009), and voxel-wise calculation of the measures fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AxD; largest eigenvalue) and radial diffusivity (RD; average of the two remaining eigenvalues) using 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 that extracted values originate from WM.

For rs-fMRI data, preprocessing consisted of 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 normalization, 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, followed by automatic identification and removal of motion artefacts using ICA-AROMA (version 0.3 beta; Pruim *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 3D_{T1w} image using Boundary-Based Registration (Greve and Fischl, 2009). Next, resulting registration parameters were concatenated to the 3D_{T1w}-to-MNI152 template registration parameters to obtain the final registration parameters.

Table 1 MRI sequence parameter settings

	Slices	TR (ms)	TE (ms)	Flip angle (°)	Matrix (mm)	Voxel size (mm)	Duration (min)
3DT _{1w}	140	9.8	4.6	8	256 × 256	0.88 × 0.88 × 1.20	4.57
DWI ^a	70	8250	80	90	128 × 128	2.00 × 2.00 × 2.00	8.48
rs-fMRI	38	2200	30	80	80 × 80	2.75 × 2.75 × 2.99 ^b	7.28

Scan protocol of whole-brain near-isotropic 3DT_{1w}-weighted (3DT_{1w}), diffusion-weighted imaging (DWI) and resting-state functional MRI T₂*-weighted MRI (rs-fMRI) on a 3T scanner at the Leiden University Medical Centre.

^a60 directions, *b* = 1000, one b0 image.

^bIncluding 10% interslice gap.

TR = repetition time; TE = echo time.

MRI feature selection

Selection of MRI features was performed as described previously (Bouts *et al.*, 2018; Feis *et al.*, 2019a).

Cortical gray matter density (GMD) and white matter density (WMD) were calculated as weighted means of their respective regional gray matter or WM probability (SPM segmentation) weighted by the probability of a voxel belonging to a specific cortical region. Cortical probabilities were derived from the 48 Harvard-Oxford probabilistic anatomical brain atlas (split into left and right), and WM probabilities were derived from the Johns-Hopkins University WM tractography atlas for 20 WM tracts. Voxels with atlas probability values under 25% were excluded. We estimated these features in native space by transforming the atlas masks to the images' native space using the inverted non-linear registration parameters. For deep gray matter regions, GMD values were calculated as the regions' volume (FIRST segmentations) divided by the total intracranial volume. This resulted in a GMD feature vector of 110 weighted mean GMD values (48 left cortical, 48 right cortical and 14 deep gray matter regions) and in a feature vector of 20 weighted mean WMD values per subject. While gray matter volume as estimated from voxel-based morphometry would also be a useful feature for prediction, we chose the simpler GMD for generalisation purposes, as it is not reliant on a study-specific template.

DTI features were calculated as weighted mean FA, MD, AxD or RD values per tract per subject. First, we projected each subject's FA, MD, AxD and RD values onto the TBSS group skeleton on a voxel-wise basis. Next, we weighted the mean values per tract by the probability of a voxel belonging to that specific tract, derived from 20 tracts of the Johns-Hopkins University WM tractography atlas. Voxels with atlas probability values under 25% were excluded. This resulted in four feature vectors of each 20 weighted mean values per subject.

To calculate functional connectivity features, all processed rs-fMRI images were combined in a temporally concatenated group-level ICA (MELODIC; Beckmann and Smith, 2004), with dimensionality fixed at 20 components and an ICA threshold of 0.99 (Smith *et al.*, 2013). This means 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 gray matter probability of that component's regions. Correlations between the mean time courses of all pairs of components were subsequently calculated. Functional connectivity was calculated as full correlations (FCor), and as sparse, L1-regularized, partial correlations (PCor) between the mean time courses of all pairs of components. Partial correlations were calculated using the graphical lasso algorithm (Friedman *et al.*, 2008). This procedure resulted in two feature vectors of each $(20 \times 19) \div 2 = 190$ (partial) correlations per subject.

Finally, we created one multimodal feature vector by concatenating all feature vectors together. As such, we had nine feature vectors in total: two 3DT_{1w} feature vectors, four DTI feature vectors, two rs-fMRI feature vectors and one multimodal imaging feature vector.

Neuropsychological assessment and cognitive feature selection

We screened global cognitive functioning by means of the Mini-Mental State Examination (MMSE; Folstein *et al.*, 1975). Experienced neuropsychologists (J.L.P., J.M.P., L.C.J.) administered neuropsychological tests within six cognitive domains: language, attention and mental processing speed, executive functioning, social cognition, memory and visuoconstruction. The language domain was assessed using the 60-item Boston Naming Test (BNT; Kaplan *et al.*, 1978), verbal Semantic Association Test (SAT; Visch-Brink *et al.*, 2005), ScreeLing phonology (Doesborgh *et al.*, 2003) and categorical fluency (Thurstone and Thurstone, 1962). The domain of attention and mental processing speed was evaluated using the Trail Making Test (TMT)-A (Army Individual Test Battery, 1944), the mean of Stroop Colour-Word Tests I and II (Stroop, 1935), Wechsler Adult Intelligence Scale III (WAIS-III) Digit Span forwards (Wechsler, 2005) and Letter Digit Substitution Test (LDST; Jolles *et al.*, 1995). We measured the executive functioning domain using TMT-B (Army Individual Test Battery, 1944), Stroop Colour-Word Test III (Stroop, 1935), WAIS-III Digit Span backwards (Wechsler, 2005), modified Wisconsin Card Sorting Test (WCST) concepts (Nelson, 1976), Letter Fluency (Thurstone and Thurstone, 1962) and

WAIS-III Similarities (Wechsler, 2005). The social cognition domain was assessed by means of Happé theory of mind cartoons, Happé non-theory of mind cartoons (Happé *et al.*, 1999) and Ekman Faces (Ekman and Friesen, 1976). We evaluated the memory domain using the immediate response and recall of the Dutch Rey Auditory Verbal Learning Test (RAVLT; Rey, 1958), and the Visual Association Test (VAT; Lindeboom *et al.*, 2002). Lastly, we tested the visuoconstruction domain by means of clock drawing (Royall *et al.*, 1998) and WAIS-III Block Design (Wechsler, 2005).

Raw test scores were added into one cognitive feature vector per domain. This resulted in a language feature vector with four features, an attention feature vector with four features, an executive feature vector with six features, a social feature vector with three features, a memory feature vector with three features, and a visuoconstruction feature vector with two features. Finally, we added these feature vectors together to create a multidomain feature vector, bringing the total number of feature vectors on seven: language, attention, execution, social, memory, visuoconstruction and multidomain.

Classifier

In order to identify future converters ($n=7$) from non-converters ($n=35$) at baseline, we trained prediction models using MRI features ($n=9$) and, separately, cognitive features ($n=7$). Feature vectors were used to train a logistic elastic net regression algorithm (Zou and Hastie, 2005; Friedman *et al.*, 2010; Schouten *et al.*, 2016; Bouts *et al.*, 2018, Feis *et al.*, 2019a). 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 (i.e. L1 penalty) and for the sum of the squared value of the coefficients (i.e. L2 penalty). This way, elastic net provides a solution for the imbalance between the large number of features and the small number of subjects. Age and sex were included in the model without penalty to ensure that estimated feature regression coefficients were conditional on subject age and sex. A prediction score of 0 represented a non-converter and 1 represented a converter.

Cross-validation

We trained our conversion prediction models in a stratified nested 7-fold cross-validation scheme to make sure that the proportion of converters and non-converters was the same in each fold (i.e. one converter and five non-converters per fold). In the outer loop, one part of the data (i.e. one of the seven folds) was set apart as a test set and served to test the generalized prediction performance of the elastic net regression model. The remaining parts (six of the seven folds) were used to train the model. Within the training set of the outer loop, we performed a nested cross-validation to optimize the model's

hyperparameters without overestimating prediction performance (Varma and Simon, 2006; Kriegeskorte *et al.*, 2009). The resulting optimal hyperparameters were used in the training set of the outer loop to train the model, and the prediction performance was then tested in the test set of the outer loop. This process was repeated seven times to make sure that each subject was part of the test set exactly once. Since the test set of the outer loop was neither used for model training, nor for parameter optimisation, we reduced the risk of overestimating the generalization performance as much as possible (Kriegeskorte *et al.*, 2009; Schouten *et al.*, 2016; Bouts *et al.*, 2018, Feis *et al.*, 2019a). The entire prediction procedure was repeated 50 times to average prediction outcome variability resulting from random partitioning in training and test folds. All prediction analyses and evaluations were implemented in R version 3.3.2 (R core 2016, GLMnet package; Friedman *et al.*, 2010).

Prediction performance

For both analyses, we quantified prediction performances using receiver operating characteristic curves. Receiver operating characteristic curves were calculated by shifting the threshold for predicting an individual as converter from 0 to 1, and plotting the true positive rate (sensitivity) versus the false-positive rate ($1 - \text{specificity}$) for each intermediate point. The area under this receiver operating characteristic curve (AUC) is a measure of prediction 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 prediction accuracy, given equal class distribution and equal penalty for false positive and false negative predictions. We averaged AUC, accuracy, sensitivity and specificity values from the 50 times repeated nested cross-validations (Schouten *et al.*, 2016; Bouts *et al.*, 2018, Feis *et al.*, 2019a).

Statistical analysis

Statistical group analyses of demographic data were performed using R (R Core 2016, Vienna, Austria). We tested for differences between converters and non-converters using unpaired t -tests (age and education), the Mann-Whitney U test [MMSE scores (0–30)] and the χ^2 test (sex distribution).

To compare MRI and cognitive prediction models' AUC values versus chance level, we used permutation tests ($N=5000$; Noirhomme *et al.*, 2014). We used the maximum t -statistic method to correct for multiple comparisons within, respectively, the nine MRI-based and the seven cognitive prediction analyses. For each permutation, we calculated the maximum absolute t -statistic within the family of tests, which resulted in a maximum t -distribution of 5000 maximum absolute t -statistics. The observed t -statistic of each analysis was then compared to this

maximum *t*-distribution in order to obtain a family-wise error rate corrected *P*-value. The alpha level required for statistical significance was set at 0.05.

Data availability

Raw data were generated at the Leiden University Medical Centre. The derived data, as well as scripts, that support the findings of this study are available from the corresponding author upon request.

Results

Demographics

Seven converters and 35 non-converters met the inclusion criteria (Table 2). At the time of MRI scan, converters did not differ from non-converters in terms of age ($P=0.85$), sex distribution ($P=0.40$) and MMSE scores ($P=0.50$). However, converters had higher levels of education than non-converters ($P=0.015$). More information on the converter group is shown in Table 3.

Prediction performance

Prediction with MRI features yielded mixed results (Table 4). The model based solely on FA features (FA model) significantly outperformed chance level with an AUC of 0.81 ($P=0.025$ versus chance level). For the FA model, subjects' prediction scores are shown in Fig. 1. Converters with a *MAPT* mutation seemed to have higher conversion prediction scores than those with a *GRN* mutation. Similarly, male mutation carriers seemed to have higher prediction scores than female mutation carriers. There seemed to be no clear correlation (Pearson's $r=0.02$) between conversion prediction scores and time from MRI to conversion. However, the above three observations should be noted with care, since the small sample size did not permit for meaningful statistical testing. The FA model's beta weights for the 50 cross-validation repeats are shown in Fig. 2. The highest beta weights were found in the forceps minor. Other WM tracts with above average beta weights included the forceps major, right corticospinal tract, right inferior longitudinal fasciculus, right anterior thalamic radiation and right uncinate fasciculus. For the two WM tracts with the largest beta weights (i.e. forceps minor and major), we plotted subjects' the mean FA value across the tract in Fig. 3. Corresponding with the beta weights, the differences between converters and non-converters in mean FA value were larger for the forceps minor than for the forceps major.

The other DTI features (i.e. MD, AxD, RD) did not outperform chance level, nor did gray and WM density features, and functional connectivity features (i.e. FCor, PCor). Concatenating all features together in a multimodal MRI model did not improve performance

Table 2 Demographics

	Converter (n = 7) ^a	Non-converter (n = 35)	P-value
Age, mean (SD) years	51.7 (8.8)	51.0 (8.4)	0.85
Sex, n (%) female	4 (57%)	26 (74%)	0.4
Education, mean (SD) years ^b	15.2 (0.75)	13.7 (2.9)	0.015
MMSE, median (range) points	29 (27–30)	30 (24–30)	0.5

^aFour *MAPT*, three *GRN*.

^bEducation values were missing for one converter.

GRN = progranulin; *MAPT* = microtubule-associated protein tau; MMSE = mini-mental state examination.

compared to the FA model (AUC = 0.63, $P=0.55$ versus chance level).

Prediction with cognitive features resulted in poor performances (Table 5). The best-performing cognitive model was based on the executive domain (AUC = 0.35). Adding multiple domains together in a multidomain cognitive model did not improve performance (AUC = 0.34).

Discussion

This study describes conversion prediction within 4 years in FTD mutation carriers using MRI-based and cognitive measures. We found that it was possible to distinguish converters from non-converters beyond chance level with an AUC of 0.81 using FA. Other MRI-based and cognitive measures did not outperform chance level.

The FA model's performance reaffirms the potential of diffusion scans for early FTD diagnosis. Group differences in FA and other DTI metrics exist in presymptomatic FTD mutation carriers compared to controls (Borroni *et al.*, 2008; Dopper *et al.*, 2014; Pievani *et al.*, 2014; Lee *et al.*, 2017; Papma *et al.*, 2017; Bertrand *et al.*, 2018, Jiskoot *et al.*, 2018a, 2019). More specifically with regard to conversion, two recent FTD-RisC studies observed how mass univariate MRI measures and MRI-based classification scores develop as subjects approach conversion, as compared to subjects who did not convert during follow-up. Jiskoot *et al.* found that gray matter atrophy and WM DTI changes become noticeable around 2 years before symptom onset, and longitudinal FA decline was primarily located in the genu of the corpus callosum (Jiskoot *et al.*, 2019), the area with the highest beta values in our current study. Feis *et al.* combined multimodal MRI to calculate a single classification score at each time point using a classification model that was trained on bvFTD patients and controls. They found that the classification scores of presymptomatic mutation carriers and non-carriers did not differ over time, but that classification scores of converters rose faster than those of non-converting mutation carriers (Feis *et al.*, 2019b). Together, these studies suggest that FTD-related deterioration on MRI accelerates when subjects are near

Table 3 Converter demographics

	Gene	Mutation	Age at MRI	Sex	Months to conversion	FTD variant	Prediction score ^a
Converter 1	GRN	S82VfsX174	65	Female	23	bvFTD	0.12
Converter 2	MAPT	P301L	54	Female	28	bvFTD	0.31
Converter 3	MAPT	P301L	57	Male	18	bvFTD	0.79
Converter 4	GRN	S82VfsX174	56	Female	11	nvPPA	0.10
Converter 5	GRN	S82VfsX174	49	Female	22	nvPPA	0.03
Converter 6	MAPT	G272V	41	Male	41 ^b	bvFTD	0.56
Converter 7	MAPT	G272V	42	Male	11	bvFTD	0.72

^aPrediction scores based on the FA prediction model range from 0 to 1, with 0 representing non-converters and 1 representing converters.

^bMRI data nearer to symptom onset were excluded due to artefacts for this subject.

bvFTD = behavioural variant FTD; FTD = frontotemporal dementia; GRN = progranulin; MAPT = microtubule-associated protein tau; MRI = magnetic resonance imaging; nvPPA = non-fluent variant primary progressive aphasia.

Table 4 MRI features' performance

MRI modality	AUC	Min-max	Sensitivity	Specificity	Accuracy	FWER Corr P-value (AUC > chance)
GMD	0.668	0.563–0.731	0.746	0.599	0.623	0.375
WMD	0.438	0.306–0.567	0.571	0.517	0.526	0.976
FA	0.812	<i>0.608–0.906</i>	<i>0.769</i>	<i>0.761</i>	<i>0.762</i>	0.025
MD	0.677	0.616–0.743	0.711	0.693	0.696	0.322
AxD	0.435	0.327–0.563	0.626	0.491	0.514	0.977
RD	0.625	0.465–0.722	0.689	0.630	0.640	0.590
FCor	0.336	0.204–0.449	0.446	0.513	0.501	1.000
PCor	0.403	0.233–0.502	0.514	0.527	0.525	0.996
Multimodal	0.626	0.547–0.710	0.557	0.781	0.744	0.552

Converters (seven presymptomatic FTD-RisC mutation carriers that developed symptoms within 4 years after assessment) versus non-converters (35 FTD mutation carriers that remained cognitively healthy after 4 years). Multimodal represents a combination of all MRI features. Min-max AUC values represent the variance across the 50 repeats. Italic: best-performing model. Bold: mean AUC significantly higher than chance level after family-wise error rate correction.

AUC = area under the receiver operating characteristic curve; FA = fractional anisotropy; FCor = full correlations between 20 ICA components; FTD-RisC = Frontotemporal dementia Risk Cohort; FWER = family-wise error rate; GMD = grey matter density; ICA = independent component analysis; MD = mean diffusivity; PCor = L1-regularized partial correlations between 20 ICA components; RD = radial diffusivity; WMD = white matter density.

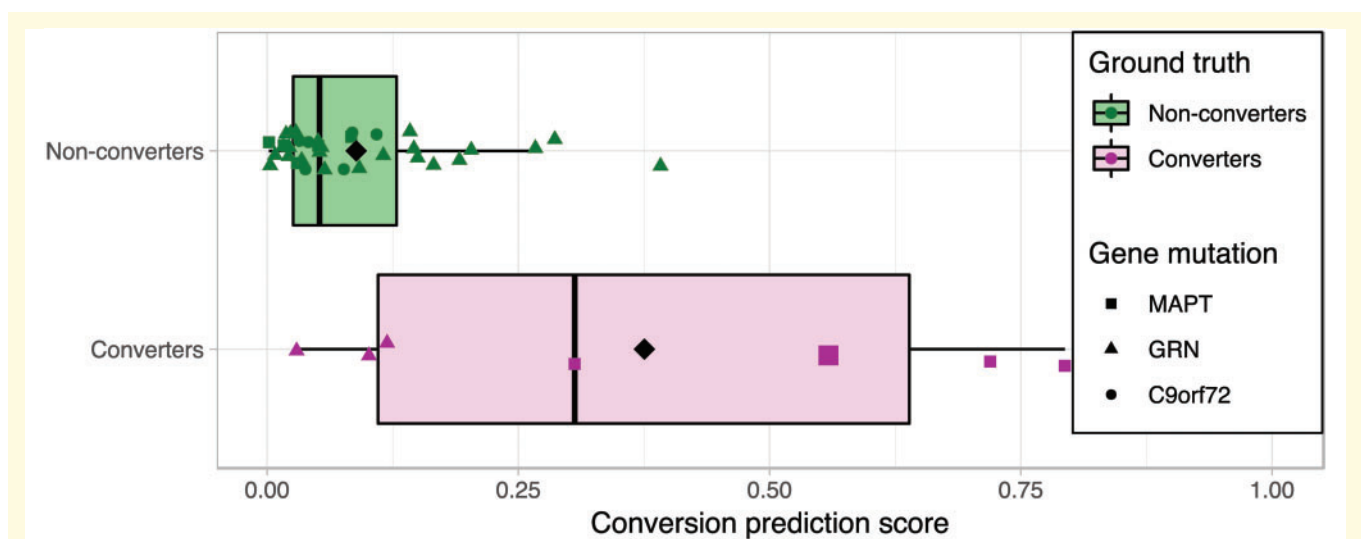


Figure 1 Converters' and non-converters' conversion prediction scores. Box and scatter plot of each subject's conversion prediction score on a scale from 0 (representing non-converter) to 1 (representing converter) after applying the FA model. Different FTD gene mutations were represented with different shapes. The converter with the longest time between MRI and conversion (i.e. 41 months) was annotated with increased size. These conversion prediction scores result in a performance of 0.81 AUC for the FA model ($P = 0.025$ versus chance level). *C9orf72* = chromosome 9 open reading frame 72; FA = fractional anisotropy; GRN = progranulin; MAPT = microtubule-associated protein tau.

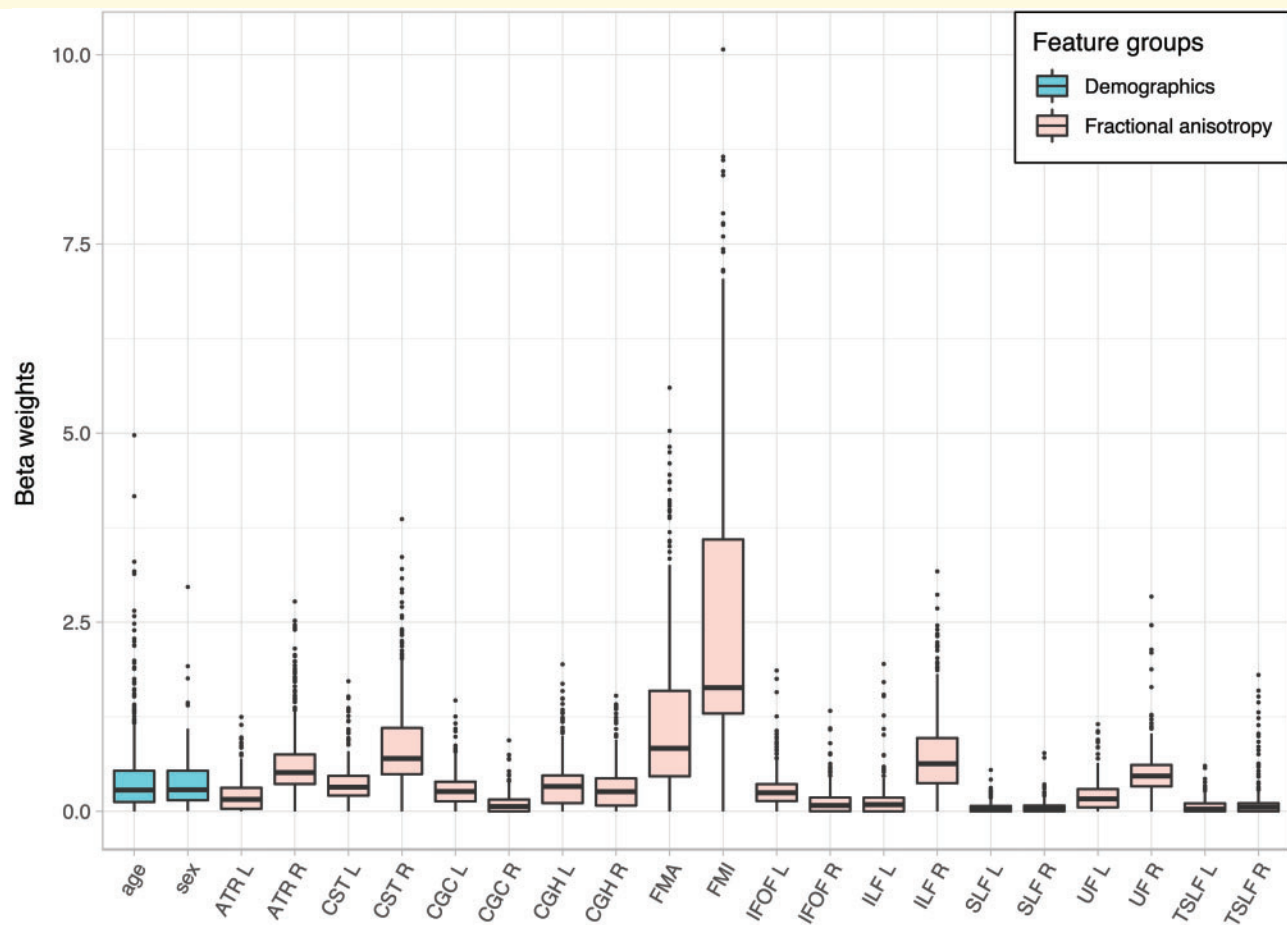
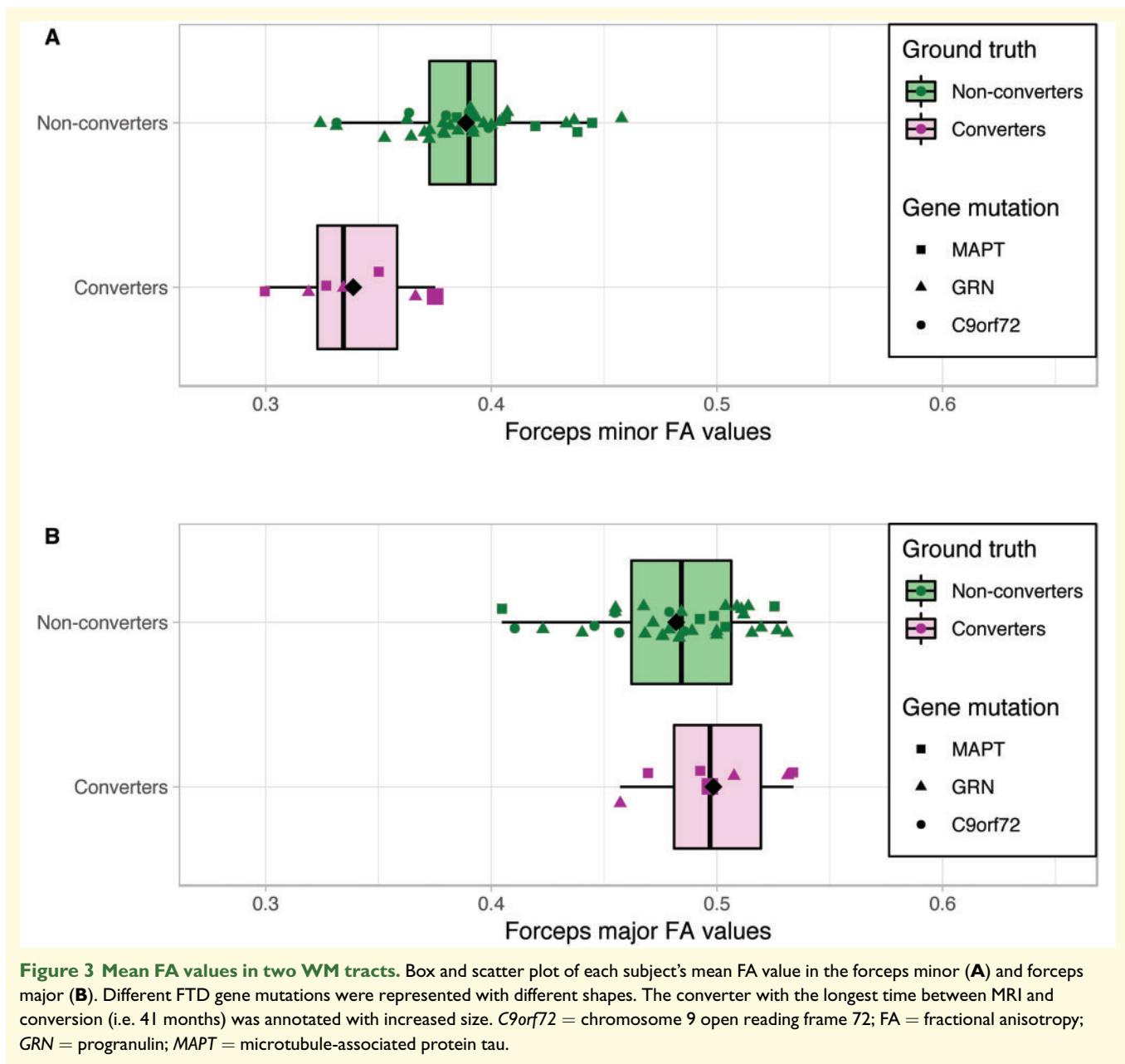


Figure 2 FA model beta weights. Box plots show the FA model's standardized beta weights for the 50 cross-validation repeats. Demographics (age and sex, in blue) were included in the model without penalty, while the FA features (red) were regularized. ATR = anterior thalamic radiation; CST = corticospinal tract; CGC = cingulum in the cingulate gyrus area; CGH = cingulum in the hippocampal area; FA = fractional anisotropy; FMA = forceps major; FMI = forceps minor; IFOF = inferior fronto-occipital fasciculus; ILF = inferior longitudinal fasciculus; L = left; R = right; SLF = superior longitudinal fasciculus; TSLF = temporal projection of the SLF; UF = uncinate fasciculus.

conversion. However, they fail to address the most important question: is it possible, using presymptomatic cross-sectional data, to predict future conversion in individual FTD mutation carriers? We expand on this by showing that individual differences in FA have predictive value for FTD onset within 4 years in FTD mutation carriers.

Within the FA model, the forceps minor feature had the highest beta weights across 50 repeats, suggesting that this area was important for prediction. Accordingly, the difference in mean FA between converters and non-converters was larger in the forceps minor than in other tracts, e.g. the forceps major. The genu of the corpus callosum (i.e. part of the forceps minor) was recently found to be the most consistent WM region in terms of DTI changes across FTD subtypes (Elahi *et al.*, 2017). As our converter group included both bvFTD and nvPPA converters, this might explain why the forceps minor had high beta weights, and could provide biological support

for our model. Forceps minor changes may therefore be a promising marker for early FTD detection, although it might not discriminate between the different FTD subtypes. Other WM tracts are affected more specifically in the different clinical syndromes. BvFTD is generally characterized by DTI changes in the anterior thalamic radiation, cingulum, and uncinate fasciculus (Zhang *et al.*, 2011; Mahoney *et al.*, 2014; Möller *et al.*, 2015; Daianu *et al.*, 2016), while the superior longitudinal fasciculus and corticospinal tract are typically affected in nvPPA (Whitwell *et al.*, 2010; Galantucci *et al.*, 2011; Omer *et al.*, 2017) and the inferior longitudinal and uncinate fasciculi show DTI changes in svPPA (Whitwell *et al.*, 2010; Galantucci *et al.*, 2011; Tu *et al.*, 2015). Future studies should investigate whether this combination of shared and different WM DTI changes between the FTD subtypes facilitate hierarchical classification. Although it is interesting to examine which features were important for prediction, care must be taken not to overinterpret



the beta weights. Contrary to explanatory regression models, the beta weights of our regularized prediction model do not designate direct relationships between the features and prediction score, nor do they reflect mean differences between the groups (Shmueli, 2010). Each feature's effect is conditional on the effects of all other features in the model, and multicollinearity between features may result in suppression of the effect of some features. For example, we cannot be sure whether the higher beta weights in the FA model for right hemispheric WM tracts point towards asymmetric effects—such as are reported in *GRN* mutation carriers (Jiskoot et al., 2018a)—or whether the corresponding regions in the left hemisphere added redundant information, and were therefore suppressed in the model.

Apart from the FA model, no other MRI- or cognition-based prediction model significantly outperformed chance. However, it would be inappropriate to draw conclusions based on the non-significant results. Due to our small sample size, it is unclear whether these features were uninformative, or whether the study was underpowered to find significant results. Grey matter atrophy occurs at a later stage than WM changes in *MAPT* and *GRN* mutation carriers (Borroni et al., 2008; Agosta et al., 2012; Rohrer and Rosen, 2013; Feis et al., 2019a; Jiskoot et al., 2019), but the question whether or not that precludes the use of gray matter features for early diagnosis requires further investigation. Similarly, we cannot comment on why the remaining diffusion features MD, AxD and RD, which often yield comparable results to FA, did

Table 5 Cognitive features' performance

Cognitive domain	AUC	Min–max	Sensitivity	Specificity	Accuracy	FWER Corr P-value (AUC > chance)
Language	0.282	0.180–0.371	0.394	0.515	0.495	0.999
Attention	0.343	0.208–0.441	0.517	0.451	0.462	0.992
<i>Executive</i>	<i>0.349</i>	<i>0.184–0.482</i>	<i>0.534</i>	<i>0.439</i>	<i>0.455</i>	<i>0.987</i>
Social	0.344	0.208–0.449	0.474	0.488	0.486	0.992
Memory	0.272	0.143–0.359	0.491	0.414	0.427	0.999
Visuoconstruction	0.288	0.186–0.424	0.429	0.477	0.469	0.999
Multidomain	0.340	0.208–0.449	0.477	0.481	0.480	0.994

Converters (seven presymptomatic FTD-RisC mutation carriers that developed symptoms within 4 years after assessment) versus non-converters (35 FTD mutation carriers that remained cognitively healthy after 4 years). Multidomain represents a combination of all cognitive features. Min–max AUC values represent the variance across the 50 repeats. *Italic*: best-performing model.

AUC = area under the receiver operating characteristic curve; FTD-RisC = Frontotemporal dementia Risk Cohort; FWER = family-wise error rate.

not outperform chance level. None of the cognitive features predicted conversion within 4 years. However, it should be noted that our selection of neuropsychological tests was arbitrary, and higher performance might be possible with different tests. Also, longitudinal cognitive assessments may be more specific and predictive of conversion than cross-sectional differences (Jiskoot *et al.*, 2016, 2018b).

Important strengths of this study include the unique population and follow-up, which enabled the prediction of conversion in genetic FTD. Our clinical follow-up demonstrated that non-converters remained asymptomatic even after 4 (and in 25 cases after 6) years, ensuring that the labels ‘converter’ and ‘non-converter’ were truly separate. Furthermore, our methods have been validated in previous studies concerning classification in AD and FTD (Schouten *et al.*, 2016; Bouts *et al.*, 2018). The most important limitation to this study was sample size. Our index group consisted of seven converters, which is a small number in neuroimaging studies, and even more so in machine learning. This resulted in performance estimates with a large degree of uncertainty, which means that relatively high performances were necessary to significantly outperform chance level. For example, the AUC needed to significantly outperform chance level after family-wise error rate correction was 0.78 for the MRI analyses, and 0.74 for cognitive analyses. Another limitation to this study was the sample’s heterogeneity. We included subjects from *MAPT*, *GRN* and *C9orf72* families in order to boost our sample size, even though there is evidence that each mutation has its own pattern of neurodegeneration over time due to different underlying pathology (Seelaar *et al.*, 2011; Whitwell *et al.*, 2012; Mann and Snowden, 2017, Jiskoot *et al.*, 2018a). Indeed, converters with an *MAPT* mutation seemed to have higher conversion prediction scores than did converters with a *GRN* mutation, though this may also be due to sex differences. Stratification was not feasible due to our small sample size, but is necessary in future studies to show whether conversion prediction is possible in all FTD gene mutations carriers or solely in *MAPT* carriers. Just as different FTD mutations have different neurodegenerative profiles, so do the different clinical syndromes

that constitute FTD (Seelaar *et al.*, 2011; Whitwell *et al.*, 2012; Rohrer and Rosen, 2013). Five subjects developed bvFTD, while two developed nfvPPA, increasing heterogeneity in our analyses. Despite these sources of heterogeneity, the FA-based model predicted conversion in 4 years beyond chance level, which is an important proof of concept that conversion prediction is possible in genetic FTD. Larger sample sizes might in the future facilitate hierarchical (Kim *et al.*, 2019) or multilabel (Raamana *et al.*, 2014; Klöppel *et al.*, 2015; Koikkalainen *et al.*, 2016; Bron *et al.*, 2017; Canu *et al.*, 2017) prediction to deal with these sources of heterogeneity. While the acquisition of a larger, similar cohort is costly and time-consuming, our current results indicate that the acquisition of such data is meaningful and may lead to more accurate conversion prediction for genetic FTD. Lastly, the timing of conversion, and therefore the time between MRI and conversion, was based on heteroanamnesic information from knowledgeable informants, and as such may not be fully accurate. However, it should be noted that an exact time of conversion is near impossible to estimate in neurodegenerative diseases due to the gradual increase in symptomatology, and this inaccuracy is unlikely to have influenced our conclusions.

To conclude, we showed that FA predicted conversion within 4 years in presymptomatic FTD mutation carriers. This proof-of-concept study underlines the potential of MRI-based prediction in genetic FTD to contribute to a reliable early-stage FTD diagnosis, and should be replicated when larger sample sizes become available to corroborate our results.

Acknowledgements

Our appreciation goes out to all FTD-RisC participants and families.

Funding

The authors of this work were supported by the Leiden University Medical Centre MD/PhD Scholarship (to R.A.F.), ZonMw programme Memorabel project 733050103, EU

Joint Programme Neurodegenerative Disease Research (JPND) PreFrontAls consortium project 733051042 (to J.C.v.S.), and Dutch Research Council (NWO) VICI grant 016-130-667 (to S.A.R.B.R.). The views expressed are those of the authors and not necessarily those of the funding sources. The funding sources were not involved in the design of the study; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Competing interests

The authors report no competing interests.

References

- Agosta F, Scola E, Canu E, Marcone A, Magnani G, Sarro L, et al. White matter damage in frontotemporal lobar degeneration spectrum. *Cereb Cortex* 2012; 22: 2705–14.
- Anderson JLR, Jenkinson M, Smith SM. Non-linear registration aka Spatial normalisation. FMRIB Technical Report TR07JA2 [Internet]. 2007. <http://www.fmrib.ox.ac.uk/analysis/techrep/tr07ja2/tr07ja2.pdf> (23 July 2019, date last accessed).
- Army Individual Test Battery. Manual of directions and scoring. Washington, DC: War Dep Adj Gen Off; 1944.
- Beckmann CF, Smith SM. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging* 2004; 23: 137–52.
- Benussi A, Padovani A, Borroni B. Phenotypic heterogeneity of monogenic frontotemporal dementia. *Front Aging Neurosci* 2015; 7: 171.
- Bertrand A, Wen J, Rinaldi D, Houot M, Sayah S, Camuzat A, et al.; for the Predict to Prevent Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis (PREV-DEMALS) Study Group. Early cognitive, structural, and microstructural changes in presymptomatic C9orf72 carriers younger than 40 years. *JAMA Neurol* 2018; 75: 236–45.
- Borroni B, Alberici A, Premi E, Archetti S, Garibotto V, Agosti C, et al. Brain magnetic resonance imaging structural changes in a pedigree of asymptomatic progranulin mutation carriers. *Rejuvenation Res* 2008; 11: 585–95.
- Bouts M, Möller C, Hafkemeijer A, van Swieten JC, Dopfer EGP, van der Flier WM, et al. Single subject classification of Alzheimer's disease and behavioral variant frontotemporal dementia using anatomical, diffusion tensor, and resting-state functional magnetic resonance imaging. *J Alzheimer's Dis* 2018; 62: 1827–39.
- Bron EE, Smits M, Papma JM, Steketee RME, Meijboom R, de Groot M, et al. Multiparametric computer-aided differential diagnosis of Alzheimer's disease and frontotemporal dementia using structural and advanced MRI. *Eur Radiol* 2017; 27: 3372–82.
- Canu E, Agosta F, Mandic-Stojmenovic G, Stojković T, Stefanova E, Inuggi A, et al. Multiparametric MRI to distinguish early onset Alzheimer's disease and behavioral variant of frontotemporal dementia. *NeuroImage Clin* 2017; 15: 428–38.
- Cohn-Hokke PE, van Swieten JC, Pijnenburg YAL, Tibben A, Meijers-Heijboer H, Kievit A. The effect of predictive testing in adult-onset neurodegenerative diseases on social and personal life. *J Genet Counsel* 2018; 27: 947–54.
- Daianu M, Mendez MF, Baboyan VG, Jin Y, Melrose RJ, Jimenez EE, et al. An advanced white matter tract analysis in frontotemporal dementia and early-onset Alzheimer's disease. *Brain Imaging Behav* 2016; 10: 1038–53.
- Doesborgh SJC, van de Sandt-Koenderman WME, Dippel DWJ, van Harskamp F, Koudstaal PJ, Visch-Brink EG. Linguistic deficits in the acute phase of stroke. *J Neurol* 2003; 250: 977–82.
- Dopfer EGP, Rombouts SARB, Jiskoot LC, den Heijer T, de Graaf JRA, de Koning I, et al. Structural and functional brain connectivity in presymptomatic familial frontotemporal dementia. *Neurology* 2014; 83: e19–26.
- Ekman P, Friesen WV. Pictures of facial affect. Consult Psychol Press Palo Alto 1976.
- Elahi FM, Marx G, Cobigo Y, Staffaroni AM, Kornak J, Tosun D, et al. Longitudinal white matter change in frontotemporal dementia subtypes and sporadic late onset Alzheimer's disease. *NeuroImage Clin* 2017; 16: 595–603.
- Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett* 2006; 27: 861–74.
- Feis RA, Bouts MJRJ, de Vos F, Schouten TM, Panman JL, Jiskoot LC, et al. A multimodal MRI-based classification signature emerges just prior to symptom onset in frontotemporal dementia mutation carriers. *J Neurol Neurosurg Psychiatry* 2019b; 90: 1207–14.
- Feis RA, Bouts MJRJ, Panman JL, Jiskoot LC, Dopfer EGP, Schouten TM, et al. Single-subject classification of presymptomatic frontotemporal dementia mutation carriers using multimodal MRI. *NeuroImage Clin* 2019a; 22: 101718.
- Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189–98.
- Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 2008; 9: 432–41.
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010; 33: 1–22.
- Friston KJ, Ashburner J, Kiebel S, Nichols T, Penny WD. Statistical parametric mapping: the analysis of functional brain images. London: Elsevier/Academic Press; 2007.
- Galantucci S, Tartaglia MC, Wilson SM, Henry ML, Filippi M, Agosta F, et al. White matter damage in primary progressive aphasia: a diffusion tensor tractography study. *Brain* 2011; 134: 3011–29.
- Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. *Neurology* 2011; 76: 1006–14.
- Greve DN, Fischl B. Robust brain image alignment using boundary-based registration. *Neuroimage* 2009; 48: 63–72.
- Happé F, Brownell H, Winner E. Acquired 'theory of mind' impairments following stroke. *Cognition* 1999; 70: 211–40.
- Harvey RJ, Skelton RM, Rossor MN. The prevalence and causes of dementia in people under the age of 65 years. *J Neurol Neurosurg Psychiatry* 2003; 74: 1206–9.
- Hogan DB, Jetté N, Fiest KM, Roberts JI, Pearson D, Smith EE, et al. The prevalence and incidence of frontotemporal dementia: a systematic review. *Can J Neurol Sci* 2016; 43: S96–109.
- Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 2002; 17: 825–41.
- Jiskoot LC, Bocchetta M, Nicholas JM, Cash DM, Thomas D, Modat M, et al.; on behalf of the Genetic Frontotemporal dementia Initiative (GENFI). Presymptomatic white matter integrity loss in familial frontotemporal dementia in the GENFI cohort: a cross-sectional diffusion tensor imaging study. *Ann Clin Transl Neurol* 2018a; 5: 1025–36.
- Jiskoot LC, Dopfer EGP, den Heijer T, Timman R, van Minkelen R, van Swieten JC, et al. Presymptomatic cognitive decline in familial frontotemporal dementia: a longitudinal study. *Neurology* 2016; 87: 384–91.
- Jiskoot LC, Panman JL, Meeter LHH, Dopfer EGP, Donker Kaat L, Franzen S, et al. Longitudinal multimodal MRI as prognostic and diagnostic biomarker in presymptomatic familial frontotemporal dementia. *Brain* 2019; 142: 193–208.

- Jiskoot LC, Panman JL, van Asseldonk L, Franzen S, Meeter LHH, Donker Kaat L, et al. Longitudinal cognitive biomarkers predicting symptom onset in presymptomatic frontotemporal dementia. *J Neurol* 2018b; 265: 1381–92.
- Jolles J, Houx PJ, van Boxtel MPJ, Ponds R. Maastricht Aging Study: determinants of cognitive aging. Maastricht: Neuropsychology Publishers; 1995.
- Josephs KA, Petersen RC, Knopman DS, Boeve BF, Whitwell JL, Duffy JR, et al. Clinicopathologic analysis of frontotemporal and corticobasal degenerations and PSP. *Neurology* 2006; 66: 41–8.
- Kaplan E, Goodglass H, Weintraub S. The Boston naming test. Philadelphia: Lea Febiger; 1978.
- Kazemi K, Noorizadeh N. Quantitative comparison of SPM, FSL, and brainsuite for brain MR image segmentation. *J Biomed Phys Eng* 2014; 4: 13–26.
- Kim JP, Kim J, Park YH, Park SB, Lee JS, Yoo S, et al. Machine learning based hierarchical classification of frontotemporal dementia and Alzheimer's disease. *NeuroImage Clin* 2019; 23: 101811.
- Klöppel S, Peter J, Ludl A, Pilatus A, Maier S, Mader I, et al. Applying automated MR-based diagnostic methods to the memory clinic: a prospective study. *J Alzheimers Dis* 2015; 47: 939–54.
- Koikkalainen J, Rhodius-Meester H, Tolonen A, Barkhof F, Tijms B, Lemstra AW, et al. Differential diagnosis of neurodegenerative diseases using structural MRI data. *NeuroImage Clin* 2016; 11: 435–49.
- Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 2009; 12: 535–40.
- Lee SE, Sias AC, Mandelli ML, Brown JA, Brown AB, Khazenon AM, et al. Network degeneration and dysfunction in presymptomatic C9orf72 expansion carriers. *NeuroImage Clin* 2017; 14: 286–97.
- Leemans A, Jones DK. The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn Reson Med* 2009; 61: 1336–49.
- Lindeboom J, Schmand B, Tulner L, Walstra G, Jonker C. Visual association test to detect early dementia of the Alzheimer type. *J Neurol Neurosurg Psychiatry* 2002; 73: 126–33.
- Lomen-Hoerth C, Anderson T, Miller B. The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology* 2002; 59: 1077–9.
- Ludolph A, Drory V, Hardiman O, Nakano I, Ravits J, Robberecht W, et al.; for The WFN Research Group On ALS/MND. A revision of the El Escorial criteria—2015. *Amyotroph Lateral Scler Front Degener* 2015; 16: 291–2.
- Mahoney CJ, Ridgway GR, Malone IB, Downey LE, Beck J, Kinnunen KM, et al. Profiles of white matter tract pathology in frontotemporal dementia. *Hum Brain Mapp* 2014; 35: 4163–79.
- Mann DMA, Snowden JS. Frontotemporal lobar degeneration: Pathogenesis, pathology and pathways to phenotype. *Brain Pathol* 2017; 27: 723–36.
- Meeter LHH, Dopfer EGP, Jiskoot LC, Sanchez-Valle R, Graff C, Benussi L, et al. Neurofilament light chain: a biomarker for genetic frontotemporal dementia. *Ann Clin Transl Neurol* 2016; 3: 623–36.
- Meyer S, Mueller K, Stuke K, Bisenius S, Diehl-Schmid J, Jessen F, et al. Predicting behavioral variant frontotemporal dementia with pattern classification in multi-center structural MRI data. *NeuroImage Clin* 2017; 14: 656–62.
- Möller C, Hafkemeijer A, Pijnenburg YAL, Rombouts SAR, van der Grond J, Dopfer E, et al. Joint assessment of white matter integrity, cortical and subcortical atrophy to distinguish AD from behavioral variant FTD: A two-center study. *NeuroImage Clin* 2015; 9: 418–29.
- Nelson HE. A modified card sorting test sensitive to frontal lobe defects. *Cortex* 1976; 12: 313–24.
- Noirhomme Q, Lesenfans D, Gomez F, Soddu A, Schrouff J, Garraux G, et al. Biased binomial assessment of cross-validated estimation of classification accuracies illustrated in diagnosis predictions. *NeuroImage Clin* 2014; 4: 687–94.
- Omer T, Finegan E, Hutchinson S, Doherty M, Vajda A, McLaughlin RL, et al. Neuroimaging patterns along the ALS-FTD spectrum: a multiparametric imaging study. *Amyotroph Lateral Scler Front Degener* 2017; 18: 611–23.
- Papma JM, Jiskoot LC, Panman JL, Dopfer EGP, den Heijer T, Donker Kaat L, et al. Cognition and gray and white matter characteristics of presymptomatic C9orf72 repeat expansion. *Neurology* 2017; 89: 1256–64.
- Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 2011; 56: 907–22.
- Pievani M, Paternicò D, Benussi L, Binetti G, Orlandini A, Cobelli M, et al. Pattern of structural and functional brain abnormalities in asymptomatic granulin mutation carriers. *Alzheimer's Dement* 2014; 10: S354–63.
- Pruim RHR, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF. ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage* 2015; 112: 267–77.
- Raamana PR, Rosen HJ, Miller B, Weiner MW, Wang L, Beg MF. Three-class differential diagnosis among Alzheimer disease, frontotemporal dementia, and controls. *Front Neurol* 2014; 5: 71.
- Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioral variant of frontotemporal dementia. *Brain* 2011; 134: 2456–77.
- Ratnavalli E, Brayne C, Dawson K, Hodges JR. The prevalence of frontotemporal dementia. *Neurology* 2002; 58: 1615–21.
- Rey A. L'examen clinique en psychologie. Paris: Presses Universitaires de France; 1958.
- Riedijk SR, Niermeijer MFN, Dooijes D, Tibben AA. Decade of genetic counseling in frontotemporal dementia affected families: few counseling requests and much familial opposition to testing. *J Genet Counsel* 2009; 18: 350–6.
- Rohrer JD, Guerreiro R, Vandrovicova J, Uphill J, Reiman D, Beck J, et al. The heritability and genetics of frontotemporal lobar degeneration. *Neurology* 2009; 73: 1451–6.
- Rohrer JD, Nicholas JM, Cash DM, van Swieten JC, Dopfer EGP, Jiskoot L, et al. Presymptomatic cognitive and neuroanatomical changes in genetic frontotemporal dementia in the Genetic Frontotemporal dementia Initiative (GENFI) study: a cross-sectional analysis. *Lancet Neurol* 2015; 14: 253–62.
- Rohrer JD, Rosen HJ. Neuroimaging in frontotemporal dementia. *Int Rev Psychiatry* 2013; 25: 221–9.
- Royall DR, Cordes JA, Polk M. CLOX: an executive clock drawing task. *J Neurol Neurosurg Psychiatry* 1998; 64: 588–94.
- Schouten TM, Koini M, de Vos F, Seiler S, van der Grond J, Lechner A, et al. Combining anatomical, diffusion, and resting state functional magnetic resonance imaging for individual classification of mild and moderate Alzheimer's disease. *NeuroImage Clin* 2016; 11: 46–51.
- Seelaar H, Kamphorst W, Rosso SM, Azmani A, Masdjedi R, de Koning I, et al. Distinct genetic forms of frontotemporal dementia. *Neurology* 2008; 71: 1220–6.
- Seelaar H, Rohrer JD, Pijnenburg YAL, Fox NC, van Swieten JC. Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review. *J Neurol Neurosurg Psychiatry* 2011; 82: 476–86.
- Shmueli G. To explain or to predict? *Statist Sci* 2010; 25: 289–310.
- Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp* 2002; 17: 143–55.
- Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, et al. Resting-state fMRI in the human connectome project. *Neuroimage* 2013; 80: 144–68.
- Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 2006; 31: 1487–505.

- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 2004; 23: S208–19.
- Stroop JR. Studies of interference in serial verbal reactions. *J Exp Psychol* 1935; 18: 643–62.
- Thurstone LLT, Thurstone TG. Primary mental abilities. Chicago: Science Research Associates; 1962.
- Tsai RM, Boxer AL. Therapy and clinical trials in frontotemporal dementia: past, present, and future. *J Neurochem* 2016; 138: 211–21.
- Tu S, Leyton CE, Hodges JR, Piguet O, Hornberger M. Divergent longitudinal propagation of white matter degradation in logopenic and semantic variants of primary progressive aphasia. *J Alzheimers Dis* 2015; 49: 853–61.
- Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* 2010; 29: 1310–20.
- Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 2006; 7: 91.
- van Swieten JC, Heutink P. Mutations in progranulin (GRN) within the spectrum of clinical and pathological phenotypes of frontotemporal dementia. *Lancet Neurol* 2008; 7: 965–74.
- Visch-Brink E, Stronks D, Denes G. SAT: Semantische Associatie Test. Lisse: Swets & Zeitlinger; 2005.
- Wang J, Redmond SJ, Bertoux M, Hodges JR, Hornberger M. A comparison of magnetic resonance imaging and neuropsychological examination in the diagnostic distinction of Alzheimer's disease and behavioral variant frontotemporal dementia. *Front Aging Neurosci* 2016; 8: 119.
- Wechsler D. WAIS-III Nederlandse bewerking. Technische Handleiding. Lisse: Harcourt Test Publishers; 2005.
- Whitwell JL, Avula R, Senjem ML, Kantarci K, Weigand SD, Samikoglu A, et al. Gray and white matter water diffusion in the syndromic variants of frontotemporal dementia. *Neurology* 2010; 74: 1279–87.
- Whitwell JL, Weigand SD, Boeve BF, Senjem ML, Gunter JL, DeJesus-Hernandez M, et al. Neuroimaging signatures of frontotemporal dementia genetics: C9orf72, tau, progranulin and sporadics. *Brain* 2012; 135: 794–806.
- Zhang Y, Schuff N, Ching C, Tosun D, Zhan W, Nezamzadeh M, et al. Joint assessment of structural, perfusion, and diffusion MRI in Alzheimer's disease and frontotemporal dementia. *Int J Alzheimers Dis* 2011; 2011: 1–11.
- Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Royal Statistical Soc B* 2005; 67: 301–20.