Understanding Atrophy Trajectories in Alzheimer’s Disease Using Association Rules on MRI Images

György J. Simon  
Dept. of Health Sciences Research, Mayo Clinic, Rochester, MN  
simon.gyorgy@mayo.edu

Peter W. Li  
Dept. of Health Sciences Research, Mayo Clinic, Rochester, MN  
lipeter@mayo.edu

Prashanthi Vemuri  
Dept. of Radiology, Mayo Clinic, Rochester, MN  
vemuri.prashanthi@mayo.edu

Clifford R. Jack Jr.  
Dept. of Radiology, Mayo Clinic, Rochester, MN  
jack.clifford@mayo.edu

ABSTRACT

Alzheimer’s disease (AD) is associated with progressive cognitive decline leading to dementia. The atrophy/loss of brain structure as seen on Magnetic Resonance Imaging (MRI) is strongly correlated with the severity of the cognitive impairment in AD. In this paper, we set out to find associations between predefined regions of the brain (regions of interest; ROIs) and the severity of the disease. Specifically, we use these associations to address two important issues in AD: (i) typical versus atypical atrophy patterns and (ii) the origin and direction of progression of atrophy, which is currently under debate.

We observed that each AD-related ROI is associated with a wide range of severity and that the difference between ROIs is merely a difference in severity distribution. To model differences between the severity distribution of a sub-population (with significant atrophy in certain ROIs) and the severity distribution of the entire population, we developed the concept of Distributional Association Rules. Using the Distributional Association Rules, we clustered ROIs into disease subsystems. We define a disease subsystem as a contiguous set of ROIs that are collectively implicated in AD. AD is known to be heterogeneous in the sense that multiple sets of ROIs may be related to the disease in a population.

We proposed an enhancement to the association rule mining where the algorithm only discovers association rules with ROIs that form an approximately contiguous volume. Next, we applied these association rules to infer the direction of disease progression based on the support measures of the association rules. We also developed a novel statistical test to determine the statistical significance of the discovered direction.

We evaluated the proposed method on the Mayo Clinic Alzheimer’s Disease Research Center (ADRC) prospective patient cohorts. The key achievements of the methodology is that it accurately identified larger disease subsystems implicated in typical and atypical AD and it successfully mapped the directions of disease progression.

The wealth of data available in Radiology gives rise to opportunities for applying this methodology to map out the trajectory of several other diseases, e.g. other neurodegenerative diseases and cancers, most notably, breast cancer. The applicability of this method is not limited to image data, as associating predictors with severity provides valuable information in most areas of medicine as well as other industries.

Categories and Subject Descriptors

H.2.8 [Information Systems]: Data Mining; J.3 [Computer Applications]: Life and Medical Sciences

General Terms: Algorithms

Keywords: spatial association analysis, continuous outcome, MRI, disease progression

1. INTRODUCTION

Alzheimer’s disease (AD) is a multi-faceted disease wherein cumulative brain pathology results in progressive cognitive decline, which ultimately leads to dementia. Of all the biomarkers that are being currently investigated and evaluated for AD, loss/atrophy of the brain structures as seen on structural Magnetic Resonance Imaging (MRI) appears to strongly correlate with the severity of cognitive impairment in AD [10]. In our previous works, we found that MRI images are extremely useful for separating AD from cognitively normal individuals as well as for staging the disease severity [19]. In addition, we recently established a data-mining method to use MRI to differentiate AD from other neurodegenerative diseases [18]. We implemented this method by considering 118 regions of interest (ROI) in the brain which characterize the amount of gray matter density in each of these regions. One key observation from this study was that the atrophy in AD is multi-modal in nature in a sense that multiple different sets of ROIs are predictive of AD. This observation is supported by recent literature [2, 8]. In this study we set out to investigate whether the multi-modal na-
ture of the atrophy patterns is related to the severity of disease using spatial association rules.

The approach we took in this study was to extract the association between the atrophy in these ROIs and clinical severity of AD. Clinical severity was measured using the Clinical Dementia Rating–sum of boxes (CDR-SB) [in this paper we simply call this CDR] which assesses global functional performance in individual subjects [4]. CDR ranges from 0 to 18 in 0.5 increments with 0 representing normal functional performance. Since the relationship between atrophy and CDR is complex, we found that the relationship between gray matter in an ROI and CDR is best described by CDR distribution across patients with significant atrophy in the ROI. Our goal was to associate atrophy regions with differences in the CDR distribution between the affected population (patients with significant atrophy in the region) and the entire population. We then use the discovered associations to decipher brain loss trajectories in AD.

Evidence indicates that the pathology of AD tends to affect contiguous volumes of ROIs and the patterns of atrophy have been shown to originate at a specific ROI and to spread further away from the origin [5]. Since the ROIs are located in a three-dimensional Euclidean space, we can use spatial constraints to discover such near-contiguous sets of ROIs. We enhanced the association rule mining algorithm to only discover rules with near-contiguous sets of ROIs.

We evaluated the proposed method on the Mayo Clinic Alzheimer’s Disease Research Center (ADRC) prospective patient cohorts. The subjects used in this study were 120 cognitively normal (controls), 81 mild cognitive impairment (MCI) subjects who later progressed to AD and 48 AD subjects confirmed pathologically. The MCI subjects represent the lower end of the CDR spectrum, while the pathologically confirmed AD subjects represent the mid and higher end of the spectrum.

The major contributions of this work were as follows.

1. We developed Distributional Association Rules, a method for associating ROIs with differences in the distribution of severity scores.
2. We enhance the association rule discovery by incorporating spatial constraints imposed by the requirement that the AD subsystems be near contiguous volumes. Our definition of the constraint allows for elongated volumes that may arise as a result of the how the disease is spreading.
3. We proposed a method for determining the direction of disease progression and a statistical test to assess the significance of the direction.
4. We used higher-order association rules to identify AD subsystems. One of the subsystems we identified is known to be consistent with atypical AD.

2. RELATED WORKS

Association rule mining has been a popular technique since its introduction in [1]. Owing to its inherent ability to handle multi-modal data, its non-greedy nature and the wealth of efficient techniques for computing the rules, it has found applications in virtually all domains, including medical imaging. Zaïane et al. [20] present a classification system for mammography images, Pan et al. [14] applied association rules to ROI detection and Sheela et al. [16] extend the well-known Classification Based on Associations (CBA) algorithm [12] to handle quantitative variables in the context of MRI image classification. While these methods are interesting, they suffer from two main drawbacks.

First, they tend to simplify the disease as a classification problem. For many diseases, just stating whether a patient is affected by the disease or not is insufficient; the degree to which the patient is affected (severity) often carries more importance. Predicting severity is difficult, because very often severity does not have a direct causal relationship with its predictors, it is merely correlated with them (often weakly and non-linearly). Therefore, we model disease severity as a distribution and we seek to discovery patient subpopulations whose severity distribution significantly deviates from the severity distribution of the entire population.

CDR is an ordinal categorical variable which lies between class labels (nominal categorical) and quantitative outcomes. Association between subpopulations and class labels have been considered from a statistical perspective by Brin using the $\chi^2$-test [6] and association between subpopulations and a continuous outcome variable has been studied by Aumann et al. [3] using the Z-test. Our approach is analogous to Aumann’s technique with differences dictated by (i) the non-continuous nature of CDR distribution, (ii) lack of homoscedasticity across the popopulation and (iii) detection of differences in the shape of the distributions.

The second drawback of the existing applications of association rule mining to the medical domain concerns the lack of incorporating known medical constraints into the mining process. Domain knowledge is mostly incorporated as features for mining or as post-discovery filtering of association rules.

2.1 Use of Spatial Constraints

A key piece of domain knowledge we have regards the pattern of how disease spreads: the pathology of AD tends to originate at a specific ROI and it spreads further away from the origin [5]. Since the ROIs in the brain are located in a 3-D Euclidean space with known distances between the ROIs, we incorporate spatial constraints into the mining process.

Although Geographic Information Systems (GIS) (see [13] for an overview) are the state of the art in incorporating spatial constraints, these systems aim to discover spatial knowledge rather than use spatial constraints to discover non-spatial (i.e., disease-related) knowledge. The most related work in using spatial constraints comes from the genomic data mining domain. Icev et al. [9] applied spatial (distance-based) constraints to find association rules involving genotypical motifs that are located close-by to each other on the genome. Our application of distance-based constraints are different in that (i) the manner in which AD spreads may result in elongated volumes of ROIs which our constraint accommodates and (ii) our definition of the distance constraint can be incorporated into the frequent pattern discovery process in a straight-forward fashion.

2.2 Interestingness Measures

As we discussed earlier, one of the main means of incorporating domain knowledge is filtering. Association rule mining algorithms tend to discover exponential number of rules and interestingness measures are used for discarding uninteresting rules. See [17] for an extensive list of interestingness
measures. Rules that do not conform to domain knowledge can be discarded as uninteresting.

A common problem with interestingness measures is that the vast majority of them is not rooted in a rigorous statistical framework. In contrast, we filter our association rules and we infer direction of disease progression based on statistical significance.

Given a statistical model and the notion that disease progression is related to causality, a causality inference technique similar to that of Schadt et al. [15] appears applicable to progression inference. Disease progression and causality, however, possess a substantial difference that renders such a maximum likelihood-based approach inapplicable. If event $A$ causes event $B$ to occur, then $A$ and $B$ tend to co-occur with high probability. On the other hand, when disease progresses from ROI $A$ to ROI $B$, $A$ and $B$ will co-occur, but not with very high probability. Our approach accounts for this crucial difference.

3. METHOD

We first provide a high-level overview of the methodology and each step will be described in detail in subsequent sections.

1. We discover distributional association rules with spatial constraints.

(a) We first extract spatial patterns from the data set. Spatial patterns consist of ROIs of significant atrophy that form near contiguous volumes in the brain.

(b) We form distributional association rules from the spatial patterns. Each spatial pattern applies to a subset of the patients in the sample, namely the patients with significant atrophy in each of these specific ROIs. The CDR values of these patients form a distribution. A distributional association rule is comprised of a spatial pattern and implies the CDR distribution of the patients the spatial pattern applies to. Distributional association rules are only interesting if they imply a CDR distribution that is different from the CDR distribution of the entire population.

In the following sections, unless otherwise noted, the term ‘association rules’ refers to distributional associations rule with spatial patterns.

2. We aggregate ROIs into subsystems to gain a more macroscopic view of the disease. Evidence exists that AD affects subsystems, which are large contiguous volumes consisting of multiple ROIs related to each other. High order association rules (association rules with at least 3 ROIs) are indicative of such sets of ROIs.

3. We infer the direction of disease progression between and within the subsystems. From the association rules consisting of pairs of ROIs, we infer whether the disease progressed from one ROI to the other. Then we combine the discovered directions of progression into a disease progression network. We enhance the interpretability of the networks by incorporating information about subsystems.

3.1 Spatial Patterns Extraction

Let $X$ denote an observation matrix of Z-scores (standard normal quantile) indicating the amount of atrophy each patient has suffered relative to our large set of controls. $X$ has 118 columns, one for each ROI and its rows correspond to patients. Let $C$ denote the vector of CDR measurements. For patient $j$, row $X_j$ contains the Z-scores and $C_j$ is his CDR score.

An item is a predicate indicating whether patient $j$ has suffered significant atrophy in ROI $i$; specifically $X_{i,j} < -1.5$. If this predicate holds, then patient $j$ is said to support item $i$. An itemset is a set of items and hence a set of predicates. The itemset holds true if all predicates corresponding to the constituent items hold. In other words, a patient supports an itemset if he supports all items in the itemset. The support (count) of an itemset $I$ is the number of patients that supports $I$. $I$ is frequent if its support exceeds a user-defined minimum support threshold $\sigma$.

**Definition 1 (Spatial Pattern).** An itemset $I = \{a, b, c\}$ is a spatial pattern with respect to a distance matrix $D$, if for all item $i$ in $I$, there exists in another item $j$ in $I$, such that $D_{i,j} < \delta$ and $i \neq j$. $D_{i,j}$ denotes the distance between ROIs $i$ and $j$; and $\delta$ is a user-specified maximum distance threshold.

Since each ROI would typically have 6 neighbors in space, we selected the 10th percentile of all distances as the threshold $\delta$.

**Lemma 1.** If two spatial patterns $I_1$, $I_2$ of the same length differ only in one item, then their super-pattern $I_1 \cup I_2$ is also a spatial pattern.

Let $I_1 = \{a, b, c\}$ and $I_2 = \{a, b, d\}$. Their super-pattern is $I = \{a, b, c, d\}$. If $I$ is a spatial pattern, then for each item $i \in I$, there is another item $j \in I$, such that $D_{i,j} < \delta$. Let us consider item $a$. Because $I_1$ is a spatial pattern, $D_{a,c} < \delta$ or $D_{a,b} < \delta$ holds. Since both $b$ and $c$ are in $I$, there is at least one item in $I$ that is within $\delta$ distance from $a$. The proof works analogously for items $b$, $c$ and $d$.

The importance of Lemma 1 lies in the candidate generation of the apriori algorithm. To form candidate itemsets of length $(k + 1)$, apriori joins two itemsets of length $k$ that differ only in a single item. Therefore the apriori algorithm will generate spatial candidate patterns of length $k + 1$ as long as the patterns of length $k$ are spatial patterns.

The goal of Spatial Pattern Extraction is to extract all frequent spatial patterns from the data set that meet the minimum support threshold $\sigma$ and the maximum distance threshold $\delta$.

We utilize a modified version of the apriori algorithm to find all spatial frequent patterns. The apriori algorithm iteratively constructs the lattice structure of the items proceeding in a depth-first manner. It first considers all items and discards the ones that have insufficient support. In the second iteration, it forms pairs of items and discards the infrequent pairs. In the $k$-th iteration, it considers the $k - 1$ patterns and combines those that differ only in one item\(^1\). The infrequent $k$-itemsets are discarded.

\(^1\)The original apriori algorithm considers two $k$-itemsets that differ only in their last item. We have consider all $k$-itemsets that differ in any one item, because not all subsets of a spatial pattern are spatial patterns.
3.2 Distributional Association Rules and their Statistical Significance

Let \( I \) be a frequent pattern and let \( P \) be the set of patients (rows of \( X \)) that support \( I \). Let \( C_P \) be the CDR values corresponding to the patients in \( P \). \( C_P \) defines a distribution of CDR values.

**Definition 2 (Distributional Association Rule).** The rule \( I \rightarrow C_P \) is a distributional association rule if and only if \( I \) is a frequent pattern and \( C_P \) is the distribution of CDR values of the patients that support \( I \). \( I \) is the antecedent and the distribution is the consequent of the rule.

More specifically, in this paper we are interested in Distributional Association Rules with spatial patterns, that is, Distributional Association Rules where \( I \) is a frequent spatial pattern.

Since both \( C \) (CDR values for the entire patient sample) and \( C_P \) represent distributions, we assess the significance of the spatial distributional association rule via the following hypothesis testing problem. The null hypothesis is that \( C \) and \( C_P \) are drawn from the same population (essentially \( C_P \) is a random sample of \( C \)); the alternative hypothesis is that \( C_P \) is associated with higher CDR values. We use the Wilcoxon rank test to compare the distributions. The Wilcoxon rank test is non-parametric, makes no distributional assumptions and does not rely on homoscedasticity.

The majority of the population—even in the sample, where we evaluated our method—does not suffer from AD. The CDR distribution of the population is therefore heavily skewed towards 0 (subjects not affected by AD have CDR 0). Distributional association rules, for which we fail to reject the null hypothesis \( H_0 \), show no evidence of being predictive of Alzheimer’s disease. Such rules are uninformative in discovering disease subsystems or in inferring disease progression and hence they are discarded.

3.3 Disease Subsystem Discovery

In order to gain a more macroscopic view of the disease we discover disease subsystems.

**Definition 3 (Disease Subsystem).** A disease subsystem is a near contiguous set of ROIs that are significantly associated with AD. Two ROIs are near contiguous if their centers are within \( \delta \) distance from each other.

In practice, we approximate disease subsystems as significant association rules (with spatial patterns). The spatial pattern in the antecedent ensures the near contiguity of the ROIs and the significant difference between CDR distributions ensures that these ROIs are related to AD.

We know that AD is heterogeneous, and its heterogeneity presents itself in the existence of multiple disease subsystems. We also know that atrophy affects the left and right hemisphere of the brain symmetrically. Therefore, we aim to discover clusters of ROIs that are (i) the antecedents of an association rule highly associated with AD (low p-value w.r.t. the Wilcoxon test) and (ii) the left-right symmetric counterparts of the ROIs form the antecedent of an association rule we discovered and (iii) they do not overlap with other subsystems.

We find these subsystems greedily in a straightforward manner. We sort the association rules on the p-value of their distributional significance and discard all association rules where the symmetric counterpart of the ROIs in the antecedent is not the antecedent of an existing association rule we discovered. We process the rules in increasing order of the p-value and add the antecedent of the rule to the set of subsystems, if it does not overlap with any previously discovered subsystems.

3.4 Inferring the Direction of Disease Progression

Consider two ROIs \( R_s \) and \( R_t \) that co-occur in association rules. If \( R_s \) is the source of progression and \( R_t \) is the target, we initially expect to see patients with atrophy in \( R_s \) only; and in later stages of the disease, we expect to see patients with atrophy both in \( R_s \) and \( R_t \). If \( R_s \) is indeed the source of the disease progression, then we expect

\[
\Pr(R_s|R_t) > \Pr(R_t|R_s),
\]

where \( \Pr(A) \) is the probability that a randomly chosen patient has Z-score < -1.5 in ROI \( A \). In this section, we develop a statistical technique to test this hypothesis.

Moreover, if \( R_s \) is the source and \( R_t \) is the target of disease progression, then \( \{R_s, R_t\} \) occurs in a later stage of the disease than \( R_s \). Thus we expect \( \{R_s, R_t\} \) to be associated with distributions that have higher severity than \( \{R_s\} \). We will use this observation as an internal check to verify that the direction inferred from Equation 1 indeed describes progression as we expected.

3.4.1 Test of Progression

We can formally infer the progression of disease through the following hypothesis testing problem.

\[
H_0 : \Pr(R_s|R_t) = \Pr(R_t|R_s) \quad (2)
\]

\[
H_A : \Pr(R_s|R_t) > \Pr(R_t|R_s) \quad (3)
\]

Under the null hypothesis

\[
\frac{\Pr(R_s, R_t)}{\Pr(R_t)} = \frac{\Pr(R_s, R_t)}{\Pr(R_s)},
\]

This is only possible if

\[
\Pr(R_s) = \Pr(R_t).
\]

If \( \Pr(R_s) = \Pr(R_t) \), then we can estimate this quantity as the pooled probability

\[
\Pr(R_p) = \frac{\Pr(R_s) + \Pr(R_t)}{2}. \quad (4)
\]

Based on the normal approximation, the 100 \( \times (1 - \alpha) \) confidence interval of \( \Pr(R_p) \) is

\[
\Pr(R_p) \pm z_{\alpha/2} \sqrt{\Pr(R_p) \left(1 - \Pr(R_p)\right) N^{-1}}, \quad (5)
\]

where \( N \) is the sample size. Consequently, the 100 \( (1 - \alpha) \) confidence interval of \( \Pr(R_s|R_t) \) as well as \( \Pr(R_t|R_s) \) under \( H_0 \) is

\[
\frac{\Pr(R_s, R_t)}{\Pr(R_t)} \pm z_{\alpha/2} \sqrt{\Pr(R_p) \left(1 - \Pr(R_p)\right) N^{-1}}. \quad (6)
\]

\( H_0 \) is rejected if \( \Pr(R_s|R_t) \) or \( \Pr(R_t|R_s) \) is outside the confidence interval in Equation 6. This coincides with \( \Pr(R_s) \)
or Pr(R_i) falling outside the confidence interval in Equation 5. Since the null distribution is symmetric around Pr(R_p), it suffices to check whether
\[
Pr(R_i) < Pr(R_p) - \frac{\sigma \sqrt{N}}{N} (1 - Pr(R_p))^N.
\]

A note on simultaneous hypothesis testing.
Since the test is carried out with respect to a large number of R_i-R_p pairs, correction for simultaneous hypothesis testing is needed. Bonferroni correction, a popular technique for correction, in this case is underestimating the true significance for the following reasons. (i) The atrophy of the brain is fairly symmetric. If a progress direction is established on one side, then it likely exists on the other side, too. (ii) If the marginally significant direction coincides with increased severity, the direction is more likely to be significant than suggested by the Bonferroni correction which is agnostic of the disease severity.

4. EVALUATION
The proposed method was evaluated on a large sample of patients coming from the Mayo Clinic Alzheimer’s Disease Research Center (ADRC) prospective patient cohorts. The sample consists of three cohorts: 120 cognitively normal, 81 subjects with mild cognitive impairment (MCI) who progressed to AD and 48 pathologically confirmed AD subjects to get pathologically clean AD patients. The MCI patients represent the lower CDR spectrum, while the pathologically confirmed AD patients represent the mid to upper range of CDR spectrum. The number of MCI vs pathologically confirmed AD subjects is approximately representative of the patient population.

Gray matter density in 118 pre-defined regions of interest was used as predictors after adjusting for gender and age, standardizing to the 120 control subjects. The CDR score is available for all 249 patients.

We used all three cohorts (the controls, MCI patients and pathologically confirmed AD patients) to discover the spatial patterns and the Distributional Association Rules. We discovered 573 association rules at a minimum support count (\(\sigma\)) of 10. 153 of these 573 rules were found to have a CDR distribution that is significantly different from the CDR distribution of the entire sample at a Bonferroni-corrected significance level of 0.05. 37 rules contained a single ROI in their antecedent, 59 rules contained a pair of ROIs and the remaining rules had three or more ROIs. The longest rule contained 5 ROIs.

4.1 Single ROI associated with AD
Rules with a single ROI represent regions of the brain significantly associated with AD. Figure 1 depicts these ROIs with the color indicating the median severity of the effected patients. More vibrant yellow signifies higher severity. The first 6 panes represent 6 slices of the brain in axial view (view from the top) and the seventh pane is a sagittal (side) view of the brain. The sagittal view has horizontal blue lines corresponding to locations of the six slices.

The regions found to be significant are mainly in the temporo-parietal and some frontal regions which are consistent with the literature.

Figure 2 shows the CDR distribution for the individual ROIs as boxplots. ROIs are displayed on the vertical axis (e.g. ROI #37 is the hippocampus); CDR is displayed on the horizontal axis. While the CDR ranges largely overlap across ROIs, the medians are different and differences in CDR score distribution across ROIs clearly exist. For example, ROIs in the medial temporal lobe (ROI #35-50), which have been implicated earlier in the disease, have lower CDR scores than the frontal lobe ROIs (ROI #3-30), which are affected later in the disease\(^2\). Such differences allow us to infer the direction of disease progression.

The large overlaps in the CDR ranges also imply that quantizing the CDR values into a relatively small number of bins would result in minimal ability to distinguish severity. Considering distributional differences is necessary.

4.2 Multi-ROI association with AD
Rules with multiple ROIs in their antecedent represent regions that commonly co-occur in AD. Due to the spatial constraint, these regions are approximately contiguous. We refer to these sets of ROIs as *subsystems*. Recall that we require subsystems to be significant (w.r.t. the Wilcoxon test) and to be symmetric: if a set of ROIs on one side form a subsystem, we require that corresponding set of ROIs on the symmetric side of the brain also forms a subsystem.

We discovered 3 such subsystems which we depicted in Figure 3. The subsystems can be discerned through their colors: the (left and right) hippocampal subsystems are colored blue, the subsystems containing the left and right insula are colored orange and the subsystems containing the left and right inferior parietal lobes are colored green. In the following, we will refer to these subsystems as A, B and C, respectively.

All of these subsystems are known to have elevated atrophy.
4.3 Inferring Disease Progression

Among the 153 significant rules, 59 had a pair of ROIs in the antecedent. These 59 rules were used to infer the direction of disease progression. We computed the significance of the disease progression (Equation 6) using only the MCI and the pathologically confirmed AD cases. Controls have no disease, hence they contribute no information to the disease progression. At significance level of .2, 28 of the 59 pairs of ROIs showed significant direction of progression.

Figure 4 depicts the patterns of AD progression we extracted from the rules. The nodes (ovals) in the figure correspond to individual ROIs and the enclosing rectangles correspond to subsystems. The black arrows going from one node (source) to the other (target) indicate the direction of the disease progression (these nodes reached significance using the test of progression). If the disease severity score of the target node is greater than the source node, the increase in the median severity scores of the patients is indicated on the black arrow. E.g. Subjects with both abnormal nodes 29 and 37 have a median severity score which is 0.51 higher than subjects with only abnormal node 37. Two nodes are connected by a gray edge when the corresponding ROIs are related but the significance of progression test failed to show significance.

We describe our findings centered around the subsystems.

In subsystem A, which originates in Nodes 37 (Left Hippocampus) and Nodes 38 (Right Hippocampus) has a disease progression pattern beginning in the hippocampus and proceeding to bilateral parahippocampal gyrus (Nodes 40 and 39) and/or bilateral fusiform (Nodes 55 and 56) with a substantial increase in the severity score of the subjects. The other ROIs there were significantly associated with hippocampal atrophy (Nodes 37 and Nodes 38) were bilateral inferior temporal ROI (Nodes 89 and 90) and bilateral amygdalae (Nodes 42 and 41). Atrophy in all the ROIs observed within this subsystem have been shown in the literature to be related to AD and the direction of disease progression concurs with the direction of tangle pathology (which is also measured by MRI) progression shown by Braak and Braak. Tangle pathology progression is used as a gold standard for AD diagnosis.

Subsystem B consisted of bilateral insulae (Nodes 29 and 30), Heschl gyrii (Nodes 79 and 80) and temporal poles (Nodes 83 and 84). The direction of disease progression was only significant on the left side with Node 29 (left insula) appearing to be the originating node for this subsystem.

Subsystem C is comprised of the bilateral inferior parietal lobes (Nodes 61 and 62), precuneus (Nodes 67 and 68) and posterior parietal lobes (Nodes 59 and 60). This result was very fascinating since there has been recent evidence that regions in this subsystem are affected early in disease and might occur in atypical cases [11].

5. SUMMARY AND DISCUSSION

In this paper we set out to address two important issues in Alzheimer’s disease: to find typical and atypical AD presentations and to study the origin and progression of the disease. Both these topics enjoy much attention in the literature lacking a definitive answer. To achieve our goal, we proposed a technique based on distributional association rule mining with spatial constraints that allows us to asso-
We associate brain structure atrophy in pre-defined regions of interest (ROIs) with distributional differences in disease severity. We successfully exploited these associations towards finding large subsystems of AD related to both typical and atypical disease presentations and also towards painting a clear picture of disease progression.

The association rules we mined allowed us to discover three large AD subsystems. All of these subsystems are noted in the literature, two corresponding to typical AD presentation and the third one related to atypical or advanced AD presentation.

We found that our methodology helped determine the direction of brain loss trajectories. In the AD patients, the hippocampus (both L and R) was the major originating node for the brain loss trajectory that affected the largest proportion of the study population. This is consistent with the existing literature that AD-related atrophy occurs earliest and most severely in the hippocampus. Also, the detected brain loss trajectory patterns i.e. the nodes downstream from the hippocampus were consistent with the pathological gold standard system for staging tangles proposed by Braak and Braak [5]. One key discovery was that there were differ-
ent sets of originating nodes in a smaller proportion of the population supporting the evidence for atypical AD seen in some patients.

While the prevalent view of AD considers the hippocampus to be the origin of the disease, Desikan et al. [7] showed that the disease originates in the neocortical regions (specifically, parietal nodes; 61, 62, 67 and 68) and atrophy in the hippocampus is a downstream process. Our data does not support the hypothesis that hippocampal atrophy would be downstream from atrophy in the parietal regions, but we found that the disease can indeed originate in this region independently from the hippocampus.

An interesting implication of the hypothesis by Scailhill et al. is that atrophy may progress along a very long path such as from subsystem C to subsystem A. Our current methodology is not well-suited for the discovery of such distant connections, but it can be extended by a second stage, where trajectory directions are sought among the ROIs that are present in known subsystems with no regard to the distance between them. Such an approach retain the capability of our methodology to reduce false discoveries. We will explore this direction in the future.

We found our methodology to be well-suited for finding contiguous disease-affected regions, associating these regions with disease severity and mapping out the disease progression directions. Many neurodegenerative diseases pose a similar problem and our methodology can be helpful in solving them. The applicability of the proposed methodology, however, is not limited to neurodegenerative diseases (e.g. breast cancer detection would benefit from our method), or even to image data. Our methodology can be viewed as a generic framework for associating disease severity with predictors, which offers valuable information in most medical domains as well as domains outside medicine.

6. REFERENCES