

# Multiscale networks in Alzheimer's disease identify brain hypometabolism as central across biological scales

Elena Lara-Simon<sup>1</sup>, Juan Domingo Gispert<sup>2</sup>, Jordi Garcia-Ojalvo<sup>1</sup>, Pablo Villoslada<sup>1,3\*</sup>, for the Alzheimer's Disease Neuroimaging Initiative<sup>1</sup>

**1** Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Barcelona, Spain

**2** Barcelona Beta Research Center, Barcelona, Spain

**3** Hospital del Mar Research Institute, Barcelona, Spain

¶Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

\* [pablo.villoslada@upf.edu](mailto:pablo.villoslada@upf.edu)

## Abstract

Alzheimer's disease encompasses multiple biological scales, spanning molecular factors, cells, tissues, and behavioral manifestations. The interplay among these scales in shaping the clinical phenotype is not yet fully comprehended. In particular, there is great interest in understanding the heterogeneity of the clinical aspects of AD in order to improve treatment and prevention, by targeting those aspects most susceptible to the disease. Here we employed a systems biology approach to address this issue, utilizing multilayer network analysis and deep phenotyping. This integrative analysis incorporated genomics, cerebrospinal fluid biomarkers, tau and amyloid beta ( $A\beta$ ) PET imaging, brain MRI data, risk factors, and clinical information (cognitive tests scores, Clinical Dementia Rating and clinical diagnosis) obtained through the ADNI collaboration. Multilayer networks were built based on mutual information between the elements of each layer and between layers. Boolean simulations allowed us to identify paths that transmit dynamic information across layers. The most prominent path that significantly predicted the average cognitive phenotype included the PET radiotracer fluorodeoxyglucose (FDG) in the posterior cingulate. Combinations of different symptomatic variables, mainly related to mental health (depression, mood swings, drowsiness) and vascular features (hypertension, cardiovascular history), were also part of the paths explaining the average phenotype. Our results show that integrating the flow of information across biological scales reveals relevant paths for AD, which can be subsequently explored as potential biomarkers or therapeutic targets.

## Author summary

Complex diseases such as Alzheimer's Disease (AD) involve a diverse array of biological processes. In our investigation, we undertook a systems biology approach to AD using network analysis and deep phenotyping within a prospective cohort of patients,

incorporating clinical, imaging, genetics, and omics assessments. The gene, molecular and imaging paths explained variation in central nervous system damage, and in metrics of disease severity. The elucidation of multilayer paths in this context provides insights into the diverse phenotypes of the disease and holds the potential to improve understanding of its pathogenesis.

## Introduction

Alzheimer's Disease (AD) is a progressive, degenerative brain disease characterized by loss of function and death of nerve cells. The disease is defined by the presence of amyloid beta ( $A\beta$ ) plaques and neurofibrillary tau tangles in the brain [1]. Abnormal deposits of these two proteins have been seen to form aggregates and inclusions, de-structuring the brain architecture. AD is the most common form of dementia, accounting for 60-80% of all cases. Accurate diagnosis is possible in vivo using biomarkers [2-4]. Although early molecular markers exist, even in plasma [5], there remains a strong interest in understanding heterogeneity at all levels in the clinical manifestations of AD, both for treatment and prevention, to identify those individuals at highest risk for the disease. To that end, a multimodal approach integrating different omics data types (genomics, proteomics and metabolomics) and imaging, appears especially useful [6]. Here we follow this approach using multilayer network analysis to represent the flow of events underlying the phenotype of AD, including gene expression, tissue damage, and clinical symptoms. The goal is to identify multimodal paths associated with specific features of AD that will help explain the observed clinical heterogeneity of the disease, and identify candidate paths for personalized interventions.

Modern complex network theory has shown to be a very useful tool for comprehending the intricate architecture of biological processes. It is not sufficient to identify and classify the system's constituent parts to fully comprehend complex biological systems; understanding the interactions between those elements is also necessary [7, 8]. However, it can be challenging to evaluate the interaction patterns and functional architecture of biological systems due to the nontrivial nature of those interactions, the limited statistical power of clinical data, and the inherent nonlinearities in the dynamics of individual elements.

In spite of those difficulties, some studies have already shown the usefulness of multilayer networks to help improve our understanding of disease pathogenesis. Multilayer networks have been used, for instance, to directly link the genomic layer with the phenotypes in different types of cancer [9]. In a different approach, a topological analysis of a bipartite network linking drugs and proteins showed an overabundance of drugs targeting the same proteins, suggesting more functional drugs for more diverse targets [10]. Multilayer network analysis has also been applied to a cohort of multiple sclerosis patients, enabling the identification of genetic, protein, and cellular paths explaining variation in central nervous system damage and disease severity metrics, and highlighting their potential as biomarkers [11]. Specifically in Alzheimer's disease, a multidimensional network framework enabled detection of the disease with 90% accuracy, revealing unique insights into disease heterogeneity through identification of similar subtypes with diverse biomarker profiles [12].

In order to better understand the relationship between endotype and phenotype, our strategy focuses on connecting the microscopic and macroscopic dimensions of AD. Our approach enables us to integrate all available biological scales through a multilayer network that allows a multi-domain analysis of a wide variety of AD features. This form of analysis is useful for identifying how the disease phenotype is manifested at different levels, so that new treatment horizons may be explored and approached. Our project is hypothesis-based, aiming at using real-world data to evaluate our hypothesis and

uncover a plausible explanation.

Our approach is divided in the following steps. First, we construct networks for each of the biological layers individually, using mutual information as the criterion to connect elements of each layer. Next we identify the connections between layers in order to obtain a multilayer network. An unbiased analysis of the network connectivity reveals a modular structure that supports the hypothesis that a hierarchy exists among the different biological layers. Finally, we identify paths and key drivers for AD phenotype expression. To that end, we use dynamical Boolean simulations to calculate the shortest paths of information transmission that lead to the phenotype layer. The variables most commonly present in those paths can be considered predictors of the phenotype, shedding light on how the different scales interact to produce this complex disease, and potentially enabling its diagnosis.

## Materials and methods

### Patients

Data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) cohort was used in order to create the network [13]. Specifically, we used data on patients between the ages of 55 and 90 for over a decade at 57 sites in the US and Canada. The cohort included 1776 participants, of which 888 were defined clinically as MCI patients, 348 were AD patients, and 540 were healthy controls. The features used to implement the multilayer network included: MRI, PET scans, CSF/plasma proteomics, genetic information, risk factors, cognitive tests and diagnosis assessment.

### Genetic layer

This layer contains genetic information such as *APOE* and *TOMM40* alleles [14], Polygenic Hazard Score (PHS) and Cumulative Incidence Rate (CIR). The genetic network, as well as the rest of the individual networks, can be seen in Fig 6, where nodes represent variables and the edges represent the mutual information between pairs of variables.

The dataset contains three variables in particular that refer to the *APOE* gene: the individual copies *APOE\_A1* and *APOE\_A2*, for which the number refers to the type of allele that forms it, and *APOE*, that represents the amount of E4 alleles present in the individual, which can be 0, 1 or 2. For example, if a subject has *APOE\_A1* = 2 (has allele E2) and *APOE\_A2* = 4 (has allele E4), then *APOE* = 1 (one E4 allele), indicating the individual is heterozygous for the E4 allele.

The polygenic hazard score is a quantity based on a combination of more than 30 genetic variants created for the quantification of AD dementia age of onset. Several SNPs were examined for their association with AD, then a stepwise Cox proportional hazard model was applied to choose the SNPs that improved the model. Finally, the vector product between the genotype for the SNPs and the AD-incidence rates provides quantitative estimates of the annualized (cumulative) incidence rate [15]. For this reason, we decided to also add the Cumulative Incidence Rate (CIR) to the genetic layer, even though it could be considered a risk factor, because it is closely related to PHS. Importantly, PHS was associated with in vivo biomarkers of AD pathology such as reduced CSF A $\beta$ 42 and elevated CSF total tau across the AD spectrum (in older healthy controls and dementia individuals) [16].

## Molecular layer

For both  $A\beta$  and tau, CSF as well as plasma samples were used in order to extract the subjects' measures. For  $A\beta$ , two types of measures were used:  $A\beta_{42}$  and the ratio  $A\beta_{40}/A\beta_{42}$ . As for tau, both tau and phosphorylated tau are measured. Other molecular measures are amyloid precursor protein (APP);  $\beta$ -secretase, an enzyme that participates in the APP to  $A\beta_{42}$  pathway, and even though it is part of the genetic information of the cells, telomere length (TL) is also considered a molecular biomarker, since it does not really act as a gene but a protection of the genetic material, thus a protection of the cells functions and regulations. There is also a variable that represents the ratio between TL and the Single Copy Gene ratio (QPCR), another measure for TL.

## PET layer

Acquisition and standardized preprocessing steps of MRI and PET data in ADNI have been reported previously and are described in detail on the ADNI website [17]. Only the variables that have the largest relevance in the development of the disease are considered. Braak stages were used for tau [18], and the Landau signature was used for FDG, which included the following regions: right and left angular gyrus, bilateral posterior cingulate, and right and left inferior temporal gyrus [19]. Finally, regions used for  $A\beta$  were: anterior cingulate cortex, isthmus cingulate cortex, posterior cingulate cortex, inferior frontal gyrus (pars opercularis, pars triangularis, and pars orbitalis), lateral orbitofrontal cortex, medial orbitofrontal cortex, middle frontal gyrus (caudal and rostral middle frontal), superior frontal gyrus, frontal pole, inferior temporal gyrus, middle temporal gyrus, superior temporal gyrus (superior temporal and transverse temporal), fusiform gyrus, entorhinal cortex, parahippocampal gyrus, lingual gyrus, lateral occipital gyrus, temporal pole, insula, inferior parietal gyrus, supramarginal gyrus, precuneus, superior parietal gyrus, precentral gyrus, postcentral gyrus, paracentral gyrus and cuneus [17].

## MRI layer

Cortical thicknesses from two signature AD regions were taken into account. First, Dickerson's signature, which refers to brain regions known to be highly atrophied due to AD, making them susceptible to thinning in subjects who might be in very early stages. Therefore, tracking the thickness of these regions years before symptoms appear could help in early detection and intervention. It encompasses all parietal as well as frontal regions and the supramarginal [20]. The second one, Jack signature, involves the entorhinal, inferior temporal, middle temporal, inferior parietal, fusiform and precuneus areas [21]. Cortical volume, thickness average, and standard deviation were used for all those regions, and also total brain gray and white matter volume as well as CSF volumes. Lastly, we also collected data from tensor based morphometry and atrophy measures.

## Phenotype layer

This is considered to be the top layer of our multilayer model. On the one hand, we have all the cognitive tests scores: Alzheimer Disease Assessment Scale (ADAS) [22], Mini-Mental State Examination (MMSE) [23], Montreal Cognitive Assessment (MOCA), the composite executive function score, composite language, memory and visuospatial scores. On the other hand, there are also two diagnosis variables, which indicate the level of dementia of the subject: the Clinical Dementia Rating (CDR) and the clinical diagnosis [24]. It defines the group the subject belongs to: healthy controls, Mild Cognitive Impairment (MCI) patients or AD patients.

## Risk Factors layer

This layer contains a variety of different aspects such as comorbidities, clinical history, demographics, depression, and more. The information in the majority of these variables is either binary (with a value of 0 or 1 representing if there is an absence or existence of a given risk factor, respectively), or categorical, like gender (*PTGENDER*), handedness (*PTHAND*) and marital status (*PTMARRY*). Additionally, we also have some quantitative measures such as birth year (*PTDOBY*), years of education (*PTEDUCAT*), and features given by scores, such as the total modified Hachinski score, which represents a sum of all modified Hachinski comorbidities. Hachinski scores are a clinical tool to differentiate types of dementia. In particular, modified Hachinski scores differentiate Alzheimer’s type dementia and other dementias [25].

## Data processing

The omics, imaging and clinical datasets were utilized to construct the multilayer network, with each dataset corresponding to a layer within the network. The datasets were scrutinized to address missing values and to determine which patients had data available for each layer. Notably, no imputation techniques were employed in this study. The patients were stratified into three groups: healthy, mild and severe, based on the clinical diagnosis variable.

## Multilayer network construction

Following the workflow proposed in [11], the first step in building the multilayer network was to construct individual networks from each of the six datasets by computing mutual information between nodes within each layer (mutual information was preferred over linear measures of correlation to account for the nonlinear nature of many biological processes). The networks within each layer were constructed separately, to highlight their inherent differences and to make use of the maximum number of available subjects for each dataset because not all subjects have data for all the layers. Once the individual networks were constructed, features between layers were connected using mutual information too, based on the information shown in Fig 8. However, not all layers were interconnected due to a predetermined hierarchy applied to the system. This resulted in a six-layer interconnected network, with each layer comprising features derived from the original six datasets. The network construction process is illustrated in Fig 5. Furthermore, a secondary network was then created by incorporating all six datasets, employing linear correlation (Pearson coefficient) to establish the edge’s directionality. This latter network was then used in the path analysis.

## Calculation of correlation for edges

Mutual information is a nonlinear measure of the dependence between numerical variables [26]. It is not limited to variables in real numbers or to linear relationships, so the mutual information is more general than Pearson’s correlation coefficient [27].

For discrete variables, mutual information is calculated using the binning method, which consists in partitioning the supports of  $X$  and  $Y$  into bins of finite size [28]:

$$I(X; Y) \approx \sum_{ij} p(i, j) \log \left( \frac{p(i, j)}{p_x(i)p_y(j)} \right) \quad (1)$$

If  $n_x(i)$  is the number of points from  $X$  that fall into the bin  $i$  (analogously for  $n_y(j)$ ),  $n(i, j)$  is the number of points that fall in the intersection, and  $N$  is the total number of

points (sample size), then:

$$p_x(i) \approx \frac{n_x(i)}{N} \quad p_y(j) \approx \frac{n_y(j)}{N} \quad p(i, j) \approx \frac{n(i, j)}{N} \quad (2)$$

which leads to (1).

On the other hand, the multilayer network used for the path analysis was constructed using Pearson correlation coefficient, that provides a measure of the strength and direction of linear relationships and it is a value between -1 and 1. Both measures were computed using the *scikit-learn* library in Python [29].

### Permutation test

We performed permutation tests to establish the statistical significance of the mutual information and the Pearson coefficient quantifiers. To that end, we randomized the data and calculated the statistics of interest for this new dataset for 1000 realizations, leading to a null distribution of the quantifier. A significance level of  $\alpha = 0.05$  was chosen, meaning that the true statistics must be larger (in the case of a one-tailed test) or larger or smaller (in the case of a two-tailed test) than 95% of the randomized measures to reject the null hypothesis.

### Connectivity

The *density* or *connectance* of a network is the fraction of observed edges to the maximum possible number of edges (without self-edges), which is  $\binom{n}{2} = \frac{1}{2}|V|(|V| - 1)$  [30].  $|V|$  is the *order* or number of nodes of a network and  $|E|$  is the *size* or number of edges of a network.

$$d = \frac{2|E|}{|V| \cdot (|V| - 1)} \quad (3)$$

For a weighted network we use an adaptation of the previous expression, changing the numerator to the sum of the edge weights.

$$d = \frac{2 \sum_{v \in V, u \in V, u \neq v} \text{weight}(u, v)}{|V| \cdot (|V| - 1)} \quad (4)$$

A subtype of multilayer network is the bipartite network: networks in which there are two types of nodes, belonging to two distinct subnetworks  $G_i$  and  $G_j$ , in such a way that edges can only connect nodes of different types. For this type of network, the density has the following definition:

$$d_{bip} = \frac{|\text{edge weights between } G_i \text{ and } G_j|}{|V_i| \cdot |V_j|} \quad (5)$$

Note that in this case the maximum number of possible connections (denominator) is  $|V_i| \cdot |V_j|$  because each node  $i$  from  $G_i$  can be connected to all nodes  $j$  from  $G_j$ .

### Path identification

#### Average shortest path length

We used Dijkstra's algorithm to solve the single-source shortest paths problem in our weighted graphs. This algorithm computes a minimal spanning tree, a tree-like structure that connects the source node (the initial node, in which the path starts) to every other node in the graph following the shortest path to each one.



This method can be used to measure a statistic of interest for characterising a network. In particular, the average shortest path length is given by the formula (6):

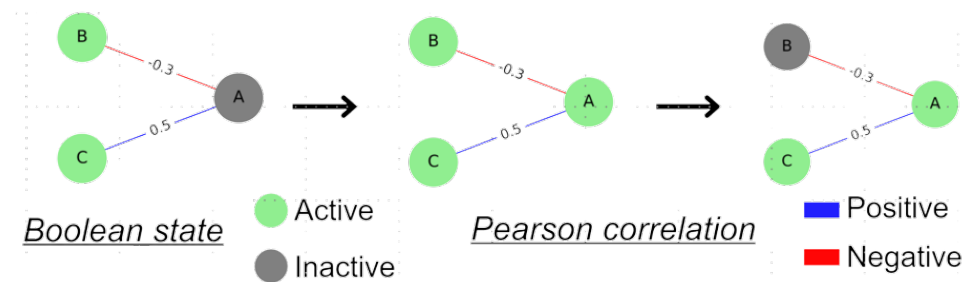
$$\alpha = \frac{\sum_{v,u \in V} d(v,u)}{|V|(|V| - 1)} \quad (6)$$

where  $V$  is the set of nodes,  $|V|$  is the total number of nodes and  $d(v,u)$  is the length of the shortest path between  $v$  and  $u$ . In the case of networks in which the weight associated with the connections is not the distance that separates the nodes, but how strong the connection between them is, we take the distance as the inverse of this weight. Then, for a weighted network, the average shortest path length gives information on how strong the connections between nodes are.

### Boolean modeling

As described before [11], the method of path identification involved constructing a combined six-layer network using Pearson correlation. Inspired by [31], Boolean simulations were then employed to analyze the flow of information across the network, with a particular focus on how perturbations affect nodes in different layers, especially those related to the phenotype. The aim was to identify variations in paths that initiate immune responses in individuals with AD.

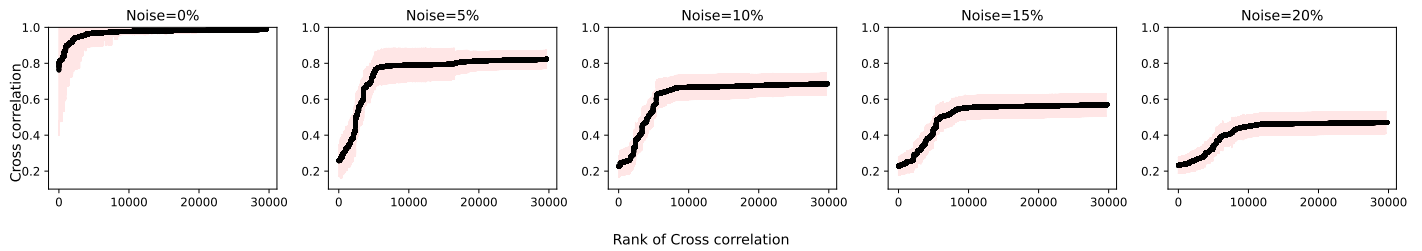
Each element in the network, belonging to one of the six layers, is assumed to be either active or inactive. The Boolean simulation starts in a random state, where each element has a 50% chance of being active or inactive. At each iteration, the activation status of the elements is updated based on the sum of their neighbors' states (Fig 1). The connections between the elements are either activating (positive) or inhibitory (negative). To determine whether a node will be active or inactive in the next iteration, each neighbour contributes a score based on the weight and sign of the corresponding Pearson correlation. The total sum of the weights of the neighbours determines the node's activation status in the subsequent iteration.



**Fig 1. Boolean dynamics.** Boolean dynamics are implemented on the networks, wherein the activation state of nodes undergoes changes determined by the cumulative sum of edge weights from their direct neighbors, taking into account the signs of connections as indicated by the Pearson coefficient: if the total sum is positive, the node becomes active, otherwise it becomes inactive. In this example, node A starts out inactive. Node B is active and is linked to node A by an edge with a negative contribution of 0.3. On the other hand, node C, which is also active, contributes to the activation of A positively with a weight of 0.5. In total, the contribution is 0.2, so node A is activated. In the next step, A is active and contributes positively to the activation of C, so C remains active. However, the contribution to node B is negative, so it becomes inactive.

The simulation was conducted for 100 steps by updating the states of the elements in each iteration. One node was selected as the input and manually switched between

active and inactive states in a defined period (i.e., 10 iterations active, then 10 iterations inactive) to analyze how perturbations propagate through the network and impact a given phenotype (output). To account for the stochastic nature of biological systems and prevent the simulation from settling into a fixed state, noise was introduced by assigning a probability for each element to change its state at each iteration. Fig 2 illustrates the effect of noise on the system. A noise level of 5% was chosen as it highlights differences in the cross-correlation of signals between nodes. In the absence of any noise, many nodes remain inactive or active for most of the simulation, resulting in high cross-correlations between nodes and masking the subtle variations in connection strength.

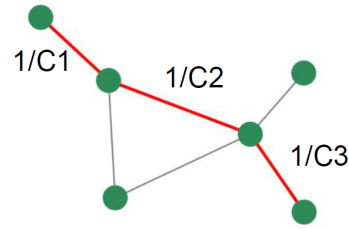


**Fig 2. Impact of noise on the cross-correlation coefficient of signals between nodes in a combined network.** In the absence of noise (0%), the majority of cross-correlation values approached 1, making it challenging to rank node pairs based on connection strength. However, with 5% noise, the cross-correlation values exhibited greater variation, enabling easier identification of paths between a selected source and target.

Following the simulations, a temporal cross-correlation function was computed between all pairs of nodes using the same measure of similarity as in [31]. The highest cross-correlation, potentially occurring at a non-zero lag time, was identified, and its reciprocal was assigned as a weight to the edges within the existing network. The high correlation values corresponded to low weights, and if there was no edge in the original network, no edge was added to the new network. The target phenotype was chosen, and the most efficient paths between it and the fixed source were identified based on the lowest path score, which was defined as the total sum of the weights (inverse maximum cross-correlations) of the edges connecting the source and target (Fig 3). The path score prioritized paths with both a minimal number of steps and high cross-correlations between nodes within the path. Dijkstra's algorithm was employed to pinpoint paths with the lowest path scores. Simulations were carried out for every conceivable pair of inputs and outputs to investigate the flow of information throughout the entire network, offering insights into the underlying pathology in AD.

To test the consistency of the results, we performed negative controls of the paths by permuting the six-layer network constructed with the Pearson correlation, as illustrated in Fig 4. This process involved swapping edges between node pairs in the network, and it was repeated 100 times. Importantly, an edge swap was only performed if it did not result in two edges connecting the same node pair. This approach maintained the original network's degree distribution. Additionally, the weights associated with each edge were permuted. The code to perform the permutation of the edges has been modified from the code published on [32]. Each edge in the original network underwent this edge swapping technique 10 times. Once the permutations were completed, the top paths for each network were identified using the same method as before.





$$\text{Path score} = 1/C1 + 1/C2 + 1/C3$$

**Fig 3. Calculation of the path score.** The cross-correlation coefficient is computed by assessing the signals for each connected node pair. A path score is then determined for all possible paths, defined as the sum of the reciprocals of the cross-correlation coefficients between consecutive node pairs along a specific path.

## Results

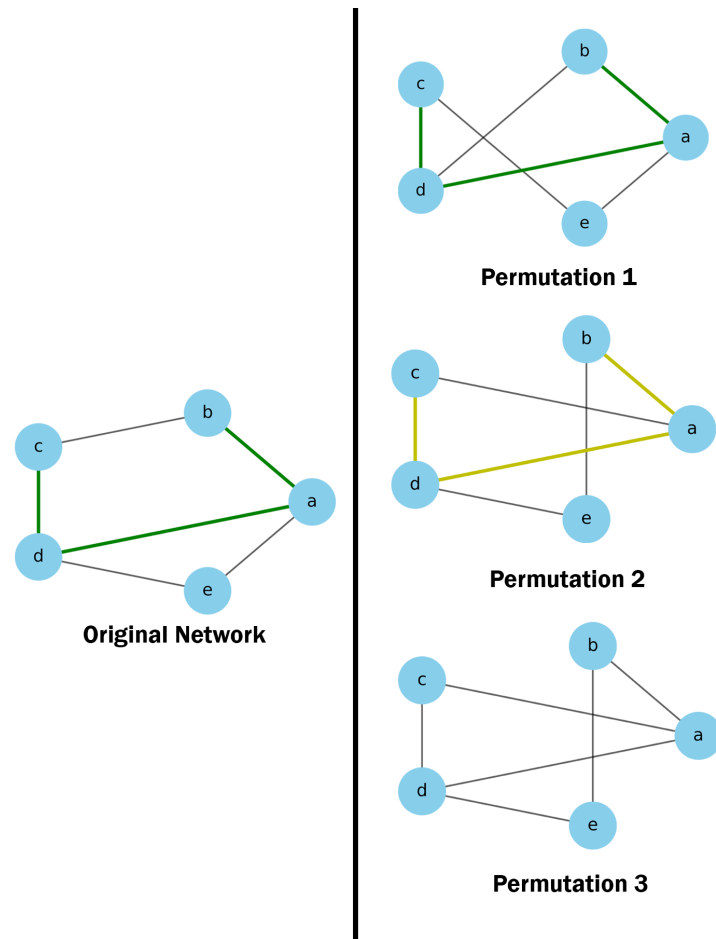
A cross-sectional study was performed by integrating data from the ADNI cohort (<https://adni.loni.usc.edu/>) at different scales: MRI, PET scans (including  $A\beta$ , tau and FDG PET), CSF/plasma proteomics, genetic information, risk factors and cognitive tests and diagnosis assessment conforming the phenotype (Fig 6). The results center around the identification of paths connecting these biological layers. The following paragraphs outline how these paths were discovered, and which sources are more strongly linked to the phenotype. The initial step provides descriptive details about the data, followed by the construction of the networks. Subsequently, Boolean simulations are executed, and ultimately, the most significant paths and nodes are chosen.

### Comprehensive phenotypic profiling: multi-omics, imaging, and clinical data in AD patients

The subjects for the study were classified, using exclusively clinical criteria, as healthy controls (n=622), subjects with MCI (n=807), or subjects with mild AD (n=533). As shown in table 1, the mean age of the 3 groups was equivalent at approximately 74 years. There were an approximately equal number of men and women in the normal control, but there were more men than women in the MCI and AD groups. More than half of the patients in the AD group were *APOE4* carriers, while the controls were less than 30%. Data corresponding to the concentration of cerebrospinal fluid biomarkers  $A\beta_{42}$ , tau and p-tau were collected from the University of Pennsylvania Alzheimer's Disease Clinical Core dataset.

**Table 1. Characteristics of the ADNI cohort**

Groups	Controls (n=622)	MCI (n=807)	AD (n=533)
Age	74 ± 7	74 ± 8	75 ± 8
Female, %	54.5	39.1	43.6
Age at disease onset	–	–	76 ± 8
APOE $\epsilon$ 4 carriers, %	29.4	48.0	64.8
$A\beta_{42}$ , pg/ml	203 ± 50	174 ± 53	143 ± 41
tau, pg/ml	65 ± 31	86 ± 52	122 ± 60
p-tau, pg/ml	32 ± 18	38 ± 21	50 ± 29



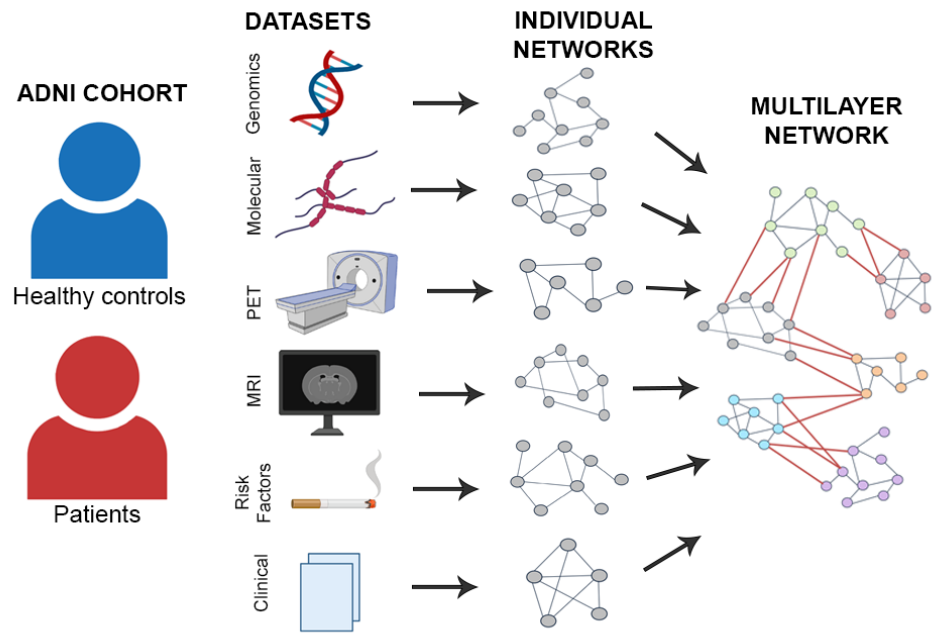
**Fig 4. Network permutation for negative controls.** In this example, in the first permutation, the edges (b,c) and (d,e) were exchanged with the edges (b,d) and (c,e). For the second permutation, the edges (a,e), (b,c) were swapped with the edges (a,c) and (b,e). In the third permutation, the top network's edge swap was applied first, followed by the middle network's edge swap: (a,b), (b,c) and (d,e) were exchanged for (a,c), (b,d) and (b,e). Three possibilities were considered when determining if the paths from the original network appeared in the permuted networks. In the first permutation, the path existed in the permuted network and was also identified as a top path. In the second permutation, the original path existed in the permuted network but was not identified as a top path. Finally, in the third permutation, the original path did not exist in the permuted network at all.

## Multilayer networks in AD

In order to create networks for each of the six layers, connections between pairs of elements within each layer were established using mutual information, as explained in [11]. Node pairs within the same layer were connected to one another with a weight equal to the normalized mutual information between them. A threshold was implemented to examine whether the correlation for a given pair was strong enough to establish an edge. Specifically, a node pair's actual mutual information value was compared to a surrogate distribution of mutual information values derived from random permutations of the data to determine the threshold.

The individual genetic, molecular and phenotype networks have few variables, due to

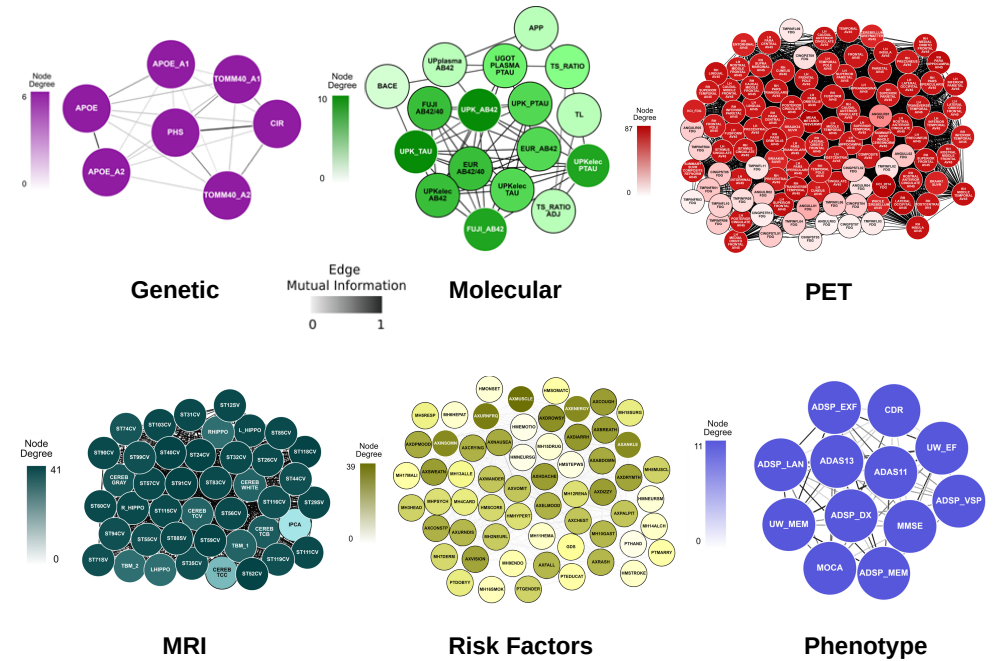
prior selection performed to retain only those features thought to have an appreciable effect on the diagnosis of AD (Fig 6). Genetic and phenotypic networks are less connected than the molecular network: the connections in the latter have a higher associated weight. In turn, the PET and MRI networks are larger in size due to the numerous brain regions contained in the dataset. The nodes in these networks represent functionally connected regions, reflecting relationships in terms of metabolic processes or neuronal activity. On the other hand, the risk factor network also has a large number of nodes, but the weight of their connections is generally low, which gives us an idea of how little dependence there is between the variables. This is to be expected given the nature of this dataset: risk factors range from the patient's gender to whether they have respiratory problems, therefore such weak correlations between variables is not surprising.



**Fig 5. Illustration of the multilayer network construction.**

Our approach to the multilayer network hierarchy is originally hypothesis-based, rather than data-based: each individual network forms a layer or level in the structure of this model, representing a different biological scale (Fig 7). The genomics layer sits at the bottom of the hierarchy and is connected with the molecular layer, constituting the microscopic substrate of the disease. The tissue imaging layers (PET and MRI) are the central functional layers, which, connected to each other and forming a bridge between the microscopic (genome and molecules) layer and the clinical output. The risk factor nodes form a loosely connected cloud around the main structure, although some of their connections are strong: this is the case of the year of birth (age) and gender of the patients. This analysis thus produced a network of six connected layers, where each layer contains features (variables) from each of the six original datasets (Fig 7). Here, connections between layers that do not conform to the proposed structure have been eliminated, as the hierarchy that we apply to the system is for representation only. In the rest of the paper we consider the global network without a predefined hierarchy, to avoid possible biases.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



**Fig 6. Individual networks.** The data from each layer is taken from the ADNI cohort and used to create networks, with nodes representing the dataset’s elements (genetics, molecular, PET, MRI, risk factors, and phenotype) and edges representing the mutual information between element pairs across all subjects. See high resolution networks at <https://dsb-lab.github.io/networks/>

First we proceeded to examine whether the connectivity of the global network reflects the biological organization into layers discussed above. To that end, node pairs were connected to each other irrespective of their layer, without a predefined hierarchy, and the *connectance matrix* between two layers  $i$  and  $j$  was calculated as follows:

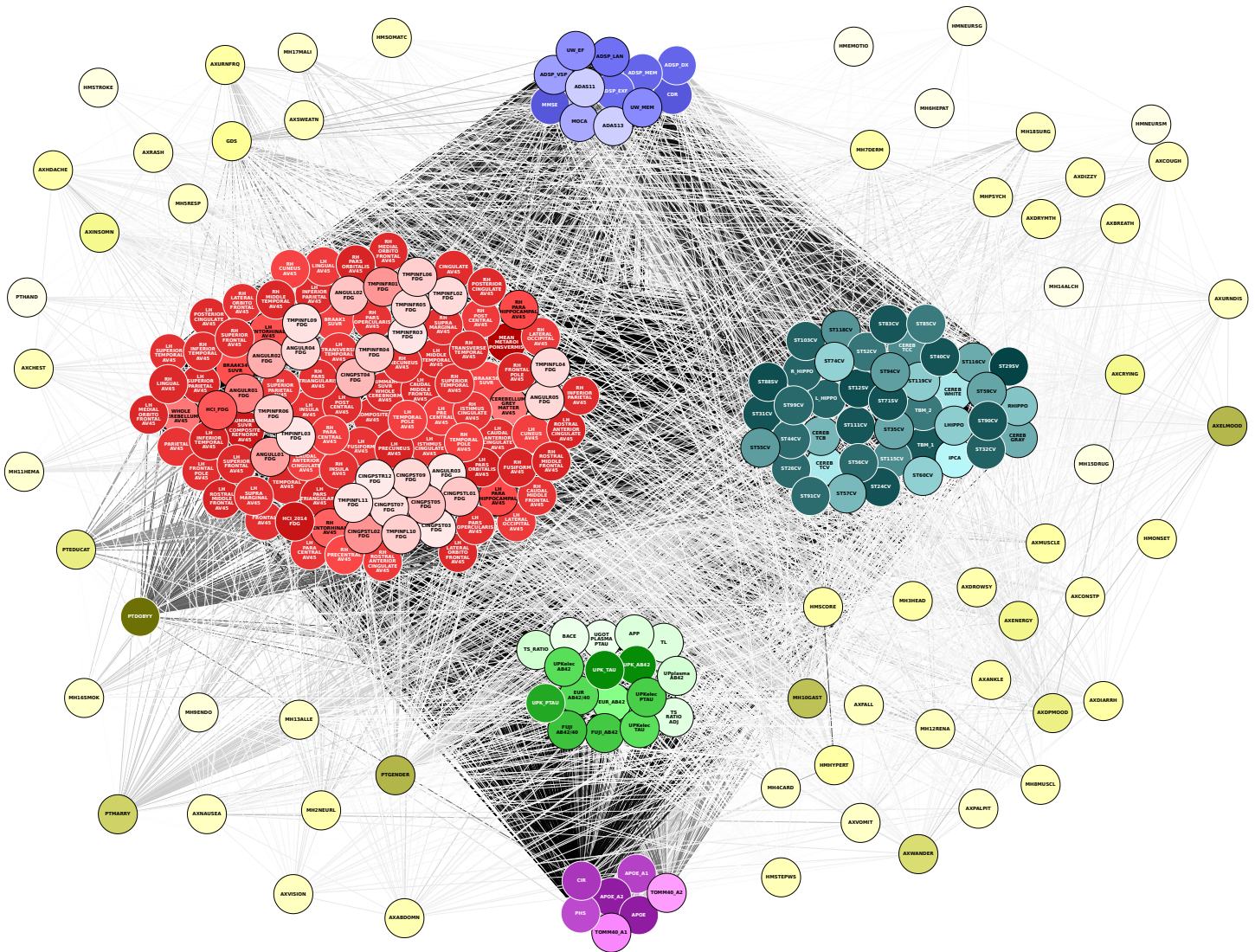
$$C_{ij} = \begin{cases} d_i, & \text{if } i = j \\ d_{ij}^{bip}, & \text{if } i \neq j \end{cases} \quad (7)$$

Here  $d$  is the standard connection density within a layer, as defined by (4) in the Materials and Methods section, and  $d_{bip}$  is the connection density of the bipartite network where each of two layers is one of the two node sets, as defined by (5). The latter measure is used because in the case of off-diagonal elements (different layers), we are only interested in how connected the nodes of one layer are to those of the other, but not with nodes of their own layer.

The resulting connectance matrix is shown in Fig. 8. A significant level of network modularity was discovered by this comparison of connections within and between layers, confirming the existence of an underlying multilayer structure: the connectivity within individual layers is higher than that between different layers. With the exception of risk factors, features within a level tend to be more correlated than between levels. The most prominent example of this is the MRI layer. The nodes in this layer are so closely related to each other because of the similarity between all the variables, representing levels of atrophy in different parts of the brain, as opposed to, for example, the nodes in the molecular layer which mostly represent concentrations of different proteins.



It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

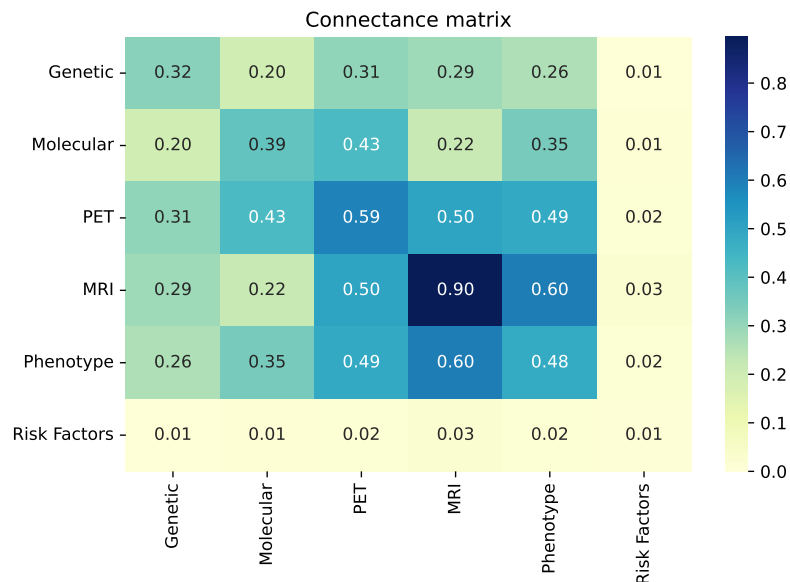


**Fig 7. Multilayer network.** Following a hierarchy that connects each layer successively, starting with the genomics layer and working up to the phenotypic (clinical) layer, individual networks are created and linked together using mutual information once again. See high resolution image at [https://dsb-lab.github.io/multilayer\\_net/](https://dsb-lab.github.io/multilayer_net/).

## Dynamic network analysis identifies paths associated with phenotype

To obtain a functional view of the information flow across layers, we aimed to integrate all six layers into paths that reflect network dynamic interactions. We constructed a single network comprising all layers using linear (Pearson) correlations that distinguish between stimulatory or inhibitory edges, depending on the correlation value being positive or negative, respectively. To identify the causal logic backbone of the network, we conducted logic (Boolean) simulations. These simulations use knowledge of activating and inhibiting relationships between nodes while ignoring the exact functional reactions between the nodes, thus providing a qualitative description of the system [31]. Nodes are either active or inactive, and their states are updated synchronously in each iteration of the simulation, depending on the activation states of

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



**Fig 8. Network densities within and between layers.** The connectance matrix was calculated using the expression in (7).

their direct neighbors and the weights of the corresponding connections. 360

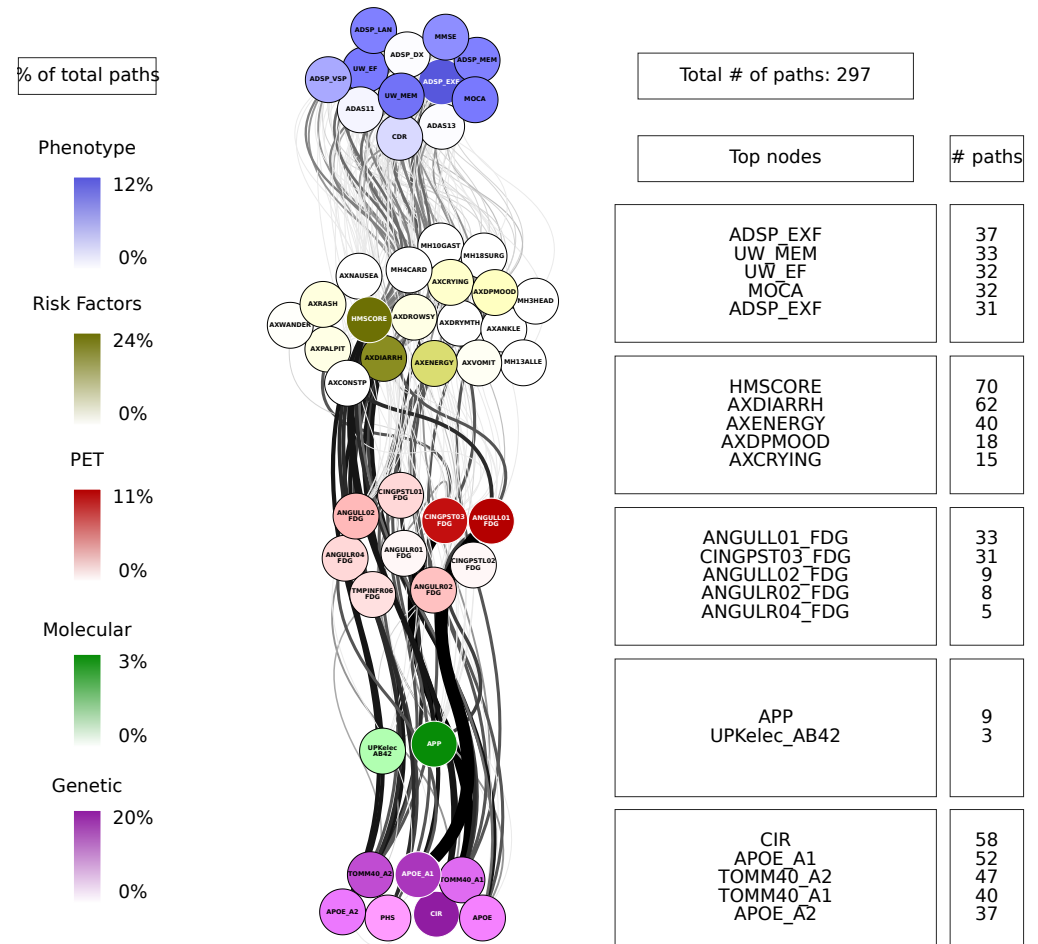
Our next objective was to investigate how dynamic changes in a specific input propagate through the network and ultimately impact a given phenotype. We achieved this by performing Boolean simulations, as described above, where the input node was periodically switched between active and inactive states. The responses of all nodes in the network were then measured by computing the temporal cross-correlation function between their time-varying state and the dynamic input signal. We then identified the paths in the network with the highest overall temporal cross-correlation between their signals, which indicate how information flows from the input to the output. The paths were chosen based on the lowest path score, which was defined as the total sum of the weights (inverse maximum cross-correlations) of the edges connecting the source and target [31]. While these paths may not necessarily reflect physical interactions among nodes, they represent groups of nodes that co-vary statistically more strongly with each other than the rest of the network. 361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373

For each combination of inputs and outputs, we selected the top ten paths with the highest joint cross-correlation values between their constituent nodes, resulting in a total of 30,000 combinations. Fig 9, 10, 11, 12 and 13 show these paths for the five inputs (genetic, molecular, PET, MRI and risk factors) and outputs (phenotype) pairs for AD patients. Darker color represents more connections among the nodes. 374  
375  
376  
377  
378

To evaluate the specificity of the Boolean simulations, we randomly permuted the network connections as described in the Materials and Methods section, to identify negative control paths that were then compared to those identified in the original networks. We focused on paths that appeared in less than 1% of the permutations. Out of the 30,000 total paths identified from AD patients, 17,877 did not appear at all in 100 realizations of the simulations in the permuted paths. The top paths (those that passed the test for negative controls) are shown in Fig 9, 10, 11, 12 and 13 for each input (genetic, molecular, PET, MRI and risk factors) and output (phenotype) pair. All variables with their acronyms are listed in the SI file. The method for path identification and network permutation is illustrated in the Materials and Methods 379  
380  
381  
382  
383  
384  
385  
386  
387  
388



It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



**Fig 9. Path analysis in AD patients from the genetic layer.** Depictions of the multilayer paths identified through Boolean simulations with the genetic layer as the starting point. The top paths, meeting criteria for negative controls, are presented for each input (genetic) - output (clinical phenotype) pair. Nodes within each layer are color-coded to reflect the node's degree, indicating the frequency of its appearance in a path as a percentage of the total paths. For detailed high-resolution paths, please refer to [https://dsb-lab.github.io/network\\_paths/](https://dsb-lab.github.io/network_paths/).

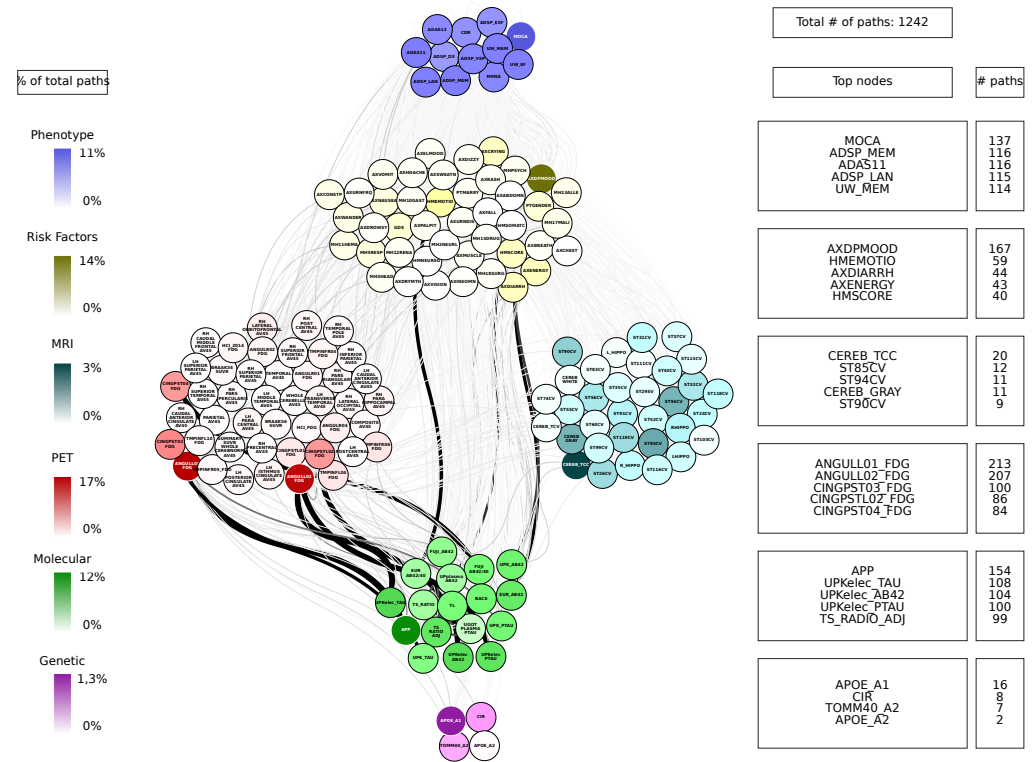
section (see Fig. 4).

## Path analysis

We now discuss the top 10 paths with the highest joint cross-correlation values. In what follows, the information flows along each path from left to right, and the perturbation starts at the first node. In the AD cases, the paths more commonly found are:

- APOE\_A1 (copy 1 of *APOE* gene) → ANGULL01\_FDG (globally normalized CMRgl from left angular gyrus) → AXRASH (rash) → any node of the phenotype layer, when the input is applied to the genetic layer (Fig. 9).
- UPKelec\_TAU (CSF total Tau using the fully automated Roche Elecsys immunoassay) → ANGULL02\_FDG → APP (CSF amyloid precursor protein) → ADSP\_VSP (harmonized composite visuospatial score), ADSP\_EXF (harmonized

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



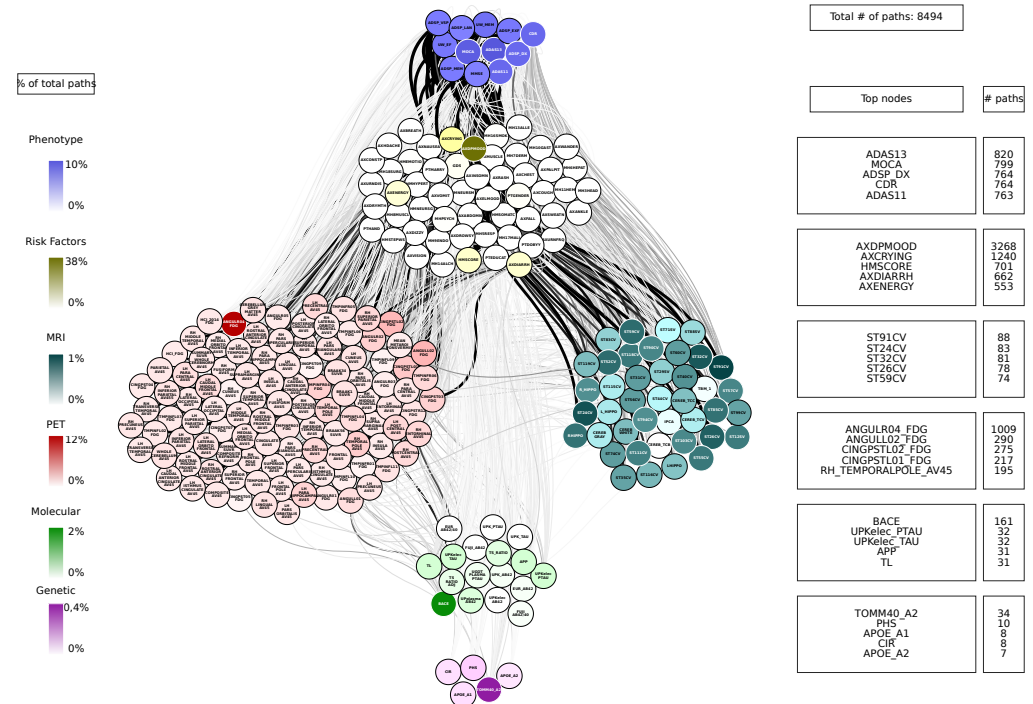
**Fig 10. Path analysis in AD patients from the molecular layer.** Depictions of the multilayer paths identified through Boolean simulations with the molecular layer as the starting point. The top paths, meeting criteria for negative controls, are presented for each input (molecular) - output (clinical phenotype) pair. Nodes within each layer are color-coded to reflect the node's degree, indicating the frequency of its appearance in a path as a percentage of the total paths. For detailed high-resolution paths, please refer to [https://dsb-lab.github.io/network\\_paths/](https://dsb-lab.github.io/network_paths/).

composite executive function score), ADSP\_MEM (harmonized composite memory score), MMSE (Mini Mental State Examination) and MOCA (Montreal Cognitive Assessment test), when the input is applied to the molecular layer (Fig. 10).

- CINGPSTR12.FDG (globally normalized CMRgl from right posterior cingulum cortex) → AXCRYING (crying) - AXDPMOOD (depressive mood) → ADSP\_EXF, when the input is applied to the PET layer (Fig. 11).
- ST44CV (cortical volume of left parahippocampal) → AXDPMOOD - GDS (Geriatric Depression Score) → MOCA, when the input is applied to the MRI layer (Fig. 12).
- AXELMOOD (elevated mood) - AXCRYING - AXDPMOOD → ADSP\_MEM, when the input is applied to the risk factor layer (Fig. 13).

When perturbing sources belonging to the genetic layer (reflecting genetic variability contributing to the risk of developing AD), we see that there is no transition through all the intermediate layers of the network: molecular nodes scarcely appear in the paths, and no node from the MRI layer appears (Fig. 9). Perturbations were directly linked to the PET and risk factor layer, whose nodes represent aspects of the patient's medical history and age, and these to changes in clinical outcomes. Perturbations of the molecular network (representing changes in protein concentration) have a direct

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

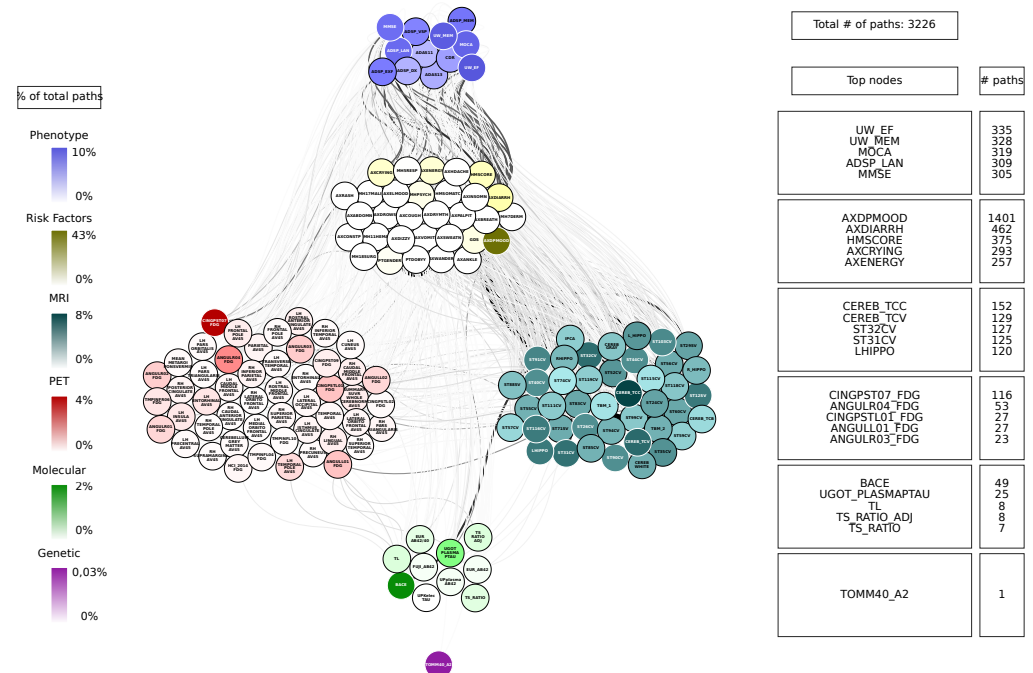


**Fig 11. Path analysis in AD patients from the PET layer.** Depictions of the multilayer paths identified through Boolean simulations with the PET layer as the starting point. The top paths, meeting criteria for negative controls, are presented for each input (PET) - output (clinical phenotype) pair. Nodes within each layer are color-coded to reflect the node's degree, indicating the frequency of its appearance in a path as a percentage of the total paths. For detailed high-resolution paths, please refer to [https://dsb-lab.github.io/network\\_paths/](https://dsb-lab.github.io/network_paths/).

influence on the PET layer, in particular on FDG signal for the angular and cingulum posterior nodes, and are less related to impact MRI nodes, although some of these represent neurodegeneration like FDG PET (Fig 10). For both genetics and molecular nodes, we found little information flow coming from the phenotype layer.

The number of paths increases considerably when we have perturbations at the imaging level (representing changes in protein concentration in the case of PET and changes in brain tissue degeneration in the case of MRI) (Figs. 11 and 12) and at the risk factor level (Fig. 13). Moreover, we can see that the influence of these paths on the phenotype is significantly higher. When the origin of the input is at the PET layer, the input nodes have little connectivity with the genetic layer nodes. In fact the PET nodes are mainly related to the risk factor layer and, through it, to the MRI nodes or directly to the phenotype nodes. The case of MRI is analogous: there are not so many connections with the deeper layers (genetic and molecular). The importance of the MRI nodes is more distributed, i.e. they appear with similar frequency in the paths, while we see that the relationship with the PET layer is quite centred on the nodes: CMRgl of the angular gyrus and posterior cingulum cortex. The indirect interaction with risk factors is also notable, highlighting some specific nodes related to symptomatology, which are discussed below. This indirect interaction is also found in the opposite direction, when the source belongs to the risk factor layer. In this case, however, we do not see a strong influence on the genetic layer. This is consistent with expectations: genes can be expected to influence the occurrence of some risk factors, but logically these factors cannot change the genetic information of the person, which is determined from birth.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



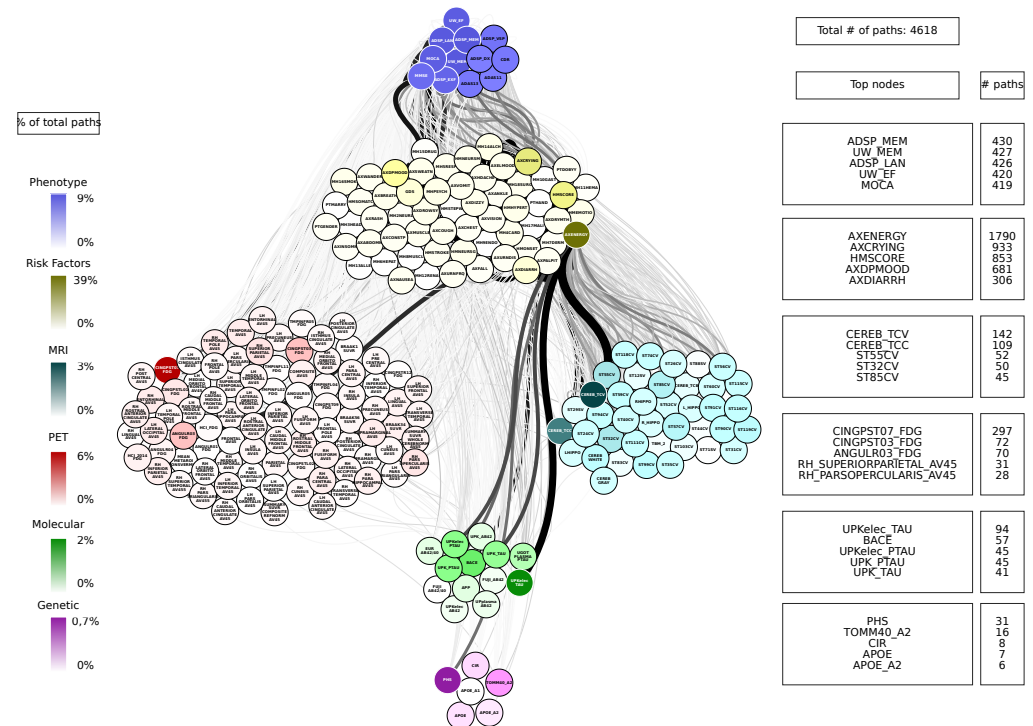
**Fig 12. Path analysis in AD patients from the MRI layer.** Depictions of the multilayer paths identified through Boolean simulations with the MRI layer as the starting point. The top paths, meeting criteria for negative controls, are presented for each input (MRI) - output (clinical phenotype) pair. Nodes within each layer are color-coded to reflect the node's degree, indicating the frequency of its appearance in a path as a percentage of the total paths. For detailed high-resolution paths, please refer to [https://dsb-lab.github.io/network\\_paths/](https://dsb-lab.github.io/network_paths/).

In general, our results show that as we move towards the highest-scale layers (the imaging layers and risk factors layer), the paths have very little tendency to return to the genetic and molecular layers, which shows the importance of considering AD as a multiscale disease, with the layers connected with different strengths and where information flows from the genetic and molecular layers, having increasing influence on the phenotype as we move up the layers.

Contrary to what might be expected, the concentrations of  $A\beta$  and tau do not appear frequently in the paths, thus their influence on the system is very low. It is also important to highlight the presence of the angular gyrus and posterior cingulum cortex FDG nodes from the PET layer in most of the paths, almost independently of the layer to which the source belongs. Also noteworthy is the influence of the patient symptoms, whose variables are found in the risk factor layer. These include: rash, crying, and elevated and depressed mood, among others.

The analysis shown above could provide information on the modifiable risk factors that can be used in preventive lifestyle modification trials. In order to detect at which levels these factors have an impact, we have depicted in Fig. 14 the 20 shortest paths that arise when the origin is a risk factor node. The main sources for these paths are drowsiness (AXDROWSY), hypertension (HMHYPERT), crying (AXCRYING), cardiovascular history (MH4CARD) and musculoskeletal pain (AXMUSCLE). It can be seen that only the risk factor layer is involved and is directly related to the target, which is the phenotype layer. In particular, other risk factor nodes related to the sources appear, for example low energy (AXENERGY), related to drowsiness and

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



**Fig 13. Path analysis in AD patients from the risk factors layer.** Depictions of the multilayer paths identified through Boolean simulations with the risk factors layer as the starting point. The top paths, meeting criteria for negative controls, are presented for each input (risk factors) - output (clinical phenotype) pair. Nodes within each layer are color-coded to reflect the node's degree, indicating the frequency of its appearance in a path as a percentage of the total paths. For detailed high-resolution paths, please refer to [https://dsb-lab.github.io/network\\_paths/](https://dsb-lab.github.io/network_paths/).

muscle pain, or depressive mood (AXDPMOOD), related to crying. No nodes from the genetic, molecular or imaging layers appear in these paths. Vascular health and muscle pain are related to executive and visuospatial function, while variable crying is also related to executive function and memory.

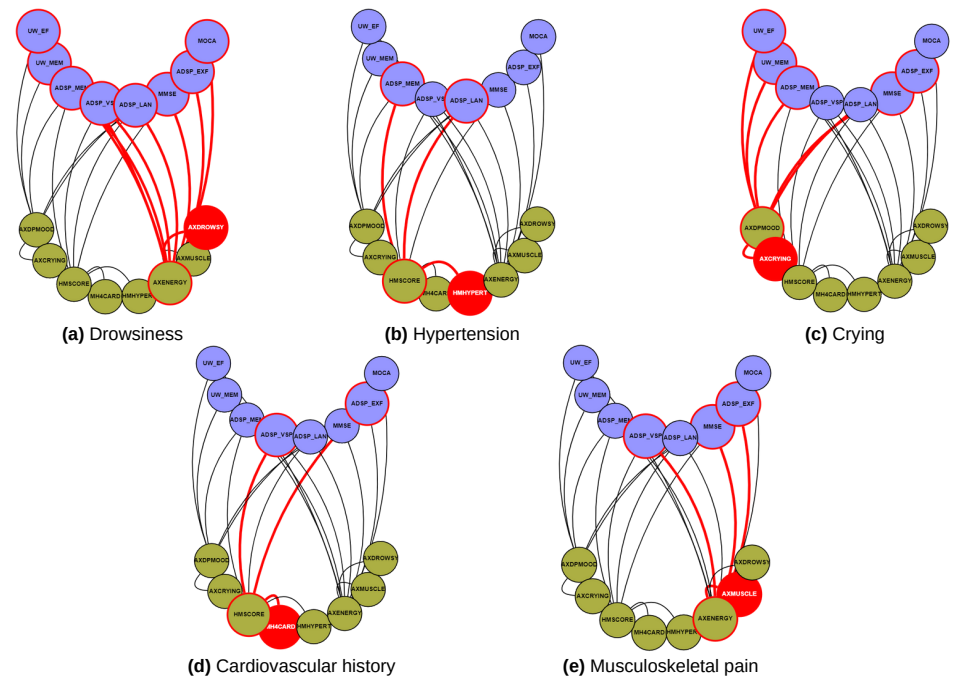
## Discussion

Our study addresses the complexity of AD from a holistic perspective, exploring the connections between the genetic, molecular, and clinical factors that underlie this pathology. Using a systems biology approach, we have integrated genomics, brain imaging, and clinical data to analyse AD. We use multilayer network analysis and deep phenotyping to unravel the complex mechanisms underlying the disease. Our research reveals significant connections between different biological features and the clinical manifestation of the disease. These findings could significantly improve our understanding of AD.

Results obtained in other studies support our finding. For instance, there is a clear influence of metabolic changes on the disease in the regions of the angular gyrus and posterior cingulate cortex [33]. The relationship of these areas with the cognitive dysfunction associated with AD is probably related to their involvement in various cognitive processes such as attention, visuospatial processing, and memory. This is



It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



**Fig 14. Selection of the top paths when the source is a risk factor node.** The top 20 shortest paths are presented for each input (risk factors) - output (clinical phenotype) pair. Nodes in the risk factor layer are shown in yellow and those in the phenotype layer in blue. The paths for a particular source out of the five chosen are shown in red: drowsiness, hypertension, crying, cardiovascular history and musculoskeletal pain. For detailed high-resolution paths, please refer to [https://dsb-lab.github.io/risk\\_paths/](https://dsb-lab.github.io/risk_paths/).

consistent with the brain hypometabolism observed in Alzheimer's disease patients [34–37]. [38] explores how restoring glucose metabolism in the hippocampus can improve cognitive function. In addition, neuronal dysfunction in these areas may contribute to the manifestation of emotional symptoms, such as depression and mood swings [39,40]. Cardiovascular history has also been associated with an increased risk of cognitive decline and dementia [41], including Alzheimer's disease and FDG hypometabolism in AD-sensitive regions [42,43]. Several observational studies have shown the potential beneficial role of antihypertensive treatment in preventing cognitive decline. However, these associations are complex and not fully elucidated [44].

Interestingly,  $A\beta$  and tau levels in the CSF are not very relevant in our paths: they do not show much direct association with cognition, probably because their effect is reflected much more in PET and MRI nodes and these are the ones that affect phenotype the most, making the effect of molecular nodes very diffuse, as other studies have pointed out [45]. However, given that APP levels in CSF represent the origin of some of the identified paths, it is interesting to focus prevention on reducing their proliferation. Therapies specifically aimed at modulating the activity of those elements have proved futile, probably because they are carried out when cognitive symptoms are already present and their effect cannot be reversed [46]. On the other hand, our data suggests that early detection of these biomarkers could allow preventive or therapeutic interventions aimed at modifying disease progression before clinical symptoms appear. Moreover, since our paths show that psychosocial symptoms such as mood and



depression can influence cognitive function and disease progression, we can expect psychosocial interventions to improve the quality of life of Alzheimer's patients by addressing these issues. Given that AD involves multiple biological and clinical pathways, multimodal therapeutic approaches can be developed that simultaneously address several aspects of the disease [47].

In the past, network techniques have shown to be very successful in offering useful insights into the complicated molecular basis of illnesses, transcending the usual viewpoints centered on single genes and pathways. Within the traditional network framework, molecules are interconnected based on their biological interactions, and by studying the structure and dynamics of these interaction networks, it becomes possible to uncover disease modules and nonlinear pathways [48]. Weighted gene co-expression networks have been used, for instance, to identify groups of genes (modules) involved in various activated pathways leading to hypertension [49] or to breast cancer and AD [50]. Recently, these approaches have evolved to encompass multiple biological layers: in [51] protein-protein interactions related to essential hypertension were studied through network analyses, also in [52] diverse biological processes such as membrane potential dynamics and signaling were studied within insulin-secreting cells. Approaches based on multilayer networks have also produced notable results in the study of cancer [53–58] and multiple sclerosis [11]. Moreover, new lines of research have been opened up by the use of network-based models, for example by suggesting a possible connection between age-related macular degeneration (nAMD) and neurodegenerative disorders such as AD, schizophrenia and Parkinson's disease [59]. Here we have applied a multilayer network analysis to represent the flow of events that underlies the phenotype of a complex disease such as AD.

Our multilayer network analysis enabled an examination of the interplay between various biological scales in Alzheimer's disease, revealing paths that connect six scales (genomics, molecular, PET, MRI, risk factors, and phenotype) through statistical associations. The analysis provides evidence for information flow across different scales, with the imaging levels (PET and MRI) emerging as particularly informative. The layers interconnect with diverse strengths, and information is modulated as it propagates across them [60, 61].

Prior research has sought to establish a direct connection between the genomic layer and phenotypes in various complex diseases, including AD [62, 63]. However, genotypes,  $A\beta$  and tau deposition, and brain metabolism alone have a limited ability to predict the phenotype [64, 65]. Our findings incorporate omics, imaging, and phenotype data, underscore the significance of conceptualizing AD as a multiscale condition. Additionally, the identified paths may serve as potential targets for future personalized medicine treatments in AD.

The data obtained from the ADNI cohort proved to be rich, covering a broad spectrum of scales. However, certain limitations were encountered during the extraction and analysis of variables. While the cohort's overall sample size was sufficient for detecting significant correlations, some specific layers, such as the genomics one, had smaller sample sizes. This limitation has an impact on both the construction of networks and the identification of paths. Additionally, the analysis had to adopt a cross-sectional approach due to inadequate follow-up in many participants. The inclusion of longitudinal data across all six layers would significantly enhance the value of future studies.

## Conclusion

In summary, this study examined the functional connections among various scales of biological data of a complex disease with a complex genetic basis, namely AD. The

approach to understanding complex biological systems through network science is a very active interdisciplinary research field that is gaining more attention nowadays. Multilayer networks offer several advantages in comparison to traditional network approaches because of their enormous potential to explore the organisation and connections of the different biological layers in both health and disease, making it a promising tool for future efforts in this area of research. This approach could be applied to other neurodegenerative diseases and autoimmune disorders.

## Supporting information

**S1 Table. Description of the variables in the Genetic dataset.**

**S2 Table. Description of the variables in the Molecular dataset.**

**S3 Table. Description of the variables in the PET dataset.**

**S4 Table. Description of the variables in the MRI dataset.**

**S5 Table. Description of the variables in the Phenotype dataset.**

**S6 Table. Description of the variables in the Risk Factors dataset.**

## Data Availability

The data used in this article was obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database. As such, ADNI investigators contributed to the design and implementation of the data collection and provided data, but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found here. Web interfaces of the individual networks, the combined six-layer network, the paths and the top risk factors paths are available at Github.

## Acknowledgments

This work was supported by project PID2021-127311NB-I00 financed by the Spanish Ministry of Science and Innovation, the Spanish State Research Agency and FEDER (MICIN/AEI/10.13039/501100011033/FEDER), by the Maria de Maeztu Programme for Units of Excellence in R&D (project CEX2018-000792-M), and by the ICREA Academia programme.

## References

1. Stelzmann RA, Norman Schnitzlein H, Reed Murtagh F. An english translation of alzheimer’s 1907 paper, “über eine eigenartige erkankung der hirnrinde”. *Clinical Anatomy*. 1995;8(6):429–431. doi:10.1002/ca.980080612.
2. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer’s disease. *Alzheimer’s & Dementia*. 2018;14(4):535–562. doi:10.1016/j.jalz.2018.02.018.

3. Tijms BM, Gobom J, Reus L, Jansen I, Hong S, Dobricic V, et al. Pathophysiological subtypes of Alzheimer's disease based on cerebrospinal fluid proteomics. *Brain*. 2020;143(12):3776–3792. doi:10.1093/brain/awaa325.
4. the Alzheimer's Disease Neuroimaging Initiative, Vogel JW, Young AL, Oxtoby NP, Smith R, Ossenkoppele R, et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nature Medicine*. 2021;27(5):871–881. doi:10.1038/s41591-021-01309-6.
5. Milà-Alomà M, Ashton NJ, Shekari M, Salvadó G, Ortiz-Romero P, Montoliu-Gaya L, et al. Plasma p-tau231 and p-tau217 as state markers of amyloid- $\beta$  pathology in preclinical Alzheimer's disease. *Nature Medicine*. 2022;28:1797–1801. doi:10.1038/s41591-022-01925-w.
6. Hasin Y, Seldin M, Lusic A. Multi-omics approaches to disease. *Genome Biology*. 2017;18(1):83. doi:10.1186/s13059-017-1215-1.
7. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*. 2011;12(1):56–68. doi:10.1038/nrg2918.
8. Zhou X, Menche J, Barabási AL, Sharma A. Human symptoms–disease network. *Nature Communications*. 2014;5(1):4212. doi:10.1038/ncomms5212.
9. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL. The human disease network. *Proceedings of the National Academy of Sciences*. 2007;104(21):8685–8690. doi:10.1073/pnas.0701361104.
10. Yildirim MA, Goh KI, Cusick ME, Barabási AL, Vidal M. Drug–target network. *Nature Biotechnology*. 2007;25(10):1119–1126. doi:10.1038/nbt1338.
11. Kennedy KE, de Rosbo NK, Uccelli A, Cellerino M, Ivaldi F, Contini P, et al. Multiscale networks in multiple sclerosis. *PLOS Computational Biology*. 2024;20(2):e1010980. doi:10.1371/journal.pcbi.1010980.
12. Avelar-Pereira B, Belloy ME, O'Hara R, Hosseini SMH, for the Alzheimer's Disease Neuroimaging Initiative. Decoding the heterogeneity of Alzheimer's disease diagnosis and progression using multilayer networks. *Molecular Psychiatry*. 2023;28(6):2423–2432. doi:10.1038/s41380-022-01886-z.
13. Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack C, Jagust W, et al. The Alzheimer's Disease Neuroimaging Initiative. *Neuroimaging Clinics of North America*. 2005;15(4):869–877. doi:10.1016/j.nic.2005.09.008.
14. Lee EG, Chen S, Leong L, Tulloch J, Yu CE. TOMM40 RNA Transcription in Alzheimer's Disease Brain and Its Implication in Mitochondrial Dysfunction. *Genes*. 2021;12(6):871. doi:10.3390/genes12060871.
15. Kauppi K, Fan CC, McEvoy LK, Holland D, Tan CH, Chen CH, et al. Combining Polygenic Hazard Score With Volumetric MRI and Cognitive Measures Improves Prediction of Progression From Mild Cognitive Impairment to Alzheimer's Disease. *Frontiers in Neuroscience*. 2018;12:260. doi:10.3389/fnins.2018.00260.
16. Tan CH, Hyman BT, Tan JJX, Hess CP, Dillon WP, Schellenberg GD, et al. Polygenic hazard scores in preclinical Alzheimer disease: Polygenic Hazard Scores in AD. *Annals of Neurology*. 2017;82(3):484–488. doi:10.1002/ana.25029.

17. Collij LE, Heeman F, Salvadó G, Ingala S, Altomare D, De Wilde A, et al. Multitracer model for staging cortical amyloid deposition using PET imaging. *Neurology*. 2020;95(11):e1538–e1553. doi:10.1212/WNL.0000000000010256.
18. Biel D, Brendel M, Rubinski A, Buerger K, Janowitz D, Dichgans M, et al. Tau-PET and in vivo Braak-staging as prognostic markers of future cognitive decline in cognitively normal to demented individuals. *Alzheimer's Research & Therapy*. 2021;13(1):137. doi:10.1186/s13195-021-00880-x.
19. Landau SM, Harvey D, Madison CM, Koeppe RA, Reiman EM, Foster NL, et al. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. *Neurobiology of Aging*. 2011;32(7):1207–1218. doi:10.1016/j.neurobiolaging.2009.07.002.
20. Dickerson BC, Stoub TR, Shah RC, Sperling RA, Killiany RJ, Albert MS, et al. Alzheimer-signature MRI biomarker predicts AD dementia in cognitively normal adults. *Neurology*. 2011;76(16):1395–1402. doi:10.1212/WNL.0b013e3182166e96.
21. Schwarz CG, Gunter JL, Wiste HJ, Przybelski SA, Weigand SD, Ward CP, et al. A large-scale comparison of cortical thickness and volume methods for measuring Alzheimer's disease severity. *NeuroImage: Clinical*. 2016;11:802–812. doi:10.1016/j.nicl.2016.05.017.
22. Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, et al. Development of Cognitive Instruments for Use in Clinical Trials of Antidementia Drugs: Additions to the Alzheimer's Disease Assessment Scale That Broaden Its Scope. *Alzheimer Disease & Associated Disorders*. 1997;11:13.
23. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". *Journal of Psychiatric Research*. 1975;12(3):189–198. doi:10.1016/0022-3956(75)90026-6.
24. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*. 1993;43(11):2412.2–2412-a. doi:10.1212/WNL.43.11.2412-a.
25. Pantoni L, Inzitari D. Hachinski's ischemic score and the diagnosis of vascular dementia: A review. *The Italian Journal of Neurological Sciences*. 1993;14(7):539–546. doi:10.1007/BF02339212.
26. Shannon CE, Weaver W. *The mathematical theory of communication*. Urbana: University of Illinois Press; 1998.
27. Li W. Mutual information functions versus correlation functions. *Journal of Statistical Physics*. 1990;60(5-6):823–837. doi:10.1007/BF01025996.
28. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. *Physical Review E*. 2004;69(6):066138. doi:10.1103/PhysRevE.69.066138.
29. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12(85):2825–2830.
30. Newman M. *Networks*. Oxford University Press; 2010. Available from: <https://academic.oup.com/book/27303>.
31. Domedel-Puig N, Rué P, Pons AJ, García-Ojalvo J. Information Routing Driven by Background Chatter in a Signaling Network. *PLoS Computational Biology*. 2011;7(12):e1002297. doi:10.1371/journal.pcbi.1002297.

32. Lasher CD. Edge swap graph; 2011-2012. GitHub Gist. Available from: <https://gist.github.com/gotgenes/2770023>.
33. Chételat G, Arbizu J, Barthel H, Garibotto V, Law I, Morbelli S, et al. Amyloid-PET and 18F-FDG-PET in the diagnostic investigation of Alzheimer's disease and other dementias. *The Lancet Neurology*. 2020;19(11):951–962. doi:10.1016/S1474-4422(20)30314-8.
34. Minoshima S, Giordani B, Berent S, Frey KA, Foster NL, Kuhl DE. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Annals of Neurology*. 1997;42(1):85–94. doi:<https://doi.org/10.1002/ana.410420114>.
35. Ewers M, Insel PS, Stern Y, Weiner MW, the Alzheimer's Disease Neuroimaging Initiative (ADNI) F. Cognitive reserve associated with FDG-PET in preclinical Alzheimer disease. *Neurology*. 2013;80(13):1194–1201. doi:10.1212/WNL.0b013e31828970c2.
36. Tsubaki Y, Akamatsu G, Shimokawa N, Takashima A, Katsube S, Sato H, et al. Examination of the relationship between PiB-PET and FDG-PET in Alzheimer's disease using random forest and two-sample t-test. *Journal of Nuclear Medicine*. 2020;61(supplement 1):3105–3105.
37. Perani D. FDG PET and cognitive symptoms of dementia. *Clinical and Translational Imaging*. 2013;1(4):247–260. doi:10.1007/s40336-013-0029-8.
38. Minhas PS, Jones JR, Latif-Hernandez A, Sugiura Y, Durairaj AS, Wang Q, et al. Restoring hippocampal glucose metabolism rescues cognition across Alzheimer's disease pathologies. *Science (New York, NY)*. 2024;385(6711):eabm6131. doi:10.1126/science.abm6131.
39. Casanova MF, Starkstein SE, Jellinger KA. Clinicopathological correlates of behavioral and psychological symptoms of dementia. *Acta Neuropathologica*. 2011;122(2):117–135. doi:10.1007/s00401-011-0821-3.
40. Lyketsos CG, Steele C, Baker L, Galik E, Kopunek S, Steinberg M, et al. Major and minor depression in Alzheimer's disease: prevalence and impact. *Journal of Neuropsychiatry and Clinical Neurosciences*. 1997;9(4):556–561. doi:10.1176/jnp.9.4.556.
41. Zeylan ME, Senyuz S, Picón-Pagès P, García-Elías A, Tajés M, Muñoz FJ, et al. Shared Proteins and Pathways of Cardiovascular and Cognitive Diseases: Relation to Vascular Cognitive Impairment. *Journal of Proteome Research*. 2024;23(2):560–573. doi:10.1021/acs.jproteome.3c00289.
42. Cortes-Canteli M, Gispert JD, Salvadó G, Toribio-Fernandez R, Tristão-Pereira C, Falcon C, et al. Subclinical Atherosclerosis and Brain Metabolism in Middle-Aged Individuals. *Journal of the American College of Cardiology*. 2021;77(7):888–898. doi:10.1016/j.jacc.2020.12.027.
43. Tristão-Pereira C, Fuster V, Oliva B, Moreno-Arciniegas A, Garcia-Lunar I, Perez-Herreras C, et al. Longitudinal interplay between subclinical atherosclerosis, cardiovascular risk factors, and cerebral glucose metabolism in midlife: results from the PESA prospective cohort study. *The Lancet Healthy Longevity*. 2023;4(9):e487–e498. doi:10.1016/S2666-7568(23)00134-4.

44. Sierra C. Hypertension and the Risk of Dementia. *Frontiers in Cardiovascular Medicine*. 2020;7. doi:10.3389/fcvm.2020.00005.
45. Wolters EE, Ossenkuppele R, Verfaillie SCJ, Coomans EM, Timmers T, Visser D, et al. Regional [18F]flortaucipir PET is more closely associated with disease severity than CSF p-tau in Alzheimer's disease. *European Journal of Nuclear Medicine and Molecular Imaging*. 2020;47(12):2866–2878. doi:10.1007/s00259-020-04758-2.
46. Mahaman YAR, Embaye KS, Huang F, Li L, Zhu F, Wang JZ, et al. Biomarkers used in Alzheimer's disease diagnosis, treatment, and prevention. *Ageing Research Reviews*. 2022;74:101544. doi:<https://doi.org/10.1016/j.arr.2021.101544>.
47. Johnson ECB, Carter EK, Dammer EB, Duong DM, Gerasimov ES, Liu Y, et al. Large-scale deep multi-layer analysis of Alzheimer's disease brain reveals strong proteomic disease-related changes not observed at the RNA level. *Nature Neuroscience*. 2022;25(2):213–225. doi:10.1038/s41593-021-00999-y.
48. Gustafsson M, Nestor CE, Zhang H, Barabási AL, Baranzini S, Brunak S, et al. Modules, networks and systems medicine for understanding disease and aiding diagnosis. *Genome Medicine*. 2014;6(10):82. doi:10.1186/s13073-014-0082-6.
49. Gargalovic PS, Imura M, Zhang B, Gharavi NM, Clark MJ, Pagnon J, et al. Identification of inflammatory gene modules based on variations of human endothelial cell responses to oxidized lipids. *Proceedings of the National Academy of Sciences*. 2006;103(34):12741–12746. doi:10.1073/pnas.0605457103.
50. Jha M, Roy S, Kalita JK. Prioritizing disease biomarkers using functional module based network analysis: A multilayer consensus driven scheme. *Computers in Biology and Medicine*. 2020;126:104023. doi:10.1016/j.combiomed.2020.104023.
51. Ran J, Li H, Fu J, Liu L, Xing Y, Li X, et al. Construction and analysis of the protein-protein interaction network related to essential hypertension. *BMC Systems Biology*. 2013;7(1):32. doi:10.1186/1752-0509-7-32.
52. Gosak M, Markovič R, Dolensšek J, Slak Rupnik M, Marhl M, Stožer A, et al. Network science of biological systems at different scales: A review. *Physics of Life Reviews*. 2018;24:118–135. doi:10.1016/j.plrev.2017.11.003.
53. Vangimalla RR, Sreevalsan-Nair J. HCNM: Heterogeneous Correlation Network Model for Multi-level Integrative Study of Multi-omics Data for Cancer Subtype Prediction. In: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Mexico: IEEE; 2021. p. 1880–1886. Available from: <https://ieeexplore.ieee.org/document/9630781/>.
54. He D, Xie L. A cross-level information transmission network for hierarchical omics data integration and phenotype prediction from a new genotype. *Bioinformatics*. 2021;38(1):204–210. doi:10.1093/bioinformatics/btab580.
55. Ha MJ, Stingo FC, Baladandayuthapani V. Bayesian Structure Learning in Multilayered Genomic Networks. *Journal of the American Statistical Association*. 2021;116(534):605–618. doi:10.1080/01621459.2020.1775611.
56. Wang C, Zhao N, Sun K, Zhang Y. A Cancer Gene Module Mining Method Based on Bio-Network of Multi-Omics Gene Groups. *Frontiers in Oncology*. 2020;10:1159. doi:10.3389/fonc.2020.01159.



57. Yurkovich JT, Tian Q, Price ND, Hood L. A systems approach to clinical oncology uses deep phenotyping to deliver personalized care. *Nature Reviews Clinical Oncology*. 2020;17(3):183–194. doi:10.1038/s41571-019-0273-6.
58. Hammoud Z, Kramer F. Multilayer networks: aspects, implementations, and application in biomedicine. *Big Data Analytics*. 2020;5(1):2. doi:10.1186/s41044-020-00046-0.
59. Latifi-Navid H, Barzegar Behrooz A, Jamehdor S, Davari M, Latifinaid M, Zolfaghari N, et al. Construction of an Exudative Age-Related Macular Degeneration Diagnostic and Therapeutic Molecular Network Using Multi-Layer Network Analysis, a Fuzzy Logic Model, and Deep Learning Techniques: Are Retinal and Brain Neurodegenerative Disorders Related? *Pharmaceuticals*. 2023;16(11):1555. doi:10.3390/ph16111555.
60. Gosak M, Markovič R, Dolensšek J, Slak Rupnik M, Marhl M, Stožer A, et al. Network science of biological systems at different scales: A review. *Physics of Life Reviews*. 2018;24:118–135. doi:https://doi.org/10.1016/j.plrev.2017.11.003.
61. Boccaletti S, Bianconi G, Criado R, del Genio CI, Gómez-Gardeñes J, Romance M, et al. The structure and dynamics of multilayer networks. *Physics Reports*. 2014;544(1):1–122. doi:https://doi.org/10.1016/j.physrep.2014.07.001.
62. Blauwendraat C, Pletnikova O, Geiger JT, Murphy NA, Abramzon Y, Rudow G, et al. Genetic analysis of neurodegenerative diseases in a pathology cohort. *Neurobiology of Aging*. 2019;76:214.e1–214.e9. doi:https://doi.org/10.1016/j.neurobiolaging.2018.11.007.
63. Liu Y, Xiao X, Liu H, Liao X, Zhou Y, Weng L, et al. Clinical characteristics and genotype-phenotype correlation analysis of familial Alzheimer’s disease patients with pathogenic/likely pathogenic amyloid protein precursor mutations. *Frontiers in Aging Neuroscience*. 2022;14. doi:10.3389/fnagi.2022.1013295.
64. Bateman RJ, Xiong C, Benzinger TLS, Fagan AM, Goate A, Fox NC, et al. Clinical and Biomarker Changes in Dominantly Inherited Alzheimer’s Disease. *New England Journal of Medicine*. 2012;367(9):795–804. doi:10.1056/NEJMoa1202753.
65. Brandt R, Hundelt M, Shahani N. Tau alteration and neuronal degeneration in tauopathies: mechanisms and models. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 2005;1739(2):331–354. doi:https://doi.org/10.1016/j.bbadis.2004.06.018.