2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 979-8-3503-3748-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/BIBM58861.2023.10385469

Interpretable Graph Convolutional Network for Alzheimer's Disease Diagnosis using Multi-Modal Imaging Genetics

1st Houliang Zhou
Computer Science and Engineering
Lehigh University
PA, USA
hoz421@lehigh.edu

2nd Yu Zhang
Bioengineering, and Electrical and Computer Engineering
Lehigh University
PA, USA
yuzi20@lehigh.edu

3th Lifang He
Computer Science and Engineering
Lehigh University
PA, USA
lih319@lehigh.edu

4th Li Shen
Biostatistics, Epidemiology and Informatics
University of Pennsylvania
PA, USA
li.shen@pennmedicine.upenn.edu

5rd Brian Y. Chen*
Computer Science and Engineering
Lehigh University
PA, USA
byc210@lehigh.edu

***Abstract*—Integrating brain images and genetic data provides a great opportunity to discover potential biomarkers for neurological disorder diagnosis. However, learning genetic information and brain network dysfunction remains a challenging task. In this paper, we propose an interpretable multi-modal imaging and genetic graph convolution network (GCN) for Alzheimer's disease diagnosis. Our genetic network uses hierarchical GCN to mimic a gene ontology-based graph of biological processes and learn the information flow in this graph. In parallel, our imaging network uses a sparse interpretable GCN with node and edge importance probabilities to learn the brain network from multi-modal images. After multi-modal fusion, the final representation guided by a cluster-based consistency constraint is used to predict the disease-related clinical measures. We evaluate our method on the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Our result shows that our imaging-genetics framework achieves superior prediction performance compared to all state-of-the-art methods. The interpretation demonstrated that the salient SNPs, and salient regions interpreted by important probabilities were significantly correlated with AD-related clinical symptoms, and considerably important for developing novel biomarkers. The code is available at https://github.com/Houliang-Zhou/IG-GCN.**

***Index Terms*—Graph convolutional network, imaging genetics, multi-modality, neuroimaging, sparse interpretation**

I. Introduction

Imaging-genetics is an emerging field that facilitates Alzheimer's disease (AD) diagnosis by merging both imaging and genetic features [1], [2]. The imaging features are extracted from the magnetic resonance imaging (MRI) or positron emission tomography (PET) modalities [3], and the genetic variants are detected from Single Nucleotide Polymorphisms (SNPs) [2]. In recent years, various methods have been proposed to analyze the imaging-genetics. On the one hand, Du et al. proposed a multitask sparse canonical correlation analysis (MTSCCALR) to detect genetic associations with imaging phenotypes in AD [1]. However, they don't consider the non-linear relationships and don't provide the intrinsic interpretation for Imaging-genetic biomarkers. On the other hand, Zuo et al. used the concrete autoencoder to identify consistent imaging genomic biomarkers [4]. However, autoencoder-based methods don't consider the information about interactions between a multitude of genetic variants that interact through various biological processes. In genetics, some studies have connected SNPs and genes to diverse biological processes [5], [6]. Gaudelet et al. built an artificial neural network with genetic risk from biological processes to predict phenotypic variables [6]. Such neural networks fail to consider the hierarchical and interconnected nature of biological processes. Ghosal et al. designed an interpretable genetic and imaging graph neural network (GUIDE) based on the gene ontology hierarchy for schizophrenia analysis [7]. However, the GUIDE ignores the connective abnormality between ROIs from images, which limits the ability of Graph convolutional network (GCN) to learn neuroimaging dysfunctions. Therefore, a GCN-based method to learn neuroimaging connectivity dysfunction and capture interconnected genetic variants from gene ontology hierarchy is a need for analyzing neurological disorders like AD.

In this paper, we propose an interpretable multi-modal imaging and genetic graph convolution framework shown in Fig. 1 to jointly regress three typical AD-related clinical scores, including Mini-Mental State Examination (MMSE), Alzheimer's Disease Assessment Score 13 (ADAS13), and Tau for Alzheimer's disease diagnosis. Our experiments show that the scores predicted by ours correlate significantly with these clinical measures. Our main contributions include:

* Corresponding author.

979-8-3503-3748-8/23/$31.00 ©2023 IEEE

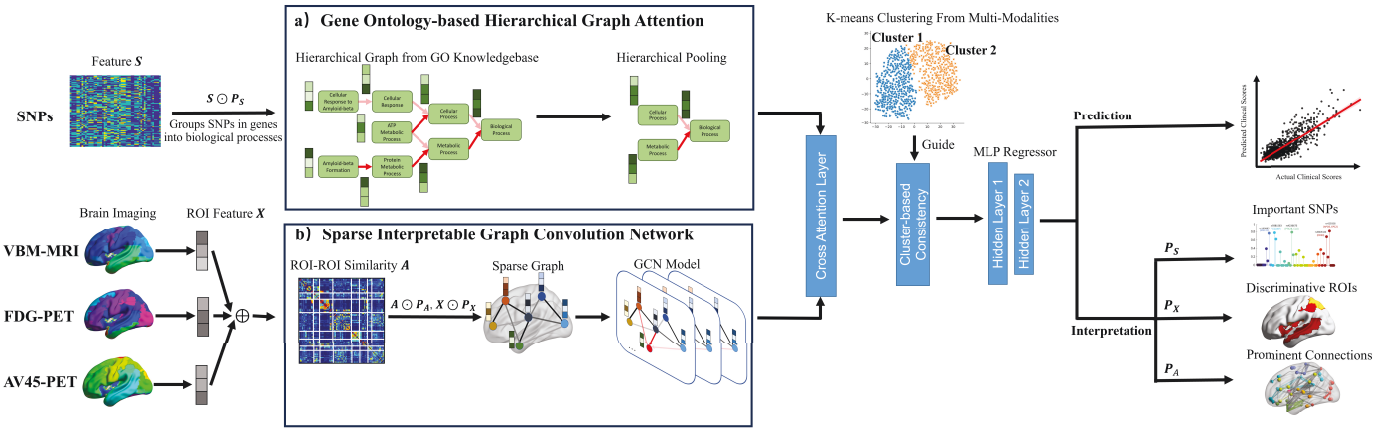


Fig. 1. An overview of the proposed framework for Alzheimer's diagnosis and imaging genetic biomarker interpretation. The box a of the framework illustrates the gene ontology-based hierarchical graph attention network with SNPs importance probability $\boldsymbol{P}_S$ to learn the SNPs information. The box b of the framework illustrates a sparse interpretable graph convolutional network with node importance probability $\boldsymbol{P}_X$ and edge importance probability $\boldsymbol{P}_A$ to learn the multi-modal brain images.

- The first application of using gene ontology-based hierarchical GCN to capture information flow through hierarchy and identify the key genetic variants in AD diagnosis.
- The first integration of multi-modal brain imaging and genetic variants by building interpretable and hierarchical GCN for imaging-genetic biomarker identification.
- The extension of interpretable GCN model with importance probabilities to discover the important SNPs and the discriminative ROIs under multi-modal imaging-genetics.

II. Methods

A. Data Acquisition

In this work, we evaluated our framework on the public Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset [8]. These brain imaging genetic data were gathered from 874 non-Hispanic Caucasian participants, including 199 HC, 524 MCI, and 151 AD subjects. The dataset includes genetic variants detected from Single Nucleotide Polymorphisms (SNPs) and brain images consisting of three modalities including structural Magnetic Resonance Imaging (VBM-MRI), 18F-fluorodeoxyglucose Positron Emission Tomography (FDG-PET), and 18F-florbetapir PET (AV45-PET) scans. In the genetic data, we included 54 SNPs collected from the neighbors of the major AD risk factor APOE based on the ANNOVAR.

B. Gene Ontology-based Hierarchical Graph on Genetics

1) Important Probabilities to Interpret Genetic Biomarkers: The top portion of Fig. 1 illustrates the gene ontology-based hierarchical graph attention module to learn genetic information. The mechanism of this graph is based on the gene ontological hierarchy [9]. Specifically, we group the SNPs into the related genes, and further map these genes into their biological processes to build the graph, where we regard the biological process as a node. Mathematically, inputs are the SNPs information $\boldsymbol{g} \in \mathbb{R}^{1\times S}$ for each subject, where S is the number of SNPs. We define the relation from SNPs to biological processes as a sparse matrix $\boldsymbol{M}_g \in \mathbb{R}^{T\times S}$, where T is the number of biological processes. In order to find the important SNPs that contribute most to disease prediction, we define the learnable SNPs important probabilities $\boldsymbol{P}_S \in \mathbb{R}^S$ with sigmoid function. We learn the projection of the important SNPs onto T graph nodes. The node signal in the graph is a d-dimensional feature vector, $\boldsymbol{X}_g \in \mathbb{R}^{T\times d}$.

$$\boldsymbol{X}_g = ReLU((\boldsymbol{M}_g \odot \boldsymbol{P}_S^T)\boldsymbol{W}_g) \tag{1}$$

where $\boldsymbol{W}_g \in \mathbb{R}^{S\times d}$ is the learnable parameters and $\odot$ denotes the Hadamard element-wise product.

2) Attention-based Hierarchical Graph Learning: We leveraged an attention-based GCN to mimic the flow of hierarchical information from biological processes. We use the attention mechanism to learn the weights that select the most discriminative edges for each node. Specifically, we define the edge weight between child node i and its parent node j at convolutional layer l as

$$e^l(i,j) = \frac{exp(\alpha([\boldsymbol{X}_g^l(i)\boldsymbol{W}^l || \boldsymbol{X}_g^l(j)\boldsymbol{W}^l]\boldsymbol{b}^l))}{\sum_{i\in child(j)} exp(\alpha([\boldsymbol{X}_g^l(i)\boldsymbol{W}^l || \boldsymbol{X}_g^l(j)\boldsymbol{W}^l]\boldsymbol{b}^l))} \tag{2}$$

where $\boldsymbol{X}_g^l(j)$ is the feature vector for node j at layer l, $\boldsymbol{W}^l \in \mathbb{R}^{d_l\times d_{l+1}}$ and $\boldsymbol{b}^l \in \mathbb{R}^{2d_{l+1}\times 1}$ are the learnable parameters, α is the activation function, and $||$ is the concatenation function. Finally, we coarsen the graph by hierarchical pooling to learn the genetic embedding.

C. Sparse Interpretable GCN on Multi-Modal Images

In neuroimages, we concatenated multiple modalities into the ROI's feature vector. We define an adjacency matrix $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ using the Gaussian similarity and feature matrix $\boldsymbol{X} \in \mathbb{R}^{N\times D}$, where N is the number of ROIs and D is the dimension of multi-modal features. To identify important subset X_s and subgraph G_s that have the greatest impact on

disease prediction, we introduce a shared multi-modal feature importance probability $\boldsymbol{P_X}$, and the edge importance probability $\boldsymbol{P_A}$ between nodes within each subject. Mathematically, we denote the important subgraph as $G_s = \boldsymbol{A} \odot \boldsymbol{P_A}$, and the important subset of multi-modal features as $X_s = \boldsymbol{X} \odot \boldsymbol{P_X}$.

Given that the diverse modalities of ROIs have varying impacts on disease prediction, the multi-modal feature importance probability is defined as $\boldsymbol{P_X} \in \mathbb{R}^{N \times D}$ and $\boldsymbol{P_X} = [\boldsymbol{p_1}, \boldsymbol{p_2}, ..., \boldsymbol{p_N}]$, where $\boldsymbol{p_i} \in \mathbb{R}^D$, $1 \leq i \leq N$, represents the feature importance probability for ith ROI. Because the ROI's multi-modal features are highly associated with the strength of their connections, we build the edge importance probability $\boldsymbol{P_A} \in \mathbb{R}^{N \times N}$ between node i and j based on the joint connection between features $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$:

$$P_{A_{i,j}} = \sigma(\boldsymbol{v}^T[\boldsymbol{x_i} \odot \boldsymbol{p_i} || \boldsymbol{x_j} \odot \boldsymbol{p_j}]) \tag{3}$$

where $\boldsymbol{p_i}$ represents the feature importance probability for ith node, $\boldsymbol{v} \in \mathbb{R}^{2D}$ denotes the learnable parameter. Later, we use cross-attention to fuse imaging and genetic embedding.

D. Loss Function

In this section, we define the mean square error loss to regress the clinical scores and determine the SNPs, ROIs, and edge importance probabilities. Specifically, we train our model and find the $\boldsymbol{P_S}$, $\boldsymbol{P_X}$ and $\boldsymbol{P_A}$ by minimizing the mean square error between the true y and the predictive output $\hat{y}$ learned by our imaging genetic framework. Assuming there is M subjects, the mean square error loss $\mathcal{L}_{regr}$ is expressed as:

$$\mathcal{L}_{regr} = \frac{1}{M} \sum_{m=1}^{M} (y_m - \hat{y}_m)^2 \tag{4}$$

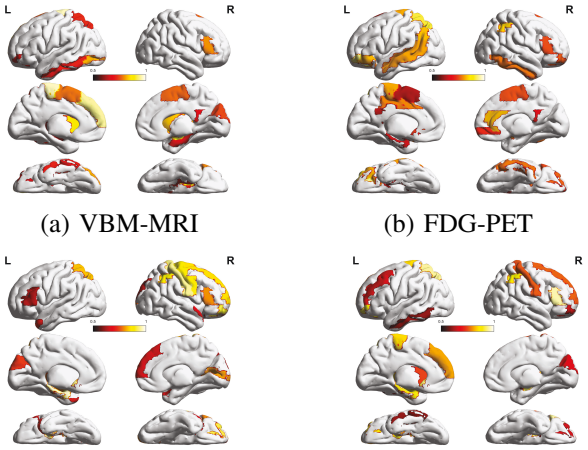
In addition, we applied the sparse loss $\mathcal{L}_{spar}$ including ℓ_1 and entropy regularization to induce the sparsity on $\boldsymbol{P_S}$, $\boldsymbol{P_X}$, and $\boldsymbol{P_A}$. We further propose the cluster-based consistency constraint to force our method to keep the latent representation of subjects within the same cluster to be similar. We conducted a k-means clustering on multi-modal features. The cluster-based consistency constraint is expressed as:

$$\mathcal{L}_{cons} = \sum_{i,j}^{n} s_{i,j} \left\| \boldsymbol{h}_i^T - \boldsymbol{h}_j^T \right\|_2^2 = tr((\boldsymbol{H})^T \boldsymbol{L} \boldsymbol{H}) \tag{5}$$

where $s_{i,j} \in [0, 1]$ is the similarity between subject i and j, $\boldsymbol{H}$ is the latent representation after cross attention layer. We note that if two subjects belong to the same cluster, $s_{i,j}$ is 1, otherwise 0. Therefore, the latent representation learned from the imaging-genetics of subjects within the same cluster will become similar. We further impose the orthogonal constraint $\mathcal{L}_{orth}$ on the latent representation $\boldsymbol{H}$. The final training objective of our proposed method is:

$$\mathcal{L} = \mathcal{L}_{regr} + \lambda_1 \mathcal{L}_{spar} + \lambda_2 \mathcal{L}_{cons} + \lambda_3 \mathcal{L}_{orth} \tag{6}$$

where λ's are the tunable hyperparameters regarded as penalty coefficients for the different loss terms. In our result, the importance probabilities learned from our method provide the interpretation regarding the important SNPs and salient ROIs in AD.



(a) VBM-MRI (b) FDG-PET

(c) AV45-PET (d) Three Modalities

Fig. 2. The interpretation of top 20 salient ROIs under multiple modalities. The bright-yellow color indicates a high score.

III. Results

A. Experimental Results

We used our proposed method to jointly regress three clinical scores including ADAS13, MMSE, and Tau. We performed the 5-fold cross-validation to examine the regression performance by employing three metrics including the Pearson correlation coefficient (Pearson's r), the coefficient of determination (R^2), and Mean Squared Error (MSE).

Table I shows the regression comparison between the state-of-the-art machine learning models and our proposed method. Our method was compared with a CCA-based method called MTSCCALR [1], a deep learning-based method called Autoencoder [4], a gene ontology-based method namely GUIDE [7], and other GCN-based methods including GCN [10] and GAT [11]. We noted that GUIDE didn't consider the brain connectivity graph from neuroimaging, and GCN and GAT used such a connectivity graph without applying important probabilities. In our result, the gene ontology-based GUIDE was better than the MTSCCALR and Autoencoder to capture the genetic variants. Brain connectivity-based GCN and GAT methods were better than GUIDE at predicting Tau protein content which built up in AD and damaged brain cells essential for learning and memory [12]. Finally, the best metrics were achieved by using our proposed method for each clinical score, indicating the superior performance of our proposed method.

B. Interpretability

To interpret the salient ROIs for predicting clinical measures, we averaged the learned feature importance probability $\boldsymbol{P_X}$ across different modalities and ranked these scores in descending order. The top 20 most salient ROIs were visualized using BrainNet Viewer [13] in Fig. 2 and identified by different modalities as well as the multiple modal analysis.

The node importance probability $\boldsymbol{P_X}$ in our model identified that the hippocampus, parietal gyrus, and olfactory gyrus were

TABLE I
REGRESSION COMPARISON BETWEEN THE STATE-OF-THE-ART MACHINE LEARNING MODELS AND OURS ON THREE CLINICAL SCORES.

Clinical Scores	Metrics	MTSCCALR	Autoencoder	GCN	GAT	GUIDE	Genetic Only	Imaging Only	Ours
ADAS13	r	0.646±0.054	0.683±0.042	0.697±0.046	0.762±0.028	0.736±0.035	0.322±0.062	0.754±0.022	0.829±0.012
	R^2	0.325±0.061	0.371±0.054	0.482±0.059	0.617±0.037	0.521±0.047	0.102±0.073	0.593±0.041	0.698±0.023
	MSE	0.128±0.015	0.124±0.013	0.117±0.014	0.097±0.009	0.103±0.011	0.145±0.031	0.099±0.010	0.077±0.003
MMSE	r	0.562±0.062	0.627±0.064	0.605±0.068	0.643±0.051	0.708±0.044	0.286±0.075	0.682±0.051	0.771±0.038
	R^2	0.205±0.049	0.269±0.053	0.243±0.057	0.358±0.032	0.462±0.037	0.095±0.051	0.421±0.028	0.609±0.017
	MSE	0.116±0.013	0.107±0.012	0.112±0.015	0.105±0.012	0.095±0.015	0.132±0.020	0.098±0.013	0.082±0.011
Tau	r	0.536±0.048	0.610±0.050	0.642±0.038	0.697±0.033	0.579±0.025	0.213±0.032	0.724±0.036	0.768±0.014
	R^2	0.263±0.052	0.316±0.064	0.357±0.046	0.458±0.041	0.266±0.037	0.051±0.038	0.475±0.047	0.531±0.021
	MSE	0.132±0.023	0.120±0.019	0.124±0.015	0.118±0.017	0.129±0.018	0.157±0.021	0.116±0.015	0.109±0.008

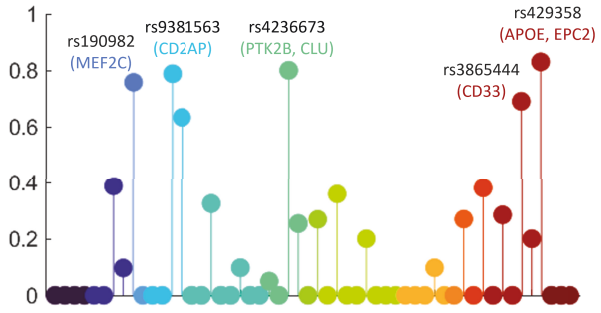


Fig. 3. The interpretation of important SNPs and their overlapping genes. Different color denotes different chromosomes.

the most salient brain regions for predicting clinical patterns. These interpretations were highly consistent with established findings on cognitive impairment in AD [3], [14]. In addition, the SNPs importance probability $\boldsymbol{P_S}$ in our model identified that the rs429358, rs4236673, and rs3865444 were the most important SNPs for predicting clinical patterns, which were consistent with the previous studies about genetic risk in AD [15]. Fig. 3 showed the important probabilities of all the SNPs with their overlapping genes. This interpretation suggested that our model can extract discriminative biological information from ontology and explore potential genetic biomarkers in AD.

IV. CONCLUSION

We have presented an interpretable multi-modal imaging and genetic GCN for predicting typical AD-related clinical scores. It extended the current interpretation of the GCN method to identify novel neurological biomarkers under multi-modal imaging and genetic analysis.

ACKNOWLEDGMENT

This work is partially supported by the National Institutes of Health (U01AG068057, U01AG-066833, R01LM013463, R01MH129694, R21MH130956, and R21AG080425), National Science Foundation (MRI-2215789 and IIS-1909879), Alzheimer's Association grant (AARG-22-972541), and Lehigh's grants under Accelerator (S00010293), CORE (001250), and FIG (FIGAWD35).

REFERENCES

[1] L. Du, K. Liu, X. Yao, S. L. Risacher, J. Han, A. J. Saykin, L. Guo, and L. Shen, "Detecting genetic associations with brain imaging phenotypes in Alzheimer's disease via a novel structured SCCA approach," *Medical Image Analysis*, vol. 61, p. 101656, 2020.

[2] R. Zhou, H. Zhou, B. Y. Chen, L. Shen, Y. Zhang, and L. He, "Attentive deep canonical correlation analysis for diagnosing alzheimer's disease using multimodal imaging genetics," in *MICCAI*. Springer, 2023, pp. 681–691.

[3] H. Zhou, Y. Zhang, B. Y. Chen, L. Shen, and L. He, "Sparse interpretation of graph convolutional networks for multi-modal diagnosis of alzheimer's disease," in *MICCAI*. Springer, 2022, pp. 469–478.

[4] Y. Zuo, Y. Wu, Z. Lu, Q. Zhu, K. Huang, D. Zhang, and W. Shao, "Identify consistent imaging genomic biomarkers for characterizing the survival-associated interactions between tumor-infiltrating lymphocytes and tumors," in *MICCAI*. Springer, 2022, pp. 222–231.

[5] H. Mi, A. Muruganujan, J. T. Casagrande, and P. D. Thomas, "Large-scale gene function analysis with the panther classification system," *Nature protocols*, vol. 8, no. 8, pp. 1551–1566, 2013.

[6] T. Gaudelet, N. Malod-Dognin, J. Sánchez-Valle, V. Pancaldi, A. Valencia, and N. Pržulj, "Unveiling new disease, pathway, and gene associations via multi-scale neural network," *PloS one*, vol. 15, no. 4, p. e0231059, 2020.

[7] S. Ghosal, Q. Chen, G. Pergola, A. L. Goldman, W. Ulrich, D. R. Weinberger, and A. Venkataraman, "A biologically interpretable graph convolutional network to link genetic risk pathways and imaging phenotypes of disease," in *ICLR*, 2021.

[8] S. G. Mueller, M. W. Weiner, L. J. Thal, R. C. Petersen, C. Jack, W. Jagust, J. Q. Trojanowski, A. W. Toga, and L. Beckett, "The alzheimer's disease neuroimaging initiative," *Neuroimaging Clinics*, vol. 15, no. 4, pp. 869–877, 2005.

[9] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig *et al.*, "Gene ontology: tool for the unification of biology," *Nature genetics*, vol. 25, no. 1, pp. 25–29, 2000.

[10] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *ICLR*, 2017.

[11] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *ICLR*, 2018.

[12] R. Medeiros, D. Baglietto-Vargas, and F. M. LaFerla, "The role of tau in alzheimer's disease and related disorders," *CNS neuroscience & therapeutics*, vol. 17, no. 5, pp. 514–524, 2011.

[13] M. Xia, J. Wang, and Y. He, "BrainNet viewer: a network visualization tool for human brain connectomics," *PloS one*, vol. 8, no. 7, p. e68910, 2013.

[14] M. M. Vasavada, J. Wang, P. J. Eslinger, D. J. Gill, X. Sun, P. Karunanayaka, and Q. X. Yang, "Olfactory cortex degeneration in alzheimer's disease and mild cognitive impairment," *Journal of Alzheimer's disease*, vol. 45, no. 3, pp. 947–958, 2015.

[15] M.-T. Lo, K. Kauppi, C.-C. Fan, N. Sanyal, E. T. Reas, V. Sundar, W.-C. Lee, R. S. Desikan, L. K. McEvoy, C.-H. Chen *et al.*, "Identification of genetic heterogeneity of alzheimer's disease across age," *Neurobiology of aging*, vol. 84, pp. 243–e1, 2019.