

VentRa. Validation study of the ventricle feature estimation and classification tool to differentiate behavioral variant frontotemporal dementia from psychiatric disorders and other degenerative diseases

Ana L. Manera¹, M.D., Mahsa Dadar^{1,2}, PhD,
Simon Ducharme^{1,2}, M.D. MSc*, D. Louis Collins¹, PhD.*

¹ McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Quebec (QC), Canada.

² Department of Psychiatry, Douglas Mental Health University Health Centre, McGill University, Montreal, Quebec (QC), Canada.

Corresponding Author Information:

Ana L. Manera, Montreal Neurological Institute, 3801 University Street, Room WB320, Montréal, QC, H3A 2B4. Email: ana.manera@mail.mcgill.ca

ABSTRACT

Introduction: Lateral ventricles are reliable and sensitive indicators of brain atrophy and disease progression in behavioral variant frontotemporal dementia (bvFTD). Here we validate our previously developed automated tool using ventricular features (known as VentRa) for the classification of bvFTD vs a mixed cohort of neurodegenerative, vascular, and psychiatric disorders from a clinically representative independent dataset.

Methods: Lateral ventricles were segmented for 1110 subjects - 14 bvFTD, 30 other Frontotemporal Dementia (FTD), 70 Lewy Body Disease (LBD), 898 Alzheimer Disease (AD), 62 Vascular Brain Injury (VBI) and 36 Primary Psychiatric Disorder (PPD) from the publicly accessible National Alzheimer's Coordinating Center dataset to assess the performance of VentRa.

Results: Using ventricular features to discriminate bvFTD subjects from PPD, VentRa achieved an accuracy of 84%, 71% sensitivity and 89% specificity. VentRa was able to identify bvFTD from a mixed age-matched cohort (i.e., Other FTD, LBD, AD, VBI and PPD) and to correctly classify other disorders as 'not compatible with bvFTD' with a specificity of 83%. The specificity against each of the other individual cohorts were 80% for other FTD, 83% for LBD, 83% for AD and 84% for VBI.

Discussion: VentRa is a robust and generalizable tool with potential usefulness for improving the diagnostic certainty of bvFTD, particularly for the differential diagnosis with PPD.

KEYWORDS

Frontotemporal dementia – Dementia – Brain Imaging – Lateral ventricles – Classification- Psychiatric disorder

ABBREVIATIONS

bvFTD= Behavioral Variant Frontotemporal Dementia

PPD= Primary Psychiatric Disorder

DBM= deformation-based morphometry

CN= normal controls

APR= anteroposterior ratio

MCI= mild cognitive impairment

AD= Alzheimer's dementia

FTD= Frontotemporal Dementia

TVV= total ventricular volume

LRFR= left-right frontal ratio

LRTR= left-right temporal ratio

NACC= National Alzheimer's Coordinating Center dataset

LBD= Lewy Body Disease

VBI= Vascular Brain Injury

T1w= T1-weighted (T1w)

TFE= Turbo Field Echo

FSE= Fast spin echo

FSPGR= fast spoiled gradient-echo

MPRAGE= Magnetization Prepared - Rapid Gradient Echo

MMSE: Mini-Mental Status Examination

CDR: Clinical Dementia Rating

CDR-SB: Clinical Dementia Rating-Sum of Boxes.

LR= Likelihood ratios

FN= False negatives

TP=True positives

FP=False positives

ROI= Region of Interest

1. INTRODUCTION

A confirmed behavioural variant frontotemporal dementia (bvFTD) diagnosis is often difficult to achieve in the clinic and heavily relies on brain imaging. This is explained in part by the high heterogeneity in heritability, pathology, and the absence of a unique molecular biomarker. In addition, the symptomatic overlap with other neurodegenerative disorders, but more importantly with non-degenerative primary psychiatric disorders (PPD) - including major depressive disorder, bipolar disorder, schizophrenia, obsessive-compulsive disorder, autism spectrum disorders and even personality disorders¹ - means that PPD often constitutes the main differential diagnosis of bvFTD². It has been reported that approximately 50% of bvFTD patients receive a prior psychiatric diagnosis (most frequently major depression), and the average diagnostic delay is up to 5-6 years from symptom onset^{1,3}. It has also been shown that patients with PPD can be wrongly diagnosed with bvFTD, particularly in community settings, preventing patients to have access to evidence-based psychiatric treatments⁴.

In previous work, using deformation-based morphometry (DBM), we showed that the ventricles are a robust metric to discriminate bvFTD from cognitively normal controls (CN)⁵. The lateral ventricles exhibited the most significant volumetric difference at baseline between bvFTD and CN and, of all structures, they showed the most significant progression of change in 1-year follow up.

In more recent work, we assessed the differences in ventricular features between bvFTD, normal controls, and other dementia cohorts and developed VentRa, the first tool to use ventricular features specifically for the differential diagnosis of bvFTD. These features included total ventricular volume (TVV), left-right frontal ventricle volume ratio (LRFR), left-right temporal ventricle volume ratio (LRTR) and the anterior-posterior ratio (APR). Using VentRa, we achieved a 10-fold cross-validation accuracy of 80% (76% sensitivity, 83% specificity) in differentiating bvFTD from all other cohorts that included normal aging, mild cognitive impairment (MCI), Alzheimer's dementia (AD), and other Frontotemporal Dementia (FTD) variants. Interestingly, using only the APR ventricular feature yields a 76% accuracy, underlying its importance in differentiating between groups⁶. The lateral ventricles are easy to reliably segment manually or when using standard publicly available tools^{7,8} and provide an estimate of the overall extent

of brain atrophy across different regions, making the VentRa ventricle-based features for bvFTD diagnosis a promising practical tool both in research and clinical settings.

In the present study, we aimed to validate VentRa on an independent cohort obtained from the National Alzheimer's Coordinating Center dataset (NACC) to differentiate bvFTD subjects from other FTD, Lewy Body Disease (LBD), AD, Vascular Brain Injury (VBI) and PPD. While the key differential diagnostic is between bvFTD variants and PPD, we assessed the performance of VentRa in data from patients with other neurodegenerative dementias (other FTD subtypes, AD, LBD, and VBD) to verify that VentRa can detect atrophy patterns specific to bvFTD and not just any neurodegenerative disease related pattern of atrophy that might also occur in other dementias. The publicly available NACC dataset was selected as the target dataset for independent validation of VentRa since it contains a large number of multi-center and multi-scanner datasets with T1-weighted images acquired with different acquisition protocols, field of strength, resolution, and field of view, allowing us to validate the generalizability of VentRa in realistic clinical settings.

2. MATERIALS AND METHODS

2.1 Participants

The data sample used for this study was obtained from the National Alzheimer's Coordinating Center's Uniform Data Set (NACC). The NACC developed a database of standardized clinical research data obtained from past and present National Institute on Aging (NIA)-funded Alzheimer's Disease Research Centers across the United States. NACC data collection has been previously described elsewhere⁹. Participants included in the database have a range of cognitive statuses: normal cognition, MCI, and demented due to various clinical diagnoses. A total of 5246 scans were obtained from NACC from the September 2019 data freeze after a data request for all participants with any dementia syndrome (i.e., bvFTD, AD, VBI, LBD and other), MCI and cognitively normal subjects who have at least one T1 weighted (T1-W) scan \pm 6 months from a clinical visit [NACC variables MRIT1=1 and NACCMRDY

between (-180 and 180)]. Figure 1 shows the flowchart for the detailed subject selection procedure. The analysis reported here used data from 1110 subjects from the six diagnostic groups of interest – bvFTD (n=14), other FTD (n=30), LBD (n=70), AD (n=898), VBI (n=62), PPD (n=36) – that passed quality control and had an available scan matching a clinical visit <12 months apart. The PPD group is composed of subjects with different psychiatric disorders: depression (n=28), bipolar disorder (n=2), anxiety disorder (n=2) and other psychiatric diseases (n=5). Of note, subjects that did not have any cognitive and/or behavioral impairment (n=51) were excluded from this samples according to the following NACC variables: 1) NACCCOGF= 0 (No impairment in cognition), 2) COGMODE= 0 (No impairment in cognition), 3) NACCBEHF= 0 (No behavioral symptoms, for the bvFTD cohort), and 4) COMPORT= 0.0 (No impairment in behavior, comportment, and personality, for the bvFTD cohort).

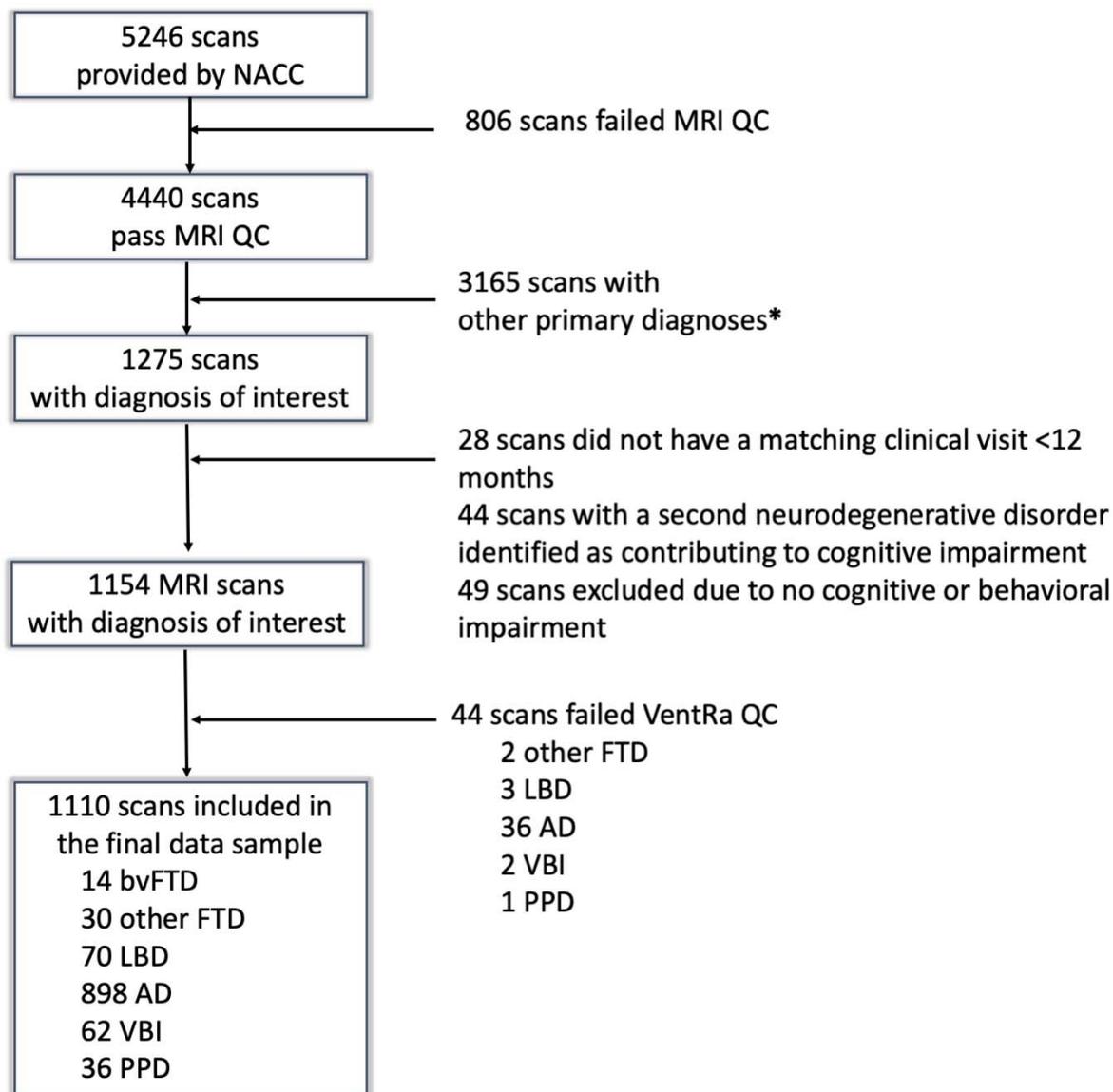


Figure 1. Flowchart showing the subject selection procedure. Abbreviations: bvFTD: behavioral variant frontotemporal dementia; FTD: Frontotemporal dementia; LBD: Lewy Body Disease; AD: Alzheimer’s dementia; VBI: Vascular Brain Injury; PPD: Primary Psychiatric Disorder. * 144 scans were excluded due to missing diagnosis, 1931 due to no cognitive impairment and, 1090 due to primary diagnoses other than the groups of interest (i.e., progressive supranuclear palsy, corticobasal degeneration, prion disease, traumatic brain injury, normal-pressure hydrocephalus, central nervous system neoplasm and other neurologic, genetic, or infectious condition).

2.2 MRI acquisition parameters

Out of all the participants selected for the final data sample, there were 529 3.0T and 449 1.5T scans acquired by either GE (N = 575), SIEMENS (N = 189) or Philips (N = 81) scanners (no data on the field strength was available for 132 scans). The scans were acquired in sagittal, coronal, and axial planes from either T1-weighted (T1w) Turbo Field Echo (TFE), Fast spin echo (FSE), fast spoiled gradient-echo (FSPGR) or Magnetization Prepared - Rapid Gradient Echo (MPRAGE) sequences. Other acquisition parameters were echo time range: 0.0013-0.02; repetition time range: 0.006-3.00; number of slices in x range: 124-512; number of slices in y range: 54-512; and number of slices in z range: 29-512; slice thickness range: 0.8 – 5; voxel size x range: 0.4-1.2; voxel size y range: 0.4-3; and voxel size z range: 0.5-7.

2.3 Statistical analysis

All statistical analyses were conducted using MATLAB (version R2021b). Differences in categorical variables between the cohorts were assessed using chi-square and continuous variables were assessed using one-way ANOVA or Kruskal-Wallis variance analysis depending on the distribution of the variables based on a normality test. Post-hoc two-sample Student's t-tests were conducted to examine clinical and ventricular differences. Results are expressed as mean \pm standard deviation and median [interquartile range] as appropriate and *p* values for all the volumetric analyses have been corrected for multiple comparisons using the Bonferroni method so that any reported adjusted *p*-value < 0.05 is statistically significant.

2.4 VentRa tool

VentRa takes as input a comma separated (.csv) file providing the path for the raw T1-weighted images as well as age and sex of the subjects, and provides as output preprocessed images along with ventricle

segmentations, QC files for the segmentations, as well as a .csv file including the diagnosis (based on the classifier trained on bvFTD vs the mixed group data) along with all the extracted ventricle features: i.e. total ventricle volume, ventricle volumes in each lobe and hemisphere, antero-posterior and left/right ratios. VentRa performs the classification task by means of a support vector machine classifier (fitsvm function from MATLAB, with default parameters: linear kernel, Sequential Minimal Optimization) using ventricular features (APR, TVV, LRF and LRTR) in combination with age and sex. This classifier was originally trained on a separate dataset⁶ and it thus tested here, unchanged, on an independent dataset.

2.4.1 Ventricular feature estimation and analyses

VentRa includes a previously validated patch-based label fusion technique to segment the lateral ventricles⁸. The method uses expert manual segmentations as priors and estimates the label of each test subject voxel by comparing its surrounding 3D voxel patch against all the patches from the training library and performing a weighted label fusion using the intensity-based distances between the patch under study and the patches in the training subjects. All resulting segmentations were visually assessed by an experienced rater blind to the clinical diagnosis, and the incomplete/inaccurate segmentations (N=44 or 3.96%) were excluded. Using the Hammers brain atlas delineating frontal, parietal, temporal, and occipital lobes separately in the left and right hemispheres^{10 11}, lateral ventricular volumes were calculated per each brain lobe and hemisphere. Using coronal coordinate y=-12mm in the stereotaxic space (i.e., registered to the template) ventricles were divided into anterior and posterior portions to obtain the ventricular APR⁶. All volumes were normalized for intracranial volume and these ratios were log-transformed to achieve normal distribution.

3. RESULTS

3.1 Demographics

Table 1 shows the demographic and cognitive testing performances for all the cohorts. The bvFTD subjects were younger than LBD, AD and VBI subjects, and the age difference was not significant between bvFTD, other FTD and PPD. There were no significant differences in years of cognitive impairment onset between the groups. There were no significant differences between bvFTD and all the other cohorts in Mini-Mental Status Examination (MMSE) ($p > 0.05$). bvFTD subjects were slightly more severely impaired than AD according to Clinical Dementia Rating-Sum of Boxes (CDR-SB) ($p = 0.05$) and significantly more impaired than other FTD ($p = 0.01$), VBI ($p < 0.001$) and PPD ($p < 0.001$). However, on average they remained mild cases (CDR 1.0) of bvFTD and at a stage relevant for differential diagnosis. Other cognitive scores can be found in Table 1.

	bvFTD (N=16)	Other FTD (N=30)	LBD (N=70)	AD (N=898)	VBI (N=62)	PPD (N=36)	<i>p</i> value
<i>Age, y</i>	62±7	66±10	74±7	75±9	79±8	69±10	<0.001
<i>Sex, male (%)</i>	9 (56%)	13 (43%)	59 (84%)	441 (49%)	39 (63%)	19 (53%)	<0.001
<i>Years from cognitive impairment onset, y</i>	2.8 [2.3-6.3]	3 [2.3-5.5]	4.6 [3.3-6.8]	4.3 [2.8-6.4]	3.9 [2.1-6]	-	0.08
<i>MMSE (n= 860)</i>	24 [21-28]	25 [21-28]	26 [20-28]	24 [20-27]	27 [24-28]	28 [26-29]	<0.001
<i>CDR Global score (n=1110)</i>	1 [0.5-1]	0.5 [0.5-1]	0.5 [0.5-1]	0.5 [0.5-1]	0.5 [0.5-1]	0.5 [0.5-0.5]	<0.001
<i>CDR-SB (n=1110)</i>	5.25 [4.5-9]	2.75 [1-4.5]	3 [2-7]	3.5 [1.5-5]	1.5 [1-2.5]	1[0.5-2]	<0.001
<i>TRAIL-A time, s (n=1041)</i>	37 [23-64]	45 [33-63]	56 [40-85]	44 [32-66]	55 [41-75]	38 [31-54]	<0.001
<i>TRAIL-B time, s (n=902)</i>	76 [66-90]	111 [81-179]	171 [111-300]	149 [99 -266]	204 [116-292]	103 [64-170]	<0.001
<i>Digits forward correct trials (n=852)</i>	6 [5 -9]	6 [4-6]	8 [7-9]	7 [6-9]	7 [6-8]	8 [6-10]	0.01
<i>Digits backwards (n=850)</i>	5 [4-7]	5 [4-6]	4 [3-5]	5 [4-6]	5 [4-6]	6 [4-7]	0.04
<i>WAIS (n=767)</i>	43 [29-46]	38 [24-48]	26 [21-35]	32 [22-42]	29 [19-34]	35 [27-45]	0.002
<i>BOSTON (n=443)</i>	20 [17-23]	22 [11-29]	24 [21-26]	22 [16-25]	23 [18-26]	26 [21-27]	0.02
<i>Semantic Fluency – Animals (n=1081)</i>	11 [4-13]	11 [6-18]	12 [7-16]	12 [9-16]	13 [11-16]	16 [13-19]	<0.001
<i>Semantic Fluency – Vegetables (n=1078)</i>	6 [3-8]	8 [4-12]	7 [1-10]	8 [5-11]	10 [7-11]	12 [9-16]	<0.001

Table 1. Demographic and clinical characteristics for all the cohorts. Values express Mean± SD / Median [interquartile range]. P value level of significance: 0.05. Abbreviations: bvFTD: behavioral variant frontotemporal dementia; FTD: Frontotemporal dementia; LBD: Lewy Body Disease; AD: Alzheimer’s dementia; VBI: Vascular Brain Injury; PPD: Primary Psychiatric Disorder; MMSE: Mini-Mental Status Examination; CDR: Clinical Dementia Rating; CDR-SB: Clinical Dementia Rating-Sum of Boxes. The bold values represent significant differences with the bvFTD cohort.

3.2 Classification task with VentRa.

3.2.1 bvFTD vs PPD

VentRa achieved an accuracy of 84%, 71% sensitivity and 89% specificity in discriminating bvFTD subjects from PPD based on ventricular features (Figure 2). Positive and negative likelihood ratios (LR) were 6.43 and 0.31 respectively.

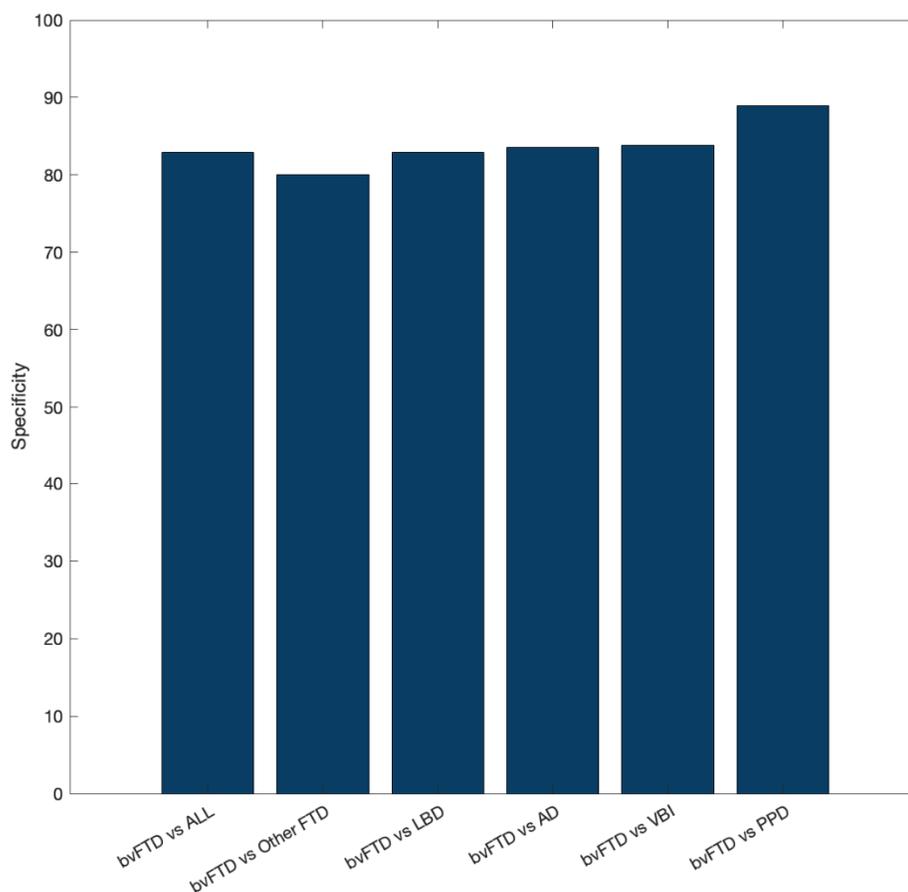


Figure 2. Specificity for the different classification tasks. Abbreviations: bvFTD: behavioral variant frontotemporal dementia; FTD: Frontotemporal dementia; LBD: Lewy Body Disease; AD: Alzheimer’s dementia; VBI: Vascular Brain Injury; PPD: Primary Psychiatric Disorder

The classification using VentRa resulted in 29% rate of false negatives (FN). As shown in Figure 3, these subjects were older than the bvFTD subjects correctly classified (63 ± 8 vs 72 ± 3 years respectively,

$p=0.06$) and they had significantly smaller ventricular APR ($p=0.01$). The FN also performed significantly better than the true positives (TP) on MMSE ($p=0.02$). Yet, the disease duration was not significantly different than true positives ($p=0.2$) nor were the severity according to CDR global score (Median CDR Global_{TP}= 1 and Median CDR Global_{FN}= 0.75, $p=0.2$) and CDR-SB (Median CDR-SB_{TP}= 6.5 and Median CDR-SB_{FN}=5, $p=0.2$). Finally, no significant differences were found in severity of behavior, compartment, and personality impairment ($p=0.4$) according to NACC variable ‘COMPORT’ where 0= No impairment; 0.5= Questionable impairment; 1= Mild impairment; 1.5= Moderate impairment; and 2= Severe impairment).

On the other hand, the false positive (FP) rate was 11%. When comparing the FP against the PPD subjects correctly diagnosed as not compatible with bvFTD (Figure 4), no significant differences were found in age ($p=0.9$), MMSE ($p=0.8$), CDR ($p=0.4$), CDR-SB ($p=0.2$) or severity of behavioral impairment (‘COMPORT’, $p=0.33$). However, the FP showed statistically significant larger APR compared to subjects correctly classified (1.26 ± 0.3 and 0.97 ± 0.3 respectively, $p=0.04$).

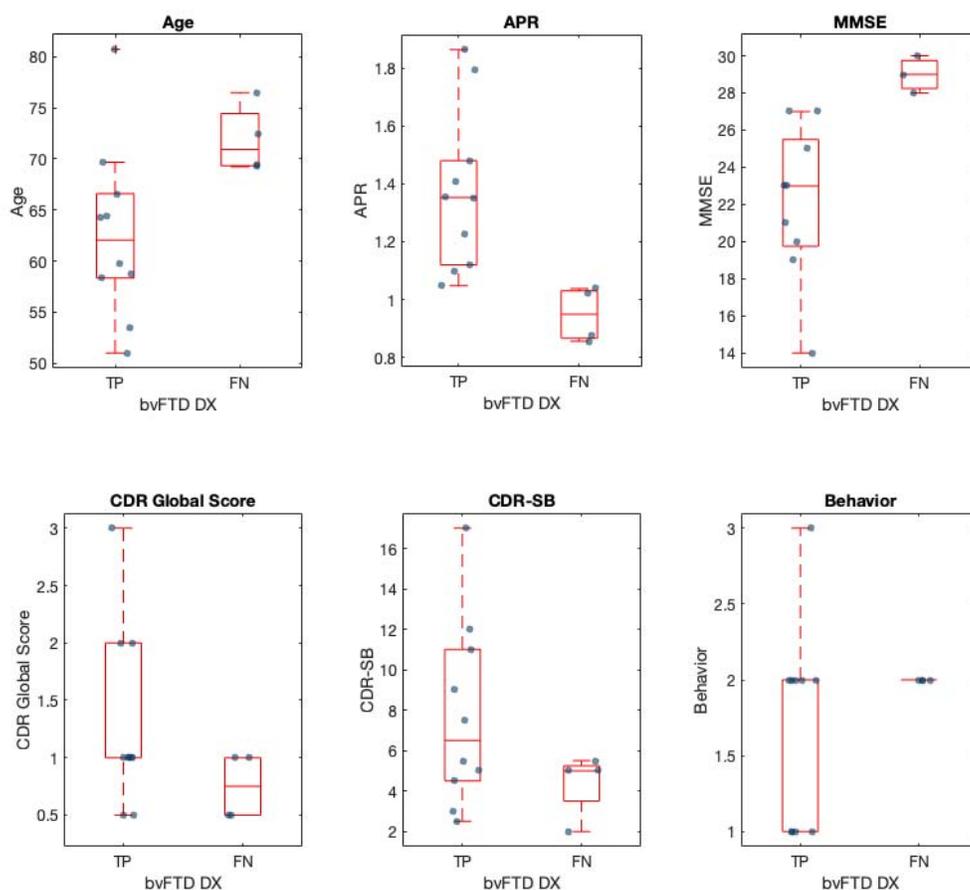


Figure 3. Boxplots showing the differences between bvFTD subjects correctly classified with VentRa (True Positives) and those classified as not compatible with bvFTD (False Negatives). Abbreviations: bvFTD: behavioral variant frontotemporal dementia; TP= True Positives; FN: False Negatives; APR: anteroposterior ratio; MMSE: Mini-Mental Status Examination; CDR: Clinical Dementia Rating; CDR-SB: Clinical Dementia Rating-Sum of Boxes. Behavior refers to a NACC variable ‘COMPORT (behavior, compartment, and personality impairment) where 0= No impairment; 0.5= Questionable impairment; 1= Mild impairment; 1.5= Moderate impairment; and 2= Severe impairment.

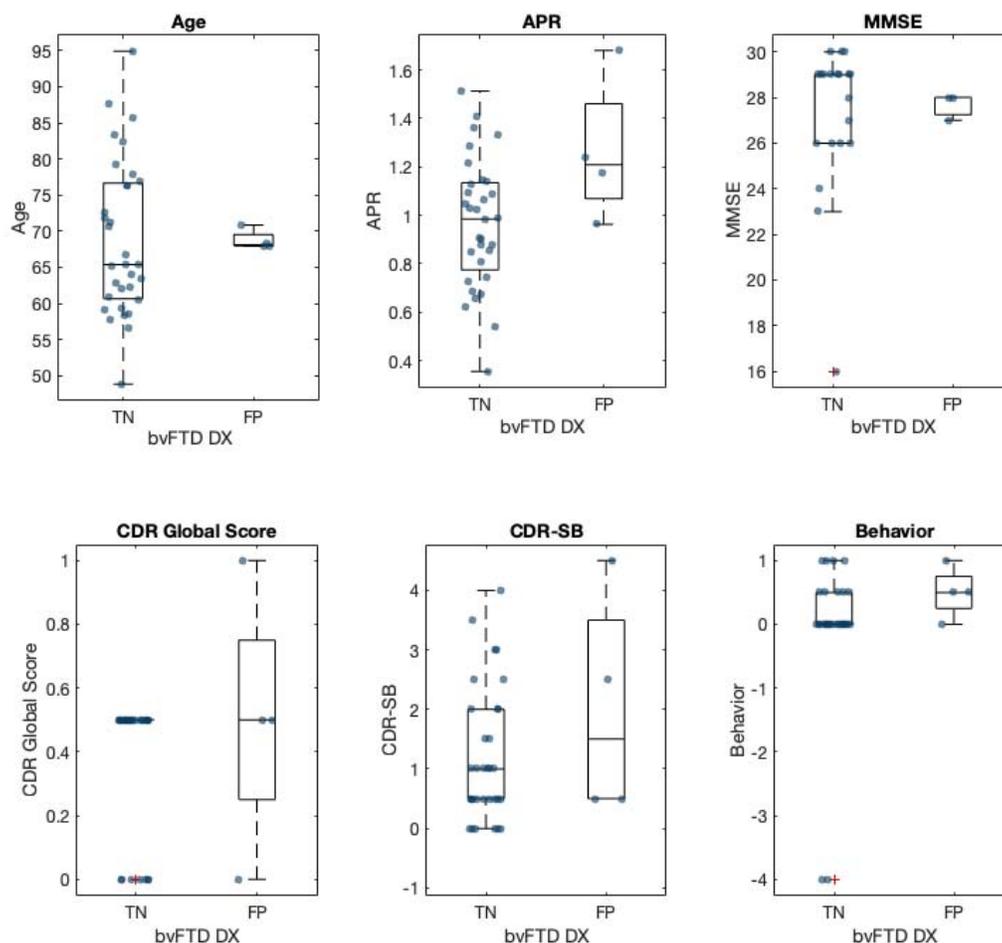


Figure 4. Boxplots showing the differences between PPD subjects incorrectly classified as bvFTD with VentRa (False Positives) and those classified as not compatible with bvFTD (True Negatives). Abbreviations: bvFTD: behavioral variant frontotemporal dementia; PPD: Primary psychiatric disorder; TN= True Negatives; FN: False Positives; APR: anteroposterior ratio; MMSE: Mini-Mental Status Examination; CDR: Clinical Dementia Rating; CDR-SB: Clinical Dementia Rating-Sum of Boxes. Behavior refers to a NACC variable ‘COMPORT (behavior, compartment, and personality impairment) where 0= No impairment; 0.5= Questionable impairment; 1= Mild impairment; 1.5= Moderate impairment; and 2= Severe impairment.

3.2.2 bvFTD vs other neurodegenerative cohorts.

Using ventricular features to identify bvFTD from a mixed age-matched cohort ('ALL'= Other FTD, LBD, AD, VBI and PPD) VentRa was able to correctly classify other disorders as 'not compatible with bvFTD' with a specificity of 83%. The specificity against each of the other individual cohorts were 80% for other FTD, 83% for LBD, 83% for AD and 84% for VBI (Figure 2)

4. DISCUSSION

In the present study, we aimed to validate our previously developed automated tool VentRa⁶ for the estimation of ventricular features on T1w MRI for the individual prediction of bvFTD within a mixed cohort of subjects from an independent dataset of clinically acquired MRIs. The main findings were: 1) VentRa was able to accurately differentiate bvFTD from PPD (84% accuracy, 71% sensitivity, 89% specificity); 2) VentRa was able to identify bvFTD from a mixed cohort with a specificity of 83%; 3) From each of the other individual cohorts (other FTD, LBD, AD and VBI), VentRa correctly identified between 80-84% as not compatible with bvFTD; 4) Our results show good positive and negative likelihood ratios proving its reliability to rule in and out bvFTD among the wider group of patients exhibiting behavioral changes such as apathy, disinhibition, and loss of empathy¹².

While current clinical diagnostic criteria for the different neurodegenerative disorders in combination with several available pathologic biomarkers (i.e. AD biomarkers in blood/cerebrospinal fluid and Positron emission tomography) perform reasonably well, the main clinical challenge in the differential diagnosis of bvFTD is still represented by PPD. Clinical criteria for bvFTD are less helpful to distinguish it from PPD^{2 13}. No bedside cognitive screening has yet been able to clearly discriminate bvFTD from PPD. Neuropsychological assessment frequently focuses more on the executive dysfunction characteristic of bvFTD though it is highly nonspecific for FTD versus PPDs¹³. Despite suggestions of its efficacy, dedicated assessment of social cognition is often not included in standard neuropsychological assessment batteries^{13 14}. Furthermore, several social cognitive instruments that have been developed for research purposes are yet to have normative data available¹⁵.

Neuroimaging is already crucial for bvFTD diagnosis; frontal and/or anterior temporal atrophy or hypometabolism increases certainty from possible to probable bvFTD in current criteria¹⁶. Yet, MRIs may not be obviously abnormal in early stages or in certain variants of bvFTD, such as the C9orf72 phenotype^{13 17}. In the Late Onset Frontal Lobe Syndrome (LOF) study, the standard visual review of structural MRI was found to be useful for the differential diagnosis of bvFTD versus PPD, however only had 70% sensitivity^{18 19}. It is hypothesized that volumetric MRI could improve this sensitivity, however there is insufficient evidence currently to formally recommend volumetric analytic techniques in clinical populations with behavioral changes¹⁵. Furthermore, while some machine learning algorithms using MRI volumetry perform well against controls and AD cohorts, less is known about the performance against PPD. Zhutovsky et. al. performed a classification of bvFTD vs PPD with different data usage approaches including clinical, MRI voxel-wise, MRI-Region of Interest (ROI), Clinical + MRI ROI, and Clinical + MRI voxel-wise techniques²⁰. In their study, a support vector machine was trained on a sample of 73 patients (18 bvFTD, 28 neurological and 27 psychiatric). Zhutovsky et. al. report cross-validation accuracies of 82.1% (76.7% sensitivity, 87.5% specificity) and 78.3% (75.9% sensitivity, 80.8% specificity) using only MRI voxel-wise and MRI ROI approaches respectively. Cross-validation can overestimate performance as the same cohort is used to select features and optimize the method's hyperparameters. In contrast, our validation using *independent* data, our method generalizes well as we achieved 84% accuracy, 71% sensitivity, 89% specificity in discriminating bvFTD subjects from PPD with VentRa. This accuracy with higher specificity relative to sensitivity is in line with reported accuracy from expert neuroradiological visual review reported in the LOF cohort¹⁸. Although the performance of VentRa was not superior, the fact that it is completely automated and performs as well as academic centre experts could improve clinical practice particularly in community areas. Further, VentRa has the potential to improve over time by including more training data.

Still, our results show a FN rate of 29%. This is likely because these FN bvFTD cases are older than the TP cases and outside of the age range of bvFTD subjects from the Frontotemporal Lobar Degeneration Neuroimaging Initiative (FTLDNI) that were originally used to train VentRa (Age ranges: $\text{bvFTD}_{\text{NIFD}} 61 \pm 6$ y, $\text{TP}_{\text{NACC}} 63 \pm 8$ y, and $\text{FN}_{\text{NACC}} 72 \pm 3$ y). Since these subjects were outside the operating range of the classifier, they were therefore more likely to be misclassified, thus underlying the importance of having training data that well represents the population of interest. It is highly probable that adding extended ages to the training dataset will likely improve the results. Not only were the FN older, but also they had better cognition (MMSE, CDR Global and CDR-SB) and their ventricular APR was smaller, we might therefore argue that there could be a mismatch between the clinical population included in NACC (including some older and milder cases, possibly with more ambiguous diagnoses) and the FTLDNI data (more advanced and clear cut cases) used to train VentRa. We also need to consider the fact that in older age FTD there is a higher prevalence of mixed pathologic diagnoses²¹, and therefore a classification algorithm is unlikely perform as well in this population.

Our study has two main strengths. First, VentRa has been trained and cross-validated on a sample composed from two different multi-center databases (FTLDNI²² and Alzheimer's Disease Neuroimaging Initiative- ANDI-²³) and its performance was tested here on a held-out database which also included multi-center and multi-scanner data from different scanner models of both 1.5T and 3T field strengths (NACC). In addition, the failure rate of the VentRa QC was very low (4%), particularly considering the variability of the MRI acquisition parameters (i.e., slice thickness range: 0.8 – 5; voxel size x range: 0.4-1.2; voxel size y range: 0.4-3; and voxel size z range: 0.5-7). This further reinforces the generalizability (i.e., external validity) of our results and ensures their applicability in a clinical scenario with different scanners and with different magnetic field strengths. This raises the confidence that VentRa would work reliably in clinical settings. Secondly, our tool is based on standard structural T1w MRI and uses a combination of easily measurable features of the lateral ventricles which can be reliably segmented using a variety of publicly available tools such as FreeSurfer in addition to the patch-based method used in

VentRa^{7 8}. Further, all the image processing tools used in this study have been established and validated for use in multi-center and multi-scanner datasets^{11 24-28}, and have been designed to minimize such center- and scanner-related differences. The ventricular features are much less likely to be impacted by such differences, particularly for the APR, where any such discrepancies would be minimized and perhaps cancelled in the ratio.

Our study is not without limitations. The number of bvFTD subjects with available and good quality scans was relatively limited. In the future, it will be important to evaluate the robustness and generalizability of our results in other bvFTD cohorts. In addition, it will be valuable to see how well VentRa would perform against a more heterogenous PPD cohort and to predict future course in patients with clinically ambiguous FTD presentations (i.e., predict who has a neurodegenerative underlying pathology versus who has PPD). Since we have made both the image processing pipeline and classification model publicly available with VentRa (<http://nist.mni.mcgill.ca/VentRa/>), this can be easily achieved by others who might have access to such datasets.

To conclude, these results prove the robustness and generalizability of VentRa, indicating the potential usefulness of this automated MRI-based tool for improving the diagnostic certainty of bvFTD. If validated in a prospective larger sample, VentRa has the potential to improve diagnostic accuracy, particularly against populations with behavioral and/or psychiatric symptoms enabling reduced diagnostic and intervention time, particularly in settings with less access to specialized neuroradiology.

Competing interests:

Dr. Manera reports no disclosures

Dr. Dadar reports no disclosures

Dr. Collins is co-founder of True Positive Medical Devices.

Dr. Ducharme receives salary funding from the Fonds de Recherche du Québec - Santé. Dr. Ducharme is the co-founder of Arctic Fox AI.

Acknowledgements

MD reports receiving research funding from the Healthy Brains for Healthy Lives (HBHL), Alzheimer Society Research Program (ASRP), and Douglas Research Centre (DRC).

DLC reports receiving research funding from Canadian Institutes of Health research, the Canadian National Science and Engineering Research Council, Brain Canada, the Weston Foundation, and the Famille Louise & André Charron.

SD receives salary funding from the Fonds de Recherche du Québec – Santé (FRQS)

The NACC database is funded by NIA/NIH Grant U24 AG072122. NACC data are contributed by the NIA-funded ADCs: P50 AG005131 (PI James Brewer, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P50 AG005138 (PI Mary Sano, PhD), P50 AG005142 (PI Helena Chui, MD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005681 (PI John Morris, MD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG008051 (PI Thomas Wisniewski, MD), P50 AG008702 (PI Scott Small, MD), P30 AG010124 (PI John Trojanowski, MD, PhD), P30 AG010129 (PI Charles DeCarli, MD), P30 AG010133 (PI Andrew Saykin, PsyD), P30 AG010161 (PI David Bennett, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG013846 (PI Neil Kowall, MD), P30 AG013854 (PI Robert Vassar, PhD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P30 AG019610 (PI Eric Reiman, MD), P50 AG023501 (PI Bruce Miller, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P30 AG028383 (PI Linda Van Eldik, PhD), P50 AG033514 (PI Sanjay

Asthana, MD, FRCP), P30 AG035982 (PI Russell Swerdlow, MD), P50 AG047266 (PI Todd Golde, MD, PhD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG049638 (PI Suzanne Craft, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG066546 (PI Sudha Seshadri, MD), P20 AG068024 (PI Erik Roberson, MD, PhD), P20 AG068053 (PI Marwan Sabbagh, MD), P20 AG068077 (PI Gary Rosenberg, MD), P20 AG068082 (PI Angela Jefferson, PhD), P30 AG072958 (PI Heather Whitson, MD), P30 AG072959 (PI James Leverenz, MD)

REFERENCES

1. Ducharme S, Price BH, Larvie M, et al. Clinical Approach to the Differential Diagnosis Between Behavioral Variant Frontotemporal Dementia and Primary Psychiatric Disorders. *Am J Psychiatry* 2015;172(9):827-37. doi: 10.1176/appi.ajp.2015.14101248 [published Online First: 2015/09/02]
2. Krudop WA, Dols A, Kerssens CJ, et al. The Pitfall of Behavioral Variant Frontotemporal Dementia Mimics Despite Multidisciplinary Application of the FTDC Criteria. *J Alzheimers Dis* 2017;60(3):959-75. doi: 10.3233/JAD-170608
3. van Vliet D, de Vugt ME, Bakker C, et al. Time to diagnosis in young-onset dementia as compared with late-onset dementia. *Psychol Med* 2013;43(2):423-32. doi: 10.1017/S0033291712001122 [published Online First: 20120528]
4. Ducharme S, Pearl-Dowler L, Gossink F, et al. The Frontotemporal Dementia versus Primary Psychiatric Disorder (FTD versus PPD) Checklist: A Bedside Clinical Tool to Identify Behavioral Variant FTD in Patients with Late-Onset Behavioral Changes. *J Alzheimers Dis* 2019;67(1):113-24. doi: 10.3233/JAD-180839 [published Online First: 2018/12/26]
5. Manera AL, Dadar M, Collins DL, et al. Deformation based morphometry study of longitudinal MRI changes in behavioral variant frontotemporal dementia. *Neuroimage Clin* 2019;24:102079. doi: 10.1016/j.nicl.2019.102079 [published Online First: 2019/12/05]
6. Manera AL, Dadar M, Collins DL, et al. Ventricular features as reliable differentiators between bvFTD and other dementias. *Neuroimage Clin* 2022;33:102947. doi: 10.1016/j.nicl.2022.102947 [published Online First: 20220122]
7. Fischl B. FreeSurfer. *Neuroimage* 2012;62(2):774-81. doi: 10.1016/j.neuroimage.2012.01.021 [published Online First: 2012/01/18]
8. Coupe P, Manjon JV, Fonov V, et al. Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *Neuroimage* 2011;54(2):940-54. doi: 10.1016/j.neuroimage.2010.09.018 [published Online First: 2010/09/21]
9. Morris JC, Weintraub S, Chui HC, et al. The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. *Alzheimer Dis Assoc Disord* 2006;20(4):210-6. doi: 10.1097/01.wad.0000213865.09806.92
10. Hammers A, Allom R, Koeppe MJ, et al. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp* 2003;19(4):224-47. doi: 10.1002/hbm.10123 [published Online First: 2003/07/23]
11. Dadar M, Maranzano J, Ducharme S, et al. Validation of T1w-based segmentations of white matter hyperintensity volumes in large-scale datasets of aging. *Hum Brain Mapp* 2018;39(3):1093-107. doi: 10.1002/hbm.23894 [published Online First: 2017/11/29]
12. Doust J. Using probabilistic reasoning. *BMJ* 2009;339
13. Pressman PS, Matlock D, Ducharme S. Distinguishing Behavioral Variant Frontotemporal Dementia From Primary Psychiatric Disorders: A Review of Recently Published Consensus Recommendations From the Neuropsychiatric International Consortium for Frontotemporal Dementia. *J Neuropsychiatry Clin Neurosci* 2021;33(2):152-56. doi: 10.1176/appi.neuropsych.20090238 [published Online First: 20210114]
14. Cotter J, Granger K, Backx R, et al. Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neurosci Biobehav Rev* 2018;84:92-99. doi: 10.1016/j.neubiorev.2017.11.014 [published Online First: 20171124]

15. Ducharme S, Dols A, Laforce R, et al. Recommendations to distinguish behavioural variant frontotemporal dementia from psychiatric disorders. *Brain* 2020;143(6):1632-50. doi: 10.1093/brain/awaa018 [published Online First: 2020/03/05]
16. Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 2011;134(Pt 9):2456-77. doi: 10.1093/brain/awr179 [published Online First: 2011/08/04]
17. Yokoyama JS, Rosen HJ. Neuroimaging features of C9ORF72 expansion. *Alzheimers Res Ther* 2012;4(6):45. doi: 10.1186/alzrt148 [published Online First: 20121116]
18. Vijverberg EG, Wattjes MP, Dols A, et al. Diagnostic Accuracy of MRI and Additional [18F]FDG-PET for Behavioral Variant Frontotemporal Dementia in Patients with Late Onset Behavioral Changes. *J Alzheimers Dis* 2016;53(4):1287-97. doi: 10.3233/JAD-160285 [published Online First: 2016/07/04]
19. Krudop WA, Dols A, Kerssens CJ, et al. Impact of Imaging and Cerebrospinal Fluid Biomarkers on Behavioral Variant Frontotemporal Dementia Diagnosis within a Late-Onset Frontal Lobe Syndrome Cohort. *Dement Geriatr Cogn Disord* 2016;41(1-2):16-26. doi: 10.1159/000441023 [published Online First: 20151017]
20. Zhutovsky P, Vijverberg EGB, Bruin WB, et al. Individual Prediction of Behavioral Variant Frontotemporal Dementia Development Using Multivariate Pattern Analysis of Magnetic Resonance Imaging Data. *J Alzheimers Dis* 2019;68(3):1229-41. doi: 10.3233/JAD-181004 [published Online First: 2019/03/26]
21. Seo SW, Thibodeau MP, Perry DC, et al. Early vs late age at onset frontotemporal dementia and frontotemporal lobar degeneration. *Neurology* 2018;90(12):e1047-e56. doi: 10.1212/WNL.0000000000005163 [published Online First: 20180216]
22. (FTLDNI) FLDNI. [Available from: <http://4rtni-ftldni.ini.usc.edu/>]
23. (ADNI) AsDNI. [Available from: adni.loni.usc.edu].
24. Manera AL, Dadar M, Collins DL, et al. Deformation based morphometry study of longitudinal MRI changes in behavioral variant frontotemporal dementia. *Neuroimage: Clinical* 2019;24 doi: <https://doi.org/10.1016/j.nicl.2019.102079>
25. Manera AL, Dadar M, Van Swieten JC, et al. MRI data-driven algorithm for the diagnosis of behavioural variant frontotemporal dementia. *J Neurol Neurosurg Psychiatry* 2021 doi: 10.1136/jnnp-2020-324106 [published Online First: 2021/03/17]
26. Zeighami Y, Ulla M, Iturria-Medina Y, et al. Network structure of brain atrophy in de novo Parkinson's disease. *Elife* 2015;4 doi: 10.7554/eLife.08440 [published Online First: 2015/09/08]
27. Boucetta S, Salimi A, Dadar M, et al. Structural Brain Alterations Associated with Rapid Eye Movement Sleep Behavior Disorder in Parkinson's Disease. *Sci Rep* 2016;6:26782. doi: 10.1038/srep26782 [published Online First: 2016/06/02]
28. Dadar M, Fonov VS, Collins DL, et al. A comparison of publicly available linear MRI stereotaxic registration techniques. *Neuroimage* 2018;174:191-200. doi: 10.1016/j.neuroimage.2018.03.025 [published Online First: 2018/03/20]

5246 scans
provided by NACC

806 scans failed MRI QC

medRxiv preprint doi: <https://doi.org/10.1101/2022.03.29.22273096>; this version posted April 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC-ND 4.0 International license](#).

4440 scans
pass MRI QC

3165 scans with
other primary diagnoses*

1275 scans
with diagnosis of interest

28 scans did not have a matching clinical visit <12 months
44 scans with a second neurodegenerative disorder identified as contributing to cognitive impairment
49 scans excluded due to no cognitive or behavioral impairment

1154 MRI scans
with diagnosis of interest

44 scans failed VentRa QC

1110 scans included in
the final data sample
14 bvFTD
30 other FTD
70 LBD
898 AD
62 VBI
36 PPD

2 other FTD
3 LBD
36 AD
2 VBI
1 PPD

medRxiv preprint doi: <https://doi.org/10.1101/2022.03.29.22273096>; this version posted April 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC-ND 4.0 International license](#).

