Automatic MRI Quantifying Methods in Behavioral-Variant Frontotemporal Dementia Diagnosis

Cajanus, Antti; Hall, Anette; Koikkalainen, Juha; Solje, Eino; Tolonen, Antti; Urhemaa, Timo; Liu, Yawu; Haanpää, Ramona M; Hartikainen, Päivi; Helisalmi, Seppo; Korhonen, Ville; Rueckert, Daniel; Hasselbalch, Steen; Waldemar, Gunhild; Mecocci, Patrizia; Vanninen, Ritva; van Gils, Mark; Soininen, Hilkka; Lötjönen, Jyrki; Remes, Anne M

Published in:
Dementia and Geriatric Cognitive Disorders Extra

DOI:
10.1159/000486849

Publication date:
2018

Document version:
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
Cajanus, A., Hall, A., Koikkalainen, J., Solje, E., Tolonen, A., Urhemaa, T., ... Remes, A. M. (2018). Automatic MRI Quantifying Methods in Behavioral-Variant Frontotemporal Dementia Diagnosis. Dementia and Geriatric Cognitive Disorders Extra, 8(1), 51-59. https://doi.org/10.1159/000486849
Automatic MRI Quantifying Methods in Behavioral-Variant Frontotemporal Dementia Diagnosis

Antti Cajanus, Anette Hall, Juha Koikkalainen, Eino Solje, Antti Tolonen, Timo Urhemaa, Yawu Liu, Ramona M. Haanpää, Päivi Hartikainen, Seppo Helisalmi, Ville Korhonen, Daniel Rueckert, Steen Hasselbalch, Gunhild Waldemar, Patrizia Mecocci, Ritva Vanninen, Mark van Gils, Hilkka Soininen, Jyrki Löjtönen, Anne M. Remes

Institute of Clinical Medicine – Neurology, University of Eastern Finland, Kuopio, Finland; Neurocenter, Neurology, Kuopio University Hospital, Kuopio, Finland; Combinostics Ltd., Tampere, Finland; Department of Radiology, Kuopio University Hospital, Kuopio, Finland; Neurocenter, Neurosurgery, Kuopio University Hospital, Kuopio, Finland; Department of Computing, Imperial College London, London, UK; Danish Dementia Research Centre, Department of Neurology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark; Section of Gerontology and Geriatrics, University of Perugia, Perugia, Italy; Medical Research Center, Oulu University Hospital, Oulu, Finland; Unit of Clinical Neuroscience, Neurology, University of Oulu, Oulu, Finland

Keywords
Frontotemporal dementia · Frontotemporal lobar degeneration · Neuroimaging · MRI · Dementia · Machine learning

Abstract
Aims: We assessed the value of automated MRI quantification methods in the differential diagnosis of behavioral-variant frontotemporal dementia (bvFTD) from Alzheimer disease (AD), Lewy body dementia (LBD), and subjective memory complaints (SMC). We also examined the role of the C9ORF72-related genetic status in the differentiation sensitivity. Methods: The MRI scans of 50 patients with bvFTD (17 C9ORF72 expansion carriers) were analyzed using 6 quantification methods as follows: voxel-based morphometry (VBM), tensor-based morphometry, volumetry (VOL), manifold learning, grading, and white-matter hyperintensities. Each patient was then individually compared to an independent reference group in order to attain diagnostic suggestions. Results: Only VBM and VOL showed utility in correctly identifying bvFTD from our set of data. The overall classification sensitivity of bvFTD with VOL + VBM achieved a total sensitivity of 60%. Using VOL + VBM, 32% were misclassified as having LBD. There was a trend of higher values for classification sensitivity of the C9ORF72 expansion carriers than...
noncarriers. **Conclusion:** VOL, VBM, and their combination are effective in differential diagnostics between bvFTD and AD or SMC. However, MRI atrophy profiles for bvFTD and LBD are too similar for a reliable differentiation with the quantification methods tested in this study.

**Introduction**

Frontotemporal lobar degeneration (FTLD) is the second most common early-onset dementing disease. FTLD is a heterogeneous group of different syndromes and also patterns of brain atrophy that vary between different phenotypes. A behavioral-variant frontotemporal dementia (bvFTD) is the most common phenotype of FTLD. Nearly one-half of the identified cases are of familial origin. While the genetic etiology of bvFTD is variable, the C9ORF72 expansion represents the most common genetic etiology, which is considered to underlie as many as over 30% of familial bvFTD cases [1, 2].

The accurate differential diagnostics of neurodegenerative diseases is crucial, not only from the perspective of clinical work, but also for the development of new treatments as well as the identification of patients suitable for clinical trials. The clinical diagnosis of bvFTD during a patient’s lifetime is based on a diagnostic battery defined by the international bvFTD criteria consortium, including clinical symptoms, neuropsychological assessment for possible bvFTD, and in addition imaging and evaluation of functional decline for probable bvFTD. The definite diagnosis would require pathological confirmation or genetic testing, which are seldom available [3, 4]. However, the diagnostics is challenging and the sensitivity of the criteria for possible bvFTD varies between 85 and 95% and for probable bvFTD between 75 and 85% in different studies [3, 5, 6]. On the other hand, the specificity of the criteria for possible bvFTD varies between 27 and 82% and for probable bvFTD between 85 and 95%, emphasizing the value of imaging in the differential diagnosis from other neurodegenerative disorders. However, visual evaluation of magnetic resonance imaging (MRI) requires an experienced neuroradiologist, and even so, the visual assessment provides only 59% sensitivity and 80% specificity from other common dementing diseases and is vulnerable for inter-rater differences [7]. It has also been proposed that different genetic forms of bvFTD are associated with different types of brain atrophy. In addition to frontal and temporal atrophy, it has been suggested that the carriers of the C9ORF72 expansion exhibit more widespread atrophy affecting the occipital lobes, cerebellum, and subcortical grey matter [8–12].

In the future, automatic decision support tools may be helpful as a part of differential diagnostics of neurodegenerative disorders. The Disease State Index (DSI) is a statistical classifier developed for clinical work as a decision support tool. It is designed to utilize multimodal data for analysis to provide a score for individual patient classification. This method has previously shown to be useful in diagnosing Alzheimer disease (AD), predicting the progression of patients with mild cognitive impairment into AD, differentiating FTLD from AD and mild cognitive impairment using multimodal data [13–16]. Its performance between different patient cohorts has also been validated [17].

The aim of this study was to evaluate the utility of the DSI using data from MRI scans only in a setting of differentiating bvFTD from AD, FTLD, Lewy body dementia (LBD) and subjective memory complaints (SMC). We performed DSI analysis individually for each patient with separate reference data, mimicking a clinical setting. We also examined if there was a difference in the sensitivity of the classification between the carriers and noncarriers of the C9ORF72 expansion.
Materials and Methods

Subjects
A total of 50 bvFTD patients diagnosed with current diagnostic guidelines were included in this study (Table 1) [3]. The cohort included men and women equally and the mean age was 62.5 years (36–79). Comorbidity with amyotrophic lateral sclerosis (bvFTD-ALS) was diagnosed in 10% (n = 5) of patients. The C9ORF72 expansion was detected in 34% (n = 17) of patients, while the remaining 66% (n = 33) were negative for the mutation. Analyses of the demographic data indicated that at the time of the MRI scan, the patients in the C9ORF72 expansion carrier group as compared to those with a negative status were significantly younger (mean age 60.5 vs. 65.5 years, respectively, p = 0.02). Otherwise there was no difference between the C9ORF72 expansion carriers and noncarriers.

SMC subjects have entered a memory clinic with subjective memory problems, but without any objective finding of a neurodegenerative process. We included these patients as a reference to evaluate if the MRI quantification methods are able to determine patients with and without neurodegenerative process.

All patients were diagnosed by an experienced neurologist specialized in memory disorders in the University Hospital of Kuopio, Finland. The ethics committee of the University Hospital of Kuopio approved the research protocol in accordance with the principles of the Declaration of Helsinki. Each patient provided written informed consent to participate in the study.

Genetic Analysis
Genomic DNA was extracted from venous blood samples using QIAamp DNA blood mini extraction kit (Qiagen). The presence of the C9ORF72 expansion was detected using repeat-primed PCR [2]. Of those testing positive for the expansion, 18 had an excess of 40 repeats, with 3 cases showing an intermediate expansion length between 10 and 40 repeats. Patients were considered negative for the C9ORF72 expansion if their number of repeats was fewer than 10 [2].

MRI Acquisition
For all patients, brain MRI was performed using either a 1.5-T or a 3-T MRI scanner: 56% (n = 28) of cases were scanned with a 1.5-T scanner, while the remaining 44% (n = 22) patients underwent a 3-T scanner. In 49 patients, both a T1-weighted 3D gradient echo sequence and a fast fluid-attenuated inversion recovery (FLAIR) sequence were obtained in MRI. The FLAIR sequence was unavailable for 1 patient. In a previous study with multiple cohorts, no notable effects of magnetic field strength between 1.5 T and 3 T for imaging biomarkers were found [18].
Image Analysis

The MRI analysis was performed with 6 fully automatic image analysis methods quantifying the following features: volume of various brain areas (VOL), local volume change using tensor-based morphometry (TBM), local grey matter changes using voxel-based morphometry (VBM), image similarity to images in the reference database using manifold-based learning (Manifold) and grading, as well as white-matter hyperintensities (WMH). All imaging biomarkers were generated using solely T1 and FLAIR images. We did the normalization by gender and age using the SMC patients as the reference group for VOL, VBM, TBM, and WMH biomarker values. We fitted a linear regression model using age and gender as explanatory variables for the MRI values, which were then transformed to match the age and gender of the patient being analyzed using these regression models [19]. We analyzed patients using each quantification method individually and all possible combinations of methods, in order to evaluate the diagnostic value of each method, which has been previously described [18].

The MRI scans of the 50 bvFTD patients were performed in Kuopio University Hospital to support the patients’ original diagnosis. The scans were originally evaluated by a radiologist specialized in neurodegenerative disorders as part of the initial diagnosing process. For this study, the scans were retrospectively collected from the patient data archive. The radiologist’s expert opinion was not used in this study, only the parameters extracted by the automatic MRI quantification methods were used to classify patients into predetermined diagnostic categories (AD, FTLD, LBD, and SMC).

Reference Data

The reference data included a total of 1,111 patients who were evaluated in 4 different centers (VU Medical Center, The Netherlands; Università degli Studi di Perugia, Italy; Rigshospitalet, Denmark; and University of Eastern Finland, Finland). Additionally, data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) were included. The reference data were completely independent of the study group. The reference data included patients pertaining to 4 major diagnostic groups as follows: AD (n = 537), FTLD (n = 154), LBD (n = 61), and SMC (n = 359). In the reference data, the FTLD group included patients with bvFTD, non-fluent-variant primary progressive aphasia, and semantic-variant primary progressive aphasia, thus we refer to the group as FTLD, not bvFTD. The reference data are described in online supplementary Table 1 (see www.karger.com/doi/10.1159/000486849 for all online suppl. material) [20].

Statistical Analysis

Both the demographic and the clinical data were compared between groups of the C9ORF72 expansion carriers and noncarriers by applying Student t tests and Pearson χ² tests. The Pearson χ² test was performed in order to explore differences in the classification sensitivity between the genetic groups. p values < 0.05 were considered as statistically significant. The statistical analyses were performed using the IBM® SPSS® Statistics software package, version 22.

The decision support tool used in this study, the DSI, generates a numerical measure for each patient, in order to evaluate the state of the disease progression on a scale of 0–1. The DSI is designed to use multimodal data in the differentiating process, but in this study we investigated the utility of features extracted only from MRI scans. By combining all available biomarkers extracted from the MRI scans with the quantification methods specified earlier, the DSI generates an individual patient’s profile. The DSI measures the fit of the patient data with reference to previously diagnosed cases from two groups. More specifically, a similarity to the reference group leads to a low DSI value, while a similarity to the study group results in a high DSI value. This pairwise comparison between two diagnostic categories is repeated for each possible pair, and finally the classifier suggests the most likely diagnosis for each
individual patient based on multiple volumetric and morphometric measures normalized by sex and age [14, 18]. In this study, 4 diagnostic classes were used: FTLD, LBD, AD, and SMC. An illustration of the study setting is presented in Figure 1.

**Results**

The DSI performed the best using only MRI parameters extracted with VBM in the classification process. In total bvFTD cohort, the DSI achieved sensitivity (56%) by using only VBM of the 6 individual quantification methods. Using VBM, the sensitivity was even higher
in patients with the C9ORF72 expansion (71%) (Table 3). Sensitivity of 44% was detected in the total group of bvFTD by using VOL. The other 4 methods (TBM, grading, WMH, and manifold) showed no utility in classification. Overall, there was a trend of higher values for classification sensitivity of the C9ORF72 expansion carriers than noncarriers. Classification sensitivities of all quantification methods are presented in Table 2.

The best combination of quantification methods was found to be VBM + VOL giving a classification sensitivity of 60% for the whole bvFTD cohort. By using the combination of VBM and VOL, the carriers of the C9ORF72 expansion were classified with 65% and noncarriers with 58% sensitivity (Table 3). However, this difference did not reach statistical significance (p = 0.37). Classification sensitivities of all possible combinations of quantification methods are presented in online supplementary Table 2.
With the combination of VOL + VBM, patients were most often misclassified as having LBD (32%), while 4% were misclassified as AD patients, and the remaining 4% as SMC subjects.

**Discussion**

The aim of the current study was to evaluate the utility of 6 automated MRI quantification methods specifically in performing differential diagnostics of bvFTD. We found that only a small subset of imaging biomarkers were useful in the classification of bvFTD. More specifically, only VBM and VOL together with grading reached a sensitivity greater than 25%, which represents the sensitivity we would achieve by complete chance in this 4-class-scenario. VBM was the best method giving a sensitivity of 56% in the whole bvFTD group. However, the sensitivity was even higher (71%) in patients with the \textit{C9ORF72} expansion. The combination of VBM + VOL improved the classification sensitivity with the \textit{C9ORF72} expansion noncarriers, while with the \textit{C9ORF72} expansion carriers the sensitivity was slightly reduced.

In a previous study utilizing identical MRI analysis methods, there was a differentiation accuracy of 62% of FTLD patients from SMC, AD, LBD, and vascular dementia patients [18]. In that study, 21% of FTLD patients were misclassified as having AD, which is more than in the present study (4%). However, in the previous study, only 5% of FTLD patients were misclassified as having LBD, which contrasts strongly with the respective value in the present study. One reason for this may be that the patient cohort differed in clinical characteristics from our cohort. In the current study, even the best performing quantification method VBM resulted in 28% of the patients being misclassified as having LBD. Among the \textit{C9ORF72} expansion carriers, the classification was better: only 18% of the carriers and 36% of the noncarriers were incorrectly classified as having LBD with VBM. Potential explanations for this result are as follows: first, at least some of the \textit{C9ORF72} expansion carriers may have been associated with overlapping neurodegenerative conditions in addition to bvFTD. Second, as the diagnosis of the \textit{C9ORF72} expansion noncarriers was solely based on the clinical evaluation of symptoms due to objective biomarkers lacking (the clinical diagnosis is considered to reach 85–95% specificity) [3, 5, 6], this patient group may have included some individuals with a clinical misdiagnosis of bvFTD. Thirdly, LBD is associated with widespread cortical and central atrophy with no specific MRI findings [21–23]. Existing evidence suggests that specifically in the \textit{C9ORF72} expansion carrier bvFTD patients, the atrophy may affect the posterior regions of the brain as well as the frontal and temporal regions [11, 12, 24]. The genetic status of the reference FTLD cases is not known and the dataset may include a number of cases with the \textit{C9ORF72} expansion. The reference data also includes other clinical phenotypes of FTLD, such as progressive aphasia and semantic dementia, in addition to bvFTD. Moreover, the confusion between LBD and bvFTD may also be associated with a long disease progression of bvFTD prior to the MRI scan, leading to a more general and widespread atrophy, which would make MRI biomarkers for bvFTD and LBD relatively similar to each other. The relatively low sample size of the study may emphasize the influence of chance and errors in the clinical diagnosis. No patients with mild cognitive impairment were included in the reference data, since the underlying neurodegenerative process at that stage is uncertain. Defining the \textit{ground truth diagnosis} for those patients would most likely produce a significant amount of false diagnoses.

It is important to note here that in a multi-class setting, the classification accuracies are much lower than those associated with pairwise comparison settings, thus rendering findings between such studies not strictly comparable with each other [25, 26]. In light of the current study, it is important to note that the multi-class setting provided herein is associated with a
higher ecological validity than that of a pairwise setting, bearing a closer resemblance to the 
real-life scenario faced by a patient suffering from a neurodegenerative disorder and under-
going a differential diagnostic procedure. However, the validation of different methods and 
diagnostic tools in several cohorts is needed to provide high quality tools.

**Conclusion**

We demonstrated high differentiation sensitivity of bvFTD from the most common 
dementing neurodegenerative disorder (AD) by using MRI quantification applications. Similarly, SMC subjects were differentiated from bvFTD with high sensitivity. However, out of the 6 image quantification methods assessed, only 2 methods, VBM and VOL, were found to be useful in the differentiation of bvFTD from other disorders. The best classification sensitivity was achieved by using the VBM method. This clearly implicates that further studies are needed to both evaluate the feasibility of various MRI quantifying methods and to optimize the set of features that are useful in the differential diagnosis of neurodegenerative dementing diseases. The most challenging differentiation was found between bvFTD and LBD, due to similarities in the MRI profiles characterizing both groups of patients. Automated MRI quantification methods and decision support tools may be helpful in the differential diagnostics of bvFTD in clinical practice. However, in the case of the more uncommon neurodegenerative disorders, such as bvFTD and LBD, differential diagnostics using only MRI methods seems to be challenging and also the diagnosis of specific clinical features and other biomarkers should be included in decision support tools to provide greater sensitivity and specificity.

**Acknowledgements**

This work received funding from the following bodies: European Union’s Seventh Frame-
work Programme for research, technological development, and demonstration under grant 
agreement No. 611005 (PredictND); VTR funding from Kuopio University Hospital; the 
Finnish Medical Foundation; the Olvi Foundation; the Finnish Alzheimer Research Assoc-
iation; the Finnish Brain Foundation; and the Päivikki and Sakari Sohlberg foundation.

**Disclosure Statement**

Equity: Juha Koikkalainen and Jyrki Lötjönen are shareholders and founders of 
Combinostics Ltd. Patents/royalties: Combinostics Ltd. owns the following IPR related to the 
paper: 1. J. Koikkalainen and J. Lötjönen. A method for inferring the state of a system, 
US7,840,510 B2, PCT/FI2007/050277. 2. J. Lötjönen, J. Koikkalainen, and J. Mattila. State 
inference in a heterogeneous system, PCT/FI2010/050545, FI20125177. The other authors 
have no conflicts of interest to disclose.

**References**

1. Majounie E, Renton AE, Mok K, Dopper EGP, Waite A, Rollinson S, et al: Frequency of the C9ORF72 hexanucle-
otide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-
sectional study. Lancet Neurol 2012;11:323–330.
2. Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, et al: A hexanucleotide repeat ex-
pansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 2011;72:257–268.
