Identifying genetic markers associated with Alzheimer’s Disease progression through image phenotyping

Danilo Pena, BS¹, Assaf Gottlieb, PhD¹, Luca Giancardo, PhD¹
¹Center for Precision Health, School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, Texas, 77030

Introduction

Alzheimer’s Disease (AD) is a detrimental neurodegenerative disorder that affects millions of people worldwide. Subjects with mild cognitive impairment (MCI) are at an increased risk of developing AD, but currently, there is a lack of clinical measurements that can inform a physician whether an at-risk patient is undergoing the prodromal phase of the neurodegenerative process. Associating imaging-based features with genomic data has proved useful for various AD-related questions (1), including associations between genomic data and brain structure (2) or brain function (3). However, validated imaging tools able to quantify the neurodegenerative process at the early stages are still lacking.

Here, we present a method that combines imaging and genomics analysis through the use of our imaging-based phenotyping algorithm (4) and available single nucleotide polymorphisms (SNPs) data to investigate associations with AD-like progression in MCI subjects. We further identify adjacent genes to AD-progression associated SNPs and verify their enrichment in pathways. Our discovered SNPs and genes can be used for risk estimation of AD progression, allowing development and administration of new preventative treatments.

Methods

Data: Whole genome sequencing data and imaging data was downloaded from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) in June 2019 and November 2018, respectively. We focused on 613 subjects from non-Hispanic Caucasian ancestry with MCI, determined according to imaging-based progression score (Table 1).

Table 1. Demographic breakdown of MCI subjects

<table>
<thead>
<tr>
<th></th>
<th>Average Thresholding Strategy</th>
<th>Median Thresholding Strategy</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Subjects</td>
<td>362</td>
<td>251</td>
<td>-</td>
</tr>
<tr>
<td>Age, years (mean (s.d.))</td>
<td>72.2 (7.3)</td>
<td>76.2 (6.9)</td>
<td>&lt;0.001 *</td>
</tr>
<tr>
<td>Sex (male, n (%))</td>
<td>211 (58%)</td>
<td>159 (63%)</td>
<td>0.24 **</td>
</tr>
<tr>
<td>Education, years (mean (s.d.))</td>
<td>15.9 (2.8)</td>
<td>15.7 (2.9)</td>
<td>0.29 *</td>
</tr>
</tbody>
</table>

*Wilcoxon Ranked Sum Test, **Fisher Exact Test

Figure 1. High-level overview of the methodology’s pipeline: pre-processing (green & yellow) and analysis (blue)
Detection of AD progression: We used our custom deep learning model, DeepSymNet (4), that identifies structural brain changes in AD subjects at two different imaging timepoints. The algorithm was benchmarked against pipelines that used both region and voxel-based analyses. Further, the model’s output was a progression metric that identified how similar a patient’s change in brain structure was to that of an AD-like pattern. Finally, we extended this metric to an external set of MCI subjects that used in this study. None of the MCI subjects used in this study were used for training DeepSymNet.

Genetic association: We excluded SNPs with low quality calls and only included SNPs that were available for all subjects. In order to minimize the hypothesis space, we focused on SNPs with known information in NCBI dbSNP, resulting in 261K SNPs (Figure 1). We used Fisher’s exact test to associate SNPs with the phenotype and applied Bonferroni correction of 0.01.

Results

We tested two progression score thresholds to identify AD-like progression (cases) and no progression (controls). Using a threshold of the median score, we identified 1 significant SNP associated with the AD-progression phenotype (rs4799404). rs4799404 resides within 50Kb of TSPYL1, previously implicated in relation with Parkinson’s Disease (PD) (5) and within 250Kb of ELP2, previously associated with Frontotemporal Dementia (6) neurodevelopmental disabilities (7). Using MAGMA gene-set analysis, we identified a significant enrichment of potassium and sodium ion channels which aligns well with the described involvement of microglial ion channels in formation of AD (8).

Using a threshold of the average score, we identified five significant SNPs. Out of these five SNPs, three are located within, or next to, the FAF1 gene. FAF1 is associated with PD cases with AD pathology (9). An additional SNP is adjacent to SFPQ whose dysregulation and dislocation might be a novel pathway in the progression of AD (10). The last significant SNPs is an intron in the RBFOX1 gene, which was found to be associated with functional neuroimaging biomarker for AD degeneration (11). We note that using the average progression score we detected over-inflation of the p-values based on the Q-Q plot, which may indicate an increase in the number of false positives, which will require more experiments to validate.

Discussion

Our methodology for association of genetic variations with an imaging-based progression-specific phenotype can identify genetic risk factors for progression from MCI to AD. Our main limitation is the small sample size, which may have reduced our detection power. In future work, we intend to include additional cohorts for training and validation.

References