



# Brain MRI Pattern Recognition Translated to Clinical Scenarios

Andreia V. Faria<sup>1\*</sup>, Zifei Liang<sup>2</sup>, Michael I. Miller<sup>3</sup> and Susumu Mori<sup>1</sup>

<sup>1</sup> Department of Radiology, Johns Hopkins University, Baltimore, MD, United States, <sup>2</sup> Department of Radiology, New York University, New York, NY, United States, <sup>3</sup> Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, United States

## OPEN ACCESS

### Edited by:

Jennifer L. Robinson,  
Auburn University, United States

### Reviewed by:

Jingyu Liu,  
Mind Research Network,  
United States  
Hyunjin Park,  
Sungkyunkwan University,  
South Korea

### \*Correspondence:

Andreia V. Faria  
afaria1@jhmi.edu

### Specialty section:

This article was submitted to  
Brain Imaging Methods,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 29 June 2017

**Accepted:** 02 October 2017

**Published:** 20 October 2017

### Citation:

Faria AV, Liang Z, Miller MI and Mori S  
(2017) Brain MRI Pattern Recognition  
Translated to Clinical Scenarios.  
*Front. Neurosci.* 11:578.  
doi: 10.3389/fnins.2017.00578

We explored the performance of structure-based computational analysis in four neurodegenerative conditions [Ataxia (AT,  $n = 16$ ), Huntington's Disease (HD,  $n = 52$ ), Alzheimer's Disease (AD,  $n = 66$ ), and Primary Progressive Aphasia (PPA,  $n = 50$ )], all characterized by brain atrophy. The independent variables were the volumes of 283 anatomical areas, derived from automated segmentation of T1-high resolution brain MRIs. The segmentation based volumetric quantification reduces image dimensionality from the voxel level [on the order of  $\mathcal{O}(10^6)$ ] to anatomical structures [ $\mathcal{O}(10^2)$ ] for subsequent statistical analysis. We evaluated the effectiveness of this approach on extracting anatomical features, already described by human experience and a priori biological knowledge, in specific scenarios: (1) when pathologies were relatively homogeneous, with evident image alterations (e.g., AT); (2) when the time course was highly correlated with the anatomical changes (e.g., HD), an analogy for prediction; (3) when the pathology embraced heterogeneous phenotypes (e.g., AD) so the classification was less efficient but, in compensation, anatomical and clinical information were less redundant; and (4) when the entity was composed of multiple subgroups that had some degree of anatomical representation (e.g., PPA), showing the potential of this method for the clustering of more homogeneous phenotypes that can be of clinical importance. Using the structure-based quantification and simple linear classifiers (partial least square), we achieve 87.5 and 73% of accuracy on differentiating AT and pre-symptomatic HD patients from controls, respectively. More importantly, the anatomical features automatically revealed by the classifiers agreed with the patterns previously described on these pathologies. The accuracy was lower (68%) on differentiating AD from controls, as AD does not display a clear anatomical phenotype. On the other hand, the method identified PPA clinical phenotypes and their respective anatomical signatures. Although most of the data are presented here as proof of concept in simulated clinical scenarios, structure-based analysis was potentially effective in characterizing phenotypes, retrieving relevant anatomical features, predicting prognosis, and aiding diagnosis, with the advantage of being easily translatable to clinics and understandable biologically.

**Keywords:** precision medicine, pattern recognition, quantitative MRI, computer aid, automated MRI analysis

## INTRODUCTION

A longtime dream of clinicians is to use computational tools for aiding decisions. Like using the spelling and grammar checkers when writing a text or Google for searching, clinical computational tools would neither define purposes nor change goals, but add a higher level of quality and speed to the results. There are three must-haves for computational-aid tools: speed, automation, and, of course, efficacy. The development of such tools for medical records and imaging, in particular, is extremely complex, involving knowledge in multiple domains. Consequently, more than two decades after the initial attempts (for review and perspectives see Orphanoudakis et al., 1996; Akgul et al., 2011; Hwang et al., 2012; Kalpathy-Cramer et al., 2015; Pinho et al., 2017; Spanier et al., 2017), no system is yet adequately suited for practical daily use. The key to translating the computational models to radiological practice is to resolve the so-called semantic gap: “the differences between image similarity on the high level of human perception and the low level of a few numbers” (Depeursinge et al., 2011). Three basic steps are involved: precise quantification, optimal feature selection and combination, and, eventually, meaningful applications and testing.

The first step, image quantification, is straightforward if one is simply interested in the intensity of a given voxel. What is not simple, however, is to extract some biological meaning from the noisy voxel-by-voxel information, which can be of the order of  $10^6$ , considering only T1-weighted images, one of the multiple MRI contrasts. There are numerous papers on voxel-based analysis (VBA) in which human involvement is eliminated on the assumption that a human being’s ability to detect abnormality is neither sensitive nor reliable. A PubMed search for “VBA,” “brain,” and “MRI” results in more than 2,300 publications in the last 10 years. These studies provide a wealth of descriptive imaging results that are usually not perceptible at an individual level and fail to be translated to clinical practice, which meanwhile, remains supported by human judgment. If we flip this approach 180° by asking: Can a computational approach describe abnormalities that agree with human perception?, we find the number of publications to be surprisingly small. A PubMed search for “structure-based analysis” or “atlas-based analysis,” “brain,” and “MRI” results in fewer than 200 publications in the last 10 years. An old strategy to replicate human perception is to group voxels in regions of interest (ROIs) and label them according to existing anatomical knowledge. For example, all the voxels associated with certain x, y, z coordinates are called “thalamus,” or “frontal lobe,” or “internal capsule,” and so on. This is what radiologists do, increasing the signal-to-noise ratio and adding a biological domain to their subjective analysis. However, objectively quantifying, structurizing, and recording the information for subsequent use is much more complex. In addition, defining ROIs in multiple subjects multidimensionally is just not feasible; precise automated tools are vital.

This structure-based analysis is linked to the second step to solve the semantic gap: the feature selection and combination. Here, two components are essential: the existing knowledge of normal and abnormal patterns and the ability to recognize these

patterns in future patients. For example, when a patient has striatum atrophy and motor disabilities, Huntington’s Disease (HD) is a possible diagnosis because physicians learned that these two features are associated with this disease. In addition to centuries of pathophysiological knowledge, what is hidden behind this apparently simple conclusion is an enormous amount of comprehension about normal variation. In order to conclude that those regions, in an individual of a certain age and gender, are smaller than expected, an analysis of multiple granularity levels (looking to the caudate, or the basal ganglia, or the deep gray matter, or the lobe, or the whole brain), and multiple image domains (volume, intensity, shapes), and finally the combination of features in different fields (clinical and imaging) are necessary. This leads to the amazing capability of pattern recognition that humans have and that machine-learning methods try to replicate.

Finally, even if we are able to quantify structures precisely in different levels and domains, to compare individual cases with large and variable normal and pathological databases, and to extract and combine important features efficiently, we still have to suit the computer-aid tools to the appropriate applications and test them. If the goal is a diagnostic-aid tool, this may be the most challenging step because the gold-standard is the clinical diagnosis, which does not necessarily reflect the actual situation. In addition, the correlation between pathology and anatomy may be weak or indirect. This is usually the case in pathologies in which the anatomical changes are subtle or happen later, or when the time course is unknown, or in those that embrace heterogeneous phenotypes. These cases are challenging and may reduce the efficiency of classification models, but they also offer an opportunity to design tools for binning a given entity into subgroups, for example, that may be of clinical relevance.

Previously, our group and others advanced in the first two steps (quantification and feature extraction). The brain quantification and segmentation accuracy improved drastically in this decade due to the advances in multi-atlas technologies (Warfield et al., 2004; Artaechevarria et al., 2009; Langerak et al., 2010; Lotjonen et al., 2010; Sabuncu et al., 2010; van Rikxoort et al., 2010; Jia et al., 2012; Wang et al., 2013), allowing use of state-of-the-art techniques for quantification and extraction of clinically meaningful image features. We confirmed the accuracy of these techniques in different populations and protocols (Liang et al., 2015). We then tested whether the structured anatomical data extracted actually captured the anatomical features that can be perceived by trained clinicians (Faria et al., 2015). In the present study, we advance to the next step and report progress on feature selection, combination, and classification, showing the potential of structure-based analysis for computer-aided decisions.

This study focused on the brain MRIs of patients with these neurodegenerative conditions: Ataxia (AT), Huntington’s Disease (HD), Alzheimer’s Disease (AD), and Primary Progressive Aphasia (PPA). Briefly, Ataxia, or more specifically, the Spinocerebellar ataxia type 6 (SCA-6) which is considered here, is an autosomal dominant disorder that is characterized by a slowly progressive cerebellar ataxia, dysarthria and nystagmus (Zhuchenko et al., 1997). The cerebellar atrophy, demonstrated

by several prior MRI studies, is a constant (Butteriss et al., 2005) and relates with clinical manifestations (Eichler et al., 2011).

HD is a progressive lethal neurodegenerative disorder characterized by movement disorders and progressive cognitive and psychological manifestations (Huntington, 1872). The anatomical hallmark of HD is striatal atrophy. Although the atrophy may start as early as 15 years before the onset of motor symptoms, and continue through the pre-manifest period (Tabrizi et al., 2009, 2012, 2013; Paulsen et al., 2014a,b), it is mostly undetectable by clinical evaluation of MRIs, at individual level, in pre-symptomatic patients. The early quantitative characterization of the atrophy, both at group and individual level, is an important piece of information for the development of disease-modifying treatments (Faria et al., 2016; Wu et al., 2016).

Alzheimer disease (AD) is a chronic neurodegenerative disease characterized by short-term memory loss in the early disease stages and progressive cognitive and functional deficits as the disease advances. It is actually not a single disease but a clinically, anatomically and biologically heterogeneous disorder encompassing a wide spectrum of cognitive and anatomical profiles (Zhang et al., 2016). Although a classical pattern of atrophy is reported for AD as a group, first noticeable in the medial temporal lobe (including hippocampus and entorhinal cortex), eventually spreading through the remainder of the brain (Apostolova et al., 2007), this pattern is not highly discriminant at individual level (Frisoni et al., 2017). In addition, the atrophy is usually clinically evident long after the cognitive deficits start. The heterogeneity of phenotypes and subtleness of early anatomical changes are extra challenges for the development of therapeutics and prognostic models.

Primary progressive aphasia (PPA) is a clinical syndrome characterized by insidious progressive language impairment that is initially unaccompanied by other cognitive deficits (Mesulam, 1982). It is caused by various neurodegenerative diseases and has a highly variable course. There are three main variants that are distinguished by their key features and supporting brain imaging characteristics, which are generally associated with distinct underlying pathologies (Gorno-Tempini et al., 2011): agrammatic (Av) is supported by left posterior frontal and (Zhuchenko et al., 1997) insular atrophy; semantic (Sv) is associated with left greater than right anterior and inferior temporal atrophy; logopenic (Lv) is associated with posterior temporal and inferior parietal atrophy (Rohrer and Rosen, 2013; Wilson et al., 2016). The identification of the variant provides some clues regarding the subsequent course (Leyton et al., 2016), and would be of great value for prognosis in the initial stages. However, the early classification is particularly challenging because the clinical deficits are common to all three variants and the anatomical changes are still clinically silent. Methods for phenotypically characterization, particularly at early phases, would be of great assistance.

The choice of these clinical entities was due to the fact that the common feature (atrophy) varies in extension and location, providing an appropriate dynamic range of abnormalities. In addition, the atrophy is mostly visible, which enables validation by qualitative human evaluation. The overall goal of this study was to test the performance of structure-based computational analysis on extracting anatomical features, already described by

human experience and a priori biological knowledge, in specific patient populations. The variables in question were the volumes of 283 structures. We showed the potential of the structure-based analysis on characterization and classification (1) when pathologies were relatively homogeneous, with evident image alterations (e.g., Ataxias); (2) when the time course was highly correlated with the anatomical changes (e.g., HD), an analogy for prediction; (3) when the pathology embraced heterogeneous phenotypes (e.g., AD) so the classification was less efficient but, in compensation, anatomical and clinical information were less redundant; and (4) when the entity was composed of multiple subgroups that had some degree of anatomical representation (e.g., Primary Progressive Aphasia), showing the potential of this method for the clustering of more homogeneous phenotypes that can be of clinical importance.

## MATERIALS AND METHODS

### Database

The overall goal was to test the performance of structure-based computational analysis in extracting anatomical features, previously described by human experience and a priori biological knowledge, in specific patient populations.

The data consisted of high-resolution T1-weighted brain MRIs (MPRAGE), for five groups of individuals: healthy individuals (controls,  $n = 208$ ), AT ( $n = 16$ ), HD ( $n = 52$ ), AD ( $n = 66$ ), and PPA ( $n = 50$ ) (Table 1). The data from healthy individuals (controls) were obtained from three sources: (1) internal datasets from Johns Hopkins University (JHU), (2) International Consortium for Brain Mapping (ICBM, [loni.usc.edu/ICBM](http://loni.usc.edu/ICBM)), and (3) the AD Neuroimaging Initiative (ADNI, [adni.loni.usc.edu](http://adni.loni.usc.edu)). The control dataset included more than 10 different protocols (including different machine manufacturers, strength of magnetic field, and resolution), thus replicating the heterogeneity encountered in clinical scenarios. Individuals with AT were from JHU and had spinocerebellar ataxia type 6 (SAC6). Individuals with HD, also from JHU, were grouped into three different stages, according to their CAG-Age Product (CAP) scores (Penney et al., 1997) and clinical symptoms: pre-symptomatic far from onset ( $n = 23$ ), pre-symptomatic close to onset ( $n = 16$ ), and early symptomatic ( $n = 13$ ). Individuals with AD, from JHU and ADNI, were diagnosed according to new clinical guidelines (Albert et al., 2011; Jack et al., 2011; McKhann et al., 2011; Sperling et al., 2011). Individuals with PPA, from JHU, were diagnosed and classified into three variants: logopenic (Lv,  $n = 18$ ), semantic (Sv,  $n = 16$ ), and agrammatic (Av,  $n = 16$ ), based on current clinical guidelines (Mesulam, 1982; Gorno-Tempini et al., 2011). All the data had previously been de-identified, and the participants consented to enrolling by written consent.

### Image Processing

In the present study, quantification of regional brain volume was performed on a structural level, which involved the mapping of each brain to 29 templates in which the structures in question had previously been labeled. The brain mapping was performed with large deformation diffeomorphic metric mapping (LDDMM)

**TABLE 1** | Demographic and protocol information.

| Group           | Sample | Age range (Years) | Mean age $\pm$ St.dev. (Years) | Male/Female | Protocols (Manufacturer, Field Strength (T), Voxel Size (mm), Sample)  |
|-----------------|--------|-------------------|--------------------------------|-------------|--|
| <b>Controls</b> | 208    | 20–95             | 57.9 $\pm$ 18.8                | 98/110      | Phillips, 1.5, 1 $\times$ 0.875 $\times$ 0.875, 28<br>Phillips, 1.5, 1.2 $\times$ 0.94 $\times$ 0.94, 9<br>Phillips, 3, 1 $\times$ 1 $\times$ 1, 7<br>Phillips, 3, 1.2 $\times$ 1 $\times$ 1, 26<br>Phillips, 3, 0.9 $\times$ 0.9 $\times$ 0.9, 52<br>Phillips, 3, 1.1 $\times$ 0.83 $\times$ 0.83, 16<br>GE, 1.5, 1.2 $\times$ 0.94 $\times$ 0.94, 6<br>GE, 3, 1.2 $\times$ 1.02 $\times$ 1.02, 6<br>Siemens, 1.5, 1.2 $\times$ 1.25 $\times$ 1.25, 7<br>Siemens, 3, 1 $\times$ 1 $\times$ 1, 28<br>Siemens, 3, 1.2 $\times$ 1 $\times$ 1, 23 |
| <b>Ataxia</b>   | 16     | 48–73             | 60.8 $\pm$ 6.8                 | 13/3        | Phillips, 3, 1.1 $\times$ 0.83 $\times$ 0.83, 16   |
| <b>HD</b>       |        |                   |                                |             |  |
| Far from onset  | 23     | 21–51             | 36.8 $\pm$ 9.7                 | 10/13       |  |
| Near to onset   | 16     | 20–55             | 45.1 $\pm$ 8.6                 | 13/3        | Phillips, 3, 0.9 $\times$ 0.9 $\times$ 0.9, 52   |
| Early symptoms  | 13     | 30–59             | 50.8 $\pm$ 7.9                 | 7/6         |  |
| <b>AD</b>       | 66     | 55–93             | 74 $\pm$ 10.5                  | 40/26       | Siemens, 3, 1.2 $\times$ 1 $\times$ 1, 27<br>Siemens, 1.5, 1.2 $\times$ 1.25 $\times$ 1.25, 7<br>Phillips, 1.5, 1.2 $\times$ 0.94 $\times$ 0.94, 7<br>Phillips, 3, 1.2 $\times$ 1 $\times$ 1, 8<br>GE, 1.5, 1.2 $\times$ 0.94 $\times$ 0.94, 9<br>GE, 3, 1.2 $\times$ 1.02 $\times$ 1.02, 8  |
| <b>PPA</b>      |        |                   |                                |             |  |
| Lv              | 18     | 51–79             | 68.3 $\pm$ 5.4                 | 10/8        | Siemens, 3, 1 $\times$ 1 $\times$ 1, 21  |
| Sv              | 16     | 57–77             | 65.5 $\pm$ 6.5                 | 11/5        | Phillips, 3, 1.2 $\times$ 1 $\times$ 1, 29   |
| Av              | 16     | 48–84             | 68.2 $\pm$ 10.7                | 9/7         |  |

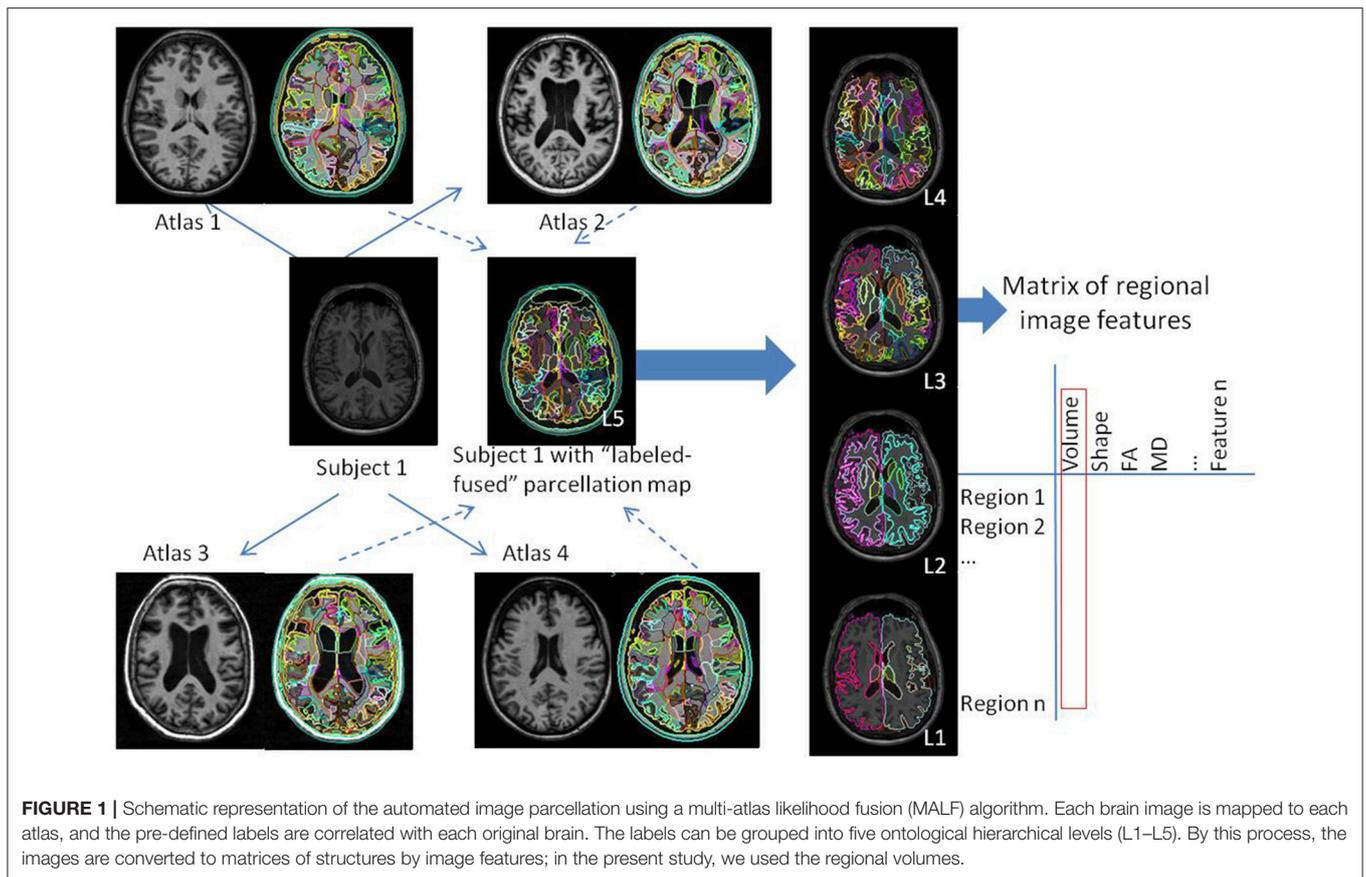
(Wang et al., 2007; Ceritoglu et al., 2009; Djamanakova et al., 2013). Inversely, the labels were warped to each subject space and then fused by a likelihood fusion algorithm, which took into account both the location and intensity information of each label (Langerak et al., 2010; Sabuncu et al., 2010; Wang et al., 2013). The details of this method, the atlas creation, and the validation in diverse protocols and anatomical phenotypes are described in our previous publications (Tang et al., 2013; Liang et al., 2015; Ma et al., 2015; Wu et al., 2016).

By this multi-atlas automated brain segmentation tool, the raw images, which consisted of more than 1 million voxels were converted to 286 structural representations, of which the volumes were measured. Based on the hierarchical relationship defined in the atlas, these structures can be combined to create five ontological levels with 8–19–53–125–286 structures respectively (Figure 1). Details of the hierarchical-ontological grouping are found in our previous publications (Djamanakova et al., 2014; Wu et al., 2016). One of the reasons for choosing the structure-based multi-level design is that the physician's analysis does not operate at the voxel level, but at the structural level, migrating

freely along the hierarchy. The choice of level is a trade-off between regional specificity and noise: in higher levels, more structures are defined and spatial specificity increases, yet noise also increases. In hypothesis-driven studies, the choice of the level depends on the interest in a given structure. In data-driven studies, the data can be analyzed using all ontological levels combined, or at each level independently. Our present analyses were performed according to the latter approach.

## Statistical Analysis and Outputs

We used partial least square—discriminant analysis (PLS-DA) to classify individuals in three different analyses: (1) AT vs. controls, (2) HD vs. controls, and (3) AD vs. controls. As many different protocols as possible were included for each analysis, yet keeping the individuals paired by age, gender, and image protocol in each group compared. The PLS-DA inputs were the regional volumes of brain structures in the five ontological levels, normalized by the intracranial volume. As the classification accuracy increased with the level of granularity and converged at level 3, the results are reported at this level. Level 3 is a medium level of



granularity, where the whole brain is segmented in lobes, deep gray matter, major deep white matter structures, ventricles, and sulci (Djamanakova et al., 2014). It matches well the radiologists reading (Faria et al., 2015), and the segmentation reproducibility is high (Djamanakova et al., 2013; Faria et al., 2015; Liang et al., 2015).

We opted for using simple linear classifiers to reduce the chance of overfitting, increase the potential for generalization of the results, and facilitate the translation to clinical practice, which is our aim, rather than the greatness of the classification. We could have obtained higher classification accuracy using more elaborate classifiers (such as a support vector machine and black-box models). Briefly, PLS is the least restrictive extension of the multiple linear regression models, therefore applicable to situations where the number of predictor variables exceeds the number of observations. As in the principal component analysis (PCA), the scores, or components, are the sets of values of linearly uncorrelated variables and the regression coefficients (loadings or weights) reflect the importance of the predictor variables in the model.

In each analysis, the samples were divided in training set, in which the classifier was built, and test set, in which the accuracy was tested. The validation in an independent test set reduces the impact of overfitting by biased variable selection and results in more realistic classification accuracy. In addition to the classifier accuracy, the outputs of interest were (1) the anatomical features

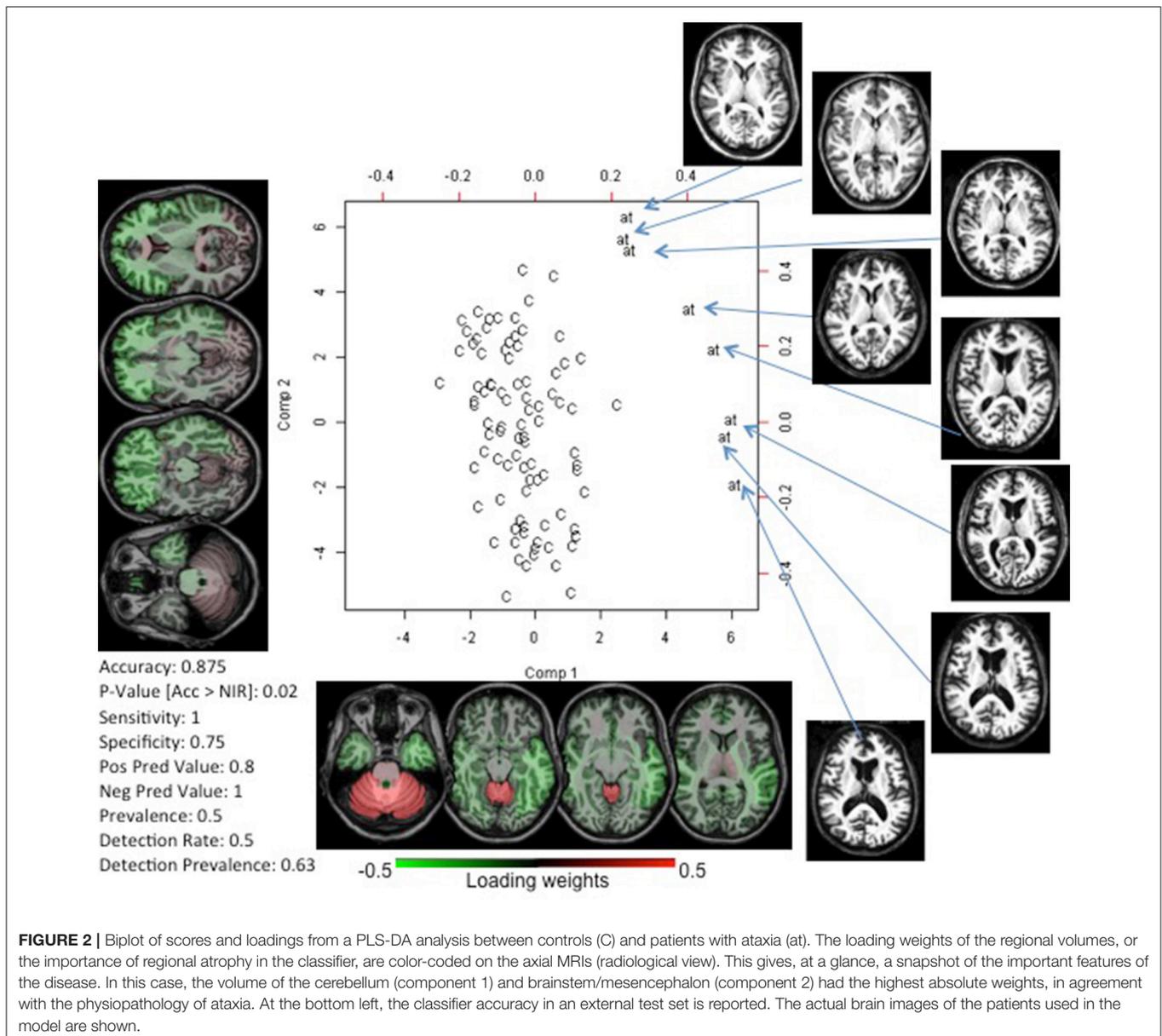
important for the classification (related to the PLS loading weights) or, in other words, the regional pattern of atrophy that characterizes each group, and (2) the individual's chances of belonging to different groups, which can be of direct importance for clinical guidance. Secondary outputs of interest are (1) the distance among individuals in the principal component space, which can be used for image retrieval of individuals with similar phenotypes, and (2) the individual z-score maps of atrophy.

In the case of PPA, we qualitatively explored a possible natural segregation among the phenotypes with PCA. The inputs were, again, the regional volumes of brain structures, normalized by the intracranial volume. We then assessed the potential of our tools on subdividing groups according to anatomical phenotype, using hierarchical clustering.

## RESULTS

### Ataxia: Extraction of Homogeneous and Noticeable Image Features

The analysis performed on 16 individuals with ataxia (8 for training, 8 for testing; Supplementary Table), and controls paired by age, gender, and image protocol achieved accuracy of 0.875 in differentiating individuals with AT from controls. **Figure 2** shows the PLS-DA plot (scores vs. loadings) and the two components used by the classifier. Component 1 is mostly responsible for

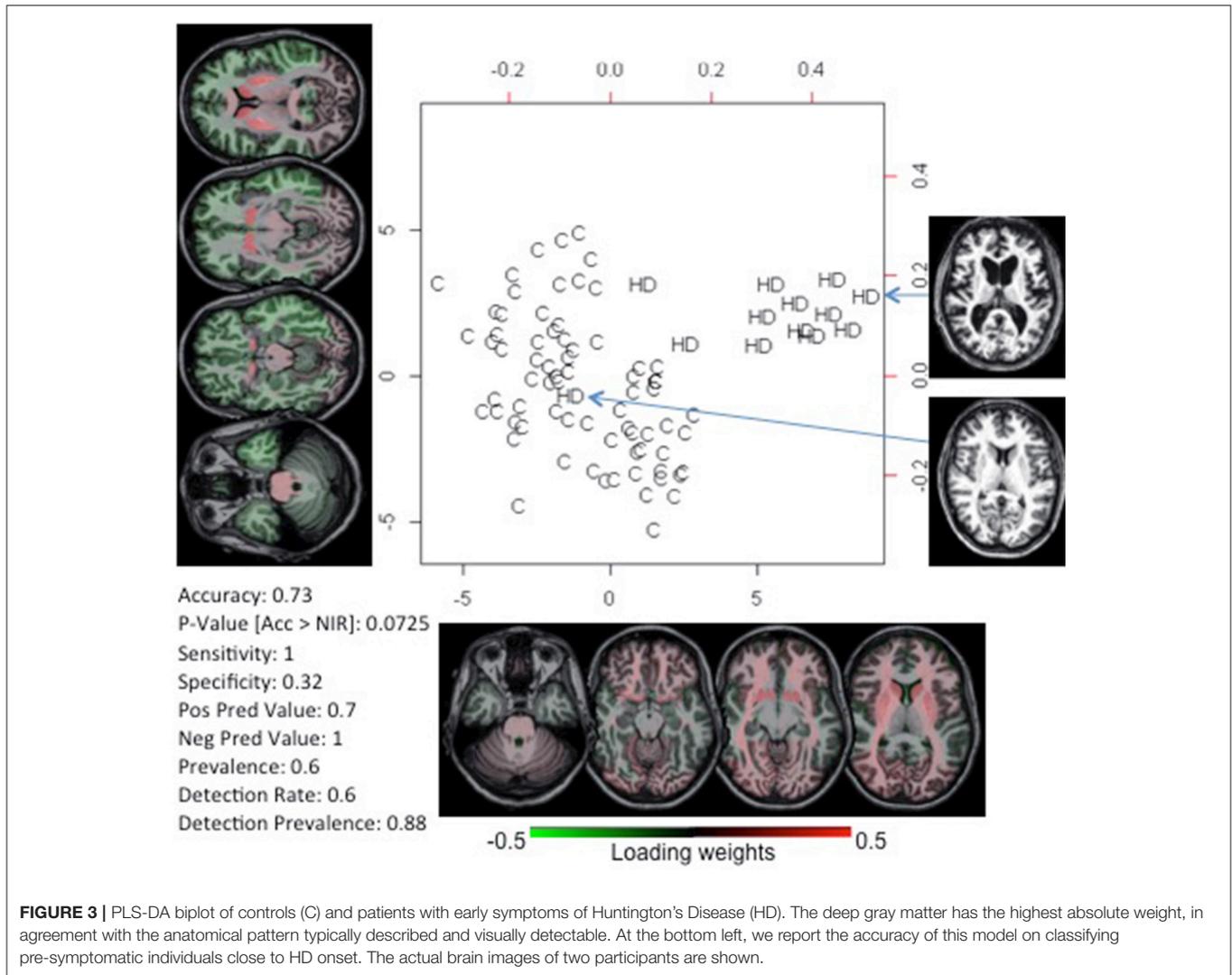


the segregation between the two groups. The cerebellum had the highest loading, i.e., the cerebellar atrophy played a major role on the classification, in agreement with the well-known and apparent cerebellar atrophy in ataxia. The highest absolute loadings of component 2 are diffusely distributed among the frontal, temporal and parietal lobes; it directly correlated with the degree of atrophy on these lobes, as measured by their volumetric z-score (Pearson rho of 0.72, 0.67, 0.61 for frontal, temporal, and parietal, respectively), and inversely correlated with age (rho = -0.77). Therefore, we infer that component 2 reflects age-related atrophy in individuals with AT.

### Huntington's Disease: Prediction

We tested whether we could correctly classify individuals with pre-symptomatic HD using the anatomic features of individuals

with early symptomatic HD. The goal was to use HD as a model to predict conversion to a specific anatomical phenotype rather than to diagnose HD, which can be done precisely by genetic tests. The classifier was built with individuals with early symptoms ( $n = 13$ ) vs. paired controls, and tested in pre-symptomatic individuals close ( $n = 16$ ) and far ( $n = 23$ ) from the onset, vs. paired controls (Supplementary Table). Again, two components were enough to create a model with 73% accuracy in classifying pre-symptomatic individuals near to disease onset (Figure 3). The highest loading weights were in the striatum, as expected, based on the disease physiopathogeny. As described by previous studies, striatum atrophy can barely be determined at the individual level on the pre-symptomatic stage, although it can be detected quantitatively, at the group level, up to 15 years before clinical onset. In addition, the early-symptomatic



HD group is anatomically heterogeneous, with some individuals presenting very clear striatum atrophy and others being very close to normal (Figure 3). This indicates that in certain disease types or at certain stages of a disease, the anatomy may not encode enough information to provide diagnosis for all patients. Regardless, we were effective enough in capturing and using this feature for the individual classification. The model did not achieve accuracy significantly higher than the by-chance for classifying pre-symptomatic individuals far from HD onset.

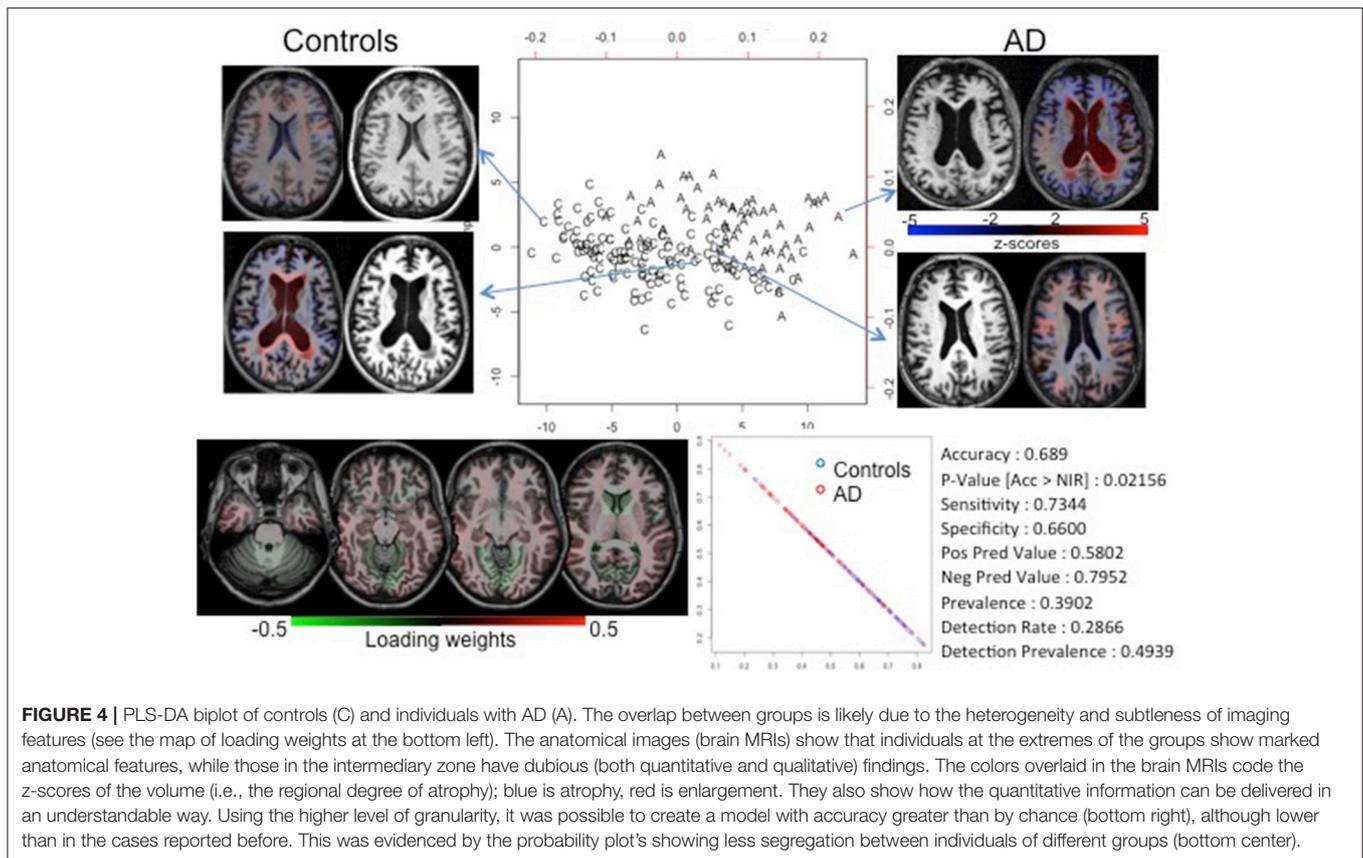
### Alzheimer's Disease: Classification of Diseases with Subtle or Heterogeneous Abnormalities

Unlike in ataxia and HD, the atrophy in most of the neurodegenerative diseases is detectable at the late stage of the disease and is regionally heterogeneous. This is the case with AD. We achieved a reasonable accuracy (69%) in diagnosing AD (model built in 33 AD individuals vs. paired controls, and tested in independent 33 AD individuals vs. paired controls;

see Supplementary Table), significantly higher than the by-chance classification. However, there was an enormous overlap among groups, as notable in the PLS-DA plot and in the probability plot that represents the chance of each individual's belonging to each group (Figure 4). The loading-weights map showed no distinguishing features; the weights are comparable and widespread, indicating that the anatomy in AD is mildly or heterogeneously affected, which can be confirmed by visual inspection of the brain MRIs.

### Primary Progressive Aphasia: Binning by Anatomical Phenotype

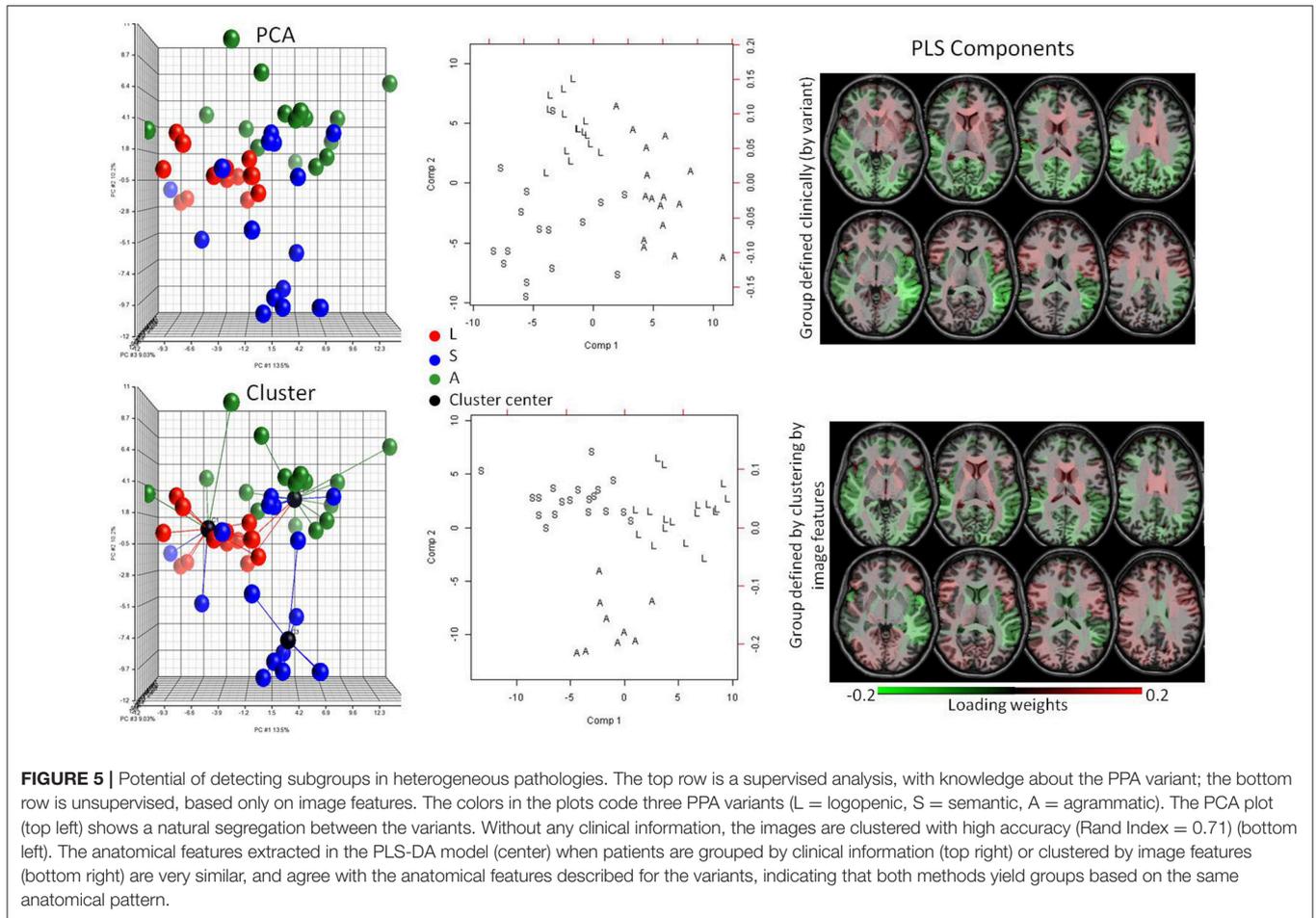
As mentioned in the previous section, increasingly therapies are targeting the early stages of neurodegenerative diseases. However, accurate diagnosis is more difficult because of the lack of clear and/or specific clinical deficits. At this stage, the initial stratification of the heterogeneous patient population is of critical importance. The difficulty arises because potential patient subgroups are degenerate both in the clinical and anatomical



domains. In this case, we are interested less in correlation between present diagnosis and anatomical features (because the diagnosis based on clinical information cannot separate important subgroups) and more in expanding the patient populations using both clinical and anatomical manifestations, potentially identifying a way to define subgroups. Binning a disease into subgroups may facilitate the design of therapies and the creation of predictive models because the subgroups may be related to specific pathological substrates, deficits or prognoses. We used PPA as a model system because of the existence of three well-known clinical variants. The knowledge of their anatomical correlates, albeit loose, could serve as our gold standard. In the PCA of the anatomical features (the regional volumes) there was a natural segregation into three clinically labeled groups (Figure 5). By clustering the data using only the anatomical features, we found groups that accurately agreed with the variant diagnosis (Rand Index = 0.71). Then, by using PLS-DA and extracting the loading weights, we confirmed that the features for automated classification according to clusters agreed with those for the classification according to clinical diagnosis. In addition, these anatomical features agreed with what is clinically defined for the variants, such as predominance of atrophy in the left temporal lobe for the Semantic variant, in the inferior parietal for the Logopenic, and in the inferior frontal lobe and the insula for the Agrammatic (Gorno-Tempini et al., 2004).

## DISCUSSION

We evaluated the performance of structure-based computational analysis on extracting anatomical features, previously described by human experience and a priori biological knowledge, in specific patient populations. Previously, we tested the robustness of our automated quantification approach against different image protocols and scanners, using subjects with different patterns and degrees of brain atrophy, and compared our conclusions with those of trained clinicians using visual analysis (Djamanakova et al., 2013; Faria et al., 2015; Liang et al., 2015). In the present study, we tested whether we could classify individuals and anatomically characterize different diseases in simulated clinical scenarios. Our database contains diverse image protocols and scanners. The demographic information taken into account by the linear classifiers include only age and gender, which are always clinically available. Although we could create better classification models by adding other clinical information, homogenizing the dataset, or using classifiers more sophisticated than PLS-DA, this would reduce the potential for generalization and translation to real clinical situations. In summary, rather than the greatness of classification, our aim was to create models robust enough to be translated to clinical practice, and at least in a first step, perform as well as clinicians in terms of extraction of important anatomical features and detection of anatomical patterns, helping to fill the semantic gap.



## Detection of Abnormal Imaging Patterns

In a disease with a clear anatomical phenotype (Ataxia), we obtained 87.5% accuracy, using a small sample size of patients in different stages of the disease. More important, the anatomical features extracted agreed with what is previously described as the hallmark of Ataxia (cerebellar and brainstem atrophy) (Klockgether et al., 1998; Schulz et al., 1999; Eichler et al., 2011; Reetz et al., 2013). The maps of the loading weights and the visual inspection of the images (Figure 2) reveal that the first component carries mostly information about the disease's anatomical phenotype, while the second component basically reflects brain atrophy directly related to age. Thus, the components extracted carry biological meaning, i.e., they contain information that can be interpreted in the light of actual medical knowledge because they both (our quantification tool and the medical knowledge) are based at the level of anatomical structures. In consequence, the classification models and the feature extraction machinery can be easily interpreted and translated to clinical practice. Although this result is purely confirmatory, the quantitative and systematic characterization of the anatomical feature in the PLS-DA space may give us an interesting clue about the patient status. For example, if there are ataxia patients who not only have the typical ataxia

feature (component 1), but also are located at an unexpected position in component 2 (i.e., accelerated whole-brain atrophy related to age), this may correlate with poor future outcomes. Thus, a quantitative approach of this type could provide new insight into diagnosis and prognosis, further facilitating research.

To investigate the prognostic value of quantitative anatomical description, we tested the classification performance in diseases where the anatomy clearly correlates with the time course, applying the classifiers in stages where the abnormal features couldn't be detected visually, at the individual level. In other words, we tested the potential for prognostic prediction using the HD population. We achieved 73% accuracy in classifying pre-symptomatic HD individuals, with a model based on features of early symptomatic HD individuals. The feature selection identified the deep gray matter as the most important region for the classification, again agreeing with the physiopathology of HD (Figure 3) (Aylward et al., 2000; Nopoulos et al., 2010; Paulsen et al., 2010, 2014a,b; Guo et al., 2012; Delmaire et al., 2013; Georgiou-Karistianis et al., 2013; Faria et al., 2016). HD is a genetic disease where the product of genetic load and age correlates very well with the time to onset (Ross et al., 2014). Therefore, one can reasonably argue that predictive models

based on imaging features are useless. The same applies to Ataxia to some extent. However, our aim was not to diagnosis HD or Ataxia. These diseases were taken as models for proof of concept because the gold standard (clinical diagnosis) is well-established. The aim was to evaluate the structure-based automated quantification approach, in terms of feature selection and robustness against heterogeneous datasets, and its potential to detect features that go beyond the artifactual noise. Particularly in HD, the potential for classifying pre-symptomatic individuals surpasses what can be done with clinical imaging analysis because the subtle abnormalities are not visually detectable at the individual level (Paulsen et al., 2008).

## Potential for Binning in More Homogeneous Phenotypes

Unlike diseases with a clear anatomical phenotype, those that embrace heterogeneous anatomical and clinical phenotypes, or subtle abnormalities, or unknown time courses, offer extra challenges for both visual and automated analysis. This is, for instance, the case with AD. To date, there are about 100 models for predicting conversion from mild cognitive impairment to AD, based on imaging. A PubMed search for “Prediction of MCI to AD conversion MRI” reveals 96 publications in the last 10 years; for more recent reviews, please see (Shaffer et al., 2013; Sanchez-Catusus et al., 2017). Either they achieve unsatisfactory accuracy, or high accuracy at the cost of overfitting, or they are late in the disease course. As a result, we can generalize by saying that there is, as yet, no effective prediction useful in clinical scenarios. **Figure 4** may offer some clues about why this happens. Our classification model achieved <70% accuracy. There is substantial overlap between controls and AD in the PLS-DA plot, and there is no predominant weight in the loadings of component 1. Visual inspection of the images reveals that both groups (control and AD) are heterogeneous in terms of atrophy pattern and degree at this age range. This explains why the individual classification, by visual radiological analysis, is also ineffective.

The source of this challenge is two-fold. First, it is possible that anatomy is not encoding enough information to characterize the pathology reliably. Second, because we do not have strongly discriminating factors, both in clinical and imaging information, the stratification of the patient population is incomplete. For example, if AD is actually a syndrome caused by multiple pathologies with multiple anatomical manifestations, AD's common anatomical features cannot be extracted. In this situation, we need to resort to different study designs, using both clinical and imaging features to stratify the population. Models such as AD provide opportunities to investigate the existence of subgroups, with certain anatomical expression, that can behave as specific entities in some clinical domains. For instance, in Ataxia (**Figure 2**) one can see a subtle spread of patients along the component that differentiates the groups (component 1). Hypothetically, this spread may reflect the effect of a correlated feature, such as disease severity. Similarly in AD or other heterogeneous disease models, there may be a non-orthogonal axis that represents an unknown variable. With regression in

this axis, it is possible to detect the subgroups that, for instance, respond differently to therapeutics, or have different prognosis.

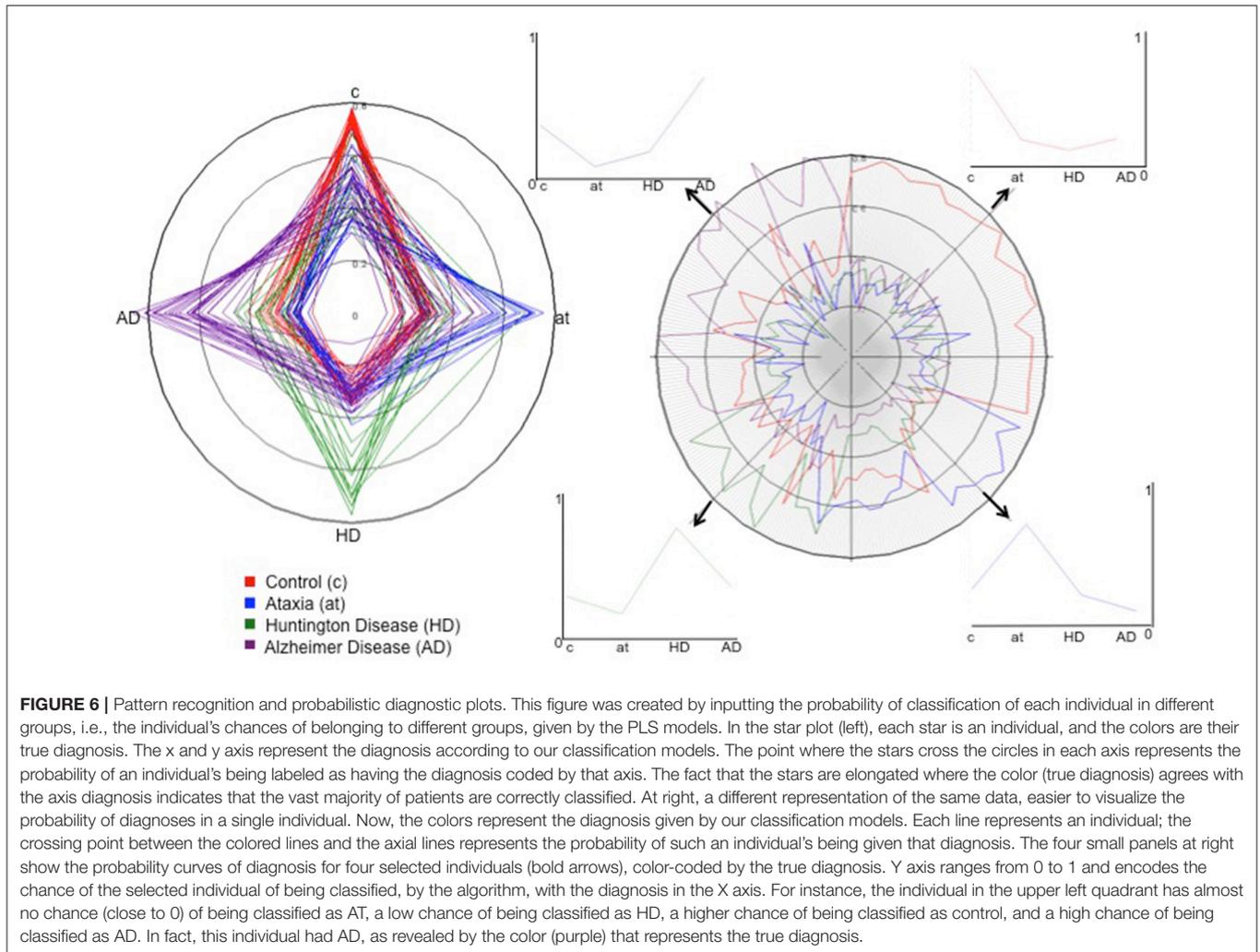
To investigate the potential of the automated structure-based quantification to binning an entity into subgroups of clinical relevance, we used individuals with PPA. PPA, a neurodegenerative clinical syndrome characterized by decline in language ability 2 years before any other cognitive deficit, is an ideal condition to investigate the clustering in sub-phenotypes, since three variants loosely correlated with underlying pathologies and with certain anatomical representation are described (Gorno-Tempini et al., 2011; Rohrer and Rosen, 2013). Although there is still no treatment for PPA, there is hope that certain therapies can be effective for specific variants (Cadorio et al., 2017). Now, suppose that the three variants are yet unknown. An unsupervised PCA plot shows a natural segregation of the data into two or three subgroups (**Figure 5**), but because the variants are hypothetically unknown, one cannot explain the data variance with clinical labels. An unsupervised hierarchical cluster shows the data divided into subgroups that correlate very well with the real variant's diagnosis. The image features selected for classification in these clusters (bottom row, **Figure 5**) agree with those selected for classification according to the real variant's diagnosis (top row, **Figure 5**) and also to those that are described as hallmarks for the variants (Turner et al., 1996; Rohrer et al., 2009; Shim et al., 2012; Zhang et al., 2013; Agosta et al., 2015; Botha et al., 2015; Bisenius et al., 2016), proving the potential of our approach to identify subgroups of clinical relevance.

## Deliverables

Subgrouping can be extrapolated to individuals, i.e., the detection of outliers in terms of anatomy may point to individuals who may be unique in additional domains. For instance, in HD (**Figure 3**) anatomical heterogeneity still remains among the genetically homogenized group, as there is at least one individual with visually normal anatomy. It is an open question if this anatomical variability has any predictive value for prognosis, to be answered by quantitative and systematic characterization of this population.

Another potential deliverable is the diagnostic probability map for each individual (**Figure 6**). Given a database large enough to contain various pathologies and the high variability of imaging protocols and age range for controls and patients, it is possible to calculate the probability of differential diagnosis for a new individual, as shown in **Figure 6**. In this example, one can reasonably argue that it is clinically improbable to have HD, Ataxia and AD as differential diagnoses. Again, these diseases were taken as proof of concept, because they all have the same basic anatomic feature (atrophy) and a clear clinical diagnosis used as the gold-standard. The concept of diagnostic probability graphics can be extended to more plausible clinical scenarios.

Finally, the potential for aiding clinical interpretation and education may be a valuable low-hanging fruit. The simple use of z-score maps (**Figure 4**) may confirm or exclude a clinical impression and speed up the radiological reading. Also, having a database big and heterogeneous enough, and coupling image and text information (such as diagnosis, prognosis, response to



treatments, etc.), it is possible to perform a direct image search, producing static reports about similar phenotypes. For example, given a new subject image, it is possible to search in a big dataset for dozens of images with similar features linked to information of clinical relevance.

## Limitations

This study is based on a single image variable, the volume. One of the greatest advantages of the structure-based approach is that it allows the combination of many other features, such as T2 contrast, diffusion tensor image indices, functional MRI correlations, metabolite concentration, and others, as we demonstrated in previous studies (Faria et al., 2012). Although there are big challenges in combining features of different domains (e.g., drawbacks on feature concatenation methods, variation among clinical protocols barring the creation of common databases for certain domains, the need for a priori knowledge of noise in order to create models for easy generalization), multi-domain structure-based analysis is a promising strategy for conditions with no single dominant discriminating feature.

An important constraint of the structure-based analysis is that any quantitative characterization and classification model will be limited by the pre-defined space. In other words, if the anatomical pattern does not respect the boundaries of a given parcellation scheme, the abnormality can be overlooked. One strategy to ameliorate this issue is to use different levels of granularity. So one can analyze the data in parcels as big as a hemisphere, or as small as a cranial nerve, which is actually smaller than the gaussian filters traditionally used for voxel-based analysis. However, if the strategy is to replicate the radiological interpretation, then structure-based analysis is intuitively a better solution because visual inspection occurs at the structural level, not at the voxel level.

## Perspectives

We explored the performance of structure-based computational analysis in simulated clinical scenarios. The pillars of this approach are automated and accurate quantification, reliability and robustness against artifactual noise, easy interpretation of selected features, and a knowledge repository that is a large database as heterogeneous as possible both in terms of

pathologies and image protocols. The deliverables are diverse, from the image quantification itself, through to the image pattern search, to the diagnostic aid. Although the impact of this method is yet to be tested, it has potential educational value, it may reduce the time for radiological reading, or it may work as second reader in locations where sub-specialized radiologists are not available. In any case, because no such tool can be directly applicable to clinical practice, any positive impact is valuable. In addition, electronic structured databases and search engines are the basis of high throughput image analysis and may represent the migration of brain MRI to the BigData era, contributing to the emergent field of Precision Medicine.

## ETHICS STATEMENT

This study was carried out with data (1) from publically available databases, and (2) from Johns Hopkins Hospital, originally collected for other studies. These studies were approved and carried out in accordance with the recommendations of the local IRB, with written informed consent from all subjects, in accordance with the Declaration of Helsinki.

## REFERENCES

- Agosta, F., Ferraro, P. M., Canu, E., Copetti, M., Galantucci, S., Magnani, G., et al. (2015). Differentiation between subtypes of primary progressive aphasia by using cortical thickness and diffusion-tensor MR imaging measures. *Radiology* 276, 219–227. doi: 10.1148/radiol.15141869
- Akgul, C. B., Rubin, D. L., Napel, S., Beaulieu, C. F., Greenspan, H., and Acar, B. (2011). Content-based image retrieval in radiology: current status and future directions. *J. Digit. Imaging* 24, 208–222. doi: 10.1007/s10278-010-9290-9
- Albert, M. S., Dekosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., et al. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers. Dement.* 7, 270–279. doi: 10.1016/j.jalz.2011.03.008
- Apostolova, L. G., Steiner, C. A., Akopyan, G. G., Dutton, R. A., Hayashi, K. M., Toga, A. W., et al. (2007). Three-dimensional gray matter atrophy mapping in mild cognitive impairment and mild Alzheimer disease. *Arch. Neurol.* 64, 1489–1495. doi: 10.1001/archneur.64.10.1489
- Artaechevarria, X., Munoz-Barrutia, A., and Ortiz-De-Solorzano, C. (2009). Combination strategies in multi-atlas image segmentation: application to brain MR data. *IEEE Trans. Med. Imaging* 28, 1266–1277. doi: 10.1109/TMI.2009.2014372
- Aylward, E. H., Codori, A. M., Rosenblatt, A., Sherr, M., Brandt, J., Stine, O. C., et al. (2000). Rate of caudate atrophy in presymptomatic and symptomatic stages of Huntington's disease. *Mov. Disord.* 15, 552–560. doi: 10.1002/1531-8257(200005)15:3<552::AID-MDS1020>3.0.CO;2-P
- Bisenius, S., Neumann, J., and Schroeter, M. L. (2016). Validating new diagnostic imaging criteria for primary progressive aphasia via anatomical likelihood estimation meta-analyses. *Eur. J. Neurol.* 23, 704–712. doi: 10.1111/ene.12902
- Botha, H., Duffy, J. R., Whitwell, J. L., Strand, E. A., Machulda, M. M., Schwarz, C. G., et al. (2015). Classification and clinicoradiologic features of primary progressive aphasia (PPA) and apraxia of speech. *Cortex* 69, 220–236. doi: 10.1016/j.cortex.2015.05.013
- Butteriss, D., Chinnery, P., and Birchall, D. (2005). Radiological characterization of spinocerebellar ataxia type 6. *Br. J. Radiol.* 78, 694–696. doi: 10.1259/bjr/73834093
- Cadorio, I., Lousada, M., Martins, P., and Figueiredo, D. (2017). Generalization and maintenance of treatment gains in primary progressive aphasia (PPA): a systematic review. *Int. J. Lang. Commun. Disord.* 52, 543–560. doi: 10.1111/1460-6984.12310
- Ceritoglu, C., Oishi, K., Li, X., Chou, M. C., Younes, L., Albert, M., et al. (2009). Multi-contrast large deformation diffeomorphic metric mapping for diffusion tensor imaging. *Neuroimage* 47, 618–627. doi: 10.1016/j.neuroimage.2009.04.057
- Delmaire, C., Dumas, E. M., Sharman, M. A., van den Bogaard, S. J., Valabregue, R., Jauffret, C., et al. (2013). The structural correlates of functional deficits in early huntington's disease. *Hum. Brain Mapp.* 34, 2141–2153. doi: 10.1002/hbm.22055
- Depeursinge, A., Fischer, B., Muller, H., and Deserno, T. M. (2011). Prototypes for content-based image retrieval in clinical practice. *Open Med. Inform. J.* 5, 58–72. doi: 10.2174/1874431101105010058
- Djamanakova, A., Faria, A. V., Hsu, J., Ceritoglu, C., Oishi, K., Miller, M. I., et al. (2013). Diffeomorphic brain mapping based on T1-weighted images: improvement of registration accuracy by multichannel mapping. *J. Magn. Reson. Imaging* 37, 76–84. doi: 10.1002/jmri.23790
- Djamanakova, A., Tang, X., Li, X., Faria, A. V., Ceritoglu, C., Oishi, K., et al. (2014). Tools for multiple granularity analysis of brain MRI data for individualized image analysis. *Neuroimage* 101, 168–716. doi: 10.1016/j.neuroimage.2014.06.046
- Eichler, L., Bellenberg, B., Hahn, H. K., Koster, O., Schols, L., and Lukas, C. (2011). Quantitative assessment of brain stem and cerebellar atrophy in spinocerebellar ataxia types 3 and 6: impact on clinical status. *AJNR Am. J. Neuroradiol.* 32, 890–897. doi: 10.3174/ajnr.A2387
- Faria, A. V., Joel, S. E., Zhang, Y., Oishi, K., van Zijl, P. C., Miller, M. I., et al. (2012). Atlas-based analysis of resting-state functional connectivity: evaluation for reproducibility and multi-modal anatomy-function correlation studies. *Neuroimage* 61, 613–621. doi: 10.1016/j.neuroimage.2012.03.078
- Faria, A. V., Oishi, K., Yoshida, S., Hillis, A., Miller, M. I., and Mori, S. (2015). Content-based image retrieval for brain MRI: an image-searching engine and population-based analysis to utilize past clinical data for future diagnosis. *Neuroimage Clin.* 7, 367–376. doi: 10.1016/j.nicl.2015.01.008
- Faria, A. V., Ratnanather, J. T., Tward, D. J., Lee, D. S., van den Noort, F., Wu, D., et al. (2016). Linking white matter and deep gray matter alterations in premanifest Huntington disease. *Neuroimage Clin.* 11, 450–460. doi: 10.1016/j.nicl.2016.02.014

## AUTHOR CONTRIBUTIONS

All the authors are accountable for all aspects of the work and approved the final version. In addition, AF and SM are responsible for the conception and design, interpretation of data, and drafting the work. ZL is responsible for data acquisition and analysis. MIM revised it critically for important intellectual content.

## ACKNOWLEDGMENTS

We are grateful to the individuals who participated in this research. We thank Dr. Argye Hillis, Dr. Christopher Ross, and Dr. Sarah Ying, for data sharing. This research was possible due to the following grants, from the National Institutes of Health: P41EB015909 (SM and MIM), R01EB017638 (MIM), and R01NS084957 (SM).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2017.00578/full#supplementary-material>

- Frisoni, G. B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., et al. (2017). Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *Lancet Neurol.* 16, 661–676. doi: 10.1016/S1474-4422(17)30159-X
- Georgiou-Karistianis, N., Scahill, R., Tabrizi, S. J., Squitieri, F., and Aylward, E. (2013). Structural MRI in Huntington's disease and recommendations for its potential use in clinical trials. *Neurosci. Biobehav. Rev.* 37, 480–490. doi: 10.1016/j.neubiorev.2013.01.022
- Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., et al. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Ann. Neurol.* 55, 335–346. doi: 10.1002/ana.10825
- Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., et al. (2011). Classification of primary progressive aphasia and its variants. *Neurology* 76, 1006–1014. doi: 10.1212/WNL.0b013e31821103e6
- Guo, Z., Rudow, G., Pletnikova, O., Codispoti, K. E., Orr, B. A., Crain, B. J., et al. (2012). Striatal neuronal loss correlates with clinical motor impairment in Huntington's disease. *Mov. Disord.* 27, 1379–1386. doi: 10.1002/mds.25159
- Huntington, G. (1872). On Chorea. *Med. Surg. Rep.* 26:2.
- Hwang, K. H., Lee, H., and Choi, D. (2012). Medical image retrieval: past and present. *Healthc. Inform. Res.* 18, 3–9. doi: 10.4258/hir.2012.18.1.3
- Jack, C. R. Jr., Albert, M. S., Knopman, D. S., McKhann, G. M., Sperling, R. A., Carrillo, M. C., et al. (2011). Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers. Dement.* 7, 257–262. doi: 10.1016/j.jalz.2011.03.004
- Jia, H., Yap, P. T., and Shen, D. (2012). Iterative multi-atlas-based multi-image segmentation with tree-based registration. *Neuroimage* 59, 422–430. doi: 10.1016/j.neuroimage.2011.07.036
- Kalpathy-Cramer, J., de Herrera, N. A. G., Demner-Fushman, D., Antani, S., Bedrick, S., and Muller, H. (2015). Evaluating performance of biomedical image retrieval systems—an overview of the medical image retrieval task at ImageCLEF 2004–2013. *Comput. Med. Imaging Graph.* 39, 55–61. doi: 10.1016/j.compmedimag.2014.03.004
- Klockgether, T., Skalej, M., Wedekind, D., Luft, A. R., Welte, D., Schulz, J. B., et al. (1998). Autosomal dominant cerebellar ataxia type I. MRI-based volumetry of posterior fossa structures and basal ganglia in spinocerebellar ataxias types 1, 2 and 3. *Brain* 121(Pt. 9), 1687–1693. doi: 10.1093/brain/121.9.1687
- Langerak, T. R., van der Heide, U. A., Kotte, A. N., Viergever, M. A., Van Vulpen, M., and Pluim, J. P. (2010). Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). *IEEE Trans. Med. Imaging* 29, 2000–2008. doi: 10.1109/TMI.2010.2057442
- Leyton, C. E., Britton, A. K., Hodges, J. R., Halliday, G. M., and Kril, J. J. (2016). Distinctive pathological mechanisms involved in primary progressive aphasias. *Neurobiol. Aging* 38, 82–92. doi: 10.1016/j.neurobiolaging.2015.10.017
- Liang, Z., He, X., Ceritoglu, C., Tang, X., Li, Y., Kutten, K. S., et al. (2015). Evaluation of cross-protocol stability of a fully automated brain multi-atlas parcellation tool. *PLoS ONE* 10:e0133533. doi: 10.1371/journal.pone.0133533
- Lotjonen, J. M., Wolz, R., Koikkalainen, J. R., Thurfjell, L., Waldemar, G., Soininen, H., et al. (2010). Fast and robust multi-atlas segmentation of brain magnetic resonance images. *Neuroimage* 49, 2352–2365. doi: 10.1016/j.neuroimage.2009.10.026
- Ma, J., Ma, H. T., Li, H., Ye, C., Wu, D., Tang, X., et al. (2015). A fast atlas pre-selection procedure for multi-atlas based brain segmentation. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2015, 3053–3056. doi: 10.1109/EMBC.2015.7319036
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. Jr., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers. Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005
- Mesulam, M. M. (1982). Slowly progressive aphasia without generalized dementia. *Ann. Neurol.* 11, 592–598. doi: 10.1002/ana.410110607
- Nopoulos, P. C., Aylward, E. H., Ross, C. A., Johnson, H. J., Magnotta, V. A., Juhl, A. R., et al. (2010). Cerebral cortex structure in prodromal Huntington disease. *Neurobiol. Dis.* 40, 544–554. doi: 10.1016/j.nbd.2010.07.014
- Orphanoudakis, S. C., Chronaki, C. E., and Vamvaka, D. (1996). I2Cnet: Content-based similarity search in geographically distributed repositories of medical images. *Comput. Med. Imaging Graph.* 20, 193–207. doi: 10.1016/S0895-6111(96)00013-4
- Paulsen, J. S., Langbehn, D. R., Stout, J. C., Aylward, E., Ross, C. A., Nance, M., et al. (2008). Detection of Huntington's disease decades before diagnosis: the Predict-HD study. *J. Neurol. Neurosurg. Psychiatr.* 79, 874–880. doi: 10.1136/jnnp.2007.128728
- Paulsen, J. S., Long, J. D., Johnson, H. J., Aylward, E. H., Ross, C. A., Williams, J. K., et al. (2014a). Clinical and biomarker changes in premanifest huntington disease show trial feasibility: a decade of the PREDICT-HD Study. *Front. Aging Neurosci.* 6:78. doi: 10.3389/fnagi.2014.00078
- Paulsen, J. S., Long, J. D., Ross, C. A., Harrington, D. L., Erwin, C. J., Williams, J. K., et al. (2014b). Prediction of manifest Huntington's disease with clinical and imaging measures: a prospective observational study. *Lancet Neurol.* 13, 1193–1201. doi: 10.1016/S1474-4422(14)70238-8
- Paulsen, J. S., Nopoulos, P. C., Aylward, E., Ross, C. A., Johnson, H., Magnotta, V. A., et al. (2010). Striatal and white matter predictors of estimated diagnosis for Huntington disease. *Brain Res. Bull.* 82, 201–207. doi: 10.1016/j.brainresbull.2010.04.003
- Penney, J. B. Jr., Vonsattel, J. P., Macdonald, M. E., Gusella, J. F., and Myers, R. H. (1997). CAG repeat number governs the development rate of pathology in Huntington's disease. *Ann. Neurol.* 41, 689–692. doi: 10.1002/ana.410410521
- Pinho, E., Godinho, T., Valente, F., and Costa, C. (2017). A multimodal search engine for medical imaging studies. *J. Digit. Imaging* 30, 39–48. doi: 10.1007/s10278-016-9903-z
- Reetz, K., Costa, A. S., Mirzazade, S., Lehmann, A., Juzek, A., Rakowicz, M., et al. (2013). Genotype-specific patterns of atrophy progression are more sensitive than clinical decline in SCA1, SCA3 and SCA6. *Brain* 136, 905–917. doi: 10.1093/brain/aww369
- Rohrer, J. D., and Rosen, H. J. (2013). Neuroimaging in frontotemporal dementia. *Int. Rev. Psychiatry* 25, 221–229. doi: 10.3109/09540261.2013.778822
- Rohrer, J. D., Warren, J. D., Modat, M., Ridgway, G. R., Douiri, A., Rossor, M. N., et al. (2009). Patterns of cortical thinning in the language variants of frontotemporal lobar degeneration. *Neurology* 72, 1562–1569. doi: 10.1212/WNL.0b013e3181a4124e
- Ross, C. A., Aylward, E. H., Wild, E. J., Langbehn, D. R., Long, J. D., Warner, J. H., et al. (2014). Huntington disease: natural history, biomarkers and prospects for therapeutics. *Nat. Rev. Neurol.* 10, 204–216. doi: 10.1038/nrneuro.2014.24
- Sabuncu, M. R., Yeo, B. T., Van Leemput, K., Fischl, B., and Golland, P. (2010). A generative model for image segmentation based on label fusion. *IEEE Trans. Med. Imaging* 29, 1714–1729. doi: 10.1109/TMI.2010.2050897
- Sanchez-Catasus, C. A., Stormezand, G. N., van Laar, P. J., De Deyn, P. P., Sanchez, M. A., and Dierckx, R. A. (2017). FDG-PET for Prediction of AD Dementia in Mild Cognitive Impairment. A review of the state of the art with particular emphasis on the comparison with other neuroimaging modalities (MRI and Perfusion SPECT). *Curr. Alzheimer. Res.* 14, 127–142. doi: 10.2174/1567205013666160629081956
- Schulz, J. B., Skalej, M., Wedekind, D., Luft, A. R., Abele, M., Voigt, K., et al. (1999). Magnetic resonance imaging-based volumetry differentiates idiopathic Parkinson's syndrome from multiple system atrophy and progressive supranuclear palsy. *Ann. Neurol.* 45, 65–74. doi: 10.1002/1531-8249(199901)45:1<65::AID-ART12>3.0.CO;2-1
- Shaffer, J. L., Petrella, J. R., Sheldon, F. C., Choudhury, K. R., Calhoun, V. D., Coleman, R. E., et al. (2013). Predicting cognitive decline in subjects at risk for Alzheimer disease by using combined cerebrospinal fluid, MR imaging, and PET biomarkers. *Radiology* 266, 583–591. doi: 10.1148/radiol.12120010
- Shim, H., Hurley, R. S., Rogalski, E., and Mesulam, M. M. (2012). Anatomical, clinical, and neuropsychological correlates of spelling errors in primary progressive aphasia. *Neuropsychologia* 50, 1929–1935. doi: 10.1016/j.neuropsychologia.2012.04.017
- Spanier, A. B., Cohen, D., and Joskowicz, L. (2017). A new method for the automatic retrieval of medical cases based on the RadLex ontology. *Int. J. Comput. Assist. Radiol. Surg.* 12, 471–484. doi: 10.1007/s11548-016-1496-y
- Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., et al. (2011). Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers. Dement.* 7, 280–292. doi: 10.1016/j.jalz.2011.03.003
- Tabrizi, S. J., Langbehn, D. R., Leavitt, B. R., Roos, R. A., Durr, A., Craufurd, D., et al. (2009). Biological and clinical manifestations of Huntington's disease in

- the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. *Lancet Neurol.* 8, 791–801. doi: 10.1016/S1474-4422(09)70170-X
- Tabrizi, S. J., Reilmann, R., Roos, R. A., Durr, A., Leavitt, B., Owen, G., et al. (2012). Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: analysis of 24 month observational data. *Lancet Neurol.* 11, 42–53. doi: 10.1016/S1474-4422(11)70263-0
- Tabrizi, S. J., Scahill, R. I., Owen, G., Durr, A., Leavitt, B. R., Roos, R. A., et al. (2013). Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data. *Lancet Neurol.* 12, 637–649. doi: 10.1016/S1474-4422(13)70088-7
- Tang, X., Oishi, K., Faria, A. V., Hillis, A. E., Albert, M. S., Mori, S., et al. (2013). Bayesian parameter estimation and segmentation in the multi-atlas random orbit model. *PLoS ONE* 8:e65591. doi: 10.1371/journal.pone.0065591
- Turner, R. S., Kenyon, L. C., Trojanowski, J. Q., Gonatas, N., and Grossman, M. (1996). Clinical, neuroimaging, and pathologic features of progressive nonfluent aphasia. *Ann. Neurol.* 39, 166–173. doi: 10.1002/ana.410390205
- van Rikxoort, E. M., Isgum, I., Arzhaeva, Y., Staring, M., Klein, S., Viergever, M. A., et al. (2010). Adaptive local multi-atlas segmentation: application to the heart and the caudate nucleus. *Med. Image Anal.* 14, 39–49. doi: 10.1016/j.media.2009.10.001
- Wang, H., Pouch, A., Takabe, M., Jackson, B., Gorman, J., Gorman, R., et al. (2013). Multi-atlas segmentation with robust label transfer and label fusion. *Inf. Process. Med. Imaging* 23, 548–559. doi: 10.1007/978-3-642-38868-2\_46
- Wang, L., Beg, F., Ratnanather, T., Ceritoglu, C., Younes, L., Morris, J. C., et al. (2007). Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the Alzheimer type. *IEEE Trans. Med. Imaging* 26, 462–470. doi: 10.1109/TMI.2006.887380
- Warfield, S. K., Zou, K. H., and Wells, W. M. (2004). Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans. Med. Imaging* 23, 903–921. doi: 10.1109/TMI.2004.828354
- Wilson, S. M., Demarco, A. T., Henry, M. L., Gesierich, B., Babiak, M., Miller, B. L., et al. (2016). Variable disruption of a syntactic processing network in primary progressive aphasia. *Brain* 139:11. doi: 10.1093/brain/aww218
- Wu, D., Ma, T., Ceritoglu, C., Li, Y., Chotiyanonta, J., Hou, Z., et al. (2016). Resource atlases for multi-atlas brain segmentations with multiple ontology levels based on T1-weighted MRI. *Neuroimage* 125, 120–130. doi: 10.1016/j.neuroimage.2015.10.042
- Zhang, X., Mormino, E. C., Sun, N., Sperling, R. A., Sabuncu, M. R., and Yeo, B. T. (2016). Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. *Proc. Natl. Acad. Sci. U.S.A.* 113, e6535–e6544. doi: 10.1073/pnas.1611073113
- Zhang, Y., Tartaglia, M. C., Schuff, N., Chiang, G. C., Ching, C., Rosen, H. J., et al. (2013). MRI signatures of brain macrostructural atrophy and microstructural degradation in frontotemporal lobar degeneration subtypes. *J. Alzheimers. Dis.* 33, 431–444. doi: 10.3233/JAD-2012-121156
- Zhuchenko, O., Bailey, J., Bonnen, P., Ashizawa, T., Stockton, D. W., Amos, C., et al. (1997). Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. *Nat. Genet.* 15, 62–69. doi: 10.1038/ng0197-62

**Conflict of Interest Statement:** SM and MM own AnatomyWorks. SM is its CEO. This arrangement is being managed by the Johns Hopkins University in accordance with its conflict-of-interest policies.

The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Faria, Liang, Miller and Mori. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.