



# Plasma protein-based identification of neuroimage-driven subtypes in mild cognitive impairment via protein-protein interaction aware explainable graph propagational network

Sunghong Park<sup>a,1</sup>, Doyoon Kim<sup>a,1</sup>, Heirim Lee<sup>b,c</sup>, Chang Hyung Hong<sup>b</sup>, Sang Joon Son<sup>b</sup>, Hyun Woong Roh<sup>b</sup>, Dokyoon Kim<sup>d,e</sup>, Yonghyun Nam<sup>d</sup>, Dong-gi Lee<sup>d</sup>, Hyunjung Shin<sup>f,g,\*</sup>, Hyun Goo Woo<sup>a,h,i,\*\*</sup>

<sup>a</sup> Department of Physiology, Ajou University School of Medicine, Suwon, 16499, Republic of Korea

<sup>b</sup> Department of Psychiatry, Ajou University School of Medicine, Suwon, 16499, Republic of Korea

<sup>c</sup> Department of Psychology, Duksung Women's University, Seoul, 01369, Republic of Korea

<sup>d</sup> Department of Biostatistics, Epidemiology & Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104, USA

<sup>e</sup> Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, 19104, USA

<sup>f</sup> Department of Industrial Engineering, Ajou University, Suwon, 16499, Republic of Korea

<sup>g</sup> Department of Artificial Intelligence, Ajou University, Suwon, 16499, Republic of Korea

<sup>h</sup> Department of Biomedical Science, Graduate School of Ajou University, Suwon, 16499, Republic of Korea

<sup>i</sup> Ajou Translational Omics Center, Research Institute for Innovative Medicine, Ajou University Medical Center, Suwon, 16499, Republic of Korea

## ARTICLE INFO

### Keywords:

Mild cognitive impairment  
Alzheimer's disease  
Vascular dementia  
Plasma protein biomarker  
Protein-protein interaction  
Graph neural network

## ABSTRACT

As an early indicator of dementia, mild cognitive impairment (MCI) requires specialized treatment according to its subtypes for the effective prevention and management of dementia progression. Based on the neuropathological characteristics, MCI can be classified into Alzheimer's disease (AD)-related cognitive impairment (ADCI) and subcortical vascular cognitive impairment (SVCI), being more likely to progress to AD and subcortical vascular dementia (SVD), respectively. For identifying MCI subtypes, plasma protein biomarkers are recently seen as promising tools due to their minimal invasiveness and cost-effectiveness in diagnostic procedures. Furthermore, the application of machine learning (ML) has led the preciseness in the biomarker discovery and the resulting diagnostics. Nevertheless, previous ML-based studies often fail to consider interactions between proteins, which are essential in complex neurodegenerative disorders such as MCI and dementia. Although protein-protein interactions (PPIs) have been employed in network models, these models frequently do not fully capture the diverse properties of PPIs due to their local awareness. This limitation increases the likelihood of overlooking critical components and amplifying the impact of noisy interactions. In this study, we introduce a new graph-based ML model for classifying MCI subtypes, called *eXplainable Graph Propagational Network* (XGPN). The proposed method extracts the globally interactive effects between proteins by propagating the independent effect of plasma proteins on the PPI network, and thereby, MCI subtypes are predicted by estimation of the risk effect of each protein. Moreover, the process of model training and the outcome of subtype classification are fully explainable due to the simplicity and transparency of XGPN's architecture. The experimental results indicated that the interactive effect between proteins significantly contributed to the distinct differences between MCI subtype groups, resulting in an enhanced classification performance with an average improvement of 10.0 % compared to existing methods, also identifying key biomarkers and their impact on ADCI and SVCI.

\* Corresponding author. Department of Industrial Engineering, Ajou University, Worldcup-ro 206, Yeongtong-gu, Suwon, 16499, Republic of Korea.

\*\* Corresponding author. Department of Physiology, Ajou University School of Medicine, Worldcup-ro 164, Yeongtong-gu, Suwon, 16499, Republic of Korea.

E-mail addresses: [shin@ajou.ac.kr](mailto:shin@ajou.ac.kr) (H. Shin), [hg@ajou.ac.kr](mailto:hg@ajou.ac.kr) (H.G. Woo).

<sup>1</sup> Sunghong Park and Doyoon Kim contributed equally to this work.

## 1. Introduction

Mild cognitive impairment (MCI) is a neurodegenerative disorder that involves cognitive impairment beyond the expected decline in memory and thinking that comes with normal aging [1]. Although the clinical symptoms of MCI are not severe enough to interfere with instrumental activities of daily living, it is an important early indicator for dementia, as up to half of patients with MCI are known to develop dementia within five years [2]. Similar as Alzheimer's disease (AD) and subcortical vascular dementia (SVD), the two major subtypes of dementia, are distinguished by the neuropathological factors associated with beta-amyloid ( $A\beta$ ) deposition and cerebrovascular degeneration [3, 4], respectively, MCI is also classified into its major subtypes, AD-related cognitive impairment (ADCI) and subcortical vascular cognitive impairment (SVC), which involve cortical amyloid burden and subcortical vascular burden, respectively [5–7]. Compared to patients with normal MCI (NMCI), patients with ADCI and SVCI entailing neuropathologic risk factors are more likely to progress to AD and SVD, respectively [8–12]. Since dementia causes irreversible brain dysfunction, it requires appropriate prevention and management from the predementia stage, MCI. Furthermore, clinical treatment for patients with MCI needs to be specialized according to the subtypes of MCI, which are distinguished based on neuropathological characteristics.

While MCI patients are primarily evaluated by cognitive function assessment, additional information such as cerebrospinal fluid (CSF) biomarker or neuroimaging test is required to classify them into ADCI or SVCI [13–16]. CSF biomarkers, including  $A\beta_{42}$ , total tau, and phosphorylated tau, are effective predictors of conversion from MCI to dementia, but they require patients to be under prolonged bed rest for several hours and are invasive, which can cause rare but serious side effects [17]. Neuroimaging tests provide objective evidence regarding the deposition of  $A\beta$  and tau tracer with positron emission tomography (PET) and the determination of hippocampal or medial temporal neurodegeneration with magnetic resonance imaging (MRI), but require high costs [18]. This approach is also limited in its ability to capture the complexity of neurodegenerative diseases, which involve a variety of biological processes, including neuroinflammation, synaptic dysfunction, and metabolic changes. In contrast, notable advancements have been made in the field of neurology with the advent of plasma protein biomarkers as an alternative method of diagnosing and evaluating the potential pathophysiology of diverse neurodegenerative disorders, including MCI and dementia. The diagnosis based on plasma protein biomarkers is less invasive and cost-effective than the use of CSF biomarkers and neuroimaging tests. The implementation of these biomarkers within the existing healthcare system can reduce the patient burden and increase access to testing. As a result, they are widely used for the diagnosis of a variety of conditions, not just neurodegenerative diseases [19–22].

Moreover, the application of machine learning has facilitated the sophisticated identification of plasma protein biomarkers associated with neurodegenerative diseases and the precise prediction of target outcomes [23–26]. However, previous studies concentrated solely on the independent effects of individual proteins, thereby neglecting the interactions between proteins. This conventional approach fails to acknowledge the collective contribution of multiple proteins with small effect sizes to the phenotype due to their interactions. The identification of plasma protein biomarkers related to diseases based on protein-protein interactions (PPIs) is crucial for understanding the molecular pathogenesis of diseases, assessing risk, and classifying diseases [27]. The advent of high-throughput technologies has led to a rapid expansion in the data available on PPIs, which now encompasses almost the entire proteome [28]. Concurrently, network-based methods that involve PPIs have gained prominence as they allow the application of machine learning techniques to identify sophisticated biomarkers [29]. Neurodegenerative diseases are a complex disorder that involves the interactions among specific molecular pathways, making the impact of

PPIs even more significant [30,31]. Therefore, there is a need for a model that takes into account PPIs within molecular pathways.

Graph convolutional network (GCN) [32] employing PPIs may have elucidated the underlying subtypes of MCI. In GCN, the features of each node are aggregated with the features of neighboring nodes via edges through a graph convolutional layer. This process is applied to the PPI network, wherein the nodes, edges, and features correspond to the proteins, PPIs, and expression values derived from blood sequencing, respectively. Therefore, the extracted features of proteins represent the interactive effect between proteins, whereby the expression value of each protein is combined with the expression values of neighboring proteins in the PPI network. This approach of representing protein interaction by applying GCN to the PPI network has been employed for a variety of tasks, including the identification of novel PPIs [33–35], the classification of disease types [36], and the prediction of cancer survivability [37]. Furthermore, when examining PPIs, it is imperative to acknowledge the significant role that indirect interactions between proteins play in disease progression. Proteins transmit signals to other proteins at a distance, collectively influencing the progression of certain diseases. This mechanism suggests that the application of machine learning to the PPI network should result in the performance of targeted tasks with a global awareness of PPIs, whereby the full range of interactive effects can be extracted by considering the entire PPI network. However, the graph convolutional operation of GCN is constrained by its local awareness of PPIs, as it aggregates features only between one-hop neighboring nodes. This limitation hinders its capacity to capture local interactions between directly connected proteins within the PPI network. To address this shortcoming, an extended GCN can be employed to aggregate features from multi-hop neighbors.

Previously, two major approaches have been employed to extend GCN for  $K$ -order feature aggregation: the multiplied convolution filter and the parallelized network architecture. The initial approach, based on the multiplied convolution filter, involves repeatedly multiplying the normalized adjacent matrix  $K$  times, thereby representing the graph convolution filter as the  $K$ -th powered single matrix. This approach was initially proposed in Simple Graph Convolution (SGC) [38], and has since been further developed in Exponential Graph Convolution (EGC) and Linear Graph Convolution (LGC) [39]. While GCN necessitates  $K$  graph convolutional layers for feature aggregation with  $K$ -hop neighboring nodes, SGC is streamlined to a single layer through the  $K$ -th powered graph convolution filter. EGC derives the graph convolution filter by combining the graph Laplacian up to the  $K$ -th power with the coefficients of the exponential power series. LGC applies the graph convolution filter achieved by a linear combination of graph Laplacian monomials up to the  $K$ -th power. Subsequently, the approach based on the parallelized network architecture comprises  $K$  graph convolution filters that extend up to  $K$ -hop neighboring nodes on the adjacent matrix. Each filter is utilized in parallel graph convolutional layers for the purpose of aggregating node features on an individual basis. The most representative method of this approach is MixHop [40], which introduces a higher-order graph convolutional architecture. This architecture has been further advanced as the Universal Graph Convolutional Network (UGCN) [41] and the Mixed-Order Graph Convolutional Network (MOGCN) [42]. MixHop, UGCN, and MOGCN represent different approaches to merge node features extracted from parallel graph convolutional layers, employing simple concatenation, an attention mechanism, and an ensemble module, respectively.

Extended versions of GCN have the ability to represent a wider range of PPIs; however, they are limited to fully capture the holistic nature of the PPI network, which in turn gives rise to several challenges. (1) *Disregarding the structural attributes of the PPI network*: the PPI network is a complex system with a hierarchical structure, comprising sub-networks. A focus on interactions between adjacent proteins may result in the overlooking of critical features across the entire network [43,44]. (2) *Missing the pivotal components within the PPI network*: hub proteins are pivotal to the interconnections between proteins and exert a significant

influence on the entire PPI network. However, these components are not always taken into account in the local PPIs [45]. (3) *Emphasizing the noisy interactions within the PPI network*: as the PPI network incorporates PPI data from various sources, it inherently includes experimentally noisy interactions. The utilization of local PPIs may result in the overestimation of the impact of noisy interactions [46–48]. While extending the architecture of the GCN-extended model and diversifying its configuration may mitigate these challenges to some extent, it represents only a provisional solution that does not achieve a comprehensive implementation of PPIs. This limitation arises from the fact that the graph convolution relies on locality, which enables the aggregation of features between adjacent nodes. To comprehensively address these challenges and accurately capture the characteristics of the PPI network, graph neural networks necessitate the use of globality-based feature aggregation.

In this study, we introduce a novel graph neural network called *eXplainable Graph Propagational Network* (XGPN). Central to our approach is the graph propagational layer, which generates a globally aggregated feature representation by diffusing the features of each node across all nodes within the graph. This enables the classification of MCI subtypes based on plasma protein biomarkers, leveraging the interactions among proteins to accurately reflect the key components and structural properties of the PPI network. The proposed method is schematically described in Fig. 1. First, we identify plasma protein biomarkers specific for ADICI and SVCI, MCI subtypes differentiated on the basis of neuroimaging. Next, we train the proposed model that utilizes the identified biomarkers to classify MCI subtypes, applying the PPI network to reflect the global interactions between biomarkers. The contributions of the proposed method are summarized as follows.

- We propose a novel method called XGPN for MCI subtype classification as it is medically important to differentiate MCI, a pre-dementia stage, into subtypes based on its neuropathological features.
- ADICI and SVCI are subtypes of MCI based on neuroimaging, and XGPN utilizes plasma protein data to distinguish between them as a less-invasive and cost-effective diagnostic tool.
- Differentially expressed proteins for ADICI and SVCI are identified as biomarkers for MCI subtype classification, and the functional effects of the biomarkers on each subtype are also analyzed.
- The identified biomarkers are utilized in the training of XGPN, where the PPI network is applied to reflect the global interactions between proteins, and explanatory results for the internals and outcomes of XGPN are also provided.

In the next section, we introduce the study materials representing information about the participants, the data collected from them. The methodology section delineates the comprehensive process of MCI subtype classification, encompassing the identification of plasma protein biomarkers and the detailed mathematical implementation of XGPN. The experimental section illustrates the predicted outcomes of

XGPN and compares them with existing methods, including an explanation of the proposed method. The last section concludes the paper with remarks on the contributions and limitations of the proposed method.

## 2. Materials

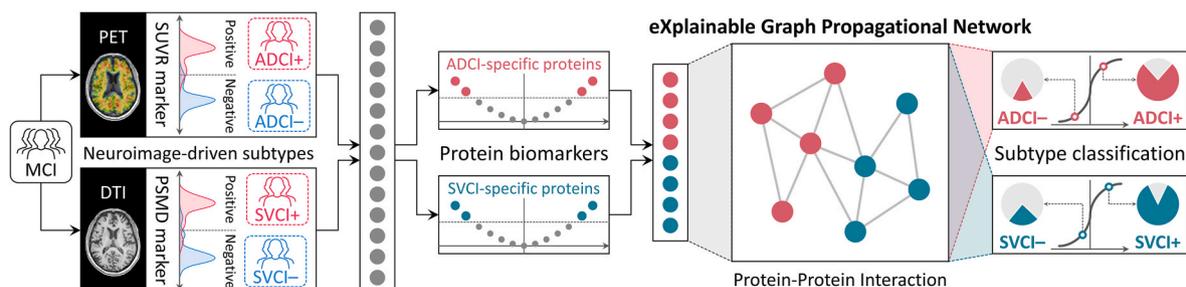
### 2.1. Study participants

Participants were recruited from the Biobank Innovations for chronic Cerebrovascular disease With ALzheimer’s disease Study (BICWALZS) at Ajou University Hospital (Suwon, Republic of Korea) [49]. Of the participants diagnosed with MCI according to the expanded Mayo Clinic criteria [50], we included 244 participants in the study cohort based on the tests for cognitive function assessment as follows: the mini-mental state examination (MMSE) of 20 or higher, the global score of clinical dementia rating (CDR) of 0.5, the CDR sum of boxes (CDR-SB) of 0.5–4, and the global deterioration scale (GDS) of 2–4. MCI subtypes for study participants were categorized based on neuroimaging-based diagnostic markers, with positron emission tomography (PET)-based standard uptake value ratio (SUVR) for ADICI diagnosis and diffusion tensor imaging (DTI)-based peak width of skeletonized mean diffusivity (PSMD) for SVCI diagnosis. There were 47 (19.3 %) participants with ADICI and 30 (12.3 %) participants with SVCI. The study participants were also divided into the discovery and validation cohorts, where the validation cohort included participants with 2-year follow-up on cognitive function assessments, and the remaining participants for whom only baseline information was available were included in the discovery cohort. As a result, there were 189 participants in the discovery cohort, including 36 (19.0 %) with ADICI and 23 (12.2 %) with SVCI, and the validation cohort included 55 participants, consisting of 11 (20.0 %) with ADICI and 7 (12.7 %) with SVCI. Demographic and clinical characteristics of the study participants are summarized in Table 1.

### 2.2. Data acquisition and processing

#### 2.2.1. Positron emission tomography (PET)

The study participants underwent  $^{18}\text{F}$ -flutemetamol PET scans by using a Discovery Ste/690 PE T/CT scanner (GE, Milwaukee, WI, USA) with an identical protocol. An intravenous bolus of  $^{18}\text{F}$ -flutemetamol was administered at a mean dose of 185 MBq. Subsequently, a 20-min PET scan was conducted, comprising four 5-min dynamic frames. The  $^{18}\text{F}$ -flutemetamol PET scans were co-registered to the individual MRI scans, which were normalized to a T1-weighted MRI template. The MRI-based co-registered  $^{18}\text{F}$ -flutemetamol PET images were normalized to the MRI template using transformation parameters. To quantify the  $^{18}\text{F}$ -flutemetamol retention, the SUVR was obtained by using the pons as a reference region. Global cortical  $^{18}\text{F}$ -flutemetamol retention was calculated from the volume-weighted average SUVR of bilateral ten cortical volumes of interest from the frontal, posterior cingulate, lateral temporal, parietal, and occipital lobes using the annotated anatomical



**Fig. 1.** Schematic description of the proposed method. Differentially expressed proteins for ADICI and SVCI, MCI subtypes distinguished by neuroimaging, are firstly identified as plasma protein biomarkers. The proposed model is then trained using the identified biomarkers for MCI subtype classification, applying the PPI network to reflect the global interactions between proteins.

**Table 1**  
Demographic and clinical characteristics of study participants.

(a) MCI subtype-wise comparison of characteristics for total participants						
Characteristics	ADCI			SVCI		
	Positive (N = 47)	Negative (N = 197)	P-value	Positive (N = 30)	Negative (N = 214)	P-value
Age, median (IQR), yr	75 (72–81)	72 (66–76)	<0.0001	75 (71–82)	72 (66–77)	0.0019
Female, No. (%)	33 (70.2)	140 (71.1)	0.9083	19 (63.3)	152 (72.0)	0.3318
MMSE, median (IQR)	26 (23–27)	26 (24–27)	0.6164	25 (23–26)	26 (24–27)	0.1144
CDR-SB, median (IQR)	2.0 (1.5–3.0)	1.5 (1.0–2.5)	0.0220	1.5 (1.0–2.5)	1.5 (1.0–2.5)	0.7257
GDS, median (IQR)	3 (2–3)	3 (2–3)	0.3360	3 (2–3)	3 (2–3)	0.8831

(b) Study cohort-wise comparison of characteristics of participants				
Characteristics	Total participants (N = 244)	Discovery cohort (N = 189)	Validation cohort (N = 55)	P-value
Age, median (IQR), yr	73 (67–77)	72 (67–77)	73 (67–77)	0.9930
Female, No. (%)	173 (70.9)	133 (70.4)	40 (72.7)	0.7361
MMSE, median (IQR)	26 (23–27)	26 (24–27)	25 (23–27)	0.1390
CDR-SB, median (IQR)	1.5 (1.0–2.5)	1.5 (1.0–2.5)	2.0 (1.0–2.5)	0.4722
GDS, median (IQR)	3 (2–3)	3 (2–3)	3 (3–3)	0.1278

Abbreviations: IQR, interquartile range; MMSE, mini-mental status examination; CDR-SB, clinical dementia rating sum of boxes; GDS, global deterioration scale.

labeling atlas. Finally, the study participants were classified as ADCI, if their SUVR values were greater than  $6.740 \times 10^{-1}$ , where the cut-off value was calculated by using the “cutoff” packages [51] implemented in R (<https://github.com/choisy/cutoff>).

### 2.2.2. Diffusion tensor imaging (DTI)

The study participants also underwent scanning of two-dimensional EPI diffusion tensor imaging (DTI) that was performed at TR/TE = 8100/87 ms, resolution =  $2.5 \times 2.5 \times 2.5 \text{ mm}^3$ , number of directions = 64, b-values = 0, and 1000 s/mm<sup>2</sup>. We quantified the impact of cerebrovascular burden by calculating a global value of PSMD [52] (<http://www.psm-marker.com>), which is a diffusion-based metric that indicates small vessel disease. The calculation of PSMD was automatically performed by using the PSMD tool (<https://github.com/miac-research/psmd>) comprising DTI skeletonization and histogram analysis. Initially, DTI scans were skeletonized using the tract-based spatial statistics process, a component of the FMRIB Software Library (FSL) [53] (<https://fsl.fmrib.ox.ac.uk>), and the FMRIB 1-mm fractional anisotropy (FA) template thresholded at an FA value of 0.2. The mean diffusivity (MD) images were then projected onto the skeleton using the FA-derived projection parameters, followed by generating FA and MD output maps. PSMD was then automatically derived from the fully processed diffusion data as the discrepancy between the 95th and 5th percentiles of the MD voxel values within the skeleton. Finally, the study participants were classified as SVCI, if their PSMD values were greater than  $4.888 \times 10^{-4}$ , which was calculated by using the “cutoff” packages.

### 2.2.3. Plasma protein assays

Plasma samples from the study participants were profiled by the Olink Proteomics using proximity extension assay (PEA) technology, and we used the Olink Target 96 Neurology and Olink Target 48 Cytokine panels in this study. The neurology panel includes the established 92 assays associated with neurobiological processes and neurological diseases (e.g., neurodevelopment, axon guidance, or synaptic function), and the cytokine panel contains the selected 45 assays with the highest relevance for studies of inflammatory diseases, or for investigation of the inflammatory processes that underly many diseases, covering key immunological processes. Quality control of the raw data was performed by using the internal and external controls in the panels. Of the 137 proteins in total, 10 proteins with higher than 10 % of missing frequency were excluded, and the imputation of missing data for the remaining 127 proteins was performed by using the k-nearest neighbor method. Subsequently, the protein expression levels were standardized by the Z-score normalization and then transformed by the logistic function for scaling. The full list of proteins in the two panels and the quality control results for them are summarized in the [Supplementary Tables S1 and S2](#).

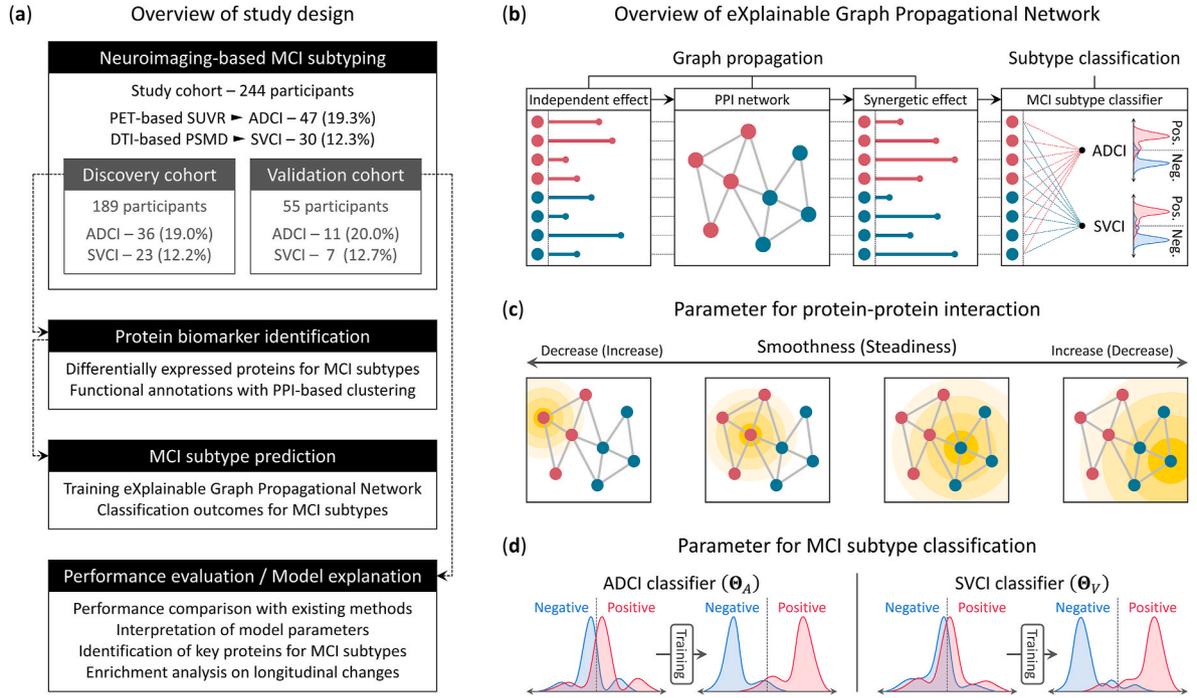
## 3. Methods

### 3.1. Overall framework

**Fig. 2** depicts the overall framework employed in this study. As illustrated in **Fig. 2(a)**, this study initially identifies differentially expressed proteins (DEPs) for each MCI subtype for prediction using plasma protein data for each MCI subtype, categorized based on neuroimaging. Subsequently, XGPN is trained on the data for the identified biomarkers to classify the MCI subtypes, followed by the performance evaluation and the model explanation. A schematic representation of XGPN is provided in **Fig. 2(b)**. As the proposed model aims to classify MCI subtypes by learning the interactive effects between proteins, XGPN was applied to the PPI network collected from the STRING database (<https://string-db.org/>) [54,55], which encompasses the global interactions between proteins. The interactive effects are extracted by propagating the independent effects across the PPI network, and in this process, the smoothness parameter controls the range of PPIs, determining the extent to which the independent effects are propagated by XGPN, as represented in **Fig. 2(c)**. A lower smoothness value indicates that lower-order PPIs, representing interactions between nearby proteins, are given higher weight in XGPN. Conversely, a higher smoothness value indicates that higher-order PPIs, reflecting interactions between distant proteins, exert a greater influence on XGPN. The smoothness is typically pre-defined; however, in XGPN, it is a learnable parameter and is optimized individually for each protein. Subsequently, XGPN combines the extracted interactive effect with the classifier parameters shown in **Fig. 2(d)**, thereby enabling the classification of MCI subtypes.

### 3.2. Protein biomarker identification

The preprocessed data from plasma samples of 189 participants in the discovery cohort is utilized to identify differentially expressed proteins (DEPs) by MCI subtypes. The significance of the subtype-wise DEPs is estimated by the linear regression analysis using “limma” [56] implemented in an R/Bioconductor (<https://www.bioconductor.org/>). Subsequently, Gene Ontology (GO) analysis is conducted on the identified plasma protein biomarkers to investigate functional annotations and gain insights into the biological meaning of the biomarkers, by using Bioinformatics Resources provided by the Database for Annotation, Visualization, and Integrated Discovery (DAVID, <https://david.ncifcrf.gov>, version 2021) [57,58].



**Fig. 2.** Overall framework for this study. (a) shows the overview of study design, and (b) schematically represents XGPN. (c) and (d) indicate the smoothness parameter for the PPI network and the classifier parameter for MCI subtype diagnosis, respectively.

### 3.3. MCI subtype prediction

#### 3.3.1. Formulated implementation

Let the data matrices for the independent effect and PPI network are denoted as  $\mathbf{X} \in \mathbb{R}^{p \times n}$  and  $\mathbf{W} \in \mathbb{R}^{p \times p}$ , respectively, where  $p$  and  $n$  are the number of plasma protein biomarkers and study participants, respectively. The interactive effect, denoted as  $\mathbf{H} \in \mathbb{R}^{p \times n}$ , is derived by propagating the independent effect to the PPI network. This process has two objectives for representing the interactive effect: *smoothness* and *steadiness*. The smoothness indicates that the features of nodes being more strongly connected are more similarly represented to each other, while the steadiness indicates that the propagated feature should not be too much different from the original feature. In the same manner as graph-based semi-supervised learning [59],  $\mathbf{H}$  can be obtained by minimizing the following quadratic function that implements those two objectives:

$$\sum_{i \sim j} \mathbf{W}^{(ij)} (\mathbf{H}^{(i)} - \mathbf{H}^{(j)})^2 + \phi \sum (\mathbf{H}^{(i)} - \mathbf{X}^{(i)})^2 \quad (1)$$

where the first and the second terms correspond to the smoothness and the steadiness, respectively,  $i \sim j$  indicates that nodes  $\mathcal{V}^{(i)}$  and  $\mathcal{V}^{(j)}$  are adjacent in  $\mathbf{W}$ , and  $\phi$  is the steadiness parameter that trades off two terms. The steadiness parameter is usually fixed to a user-specified value so that all proteins have the same steadiness; rather, the proposed method aims to train the steadiness and further optimize for each protein individually. To implement this, (1) is transformed as follows,

$$\sum_{i \sim j} \mathbf{W}^{(ij)} (\mathbf{H}^{(i)} - \mathbf{H}^{(j)})^2 + \sum \phi^{(i)} (\mathbf{H}^{(i)} - \mathbf{X}^{(i)})^2 \quad (2)$$

where the steadiness parameter for each protein is defined as  $\phi = (\phi^{(1)}, \dots, \phi^{(i)}, \dots, \phi^{(r)}) \in \mathbb{R}^r$ . Very often, the quadratic problem of (2) is represented in terms of matrix,

$$\min_{\mathbf{H}} \mathbf{H}^T \mathbf{L} \mathbf{H} + (\mathbf{H} - \mathbf{X})^T \Phi (\mathbf{H} - \mathbf{X}) \quad (3)$$

where  $\Phi$  is the matrix of steadiness parameters defined as  $\Phi = \text{Diag}(\phi)$ , and  $\mathbf{L}$  is the normalized graph Laplacian defined as  $\mathbf{L} = \mathbf{I}_p -$

$\mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}$  with the identity matrix  $\mathbf{I}_p \in \mathbb{R}^{p \times p}$  and the diagonal matrix  $\mathbf{D} = \text{Diag}(\mathbf{D}^{(i)})$  from  $\mathbf{D}^{(i)} = \sum_j \mathbf{W}^{(ij)}$ . The solution of (3) is obtained in the closed form as below.

$$\mathbf{H} = (\Phi + \mathbf{L})^{-1} \Phi \mathbf{X} \quad (4)$$

Subsequently,  $\mathbf{H}$  in (4) is combined with the classifier vectors, denoted as  $\Theta_{AD}$  and  $\Theta_{SV}$ , and then, the risk probabilities for AD CI and SV CI, represented as  $\mathbf{P}_{AD}$  and  $\mathbf{P}_{SV}$ , are derived by applying the logistic function as below.

$$\mathbf{P}_* = \frac{1}{1 + e^{-\Theta_* \mathbf{H}}} \quad (* : AD, SV) \quad (5)$$

#### 3.3.2. Optimization procedure

By denoting the actual diagnosis as  $\mathbf{Y}_* \in \mathbb{R}^{1 \times n}$ , which includes binary elements indicating whether each MCI subtype or not as 1 or 0, respectively, the parameters are trained to minimize the binary cross-entropy loss  $\mathcal{L}_*$  as represented below.

$$\mathcal{L}_* = \frac{1}{n} \{ (\mathbf{Y}_*^T \log \mathbf{P}_*) + (\mathbf{1}_n - \mathbf{Y}_*)^T \log(\mathbf{1}_n - \mathbf{P}_*) \}$$

Therefrom, the objective function for XGPN is defined as follows:

$$\text{argmin}_{\phi, \Theta_{AD}, \Theta_{SV}} \mathcal{L}_{AD} + \mathcal{L}_{SV} + \delta \mathcal{R} \quad (6)$$

where  $\mathcal{R} = \|\phi\|_2^2 + \|\Theta_{AD}\|_2^2 + \|\Theta_{SV}\|_2^2$  is the L2 regularization term to penalize the complexity with the positive coefficient  $\delta$ . The objective function is optimized by the gradient descent method [60].

**Minimization over  $\Theta_*$ :** to find gradient *w.r.t.* the classifier matrix  $\Theta_*$ , the terms in (6) are firstly derived as below.

$$\frac{\partial \mathcal{L}_*}{\partial \Theta_*} = \frac{1}{n} \mathbf{H} (\mathbf{P}_* - \mathbf{Y}_*)^T, \quad \frac{\partial \mathcal{R}}{\partial \Theta_*} = 2\delta \Theta_*$$

Then, the gradient of  $\Theta_*$  is obtained as follows.

$$\nabla \Theta_* = \frac{1}{n} \mathbf{H}(\mathbf{P}_* - \mathbf{Y}_*)^T + 2\delta \Theta_*$$

**Minimization over  $\phi$ :** to find the gradient *w.r.t.*  $\phi$ , the derivative of  $\mathcal{L}_*$  *w.r.t.*  $\mathbf{H}$  and the derivative of  $\mathbf{H}$  *w.r.t.*  $\Phi$  are firstly obtained as follows:

$$\frac{\partial \mathcal{L}_*}{\partial \mathbf{H}} = \frac{1}{n} \Theta_* (\mathbf{P}_* - \mathbf{Y}_*), \quad \frac{\partial \mathbf{H}}{\partial \Phi} = \mathbf{X}^T (\Phi + \mathbf{L})^{-1} \left\{ \mathbf{I}_p - (\Phi + \mathbf{L})^{-1} \Phi \right\}.$$

Then, the derivative of  $\mathcal{L}$  *w.r.t.*  $\Phi$  is presented as below.

$$\frac{\partial \mathcal{L}}{\partial \Phi} = \frac{\partial (\mathcal{L}_{AD} + \mathcal{L}_{SV})}{\partial \Phi} = \frac{1}{n} \left\{ \Theta_{AD} (\mathbf{P}_{AD} - \mathbf{Y}_{AD}) + \Theta_{SV} (\mathbf{P}_{SV} - \mathbf{Y}_{SV}) \right\} \times \mathbf{X}^T (\Phi + \mathbf{L})^{-1} \left\{ \mathbf{I}_p - (\Phi + \mathbf{L})^{-1} \Phi \right\} \quad (7)$$

Next, the expression of  $\Phi$  is transformed by using an  $r$ -dimensional row vector  $\mathbf{1}_p$  with all elements being 1, as

$$\Phi = \text{Diag}(\phi) = \phi \mathbf{1}_p \odot \mathbf{I}_p,$$

and therefrom, the derivative of  $\Phi$  *w.r.t.*  $\phi$  is indicated as below.

$$\frac{\partial \Phi}{\partial \phi} = \mathbf{I}_p \mathbf{1}_p^T \quad (8)$$

Finally, by combining (7) with (8), the derivative of  $\mathcal{L}$  *w.r.t.*  $\phi_*$  is obtained as

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \phi} &= \frac{\partial \mathcal{L}}{\partial \Phi} \odot \frac{\partial \Phi}{\partial \phi} \\ &= \frac{1}{n} \left\{ \Theta_{AD} (\mathbf{P}_{AD} - \mathbf{Y}_{AD}) + \Theta_{SV} (\mathbf{P}_{SV} - \mathbf{Y}_{SV}) \right\} \times \mathbf{X}^T (\Phi + \mathbf{L})^{-1} \left\{ \mathbf{I}_p \right. \\ &\quad \left. - (\Phi + \mathbf{L})^{-1} \Phi \right\} \odot \mathbf{I}_p \mathbf{1}_p^T, \end{aligned}$$

and then, the gradient *w.r.t.*  $\phi$  is derived as follows.

$$\nabla \phi = \frac{1}{n} \left\{ \Theta_{AD} (\mathbf{P}_{AD} - \mathbf{Y}_{AD}) + \Theta_{SV} (\mathbf{P}_{SV} - \mathbf{Y}_{SV}) \right\} \times \mathbf{X}^T (\Phi + \mathbf{L})^{-1} \left\{ \mathbf{I}_p - (\Phi + \mathbf{L})^{-1} \Phi \right\} \odot \mathbf{I}_p \mathbf{1}_p^T + 2\delta \phi$$

## 4. Results

### 4.1. Differentially expressed proteins for MCI subtypes

To elucidate the underlying molecular characteristics of the MCI subtypes, DEP analysis was conducted on the 127 plasma proteins of the discovery cohort. As shown in Fig. 3(a) and (b), 29 proteins were identified as plasma protein biomarkers for MCI subtypes, comprising 5 and 24 proteins for ADCI and SVCI, respectively, and the list of proteins for the two subtypes is provided in the [Supplementary Table S3](#). For ADCI, NTRK3 and MDGA1 were the upregulated proteins, while LAIR2, FCRL2, and CCL19 were downregulated. NTRK3 and MDGA1 have been revealed to be involved in the central nervous system [61] and the amyloid precursor protein [62], respectively, and thereby, they have been associated with Alzheimer's disease and other neuropsychiatric disorders [63,64]. FCRL2 has previously been implicated in AD pathophysiology through its essential role in immune pathways [65,66], and recent studies have identified LAIR2 and CCL19 as AD-related

biomarkers significantly related to cognitive function [67,68]. On the other hand, there were 21 upregulated proteins for SVCI, the most remarkable being MMP12, which is an emerging biomarker as an early indicator of cerebral small vessel disease [69], while BCAN, NCAN, and CNTN5 were downregulated proteins for SVCI. BCAN and NCAN were demonstrated as diagnostic biomarkers for VD [70], and CNTN5 was found to be significantly associated with brain vascular burden [71]. Furthermore, as illustrated in Fig. 3(c), a comparison of the ADCI and SVCI groups revealed that 5 proteins exhibited higher expression levels in the ADCI group than in the SVCI group, while 24 proteins demon-

strated lower expression levels in the ADCI group than in the SVCI group. These findings suggest that the identified plasma protein biomarkers are consistent with clinical studies on ADCI and SVCI and include proteins that are already well recognized as significant biomarkers as well as those that are emerging as novel biomarkers, and thus, the identified 29 plasma proteins may be useful biomarkers for classifying MCI subtypes.

### 4.2. Functional annotations for the identified biomarkers

To characterize the functional annotations of proteins specific to MCI subtypes, we first performed the PPI network-based clustering on the identified biomarkers, and as shown in Fig. 4(a), the biomarkers were grouped into three clusters, where the cluster A, B, and C contain 8, 13, and 8 proteins, respectively. Subsequently, as illustrated in Fig. 4(b) and (c), GO analysis was conducted to ascertain the primary functions of each cluster, and the PPI subnetworks for protein clusters were examined. The protein cluster A was found to be associated with general neurological functions, including neurogenesis and nervous system development, and within the cluster, MDGA1 was the hub node, interacting with all other proteins, followed by BCAN, CNTN5, and NTRK3.

These four proteins showed remarkable significance for ADCI and SVCI by DEP analysis. For the protein cluster B, immunity-associated functions, including cytokine activity and inflammatory response, were dominant, with CCL3, CXCL9, and IL17A being the hub nodes of the PPI subnetwork. These were SVCI-specific upregulated proteins that were found to be related to the blood-brain barrier and significantly associated with cerebral small vessel disease [72–74]. The protein cluster C was identified as playing a role in signal transduction between the other clusters. As a result, PPI-based clustering of the identified biomarkers revealed distinct functional annotations for each cluster, and three main functions of the biomarkers were identified: neurological function, immunity-associated function, and signal transduction.

### 4.3. Classification outcomes for MCI subtypes

Fig. 5 depicts the outcomes of MCI subtype classification for the validation cohort, as predicted by the proposed method. As shown in Fig. 5(a) and (b), the averages of predicted ADCI and SVCI risks for all 55 participants were 0.4026 and 0.3282, respectively, representing that the ADCI risk indicated 1.23 times higher than the SVCI risk on average.

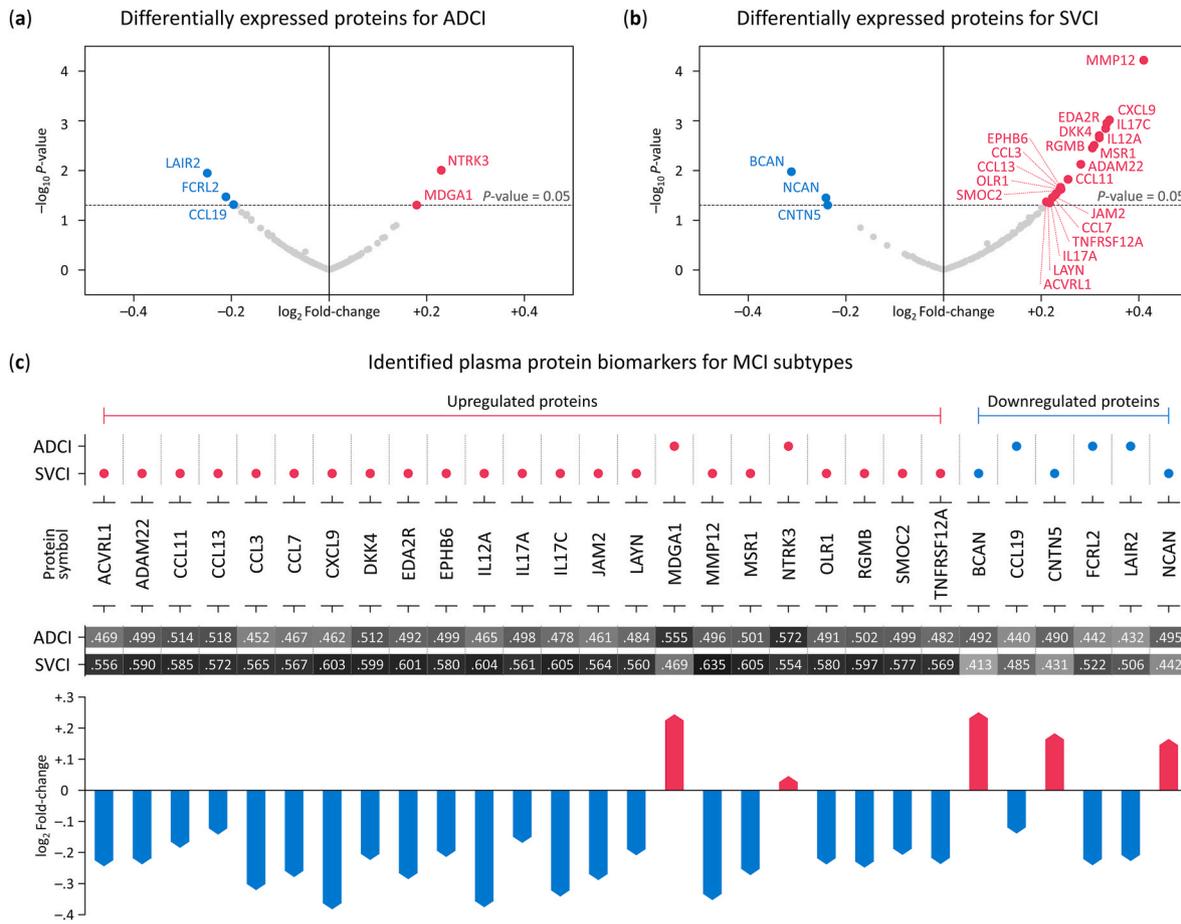


Fig. 3. Differentially expressed proteins for MCI subtypes. (a) and (b) represent differentially expressed proteins for ADCI and SVCI, respectively, and (c) compares the identified biomarkers across MCI subtypes.

Comparing the outcomes by predicted diagnoses, the averages for the ADCI-negative and ADCI-positive groups were 0.3525 and 0.6033, respectively, and the averages for the SVCI-negative and SVCI-positive groups were 0.3031 and 0.4999, respectively, with  $P$ -values of  $2.29 \times 10^{-6}$  and  $5.21 \times 10^{-5}$  for the group-wise differences for ADCI and SVCI, respectively. In addition, as illustrated in Fig. 5(c) and (d), we validated the predicted risks for ADCI and SVCI by comparing them to actual diagnosis and neuroimaging markers. The predicted risks for both MCI subtypes had an AUROC performance of 0.8806 on average, with an AUROC for ADCI of 0.9104, which is 1.07 times higher than the AUROC for SVCI of 0.8507. The predicted risks also showed an average Pearson correlation of 0.5757 with the neuroimaging markers, with ADCI indicating a correlation of 0.6382 for SUVR, which was 1.24 times higher than SVCI, which represented a correlation of 0.5132 for PSMD. As a result, the classification of MCI subtypes by the proposed method not only accurately matched the binarized actual diagnosis, but also significantly correlated with neuroimaging-based diagnostic markers represented as continuous values.

#### 4.4. Enrichment analysis on the predicted outcomes

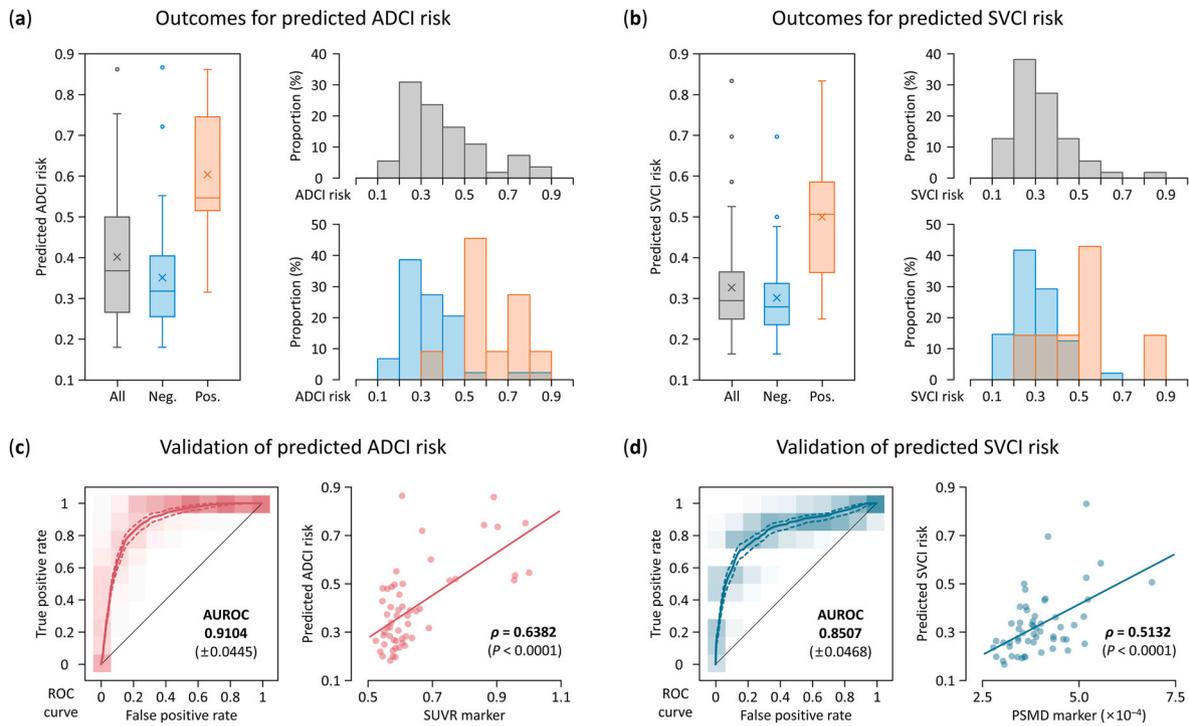
Based on the predicted outcomes for MCI subtypes, the two-year changes in cognitive function exhibited by participants in the validation cohort were analyzed as shown in Fig. 6. At first, the longitudinal changes in MMSE, CDR-SB, and GDS were compared between the positive and negative groups for ADCI or SVCI, with 15 and 40 participants included in the positive and negative groups, respectively. As illustrated in Fig. 6(a), in the negative group, the averages for MMSE, CDR-SB, and GDS at baseline were 25.2, 1.8, and 2.8, respectively, and at follow-up

were 25.7, 2.0, and 2.9, respectively, all of which were not significantly different with an average  $P$ -value of 0.3736. In contrast, the positive group showed a significant difference with an average  $P$ -value of 0.0211 for the longitudinal changes in the three measures, where the averages for MMSE, CDR-SB, and GDS at baseline were 24.2, 2.3, and 3.3, respectively, and at follow-up were 21.1, 3.4, and 4.0, respectively, all of which indicated remarkable changes. In Fig. 6(b), we further compared the follow-up results for cognitive function between ADCI-positive and SVCI-positive groups. For the ADCI-positive group, the averages of MMSE, CDR-SB, and GDS were 20.5, 3.5, and 4.1, respectively, compared to 23.2, 2.7, and 3.2, respectively, for the SVCI-positive group, where the ADCI-positive group showed an average 23.3 % decline in cognitive function compared to the SVCI-positive group. These findings are consistent with clinical studies demonstrating that the amyloid burden contributes more to cognitive decline than does the vascular burden [75–77], suggesting that the predicted outcomes by the proposed method well differentiate ADCI and SVCI.

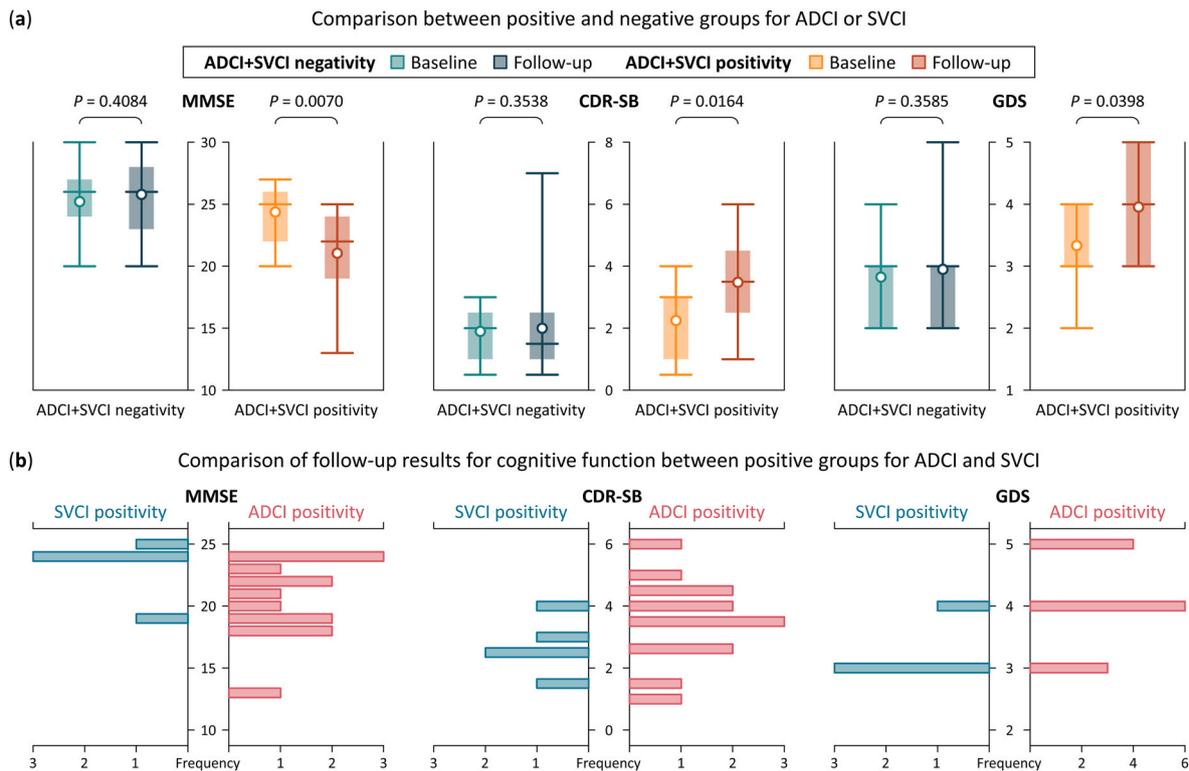
#### 4.5. Performance comparison with existing methods

The performance of XGPN was evaluated by comparing it to seven different methods: GCN [32], SGC [38], EGC [39], LGC [39], MixHop [40], UGCN [41], and MOGCN [42]. The model architectures for the comparison methods were configured by referencing the best performance reported in each paper. Therefore, the maximum order of interaction between proteins that reflected in those methods were 2 for GCN and SGC, 3 for EGC, 4 for MixHop and UGCN, 5 for LGC, and 6 for MOGCN. The model performance was measured by area under receiving operating characteristic curve (AUROC), area under precision-recall





**Fig. 5.** Classification outcomes for MCI subtypes. (a) and (b) depict the predicted risks by the proposed method for ADCI and SVCI, respectively, and their comparison by diagnostic group. (c) and (d) show the validation of the predicted risks for ADCI and SVCI, respectively, by comparing them with the actual diagnosis and neuroimaging markers.



**Fig. 6.** Enrichment analysis on the predicted outcomes. We analyzed the changes in cognitive function over two years exhibited by participants in the validation cohort. (a) shows the comparison between positive and negative groups for ADCI or SVCI, and (b) indicates the comparison of follow-up results between ADCI positive and SVCI positive groups.

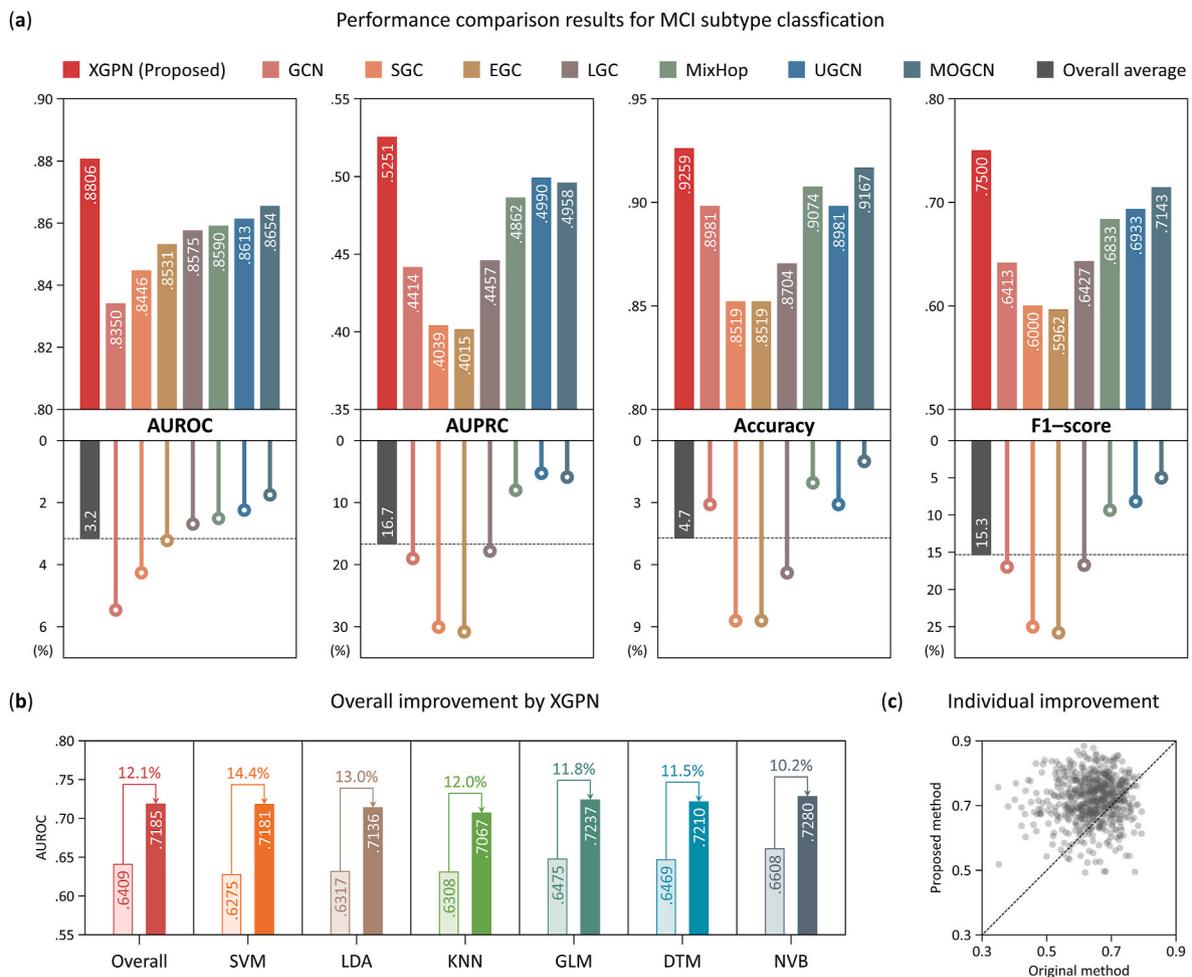


Fig. 7. Performance comparison with existing methods. (a) compares the performance of the proposed method with that of existing methods for the classification of MCI subtypes. (b) and (c) evaluate the discriminative power of the proposed method using six traditional machine learning algorithms.

average AUROC performance of 0.6409 observed when training with the independent effect. As depicted in Fig. 7(c), by comparing the individual AUROC improvement by training the interactive effect, where a point in the scatter plot located above the diagonal line indicates that the model on the vertical axis performs better, the results demonstrate that the majority of the dots lie above the diagonal line. Consequently, the interactive effect by XGPN indicates a significant difference between all MCI subtype groups, and this discriminative power contributes meaningfully to the diagnosis of MCI subtypes using other algorithms.

#### 4.6. Interpretation of model parameters

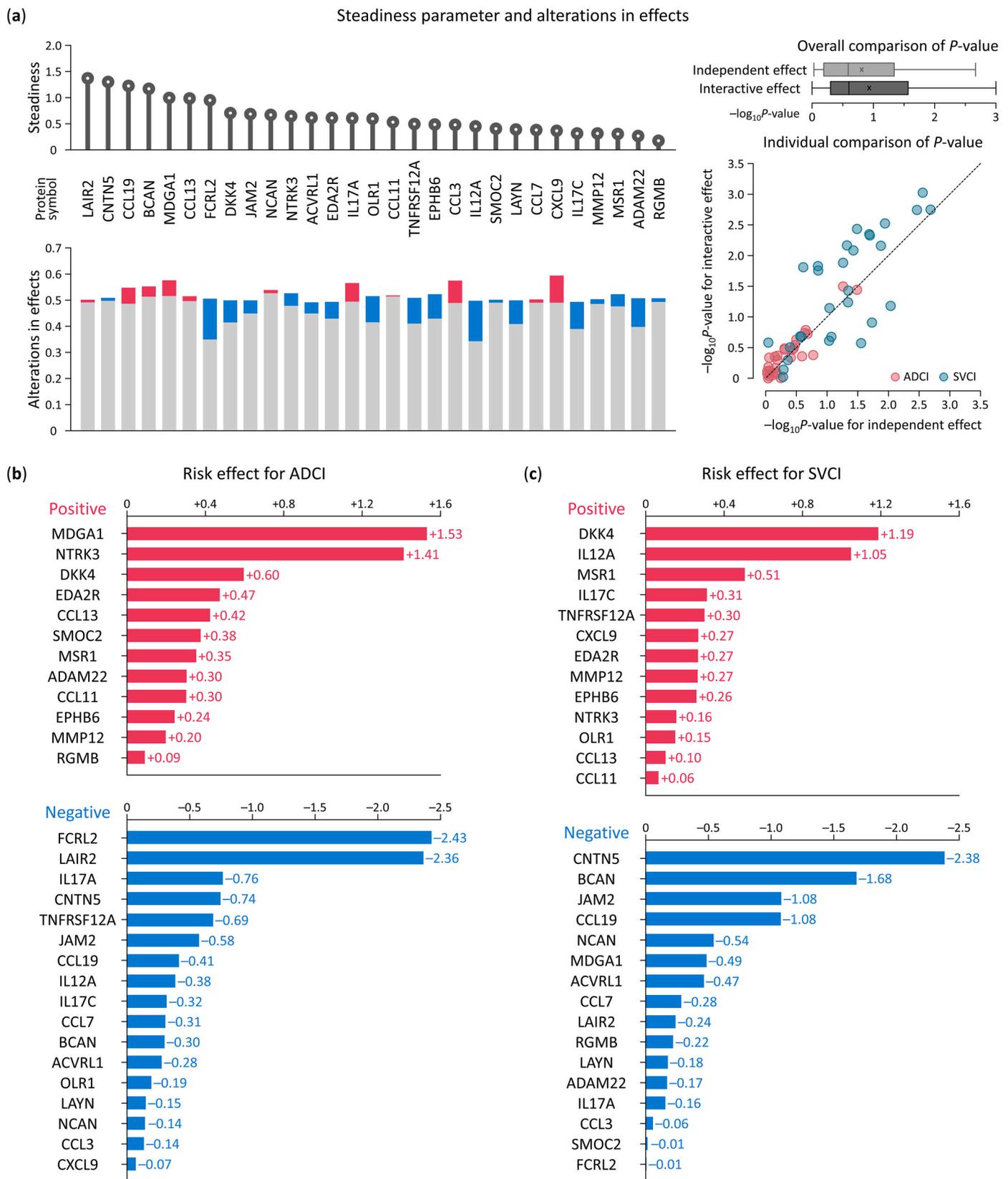
The interpretations of the steadiness and classifier parameters in XGPN are presented in Fig. 8. At first, the results for the steadiness parameter are displayed in Fig. 8(a). Following the optimization of the steadiness parameters for the plasma protein biomarkers (all of which had initial values of 1), 7 of the 29 biomarkers exhibited an increase in their steadiness, reaching an average of 1.2908, while the remaining 22 biomarkers exhibited a decrease in steadiness, reaching an average of 0.6269, resulting in an overall average of 0.7871. The overall average of the extracted interactive effect was 0.4759, representing a 5.48 % decrease compared to the average of the independent effect, which was 0.5035, where the difference between the two effects was statistically significant ( $P$ -value = 0.0338), with 11 and 18 biomarkers exhibiting increases and decreases, respectively. For each effect, a comparison between the positive and negative groups for ADCI and SVCI was performed, and the overall average of  $-\log_{10}P$ -value for the interactive

effect of the identified biomarkers was 0.93, which was 14.8 % higher than the value of 0.81 for the independent effect. An individual comparison of the biomarkers for each MCI subtype revealed that 19 and 20 biomarkers exhibited the more significant group-wise differences with interactive effects than with the independent effects for ADCI and SVCI, respectively. These results substantiate the advantage of XGPN that incorporating the global PPI into the independent effects enhances the discriminatory power for classifying MCI subtypes.

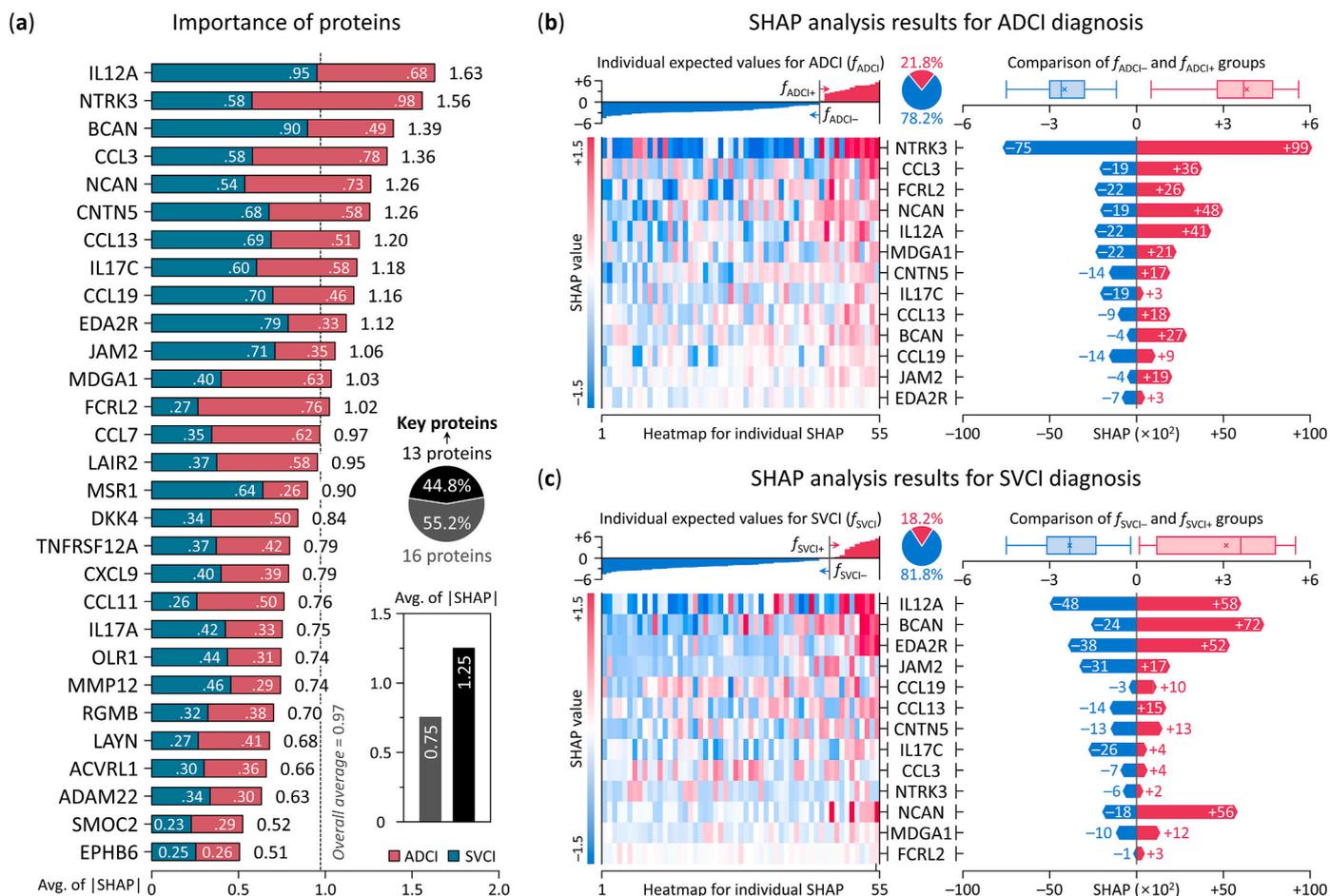
Next, the classifier parameters for ADCI and SVCI are presented in Fig. 8(b) and (c), respectively. For the ADCI classification, 12 proteins were found to exert a positive risk effect, while the remaining 17 demonstrated a negative one. MDGA1 and NTRK3, which were identified as upregulated proteins for ADCI, exhibited markedly higher risk effects in comparison to other proteins. In contrast, FCRL2 and LAIR2, identified as downregulated proteins for ADCI, exhibited notable negative risk effects. Furthermore, for the classification of SVCI, 13 and 16 proteins exhibited positive and negative risk effects, respectively. Among the proteins with positive risk effects, DKK4 and IL12A, identified as upregulated proteins for SVCI, were particularly noteworthy, whereas CNTN5 and BCAN, identified as downregulated proteins for SVCI, were the most prominent among the proteins with negative risk effects.

#### 4.7. Explanation of MCI subtype classification

The interactive effect of XGPN was explained by using the SHapley Additive exPlanations (SHAP) [78], and the SHAP values for each



**Fig. 8.** Interpretation of model parameters. (a) shows the steadiness parameter and the comparison of the interactive and independent effects. (b) and (c) presents the classifier parameters for ADCI and SVCI, respectively.



**Fig. 9.** Explanation of MCI subtype classification. (a) illustrates the importance of plasma protein biomarkers in classifying MCI subtypes, as determined by calculating SHAP values. Subsequently, protein biomarkers with higher importance than the overall average were selected as key proteins. To identify the effects of key proteins on MCI subtype classification, SHAP values were examined by comparing individual expected values into positive and negative groups for AD CI and SV CI in (b) and (c), respectively.

biomarker was derived by applying XGBoost model trained on the discovery cohort to the validation cohort. Initially, as illustrated in Fig. 9 (a), the assessment of the importance of each biomarker was conducted by using absolute SHAP values. For the classification of AD CI, NTRK3 was identified as the most significant protein, followed by CCL3 and NCAN. NTRK3 has been demonstrated to play a role in the proliferation and differentiation of neurons during embryonic development and their subsequent growth and survival in the adult nervous system [79,80], and thereby, it has been revealed to be associated with Alzheimer’s disease and other neuropsychiatric disorders [63,81]. CCL3 and NCAN also have been revealed to be associated with neurodegeneration and to play a role in AD-related pathology [82,83]. For the classification of SV CI, IL12A was identified as the most significant protein, followed by BCAN and EDA2R. IL12A has been demonstrated to significantly contribute to neuroinflammatory pathways associated with the pathogenesis of cerebral small vessel disease [84,85]. BCAN and EDA2R also have been found to be functionally related to the nervous system and significant biomarkers for vascular dementia [70,86]. Subsequently, we summed the importance of each protein to AD CI and SV CI and then selected 13 biomarkers whose values exceeded the overall average as key proteins for MCI subtype classification.

Furthermore, in order to elucidate the impact of key proteins on the classification of MCI subtypes, SHAP values were analyzed by categorizing them into subtypes. The outcomes included the individual expected value and a comparison of SHAP values between groups based on their positivity and negativity. As shown in Figs. 9(b), 12 participants

indicated positive expected values for AD CI ( $f_{AD CI+}$ ), accounting for 21.8 % of the validation cohort, while the remaining 43 participants exhibited negative expected values for AD CI ( $f_{AD CI-}$ ). A comparison of the contribution of key proteins in the  $f_{AD CI+}$  and  $f_{AD CI-}$  groups was further conducted by calculating the mean of the SHAP values by group for each biomarker. The results indicated that NTRK3, NCAN, and IL12A were the proteins with the most significant differences between the two groups, followed by CCL3, FCRL2, and MDGA1. Furthermore, as illustrated in Figs. 9(c), 10 participants indicated positive expected values for SV CI ( $f_{SV CI+}$ ), representing 18.2 % of the validation cohort, while the remaining 45 participants exhibited negative expected values for SV CI ( $f_{SV CI-}$ ). Remarkable differences observed between the  $f_{SV CI+}$  and  $f_{SV CI-}$  groups with regard to IL12A, BCAN, EDA2R, JAM2 and NCAN.

### 5. Conclusion

In this study, we proposed XGPN for MCI subtype classification based on plasma protein biomarkers. One of the pronouncing features of the proposed method is to perform classification based on the interactive effects between proteins. In this process, XGPN propagates the independent effects of the identified biomarkers throughout the PPI network, with both local and global awareness of PPIs, which is individually optimized for each biomarker. As the other pronouncing feature, the proposed method provides strong explainability for outcomes. XGPN is comprised of a single function for extracting interaction effects and a fully-connected layer for classifying MCI subtypes. This transparent

architecture facilitates for the straightforward interpretation of the model parameters and the clear explanation of the predicted outcomes. Accordingly, XGPN demonstrated the capacity to achieve string explainability without sacrificing performance. The experimental results on the BICWALZS dataset for the Korean population demonstrated the aforementioned features of the proposed method, by indicating that XGPN outperformed existing methods for MCI subtype classification while providing a detailed understanding of the impact of the identified biomarkers. Furthermore, our findings were consistent with those of previous clinical studies, suggesting that XGPN may serve as a valuable adjunct in medical decision-making for patients in the predementia stage.

Here are some remarks on the method we proposed. First, XGPN can be applied to a variety of domains where interactions between entities exert a significant influence on the target outcome. XGPN utilized the PPI network that has already been constructed in the STRING database. This approach enhances the generalizability of XGPN by ensuring the graph representation is independent of the training data. Accordingly, XGPN can be an effective method in other domains, such as chemical interactions, brain connectivity, and social networks, as there are various publicly accessible databases that provide information regarding interactions between entities. Furthermore, the global awareness of interactions in XGPN will be beneficial for work in those domains, and as our experimental results demonstrated, the interactive effect extracted by XGPN can enhance the performance of any classifier. Second, the utilization of data on a larger number of plasma proteins enables XGPN to classify MCI subtypes more precisely. In this study, the identification of plasma biomarkers was limited to proteins included in the neurology and cytokine panels. This limitation can be alleviated by using other panels focused on the assessment of inflammation, cell regulation, and metabolism. We expect that the incorporation of additional panels of proteomic assays may facilitate further advancements in the classification of MCI subtypes based on plasma protein biomarkers. Third, the utility of XGPN could be enhanced by additional validation with external cohort datasets. The findings presented in this study are limited to the cohort dataset for the Korean population and were derived through internal validation only, as there were few cohorts or public data that included both blood and neuroimaging samples focusing on neurodegenerative diseases such as MCI and dementia. A further demonstration of the utility of XGPN will be provided by applying it to multiple cohort datasets of various ethnicities. This is a high priority for future research. Fourth, the enhancement of robustness for neuroimage-based diagnostic markers is a promising avenue for future research. Although the pipeline for neuroimage processing applied in this study is sufficiently reliable, it is possible that SUVR and PSMD may be sensitive to noise present in the PET and DTI scans. Accordingly, it can be enhanced to provide greater robustness of data, thus complementing preprocessing steps or regularization techniques, by applying well-established previous works [87–89]. Lastly, the optimization of protein interactions can lead to further improvement of XGPN. In this study, XGPN employed the combined score of PPI, which simply multiplies the scores of the seven types of interactions. However, as the contribution of each type to classifying MCI subtypes may vary, the individual optimization of each PPI type may enhance the technical sophistication of XGPN. In addition, the incorporation of a greater number of proteins will result in an expansion of the PPI network applied to XGPN, which will consequently elevate the computational expense. While this issue was not a concern in the present study, future research will encompass the optimization of the large-scale graph.

#### CRediT authorship contribution statement

**Sunghong Park:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Doyoon Kim:** Writing – review & editing, Writing – original draft,

Visualization, Investigation, Formal analysis, Conceptualization. **Heirim Lee:** Software, Formal analysis. **Chang Hyung Hong:** Resources, Investigation, Funding acquisition. **Sang Joon Son:** Resources, Investigation. **Hyun Woong Roh:** Resources, Investigation. **Dokyoon Kim:** Investigation, Funding acquisition. **Yonghyun Nam:** Investigation. **Dong-gi Lee:** Investigation. **Hyunjung Shin:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Funding acquisition, Conceptualization. **Hyun Goo Woo:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Funding acquisition, Conceptualization.

#### Data availability

The data used in this study are available upon request from the BICWALZS consortium biobank (<http://www.bicwalzs.com>) as well as the National Biobank of Korea (<https://biobank.nih.go.kr>) supporting the Korea Biobank Project.

#### Funding

This study was conducted with biospecimens and data from the consortium of the Biobank Innovations for Chronic cerebrovascular disease With Alzheimer's disease Study (BICWALZS) funded by the Korea Disease Control and Prevention Agency (#6637-303). This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE), Republic of Korea (NRF-2022R1A6A3A01086784), the BK21 FOUR program of the NRF funded by the MOE (NRF-5199991014091), and Ajou University Research Fund. This study was also supported by the NRF grants funded by the Ministry of Science and ICT (MSIT), Republic of Korea (NRF-2019R1A5A2026045, NRF-2021R1A2C2003474, and NRF-RS-2022-001653), the Institute of Information & communications Technology Planning & Evaluation (IITP) grants funded by the MSIT (IITP-2024-No.RS-2023-00255968 for the Artificial Intelligence Convergence Innovation Human Resources Development and No. 2022-0-00653), the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by Ministry of Health and Welfare (MOHW), Republic of Korea (HR21C1003), a grant of 'Korea Government Grant Program for Education & Research in Medical AI' through the KHIDI funded by the Korea government (MOE and MOHW), and a grant funded by the National Institutes of Health, USA (R01 AG071470).

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgement

The data analysis operations were supported by KREONET (Korea Research Environment Open NETWORK), managed and operated by the Korea Institute of Science and Technology Information (KISTI).

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.combiomed.2024.109303>.

#### References

- [1] R.C. Petersen, G.E. Smith, S.C. Waring, R.J. Ivnik, E.G. Tangalos, E. Kokmen, Mild cognitive impairment: clinical characterization and outcome, *Arch. Neurol.* 56 (1999) 303–308.
- [2] R.C. Petersen, O. Lopez, M.J. Armstrong, T.S. Getchius, M. Ganguli, D. Gloss, G. S. Gronseth, D. Marson, T. Pringsheim, G.S. Day, Practice guideline update

- summary: mild cognitive impairment: report of the guideline development, dissemination, and implementation subcommittee of the American academy of neurology, *Neurology* 90 (2018) 126–135.
- [3] V.L. Villemagne, S. Burnham, P. Bourgeat, B. Brown, K.A. Ellis, O. Salvado, C. Szoeke, S.L. Macaulay, R. Martins, P. Maruff, Amyloid  $\beta$  deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study, *Lancet Neurol.* 12 (2013) 357–367.
- [4] R.N. Kalaria, Neuropathological diagnosis of vascular cognitive impairment and vascular dementia with implications for Alzheimer's disease, *Acta Neuropathol.* 131 (2016) 659–685.
- [5] J.S. Meyer, G. Xu, J. Thornby, M.H. Chowdhury, M. Quach, Is mild cognitive impairment prodromal for vascular dementia like Alzheimer's disease? *Stroke* 33 (2002) 1981–1985.
- [6] J.S. Lee, H. Lee, S. Park, Y. Choe, Y.H. Park, B.K. Cheon, A. Hahn, R. Ossenkoppele, H.J. Kim, S. Kim, Association between APOE  $\epsilon$ 2 and A $\beta$  burden in patients with Alzheimer-and vascular-type cognitive impairment, *Neurology* 95 (2020) e2354–e2365.
- [7] M.J. Lee, S.W. Seo, D.L. Na, C. Kim, J.H. Park, G.H. Kim, C.H. Kim, Y. Noh, H. Cho, H.J. Kim, Synergistic effects of ischemia and  $\beta$ -amyloid burden on cognitive decline in patients with subcortical vascular mild cognitive impairment, *JAMA Psychiatr.* 71 (2014) 412–422.
- [8] S. Landau, D. Harvey, C. Madison, E. Reiman, N. Foster, P. Aisen, R.C. Petersen, L. Shaw, J. Trojanowski, C. Jack Jr., Comparing predictors of conversion and decline in mild cognitive impairment, *Neurology* 75 (2010) 230–238.
- [9] A. Ciarmiello, A. Tartaglione, E. Giovannini, M. Riondato, G. Giovacchini, O. Ferrando, M. De Biasi, C. Passera, E. Carabelli, A. Mannironi, Amyloid burden identifies neuropsychological phenotypes at increased risk of progression to Alzheimer's disease in mild cognitive impairment patients, *Eur. J. Nucl. Med. Mol. Imag.* 46 (2019) 288–296.
- [10] H.-J. Kim, J.S. Oh, J.-S. Lim, S. Lee, S. Jo, E.-N. Chung, W.-H. Shim, M. Oh, J. S. Kim, J.H. Roh, The impact of subthreshold levels of amyloid deposition on conversion to dementia in patients with amyloid-negative amnesic mild cognitive impairment, *Alzheimer's Res. Ther.* 14 (2022) 93.
- [11] G.B. Frisoni, S. Galluzzi, L. Bresciani, O. Zanetti, C. Geroldi, Mild cognitive impairment with subcortical vascular features: clinical characteristics and outcome, *J. Neurol.* 249 (2002) 1423–1432.
- [12] M. Zanetti, C. Ballabio, C. Abbate, C. Cutaia, C. Vergani, L. Bergamaschini, Mild cognitive impairment subtypes and vascular dementia in community-dwelling elderly people: a 3-year follow-up study, *J. Am. Geriatr. Soc.* 54 (2006) 580–586.
- [13] M.S. Albert, S.T. DeKosky, D. Dickson, B. Dubois, H.H. Feldman, N.C. Fox, A. Gamst, D.M. Holtzman, W.J. Jagust, R.C. Petersen, 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, *Focus* 11 (2013) 96–106.
- [14] W. Liu, H. Lin, X. He, L. Chen, Y. Dai, W. Jia, X. Xue, J. Tao, L. Chen, Neurogranin as a cognitive biomarker in cerebrospinal fluid and blood exosomes for Alzheimer's disease and mild cognitive impairment, *Transl. Psychiatry* 10 (2020) 125.
- [15] B.M. Tijms, E.M. Vromen, O. Mjaavatten, H. Holstege, L.M. Reus, S. van der Lee, K. E. Wesenhausen, L. Lorenzini, L. Vermunt, V. Venkatraghavan, Cerebrospinal fluid proteomics in patients with Alzheimer's disease reveals five molecular subtypes with distinct genetic risk profiles, *Nature aging* 4 (2024) 33–47.
- [16] J. Kim, J. Kim, Y.-h. Park, H. Yoo, J.P. Kim, H. Jang, H. Park, S.W. Seo, Distinct spatiotemporal patterns of cortical thinning in Alzheimer's disease-type cognitive impairment and subcortical vascular cognitive impairment, *Commun. Biol.* 7 (2024) 198.
- [17] C. Ritchie, N. Smailagic, A.H. Noel-Storr, Y. Takwoingi, L. Flicker, S.E. Mason, R. McShane, Plasma and cerebrospinal fluid amyloid beta for the diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI), *Cochrane Database Syst. Rev.* 6 (2014) 1–80.
- [18] C.R. Jack Jr., D.A. Bennett, K. Blennow, M.C. Carrillo, B. Dunn, S.B. Haeblerlein, D. M. Holtzman, W. Jagust, F. Jessen, J. Karlawish, NIA-AA research framework: toward a biological definition of Alzheimer's disease, *Alzheimer's Dementia* 14 (2018) 535–562.
- [19] T.A. Addona, X. Shi, H. Keshishian, D. Mani, M. Burgess, M.A. Gillette, K. R. Clauser, D. Shen, G.D. Lewis, L.A. Farrell, A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease, *Nat. Biotechnol.* 29 (2011) 635–643.
- [20] J. Huang, M. Khademi, L. Fugger, Ö. Lindhe, L. Novakova, M. Axelsson, C. Malmström, C. Constantinescu, J. Lycke, F. Piehl, Inflammation-related plasma and CSF biomarkers for multiple sclerosis, in: *Proceedings of the National Academy of Sciences*, vol. 117, 2020, pp. 12952–12960.
- [21] M.P. Davies, T. Sato, H. Ashoor, L. Hou, T. Liloglou, R. Yang, J.K. Field, Plasma protein biomarkers for early prediction of lung cancer, *EBioMedicine* 93 (2023).
- [22] Y. Jiang, X. Zhou, F.C. Ip, P. Chan, Y. Chen, N.C. Lai, K. Cheung, R.M. Lo, E. P. Tong, B.W. Wong, Large-scale plasma proteomic profiling identifies a high-performance biomarker panel for Alzheimer's disease screening and staging, *Alzheimer's Dementia* 18 (2022) 88–102.
- [23] F. Gao, X. Lv, L. Dai, Q. Wang, P. Wang, Z. Cheng, Q. Xie, M. Ni, Y. Wu, X. Chai, A combination model of AD biomarkers revealed by machine learning precisely predicts Alzheimer's dementia: China Aging and Neurodegenerative Initiative (CANDI) study, *Alzheimer's Dementia* 19 (3) (2022) 749–760.
- [24] C.S. Eke, E. Jammeh, X. Li, C. Carroll, S. Pearson, E. Ifeachor, Early detection of Alzheimer's disease with blood plasma proteins using Support vector machines, *IEEE journal of biomedical and health informatics* 25 (2020) 218–226.
- [25] P. Kivisäkk, C. Magdamo, B.A. Trombetta, A. Noori, Y.k.E. Kuo, L.B. Chibnik, B. C. Carlyle, A. Serrano-Pozo, C.R. Scherzer, B.T. Hyman, Plasma biomarkers for prognosis of cognitive decline in patients with mild cognitive impairment, *Brain Communications* 4 (2022) fca155.
- [26] S.-I. Chiu, L.-Y. Fan, C.-H. Lin, T.-F. Chen, W.S. Lim, J.-S.R. Jang, M.-J. Chiu, Machine learning-based classification of subjective cognitive decline, mild cognitive impairment, and Alzheimer's dementia using neuroimage and plasma biomarkers, *ACS Chem. Neurosci.* 13 (2022) 3263–3270.
- [27] X. Wu, L. Chen, X. Wang, Network biomarkers, interaction networks and dynamical network biomarkers in respiratory diseases, *Clin. Transl. Med.* 3 (2014) 1–7.
- [28] D.-H. Le, Machine learning-based approaches for disease gene prediction, *Briefings in functional genomics* 19 (2020) 350–363.
- [29] S.K. Ata, M. Wu, Y. Fang, L. Ou-Yang, C.K. Kwok, X.-L. Li, Recent advances in network-based methods for disease gene prediction, *Briefings Bioinf.* 22 (2021) bbaa303.
- [30] J. Vélez, F. Lopera, D. Sepulveda-Falla, H. Patel, A. Johar, A. Chuah, C. Tobon, D. Rivera, A. Villegas, Y. Cai, APOE\*  $\epsilon$ 2 allele delays age of onset in PSEN1 E280A Alzheimer's disease, *Mol. Psychiatr.* 21 (2016) 916–924.
- [31] J.F. Arboleda-Velasquez, F. Lopera, M. O'Hare, S. Delgado-Tirado, C. Marino, N. Chmielewska, K.L. Saez-Torres, D. Amarnani, A.P. Schultz, R.A. Sperling, Resistance to autosomal dominant Alzheimer's disease in an APOE3 Christchurch homozygote: a case report, *Nat. Med.* 25 (2019) 1680–1683.
- [32] T.N. Kipf, M. Welling, Semi-supervised Classification with Graph Convolutional Networks, 2016 arXiv preprint arXiv:1609.02907.
- [33] L. Liu, Y. Ma, X. Zhu, Y. Yang, X. Hao, L. Wang, J. Peng, Integrating sequence and network information to enhance protein-protein interaction prediction using graph convolutional networks, in: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2019, pp. 1762–1768.
- [34] E. Nasiri, K. Berahmand, M. Rostami, M. Dabiri, A novel link prediction algorithm for protein-protein interaction networks by attributed graph embedding, *Comput. Biol. Med.* 137 (2021) 104772.
- [35] Z. Gao, C. Jiang, J. Zhang, X. Jiang, L. Li, P. Zhao, H. Yang, Y. Huang, J. Li, Hierarchical graph learning for protein-protein interaction, *Nat. Commun.* 14 (2023) 1093.
- [36] R. Ramirez, Y.-C. Chiu, A. Hererra, M. Mostavi, J. Ramirez, Y. Chen, Y. Huang, Y.-F. Jin, Classification of cancer types using graph convolutional neural networks, *Frontiers in physics* 8 (2020) 203.
- [37] H. Cai, Y. Liao, L. Zhu, Z. Wang, J. Song, Improving cancer survival prediction via graph convolutional neural network learning on protein-protein interaction networks, *IEEE Journal of Biomedical and Health Informatics* 28 (2) (2023) 1134–1143.
- [38] F. Wu, A. Souza, T. Zhang, C. Fifty, T. Yu, K. Weinberger, Simplifying graph convolutional networks, in: *International Conference on Machine Learning*, PMLR, 2019, pp. 6861–6871.
- [39] L. Pasa, N. Navarin, W. Erb, A. Sperduti, Empowering simple graph convolutional networks, *IEEE Transact. Neural Networks Learn. Syst.* 35 (4) (2023) 4385–4399.
- [40] S. Abu-El-Hajja, B. Perozzi, A. Kapoor, N. Alipourfard, K. Lerman, H. Harutyunyan, G. Ver Steeg, A. Galstyan, Mixhop: higher-order graph convolutional architectures via sparsified neighborhood mixing, in: *International Conference on Machine Learning*, PMLR, 2019, pp. 21–29.
- [41] D. Jin, Z. Yu, C. Huo, R. Wang, X. Wang, D. He, J. Han, Universal graph convolutional networks, *Adv. Neural Inf. Process. Syst.* 34 (2021) 10654–10664.
- [42] J. Wang, J. Liang, J. Cui, J. Liang, Semi-supervised learning with mixed-order graph convolutional networks, *Inf. Sci.* 573 (2021) 171–181.
- [43] A.-L. Barabási, Z.N. Oltvai, Network biology: understanding the cell's functional organization, *Nat. Rev. Genet.* 5 (2004) 101–113.
- [44] J.-D.J. Han, N. Bertin, T. Hao, D.S. Goldberg, G.F. Berriz, L.V. Zhang, D. Dupuy, A. J. Walhout, M.E. Cusick, F.P. Roth, Evidence for dynamically organized modularity in the yeast protein-protein interaction network, *Nature* 430 (2004) 88–93.
- [45] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, Lethality and centrality in protein networks, *Nature* 411 (2001) 41–42.
- [46] C. Von Mering, R. Krause, B. Snel, M. Cornell, S.G. Oliver, S. Fields, P. Bork, Comparative assessment of large-scale data sets of protein-protein interactions, *Nature* 417 (2002) 399–403.
- [47] G.D. Bader, C.W. Hogue, Analyzing yeast protein-protein interaction data obtained from different sources, *Nat. Biotechnol.* 20 (2002) 991–997.
- [48] M.P. Stumpf, T. Thorne, E. De Silva, R. Stewart, H.J. An, M. Lappe, C. Wiuf, Estimating the size of the human interactome, *Proc. Natl. Acad. Sci. USA* 105 (2008) 6959–6964.
- [49] H.W. Roh, N.-R. Kim, D.-g. Lee, J.-Y. Cheong, S.W. Seo, S.H. Choi, E.-J. Kim, S. H. Cho, B.C. Kim, S.Y. Kim, Baseline clinical and biomarker characteristics of biobank innovations for chronic cerebrovascular disease with Alzheimer's disease study: BiCOWALS, *Psychiatry Investigation* 19 (2022) 100.
- [50] B. Winblad, K. Palmer, M. Kivipelto, V. Jelic, L. Fratiglioni, L.O. Wahlund, A. Nordberg, L. Bäckman, M. Albert, O. Almkvist, Mild cognitive impairment—beyond controversies, towards a consensus: report of the international working group on mild cognitive impairment, *J. Intern. Med.* 256 (2004) 240–246.
- [51] N. Trang, M. Choisy, T. Nakagomi, N. Chinh, Y. Doan, T. Yamashiro, J. Bryant, O. Nakagomi, D. Anh, Determination of cut-off cycle threshold values in routine RT-PCR assays to assist differential diagnosis of norovirus in children hospitalized for acute gastroenteritis, *Epidemiol. Infect.* 143 (2015) 3292–3299.
- [52] E. Baykara, B. Gesierich, R. Adam, A.M. Tuladhar, J.M. Biesbroek, H.L. Koek, S. Ropele, E. Jouvent, A.s.D.N. Initiative, H. Chabriat, A novel imaging marker for small vessel disease based on skeletonization of white matter tracts and diffusion histograms, *Ann. Neurol.* 80 (2016) 581–592.
- [53] M. Jenkinson, C.F. Beckmann, T.E. Behrens, M.W. Woolrich, S.M. Smith, Fsl, *Neuroimage* 62 (2012) 782–790.

- [54] C.v. Mering, M. Huynen, D. Jaeggi, S. Schmidt, P. Bork, B. Snel, STRING: a database of predicted functional associations between proteins, *Nucleic Acids Res.* 31 (2003) 258–261.
- [55] D. Szklarczyk, R. Kirsch, M. Koutrouli, K. Nastou, F. Mehryary, R. Hachilif, A. L. Gable, T. Fang, N.T. Doncheva, S. Pyysalo, The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest, *Nucleic Acids Res.* 51 (2023) D638–D646.
- [56] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, *Nucleic Acids Res.* 43 (2015) e47, e47.
- [57] B.T. Sherman, M. Hao, J. Qiu, X. Jiao, M.W. Baseler, H.C. Lane, T. Imamichi, W. Chang, DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update), *Nucleic Acids Res.* 50 (2022) W216–W221.
- [58] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, *Nat. Protoc.* 4 (2009) 44–57.
- [59] O. Chapelle, B. Scholkopf, A. Zien, Semi-supervised learning (Chapelle, O. et al., eds.; 2006)[book reviews], *IEEE Trans. Neural Network.* 20 (2009) 542, 542.
- [60] L. Bottou, F.E. Curtis, J. Nocedal, Optimization methods for large-scale machine learning, *SIAM Rev.* 60 (2018) 223–311.
- [61] N. Mehterov, D. Minchev, M. Gevezova, V. Sarafian, M. Maes, Interactions among brain-derived neurotrophic factor and neuroimmune pathways are key components of the major psychiatric disorders, *Mol. Neurobiol.* 59 (2022) 4926–4952.
- [62] J. Kim, S. Kim, H. Kim, I.-W. Hwang, S. Bae, S. Karki, D. Kim, R. Ogelman, G. Bang, J.Y. Kim, MDGA1 negatively regulates amyloid precursor protein-mediated synapse inhibition in the hippocampus, *Proc. Natl. Acad. Sci. USA* 119 (2022) e2115326119.
- [63] E. Savaskan, F. Müller-Spahn, G. Olivieri, S. Bruttel, U. Otten, C. Rosenberg, C. Hulette, C. Hock, Alterations in trk A, trk B and trk C receptor immunoreactivities in parietal cortex and cerebellum in Alzheimer's disease, *Eur. Neurol.* 44 (2000) 172–180.
- [64] R.F. Hillary, D.L. McCartney, S.E. Harris, A.J. Stevenson, A. Seeboth, Q. Zhang, D. C. Liewald, K.L. Evans, C.W. Ritchie, E.M. Tucker-Drob, Genome and epigenome wide studies of neurological protein biomarkers in the Lothian Birth Cohort 1936, *Nat. Commun.* 10 (2019) 3160.
- [65] D. Cohen, A. Pillozzi, X. Huang, Network medicine approach for analysis of Alzheimer's disease gene expression data, *Int. J. Mol. Sci.* 21 (2020) 332.
- [66] V. Fominykh, A.A. Shadrin, P.P. Jaholkowski, S. Bahrami, L. Athanasios, D. P. Wightman, E. Uffelmann, D. Posthuma, G. Selbæk, A.M. Dale, Shared genetic loci between Alzheimer's disease and multiple sclerosis: crossroads between neurodegeneration and immune system, *Neurobiol. Dis.* 183 (2023) 106174.
- [67] Y.J. Li, K. Nuytemans, J.O. La, R. Jiang, S.H. Slifer, S. Sun, A. Naj, X.R. Gao, E. R. Martin, Identification of novel genes for age-at-onset of Alzheimer's disease by combining quantitative and survival trait analyses, *Alzheimer's Dementia* 19 (2023) 3148–3157.
- [68] J. Chen, M.F. Doyle, Y. Fang, J. Mez, P.K. Crane, P. Scollard, C.L. Satizabal, M. L. Alosco, W.Q. Qiu, J.M. Murabito, K.L. Lunetta, ADSPData Harmonization Consortium Cognitive Harmonization Core, Peripheral inflammatory biomarkers are associated with cognitive function and dementia: framingham Heart Study Offspring cohort, *Aging Cell* 22 (2023) e13955.
- [69] I. Hristovska, A. Kumar, A.P. Binette, D. van Westen, S. Janelidze, E. Stomrud, S. Palmqvist, R. Ossenkoppele, N. Mattsson-Carlgen, J.W. Vogel, Identification of distinct and shared biomarkers in cerebral small vessel disease (SVD) through proteomic profiling of cerebrospinal fluid, *Alzheimer's Dementia* 19 (2023) e082927.
- [70] K. Minta, G. Brinkmalm, E. Portelius, P. Johansson, J. Svensson, P. Kettunen, A. Wallin, H. Zetterberg, K. Blennow, U. Andreasson, Brevican and neurocan peptides as potential cerebrospinal fluid biomarkers for differentiation between vascular dementia and Alzheimer's disease, *J. Alzheim. Dis.* 79 (2021) 729–741.
- [71] S.E. Harris, S.R. Cox, S. Bell, R.E. Marioni, B.P. Prins, A. Pattie, J. Corley, S. Muñoz Maniega, M. Valdés Hernández, Z. Morris, Neurology-related protein biomarkers are associated with cognitive ability and brain volume in older age, *Nat. Commun.* 11 (2020) 800.
- [72] R. Bhagat, S. Marini, J.R. Romero, Genetic considerations in cerebral small vessel diseases, *Front. Neurol.* 14 (2023) 1080168.
- [73] H. Gao, E.L. Findeis, L. Culmone, B. Powell, J. Landschoot-Ward, A. Zacharek, T. Wu, M. Lu, M. Chopp, P. Venkat, Early therapeutic effects of an Angiotensin-1 mimetic peptide in middle-aged rats with vascular dementia, *Front. Aging Neurosci.* 15 (2023) 1180913.
- [74] W. Cai, X. Chen, X. Men, H. Ruan, M. Hu, S. Liu, T. Lu, J. Liao, B. Zhang, D. Lu, Gut microbiota from patients with arteriosclerotic CSVD induces higher IL-17A production in neutrophils via activating ROR $\gamma$ t, *Sci. Adv.* 7 (2021) eabe4827.
- [75] R.Y. Lo, A.E. Hubbard, L.M. Shaw, J.Q. Trojanowski, R.C. Petersen, P.S. Aisen, M. W. Weiner, W.J. Jagust, A.S.D.N. Initiative, Longitudinal change of biomarkers in cognitive decline, *Arch. Neurol.* 68 (2011) 1257–1266.
- [76] B.S. Ye, S.W. Seo, J.-H. Kim, G.H. Kim, H. Cho, Y. Noh, H.J. Kim, C.W. Yoon, S.-y. Woo, S.H. Kim, Effects of amyloid and vascular markers on cognitive decline in subcortical vascular dementia, *Neurology* 85 (2015) 1687–1693.
- [77] A. Soldan, C. Pettigrew, Q. Cai, M.-C. Wang, A.R. Moghekar, R.J. O'Brien, O. A. Selnes, M.S. Albert, B.R. Team, Hypothetical preclinical Alzheimer disease groups and longitudinal cognitive change, *JAMA Neurol.* 73 (2016) 698–705.
- [78] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, *Adv. Neural Inf. Process. Syst.* 30 (2017).
- [79] S. Beltaifa, M.J. Webster, D.L. Ligons, R.J. Fatula, M.M. Herman, J.E. Kleinman, C. S. Weickert, Discordant changes in cortical TrkC mRNA and protein during the human lifespan, *Eur. J. Neurosci.* 21 (2005) 2433–2444.
- [80] H. Fukumitsu, M. Ohtsuka, R. Murai, H. Nakamura, K. Itoh, S. Furukawa, Brain-derived neurotrophic factor participates in determination of neuronal laminar fate in the developing mouse cerebral cortex, *J. Neurosci.* 26 (2006) 13218–13230.
- [81] C. Weickert, D. Ligons, T. Romanczyk, G. Ungaro, T. Hyde, M. Herman, D. Weinberger, J. Kleinman, Reductions in neurotrophin receptor mRNAs in the prefrontal cortex of patients with schizophrenia, *Mol. Psychiatr.* 10 (2005) 637–650.
- [82] E. Cudaback, Y. Yang, T.J. Montine, C.D. Keene, APOE genotype-dependent modulation of astrocyte chemokine CCL3 production, *Glia* 63 (2015) 51–65.
- [83] S. Mravinacová, V. Alanko, S. Bergström, C. Bridel, Y. Pijnenburg, G. Hagman, M. Kivipelto, C. Teunissen, P. Nilsson, A. Mattson, CSF protein ratios with enhanced potential to reflect Alzheimer's disease pathology and neurodegeneration, *Mol. Neurodegener.* 19 (2024) 15.
- [84] Z. Pu, X. Bao, S. Xia, P. Shao, Y. Xu, Serpine1 regulates peripheral neutrophil recruitment and acts as potential target in ischemic stroke, *J. Inflamm. Res.* (2022) 2649–2663.
- [85] A. Cifù, F. Janes, C. Mio, R. Domenis, M.E. Pessa, R. Garbo, F. Curcio, M. Valente, M. Fabris, Brain endothelial cells activate neuroinflammatory pathways in response to early cerebral small vessel disease (CSVD) patients' plasma, *Biomedicines* 11 (2023) 3055.
- [86] Y. Guo, J. You, Y. Zhang, W.-S. Liu, Y.-Y. Huang, Y.-R. Zhang, W. Zhang, Q. Dong, J.-F. Feng, W. Cheng, Plasma proteomic profiles predict future dementia in healthy adults, *Nature Aging* 4 (2024) 247–260.
- [87] S. Mohanty, S.P. Dakua, Toward computing cross-modality symmetric non-rigid medical image registration, *IEEE Access* 10 (2022) 24528–24539.
- [88] M.Y. Ansari, Y. Yang, S. Balakrishnan, J. Abinadah, A. Al-Ansari, M. Warfa, O. Almkhad, A. Barah, A. Omer, A.V. Singh, A lightweight neural network with multiscale feature enhancement for liver CT segmentation, *Sci. Rep.* 12 (2022) 14153.
- [89] M.Y. Ansari, Y. Yang, P.K. Meher, S.P. Dakua, Dense-PSP-UNet: a neural network for fast inference liver ultrasound segmentation, *Comput. Biol. Med.* 153 (2023) 106478.