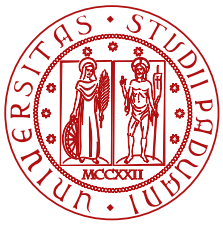


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

UNIVERSITY OF PADUA

DEPARTMENT OF INFORMATION ENGINEERING

MASTER'S DEGREE IN BIOENGINEERING

Integrating morphometric similarity networks with imaging transcriptomics

Supervisor
Prof. Veronese Mattia

Co-Supervisor
Dott. Barzon Leonardo

Candidate
Tomasella Francesco

Student ID
2121661

Academic Year 2025-2026

Graduation Date 23/04/2026

From You, 5 Years Ago
To You, 5 Years From Now

Abstract

Neurodegenerative diseases represent a heterogeneous group of progressive neurological disorders whose prevalence and socioeconomic burden are expected to rise with increasing global life expectancy. While traditional neuroimaging approaches are capable of detecting macroscopic structural alterations, they are limited in resolving the underlying microscale cellular and molecular mechanisms.

This work proposes and validates a computational pipeline integrating Morphometric Similarity Networks (MSNs) with imaging transcriptomics to bridge whole-brain structural topology and gene expression signatures in neurodegenerative disease cohorts. As a case study, it focuses on Parkinson's disease (PD), which is increasingly recognized as a brain-wide structural network disorder rather than a focal dopaminergic pathology.

After an exploratory mass univariate analysis, individualized MSNs were constructed from T1-weighted MRI data for 212 subjects (106 PD patients, 106 healthy controls) from the Parkinson's Progression Markers Initiative (PPMI), using five cortical morphometric features extracted via FreeSurfer and parcellated across the 68-region Desikan-Killiany atlas. Pairwise structural similarity was quantified through multivariate Euclidean distance, and node strength was derived as the primary topological metric. Group-level differences were assessed using a General Linear Model controlling for age, sex, estimated total intracranial volume and scanner model, with Benjamini-Hochberg FDR correction applied across all regions. The resulting spatial phenotype map of beta coefficients was cross-referenced with Allen Human Brain Atlas microarray gene expression data via Partial Least Squares regression, with spatial autocorrelation addressed through spin-test permutations using the Váša algorithm. Finally, gene set enrichment analyses were performed.

Mass univariate and node strength analyses revealed subtle but directionally consistent structural alterations in PD, confirming the pipeline's sensitivity to detect low-effect signals. Hub decoupling emerged as the dominant topological reorganization pattern, affecting 32.4% of regions and aligning with established hub vulnerability hypotheses. Although global PLS permutation testing did not reach significance, transcriptomic analysis identified LRRK2 as a significant canonical PD gene correlate, while Gene Set Enrichment Analysis consistently implicated neu-

roinflammatory pathways, glial cell signatures and excitatory neuronal enrichment as molecular correlates of the observed structural reorganization. This confirmed coherent integration between structural and transcriptomic layers. Secondary application of the full pipeline to a frontotemporal dementia cohort confirmed the framework's generalizability, yielding more pronounced results consistent with the greater cortical involvement characteristic of FTD.
This work delivers a modern, scalable and reproducible computational framework, providing a foundation for future integrative investigations of structural connectome alterations and their molecular underpinnings across neurodegenerative conditions.

Contents

Chapter 1

Introduction

The prevalence and socioeconomic burden of neurodegenerative diseases (NDDs) are expected to rise proportionately to the increase of global life expectancy. Understanding the structural and molecular underpinnings of these conditions requires computational frameworks capable of integrating neuroimaging and transcriptomic data at scale.

Parkinson's disease, in particular, emerges as the second most common neurodegenerative illness, with its incidence expanding rapidly and representing an ideal case study given the breadth of available neuroimaging cohorts. While it was initially categorized as a focal neurological disorder, mainly linked to its clinically defined motor symptoms such as resting tremors, b radykinesia and rigidity, recent research has shifted the view on Parkinson's disease. The characteristic degeneration of the dopaminergic neurons in the substantia nigra remains as a common trait between the two paradigms, but now the disease is also seen as a highly complex, brain-wide and multi-systemic network disorder[1].

This connectivity-centered view of the disorder is supported by its physical progression path, where pathogenic misfolded proteins like α-synuclein are thought to propagate along the brain's structural connectome in a prion-like manner. Consequently, the disease can be categorized in stages, which correspond to its progression through the brain structural pathways, from white matter to the cerebral cortex. It also helps in linking structural alterations to changes in symptoms like manifestation of cognitive decline and eventually dementia.

To understand the structural connectome disruptions, researchers have traditionally used neuroimaging approaches like diffusion-weighted imaging (DWI) tractography and structural covariance networks (SCNs). However, these methods carry inherent limitations, prompting the need for a framework capable of analyzing the cortex through detailed, single-subject analyses.

To do this, Morphometric Similarity Networks (MSNs) were introduced as a multidimensional approach for cortical mapping, capable of constructing individualized and biologically mean-

ingful whole-brain anatomical networks [2].
These networks have been shown to mirror spatial patterns of gene expression and capture known cortical cytoarchitecture. Despite this, traditional macroscopic MRI resolution only offers an indirect proxy of the actual underlying microscopic pathology. To bridge the two different scales, the field of imaging transcriptomics has emerged, by establishing a quantitative framework mapping spatially resolved gene expression profiles to phenotypes derived from neuroimaging pipelines [3].

The primary objective of this work is to create and validate a computational framework integrating MSN analysis with imaging transcriptomics to track topological alterations across different clinical cohorts, and subsequently investigate their spatial covariance with brain transcriptomic data. The full pipeline aims to do this by: constructing individualized networks through multivariate euclidean distance from regional morphometric data, evaluating the topological alterations in networks through the analysis of node strength, employing generalized linear models to quantify group differences between cohorts, performing partial least square regression to evaluate their spatial covariance with gene expression data and finally performing enrichment analysis on the resulting genes. This work also uses statistically rigorous correction techniques like spin testing to account for the spatial covariance of brain regions, essential to not incur into false positive associations.

As the present work aims to establish a structured pipeline that can be replicated and easily followed, it is here applied and validated with data sourced from the Parkinson's Progression Markers Initiative (PPMI), while also being tested and providing results for frontotemporal dementia subjects. This computational tool helps in uncovering abnormalities in the structural connectome and improve our comprehension of the underlying cellular pathology, potentially advancing our knowledge of poorly understood neurodegenerative diseases.

Chapter 2

Background

2.1 Neurodegenerative disease

As life expectancy increases throughout the world, the prevalence and burden of neurodegenerative diseases is expected to rise[4]. Neurodegenerative diseases, in short NDDs, can be classified as a heterogeneous group of neurological disorders that give rise to a progressive loss of neurons in the central nervous system (CNS) or peripheral nervous system (PNS). Their symptoms mainly include impaired memory, cognition, behavior and sensory and/or motor function, and are related to the collapse of the structure and function of the neuronal networks constituting the brain.
Among all the existing neurodegenerative diseases, the most prevalent ones are: Alzheimer's disease (AD), Parkinson's diseases (PD), frontotemporal dementia (FTD), amyotrophic lateral sclerosis (ALS), Huntington disease (HD), prion disease (PrD) and multiple sclerosis[5]. While our understanding of NDDs is still limited, with their pathogenesis still mostly unknown, their emergence has been strongly linked with inflammation, neuronal networks dysfunction, aging, DNA and RNA defects and altered energy metabolism[6]. The impact of pathologies like Alzheimer's disease, Parkinson's disease and dementias is enormous, as they are not only limited to the affected individuals, but present a significant challenge to healthcare systems and societies as well[7]. Therefore, there is an urgent need to explore every possible research avenue and investigation to learn more about these conditions.

2.1.1 Parkinson's disease

Parkinson's disease is the second most common neurodegenerative illness, with a growing number of people all around the globe being affected by this condition[8]. In 2016, 6.1 million people were living with the disease, and its incidence and prevalence have continued to rapidly

rise throughout the years [9][10]. Unique to a neurodegenerative disease, its duration can span decades, amplifying even more the difficulty in recognizing the earliest stages of the disease. This leads to an average delay between the first noticeable symptoms and the timing of the diagnosis of 10 years. The typical presentation includes a slow progression with accumulating disability for affected individuals, with the disease also having profound consequences for caregivers and society, with a mounting socioeconomic burden[11]. Initially, Parkinson's Disease was viewed as a focal neurological disorder, clinically recognized via its most common motor symptoms that include progressive involuntary resting tremors, bradykinesia (slowness of movement) and rigidity[1]. However, the clinical spectrum also contains many less visible components, including non-motor features, such as cognitive decline, depression, and pain. These non-motor features contribute substantially to the impairment experienced by the patients. Parkinson's disease's highly localized perspective was supported by its primary pathological signature: the progressive degeneration of dopaminergic neurons in the substantia nigra (SN), a basal ganglia structure located in the midbrain, which in turn leads to severe damage of the nigrostriatal dopaminergic system[12]. The underlying pathophysiology of this neuronal death has been linked to the toxic accumulation of misfolded protein aggregates, in particular α-synuclein, which has been strongly associated to the rare familial forms of PD. While α-synuclein in its normal state is highly expressed in the brain at presynaptic terminals, in order to modify cellular activity by interacting with proteins, in Parkinson's disease it assumes an abnormal insoluble fibril configuration, which was discovered to aggregate into pathological inclusions called Lewy bodies. This accumulation in the dopaminergic neurons of the substantia nigra disrupts several crucial biological processes including oxidative stress, neuroinflammation, as well as mitochondrial, lysosomal and endosomal function, and proteostasis, ultimately leading to cell death[13]. Despite this, clinical and scientific research has shifted the perspective on Parkinson's disease over the years: rather than a localized pathology of the substantia nigra, it is now recognized as a highly complex, multi-systemic neurodegenerative disorder. This is motivated by the fact that, alongside the more classic motor symptoms, patients also experience a wide range of non-motor manifestations, including mood and cognitive decline, sleep disturbances and olfactory impairment. Notably, these symptoms can appear earlier in the disease process, even before the onset of motor symptoms, suggesting the involvement of structural and functional network alterations that extend well beyond isolated nigrostriatal dopamine projections[1][14]. Recent research has thus shifted toward a connectivity-centered view of the disease's spread, hypothesizing that pathogenic α-synuclein propagates along the neuronal connectome in a prion-like manner [15][8], a process that may be further accelerated by maladaptive immune and inflammatory gut responses[16]. The physical progression of PD pathology demands for a distinction between the effect of the disease in cortical and non-cortical regions, with the latter representing the epi-

center of the disease. According to the now well-established Braak staging hypothesis[17], this pathology begins deep in the lower brainstem, then progresses into the midbrain before moving to the subcortical areas. During this early phase of the disease, structural and functional damage is usually confined to these non-cortical networks, with neuroimaging revealing that early structural disconnectivity in these areas drives classic motor deficits[18]. With the advancement of the disease, it does not remain confined to the subcortex, but propagates upward and outward from the substantia nigra, ultimately reaching and ecompassing the cerebral cortex [19]. Figure 2.1 depicts the Braak staging hypothesis. Magnetic Resonance Imaging studies have demonstrated that the rate and severity of brain atrophy and focal cortical thinning in an area of the cerebral cortex is directly proportional to how functionally and structurally connected that area is to the initial substantia nigra epicenter. This suggests that regions with higher connectivity density experience greater disease exposure, resulting in more pronounced cortical atrophy[15].

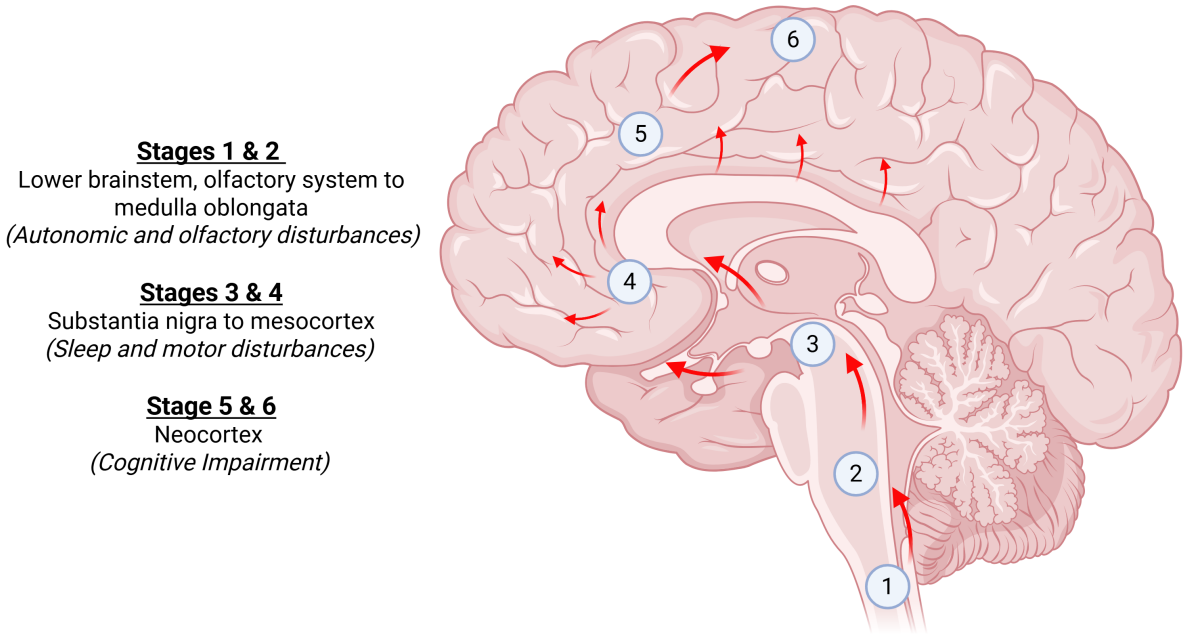


Figure 2.1: Visualization of the Braak staging hypothesis for Parkinson's disease.

The transition from non-cortical to cortical involvement marks a critical clinical turning point, because while non-cortical damage primarily manifests as motor dysfunction, the spread of the disease pathology in cortical regions such as frontal and somatosensory regions signals the onset of severe non-motor symptoms. This progressive cortical atrophy finally leads to ex-

ecutive dysfunction, behavioral disorders and eventually dementia, potentially in synergy with copathologies like amyloid-beta aggregation, that are also present in Alzheimer's disease[16]. This paradigm shift, from a focal dopaminergic loss to a brain-wide network disorder, has been strengthened thanks to advances in neuroimaging. Techniques like Diffusion Tensor Imaging (DTI), resting-state functional MRI (fMRI) and positron emission tomography (PET) scans, together with the mathematical framework of graph theory, have empowered researchers to map these complex brain-wide disruptions [20][18]. Understanding these alterations requires integrating knowledge coming from two deeply interconnected but distinct levels: the structural or physical networks and the functional networks, involving communication and metabolic activity across those pathways. In addition, analyzing these networks, while shedding light on the unknown underlying disease mechanisms, could also provide a way to overcome traditional clinical challenges by developing new objective biomarkers[1].

Structural networks in particular are of vital importance to investigate links between functional activity and pathophysiology, providing an insight in the progression and spread of the disease. In non-cortical structures, deformation-based morphometry reveals pronounced, non-random structural atrophy across the basal ganglia, midbrain, thalamus and limbic structures. The rate of volume loss appears to be directly proportional to its anatomical distance from the substantia nigra, supporting the network-spread idea of toxic alpha-synuclein [8]. At the nodal level, the subcortex demonstrates structural failure with the amygdala, pallidum, caudate, and putamen showing significantly decreased nodal local efficiency and clustering coefficients. This indicates a severe disruption of localized information processing and segregation within the basal ganglia structural subgraphs.

Cortical structures also undergo thinning and atrophy, with longitudinal studies revealing severe thinning in the left occipital lobe, bilateral frontal lobes and the right somatomotor-sensory cortex. For these areas, deformation-based morphometry also detects brain atrophy localized to bilateral temporoparietal regions. Structural cortical atrophy maps closely onto specific functional intrinsic brain networks, with regions belonging to the limbic, frontoparietal and ventral attention networks, as well as anterior parts of the default mode network (DMN), demonstrating the greatest degree of cortical thinning [15]. The structural networks of the cortex, usually derived from correlation analyses, undergo significant nodal reorganization: PD patients experience a disruption and loss of network hubs in the parietal, occipital, and temporal cortices. In response, there is a generation of new structural hubs in the frontal lobes, likely representing a compensatory upregulation of dopaminergic transmission [19]. Finally, PD brains also show degeneration of interhemispheric axonal connections between frontal areas and altered connectivity in the uncinate fasciculus, which connects limbic regions to the frontal cortex [20]. While the worsening of motor symptoms like akinesia and rigidity directly correlates with greater

deformation and brain atrophy in the contralateral middle temporal gyrus, tremor arises from a completely different subnetwork[21], which could also explain their different responses to dopaminergic pharmacotherapy[22]. Cognitive symptoms are instead characterized by reduced metabolic network activity and cortical thinning, especially in the left frontal cluster. Cortical thinning is also amplified by other dysfunctions like amyloid-beta deposition [12].

The brain's structural connectome is now recognized as the physical conduit for propagation of toxic misfolded proteins, such as α-synuclein, and advanced neuroimaging studies combined with graph theory have provided direct evidence supporting the network-spread hypothesis [8]. However, while traditional structural metrics like DTI have successfully mapped white matter disconnectivity and localized volume loss, the neurodegenerative process in PD alters multiple morphological properties of the brain simultaneously, including cortical thickness, gray matter volume, and surface area [15]. Analyzing these features in isolation often limits our ability to see the full picture of structural network reorganization[19]. To fully characterize how the parkinsonian brain structurally reorganizes, more integrative connectomic approaches are required: this necessitates a shift toward advanced structural networks analysis, so that researchers can unlock a much deeper, integrated map of the brain to study its structural vulnerability and compensatory reorganization.

2.2 Morphometric Similarity Networks

Mapping the structural connectome of the human brain is of great importance in the neuroscience field and is essential for understanding how anatomical organization affects cognitive function. Through the decades, scientists have tried to map these large-scale networks of connected regions in living humans. Traditionally, this has been attempted through two main neuroimaging approaches: diffusion-weighted imaging (DWI) tractography and structural covariance network (SCN) analysis[23][2]. While both of them have advanced our understanding of brain topology, they also possess methodological and biological limitations that require better approaches for more informative connectomics. DWI tractography, which works by reconstructing the trajectory of axonal tracts based on the principal directions of water molecule diffusion, is mainly limited to reproducing white matter connections, particularly sensitive to noise and struggles with long-distance projections. Instead, SCN approaches connectivity through a different lens, utilizing simpler, macrostructural measurements to reconstruct whole-brain networks, starting from a measure of a single morphometric feature (usually cortical thickness) at multiple brain regions. The network is then formed by estimating the covariance or correlation of this single feature between all possible pairs of regions. The main limitation of this approach is that typically the correlation is computed across several participants, obtaining a population-level network.

This hinders the use of SCN for personalized, single-subject analysis, making it unsuitable to capture the complex, multidimensional nature of each individual's brain anatomy [24][25]. To overcome the limitations of both DWI and SCNs, Morphometric Similarity Networks (MSNs) were introduced, in order to offer a different, multidimensional approach to human cortical network mapping [2].

Unlike SCNs, which estimate the inter-regional correlation of a single macro-structural feature across multiple individuals, MSNs estimate the correlation between multiple macro and microstructural MRI variables within a single subject. This allows the construction of individualized and meaningful whole-brain anatomical networks by extracting an array of features for every brain region. The most common cortical properties used for MSNs calculations include cortical thickness, surface area, gray matter volume, curvature indices and myelination markers [26]. The theoretical foundation of MSNs is rooted in established neurobiology, with histological evidence from non-human primates confirming the structural model of connectivity, demonstrating that a stronger axo-synaptic connectivity between cortical regions is linked to greater cytoarchitectonic similarity [27][28][2]. Because MRI metrics serve as proxy markers for distinct underlying histological properties, brain regions exhibiting high morphometric similarity are very likely to be axonally connected[29].

To demonstrate MSN's superior biological validity and anatomical accuracy compared to SCNs and DWI tractography, studies have compared how well these networks align with classical cytoarchitectonic classifications. The results show that MSNs capture known cortical cytoarchitecture, align with true axonal tract-tracing connectivity and also mirror spatial patterns of gene co-expression, specifically for genes enriched in neuronal functions and those overexpressed in the supragranular layers of the cortex. This makes them a powerful tool for bridging macroscopic structure and molecular mechanisms [30][2]. The standard methodology for constructing a MSN typically follows a systematic, multi-step pipeline. First, the cerebral cortex is parcellated into a defined set of regions, the number of which is determined by the specific cortical atlas employed. The most frequently used one, which represents the foundation for MSNs, is the 68-region Desikan-Killiany (DK) Atlas [31]. However, because standard DK regions vary significantly in size, it can also be further sub-parcellated into 308 spatially contiguous regions of approximately equal volume using a backtracking algorithm [7]. Another atlas with widespread use is the Destrieux Atlas (DA) [32], which provides an automated classification of 74 regions per hemisphere and can also be utilized for high-resolution MSN construction [33]. Both the DK atlas and the DA atlas are based on sulcal and gyral landmarks. To validate the replicability of MSN topologies across different spatial definitions, some studies have also used functionally-derived parcellations like the Schaefer 400 atlas [34] and more biologically accurate ones like the von Economo atlas [35]. For each regional node coming from the parcellation of choice,

a multitude of morphometric features are extracted, with each Region of Interest (ROI) being represented by an individual value. MSN approaches use up to 10 features, obtained from multiple acquisition tecnhiques. T1w MRI allows to derive surface-based macrostructural metrics like cortical thickness (CT), surface area (SA), gray matter volume (GM), intrinsic and Gaussian curvature (IC, GC), mean curvature (MC), curvature index (CI) and folding index (FI), while other imaging modalities like diffusion tensor imaging and magnetization transfer imaging (MTI) enable the measurement of volume-based diffusion and magnetization metrics like fractional anisotropy (FA), mean diffusivity (MD) and magnetization transfer (MT) [26][2][25]. Because these features are measured on vastly different numerical scales, each feature vector is often mathematically normalized to produce distributions comparable in value. Following normalization, the morphometric similarity between each possible pair of regions is commonly calculated using the Pearson product-moment correlation coefficient. Other researchers propose the use of multivariate euclidean distance, which offers a more granular approach, to compute a distance measure between cortical areas [36]. This raw distance then undergoes a conversion into a similarity measurement through mathematical functions. Another advanced framework is the Morphometric INverse Divergence (MIND) [37], which estimates inter-areal similarity by applying the symmetric Kullback-Leibler (KL) divergence to the multivariate distributions of morphometric features at the vertex level, capturing similarities that regional averages might miss[38]. After computing the similarity between brain regions, a symmetrical matrix representing the individual's structural connectome is generated, which can optionally be thresholded before proceeding to the next step of topological analysis [39]. Figure 2.2 highlights the framework for Morhometric Similarity Network analysis as proposed by Seidlitz and his colleagues in 2018 [2].

Rather then being static, the MSN framework is highly adaptable, allowing researchers to develop custom pipelines with specific similarity metrics and features tailored to their clinical questions. While the original, comprehensive MSNs rely on multimodal images, acquiring these sequences is time-consuming, expensive and often unsuitable for pediatric populations, neonates or clinical groups who cannot tolerate long scans and are prone to head motion artifacts. So, instead of calculating MSNs combining T1-weighted, T2-weighted, and diffusion-weighted imaging to extract up to 10 morphometric features, the use of a restricted set of features derived solely from standard T1w MRI has become highly appealing and helps in making the MSN framework more broadly applicable. This way T1w-only models can be applied retrospectively to vast amounts of legacy clinical data where advanced imaging sequences were never acquired. As a result, researchers have validated T1w-restricted MSNs that typically extract between 5 and 7 macrostructural features per region which include cortical thickness, surface area, gray matter volume, mean curvature, Gaussian (intrinsic) curvature, folding index and curvature index

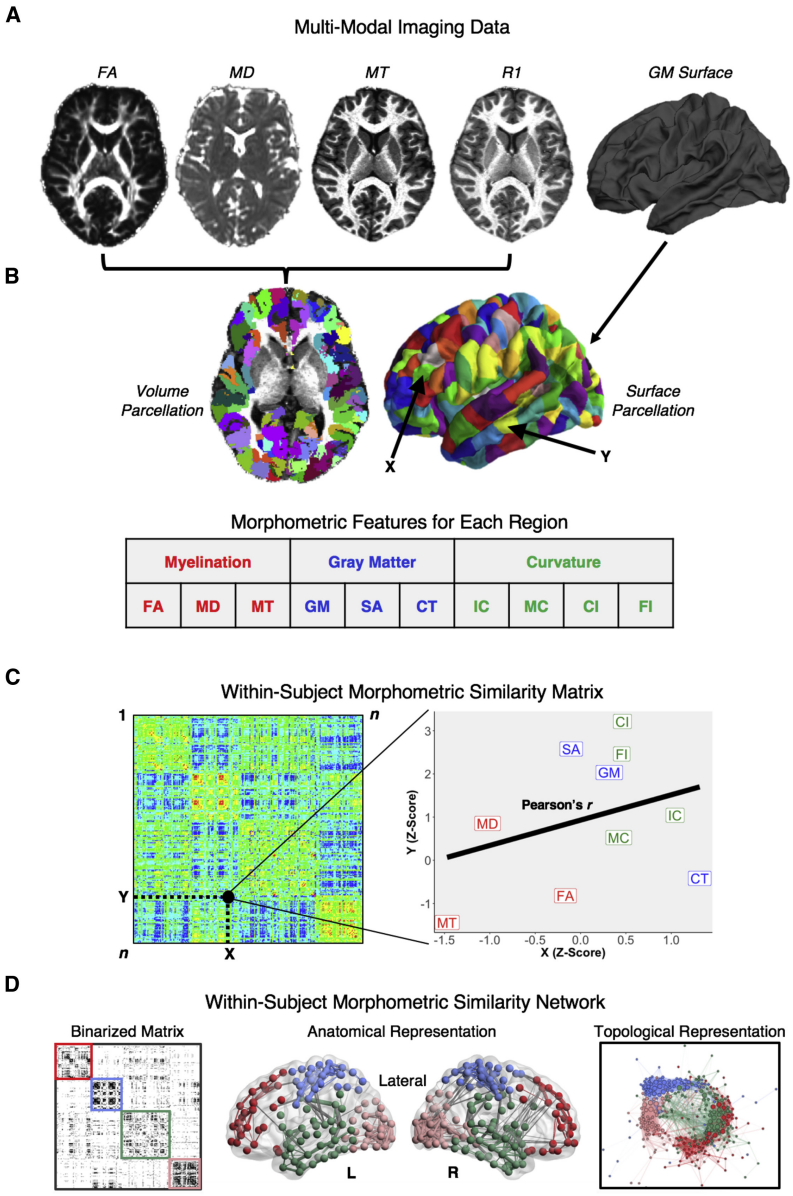


Figure 2.2: MSN framework proposed by Seidlitz et. al, 2018 [2]. (A) Collection of multiple MRI parameters. (B) Parameters mapped to parcellation template (C) Computation of the Morphometric Similarity Matrix through Pearson's correlation. (D) Thresholding and Topological Analysis.

[39][26][40]. These reduced-feature models successfully approximate the topological architecture of full multimodal MSNs and exhibit high test-retest reliability [25][26]. The main drawback of this T1w restriction is that removing myelo-architectural features (like fractional anisotropy or magnetization transfer) alters the resulting nodal-level topology of the network. Because different MRI modalities capture fundamentally different biological properties, dropping features systematically produces a different neuroimaging phenotype. T1w-restricted models inevitably exhibit lower estimation precision, meaning there is a greater standard deviation of edge weights across participants compared to full multimodal MSNs [2]. Furthermore, some studies indicate

that T1w-only MSNs may possess limited predictive validity for certain tasks; for example, they have sometimes failed to replicate the strong predictions of generalized cognitive ability (IQ) that multimodal MSNs achieve [25].

The implementation of MSNs unlocks extensive possibilities for neuroinformatics, and by using graph theory researchers can profile the complex topological architecture of the brain through the evaluation of measures such as nodal degree, small-worldness, modularity, clustering coefficients, and local efficiency [38][39]. Furthermore, analyzing individual deviations in MSN topology has demonstrated vast clinical and biological utility, as histologically similar cortical areas have a higher probability of being axonally connected [29][41].
In relation to neurological diseases, MSNs can serve as biomarkers to track Alzheimer's, Autism, Depression, and Insomnia [25][36]. When paired with machine learning, they can uncover hidden neuroanatomical subtypes, such as in Early-Onset Schizophrenia, and thus map personalized disease progression[42]. MSNs can also predict brain abnormalities linked to childhood obesity[43] and bridge brain structure and transcriptomics. By overlapping individual MSN maps with brain-wide gene expression databases (such as the Allen Human Brain Atlas), researchers can pinpoint the transcriptomic vulnerabilities driving macroscopic brain changes. For example, morphometric similarity reductions in AD have been directly correlated with the downregulation of genes enriched for synaptic signaling and neuronal development[30].

2.2.1 The Desikan-Killiany Atlas

The quantification of structural and functional variations in the human cerebral cortex is the basis of modern neuroimaging research, due to the cortex's topological complexity, morphologically variable structure and profound inter-subject heterogeneity. This amount of variance poses a significant challenge in the processing of neuroimaging data, complicating both standardization and statistical analysis across different populations. Historically, the quantification of cortical features like cortical thickness mainly relied on manual tracing and cytoarchitectonic structures[44]. This approach was time-expensive and was still affected by subjectivity even though it was performed by experts. The time constraint in particular rendered large-scale studies practically unfeasible. To overcome these limitations, automated computational pipelines were developed, in order to make the segmentation and parcellation steps of neuroimaging studies faster and reproducible. For structural MRI scans in particular, thanks to its deep integration in the FreeSurfer processing pipeline [45], the Desikan-Killiany (DK) cortical atlas has become one of the most widely used parcellations. Officially introduced in 2006, it was specifically built for subdviding the human cerebral cortex into standard, gyral-based regions of interest (ROIs) [31]. The foundational ground truth of the DK atlas was established through rigorous manual labeling, defining and outlining gyral structures and thus including deeper segments of the

cortex.

This process ultimately yielded 34 different regions per hemisphere, with a total of 68 labeled regions for each brain. To align the parcellation with classical neuroanatomical nomenclature and facilitate macroscopic analysis, these ROIs are mapped to the major cerebral lobes: the frontal lobe, the parietal lobe, the temporal lobe, the occipital lobe and also include the insula and subdivisions of the cingulate cortex. The latter can either be considered separately or can be included to follow the main 4 lobes grouping. Specific region names are outlined in table 2.1, while a visual representation of the parcellation on a cortical surface is presented in Figure 2.3. Notably, the corpus callosum was excluded a priori, since it is fundamentally a deep white matter structure, not cortex, and was only included in the manual tracing phase to explicitly constrain the boundaries of adjacent cortical regions.

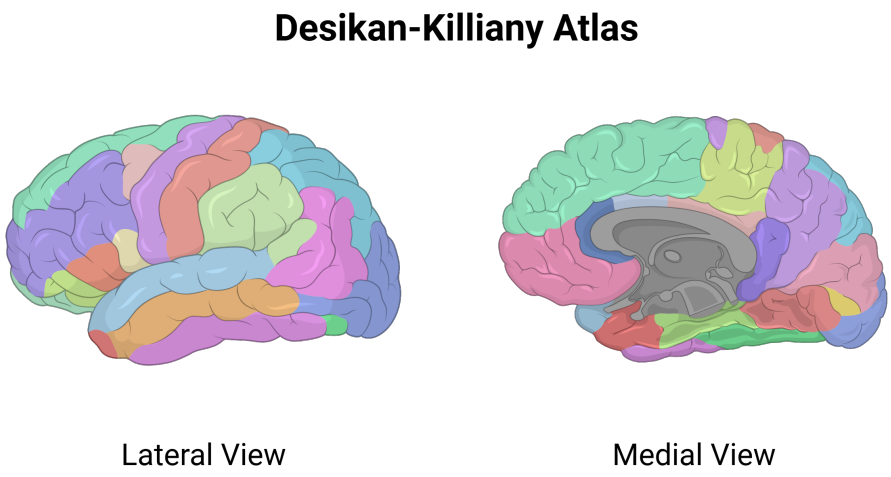


Figure 2.3: Desikan-Killiany parcellation visualization. The brain is presented both through a lateral view and a medial view, to show structures hidden inside the brain's gyri.

While nowadays the DK atlas is the undisputed default parcellation scheme in a vast majority of neuroimaging processing pipelines, it is frequently compared to alternatives, like the Destrieux atlas. The latter divides the cortex into separate gyral and sulcal segments based on precise curvature thresholds, resulting in a different spatial granularity compared to the DK atlas. Instead of 34 regions per hemisphere, the Destrieux atlas produces 74 regions per hemisphere, with a total of 148 regions [32]. The choice between the two involves an important trade-off between spatial granularity, statistical power, and computational efficiency, resulting in the the DK atlas usually being preferred. An improved version of the DK atlas was also intro-

Table 2.1: Desikan-Killiany atlas regions

Major subdivisions	**Regions (per Hemishpere)**
Frontal Lobe	Superior Frontal
	Rostral Middle Frontal
	Caudal Middle Frontal
	Pars Opercularis
	Pars Triangularis
	Pars Orbitalis
	Lateral Orbitofrontal
	Medial Orbitofrontal
	Precentral
	Paracentral
	Frontal Pole
Parietal Lobe	Superior Parietal
	Inferior Parietal
	Supramarginal
	Postcentral
	Precuneus
Temporal Lobe	Superior Temporal
	Middle Temporal
	Inferior Temporal
	Banks of the Superior Temporal Sulcus (STS)
	Fusiform, Transverse Temporal
	Entorhinal
	Temporal Pole
	Parahippocampal
Occipital Lobe	Lateral Occipital
	Lingual
	Cuneus
	Pericalcarine
Cingulate Cortex	Rostral Anterior Cingulate (Frontal)
	Caudal Anterior Cingulate (Frontal)
	Posterior Cingulate (Parietal)
	Isthmus of the Cingulate (Parietal)
Insula	Insula

duced, the Desikan-Killiany-Tourville (DKT) parcellation, that is a modified, more conservative scheme that includes 31 regions per hemisphere compared to the original 34 [46]. The Desikan-Killiany cortical atlas, by successfully bridging the methodological gap between highly rigorous, anatomically valid manual tracing and high-throughput, computationally scalable probabilistic automation, systematically resolves one of the most critical bottlenecks in macroscopic structural brain imaging. Its deep integration into the FreeSurfer processing pipeline provides researchers with unparalleled spatial consistency, statistical power and generalizability across different clinical populations. As network neuroscience progresses, the Desikan-Killiany atlas ensures that modern, multi-modal human brain mapping remains fundamentally grounded in a robust, standardized and reproducible neuroanatomical coordinate framework.

2.3 Imaging Transcriptomics

For decades, noninvasive neuroimaging modalities, such as magnetic resonance imaging (MRI) and positron emission tomography (PET), have been the foundation for the exploration of the human brain in vivo. These tools have provided extensive anatomical characterization of cortical atrophy and functional network disruptions both in healthy aging and across a different array of neuropsychiatric and neurodegenerative disorders[47]. However, a fundamental limitation of traditional neuroimaging remains its macroscopic resolution, which only offers an indirect proxy of the underlying microscopic and cellular pathology[48]. In order to overcome this inferential gap, the field of imaging transcriptomics has emerged. Imaging transcriptomics establishes a quantitative, multiscale framework that maps spatially resolved gene expression profiles onto neuroimaging-derived phenotypes, within a shared stereotaxic space [3]. This approach enables researchers to look into the way microscopic molecular pathways shape the macroscale structural and functional architecture of the brain, as well as how regional vulnerabilities manifest in disease [47]. The whole field of imaging transcriptomics heavily relies on the existence of anatomically comprehensive, high-throughput transcriptomics databases, with the Allen Human Brain Atlas (AHBA) emerging as the most extensively used one [49]. It provides a high-resolution, spatial map of the adult human brain transcriptome, integrating microarray expression data for over 20,000 genes sampled from 3,702 spatially localized tissue structures across six post-mortem donor brains. Considering that the donor brains underwent ex-vivo MRI prior to tissue dissection, the precise stereotaxic coordinates of each microarray sample can be mapped directly to standardized neuroimaging parcellations[3]. While the AHBA is unparalleled in its spatial density, it is limited by its small sample size (six total donors), adult-only demographic and left-hemisphere focused data. Consequently, other databases can be used, such as BrainSpan[50] to track gene expression across development or PsychENCODE[51] to explore

expression data specific to psychiatric conditions. Single-cell RNA sequencing (scRNA-seq) datasets are also becoming increasingly integrated in order to determine the neuronal cell-type specificity of the genes identified in brain-wide spatial analyses[52]. The process of translating raw AHBA microarray samples into a standardized region-by-gene expression matrix for neuroimaging integration is complex, computationally demanding and highly sensitive to methodological choices [53].

The standard processing pipeline generally requires six critical decision phases: (1) verifying and updating probe-to-gene annotations; (2) filtering probes to remove background noise; (3) selecting a representative probe when multiple probes map to the same gene; (4) assigning tissue samples to parcellated brain regions using distance thresholds; (5) normalizing expression measures across genes and samples to account for batch effects and substantial inter-donor variability; (6) finalizing the consistent gene matrix [54].

Since each one of these steps requires the selection of some parameters, researchers can theoretically generate hundreds of thousands of distinct processing pipelines, leading to issues in reproducibility. To address this, some standardized and open source toolboxes were developed, most notably abagen [53], which streamlines AHBA data processing, standardizes default parameters and ensures results reproducibility. Once a parcellated region-by-gene matrix is created, statistical associations with neuroimaging metrics are computed through three main strategies: Regional Gene Expression (RGE), which correlates local gene profiles with imaging traits using correlations or regression models; Correlated Gene Expression (CGE), which links region-to-region transcriptional similarity with brain connectivity networks; and Gene Co-expression (GCE), which clusters genes into spatial modules to map onto regional imaging variations [3].

Another fundamental challenge in the field of imaging transcriptomics is the presence of spatial autocorrelation in both neuroimaging maps and gene expression atlases. Brain regions that are physically close tend to share morphological features, functional properties, and transcriptional profiles, with gene expression correlation decaying exponentially as physical separation increases[54].

As traditional parametric statistical tests assume the independence of observations, applying them directly to spatially autocorrelated data severely inflates type I error (false positive) rates, yielding overly optimistic estimates of gene-imaging associations. To ensure rigorous statistical inference, researchers must then deploy spatially constrained null models, of which there are two primary classes: spatial permutation models (spin tests) and parametrized spatial null models. The former consist in projecting parcellated cortical maps onto a sphere and applying random rotations before mapping them back onto the original surface, while the latter involve the generation of null maps with randomized topography and the original spatial autocorrelation structure.

Once significantly associated genes are identified using spatial null models, they undergo Gene Category Enrichment Analysis (GCEA) to determine biological mechanisms. This process can be done with either Gene Set Enrichment Analysis (GSEA) or Over Representation Analysis (ORA). These analyses rely on a reference gene library, most commonly drawing from Gene Ontology, established biological pathways, or cell-type signatures. Standard GCEA algorithms also suffer from false-positive inflation due to genes co-expression, leading to a need for testing against phenotype-based spatial nulls [48][54].

In neurobiology, imaging transcriptomics has been able to link macroscale brain architecture to its biological origins, mapping regional gene expression directly to cortical hierarchies[55]. Furthermore, this approach validates in vivo scans by matching MRI marker distributions to their target gene expression signatures, for example aligning MRI myelination markers with oligodendrocytes signatures[47].

Macroscopic structural deviations observed in psychiatric conditions also strongly reflect microscale vulnerabilities. In schizophrenia gray matter volume and connectivity changes correlate with genes regulating synapses and neurodevelopment [56], while in Major Depressive Disorder (MDD) and Autism Spectrum Disorder (ASD) differences in cortical thickness and morphometric similarity have been linked to the dysregulation of genes related to immune response and apoptosis [57][58]. In neurodegenerative diseases like Alzheimer's and Parkinson's disease, imaging transcriptomics not only supports the network diffusion model of the pathology, but also correlates regional vulnerability to progressive atrophy to the spatial expression of genes mediating synaptic protein propagation. For example, the regional expression of the MAPT gene (which encodes the tau protein) predicts the progressive loss of network hubs in PD, while immune-related risk genes dictate the epicenter and initiation sites of neurodegeneration [59] [60] [61][62].

Imaging Transcriptomics faces distinct limitations, despite its transformative impact. Because of its reliance on the Allen Human Brain Atlas, it suffers from the limited sample size and heterogeneity of the available adult brains. In addition, bulk microarray tissue sampling obscures precise single-cell granularity, rendering algorithmic cell-deconvolution a necessary step [63].

To overcome these limitations, the field is advancing toward unified spatial imaging-transcriptomics. Emerging paradigms now physically integrate ultra-high-resolution ex vivo MRI with next-generation Spatially Resolved Transcriptomics (ST) technologies derived from the exact same brain tissue. Using sophisticated non-linear registration pipelines, these frameworks establish direct spot-to-voxel correspondence [64]. By directly co-registering neuroimaging voxels with high-resolution transcriptomic spots, these next-generation paradigms eliminate the need for cross-cohort atlas inference, offering an unprecedented, unified bridge to decode

the molecular etiology of brain structure and disease.

Chapter 3

Materials and Methods

3.1 Study participants: the PPMI dataset

The methodological framework proposed in this work was tested on MRI data obtained from the Parkinson's Progression Markers Initiative (PPMI), with both healthy controls and Parkinson's disease patients included in the evaluation. The Parkinson's Progression Markers Initiative is an observational, longitudinal and multicenter study aimed at identifying PD progression biomarkers in order to accelerate therapy development and improve our understanding of the disease [65]. The complete PPMI dataset is comprised of data derived from different types of cohorts, ranging from early-onset PD to late-stage PD, prodromal and SWEDD (scans without evidence for dopaminergic deficit), with healthy controls usually matched accordingly. Enrollment typically occurs at the earliest detectable stages of the disease[66]. To help the discovery of possible new and effective biomarkers, several types of data are included together: clinical data, neuroimaging data and biological specimens. Upon enrollment and during follow-up, all participants of the study undergo DaTSCAN (Dopamine Transporter Scan) and MRI scans acquisitions, blood, urine, CSF and genetic specimen collection, and clinical assessment including motor, cognitive, sleep and olfactory testing [67]. Ultimately, biomarker experts can leverage this comprehensive data to investigate new potential links and underlying mechanisms that could help advance our knowledge in disease etiology and progression [66].

For this work, the first step consisted in downloading clinical and imaging biomarkers from the PPMI website. The retrieved sample included baseline data for PD patients and healthy controls, with a total of number of N = 953 subjects retrieved for further screening. After the acquisition step was performed, participants' data were subjected to an additional phase of selection. Inclusion in the final dataset required high-quality T1-weighted MRI scans acquired on a 3 Tesla (3T) scanner. To mitigate potential confounding effects arising from hardware heterogeneity [68], scanners with low subject representation were excluded from the analysis, and

Group	n	Age (Mean ± SD)	Sex (M/F)
HC	106	60.1 ± 10.7	M = 66, F = 40
PD	106	61.2 ± 5.9	M = 66, F = 40

Table 3.1: Demographic breakdown of the final dataset utilized in this study.

only the most used ones across the entire cohort were kept to keep variance under control.

The final sample comprised n = 212 subjects, further partitioned in 106 healthy controls (HC) and 106 PD patients. The two groups were approximately matched for age and sex. The full composition of the dataset is shown in Table 3.1, while details regarding the scanner manufacturer, specific model and subject number allocations are reported in Table 3.2.

	SIEMENS				
	Prisma	Prisma_fit	Verio	Skyra	TrioTim
Subjects	13	9	32	14	144

Table 3.2: Scanner manufacturer, models and subject distribution inside the final dataset.

3.2 Methodological Framework

The full methodological framework, explained in detail in this section, is presented in Figure 3.1.

3.2.1 Data preprocessing

Regional morphometric data were extracted from each structural MRI scan using a FreeSurfer preprocessing pipeline (FreeSurfer v7.3.2) [45]. FreeSurfer is an open-source neuroimaging software suite for automated processing, analysis and visualization of human brain MR images, widely established for cortical surface reconstruction and inspection [69]. FreeSurfer's `recon-all` command, which runs the standard cortical reconstruction and parcellation pipeline, was used to obtain for each subject the cortical parcellation statistics files for both brain hemispheres: `lh.aparc.stats` and `rh.aparc.stats`. The resulting regional statistics were retrieved as numerical data for each one of the 68 regions of the Desikan-Killiany Atlas. The full set of recovered features included: Surface Area (in mm^2), Gray Matter Volume (in mm^3), Average Cortical Thickness (in mm), Thickness Standard Deviation (mm), Integrated Rectified Mean Curvature (in mm^{-1}), Integrated Rectified Gaussian Curvature (in mm^{-2}), Folding Index (unitless) and Intrinsic Curvature Index (unitless). In addition, the Estimated Total Intracranial

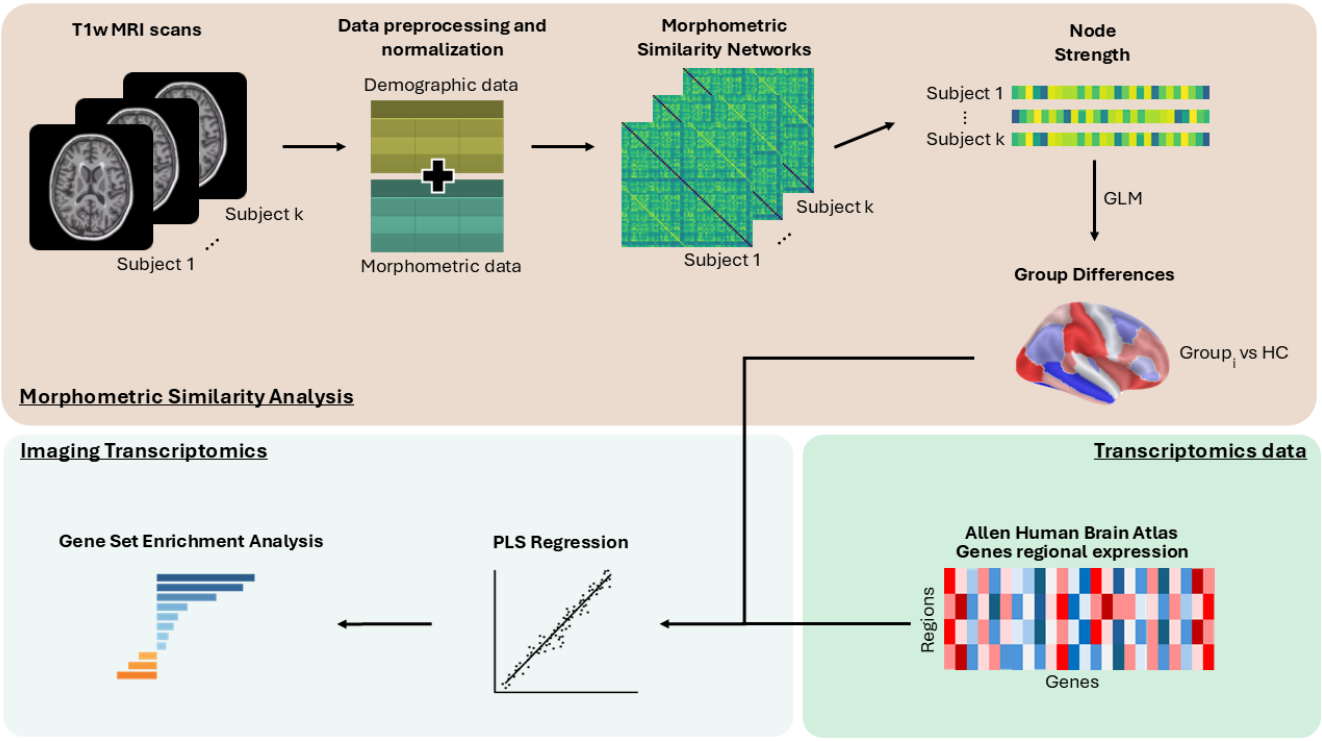


Figure 3.1: Full pipeline used in this work, from the construction of Morphometric Similarity Networks to the connection with Allen Human Brain Atlas transcriptomics data to perform enrichment analysis and bridge micro and macroscale information.

Volume (eTIV), another value measured in mm^3, automatically computed by the FreeSurfer pipeline and characteristic of each subject's brain, was extracted for each individual and provided in those same files. Of the 8 total cortical morphometric features obtained from the T1w MRI scans, only 5 were retained, as done in previous studies[70]: Surface Area, Gray Matter Volume, Average Cortical Thickness, Integrated Rectified Mean Curvature and Integrated Rectified Gaussian Curvature. Morphometric data were then merged with subjects' demographic data, including their assigned clinical group, age, sex and model of MRI scanner for subsequent analyses.

3.2.2 Data Normalization

Once the data were retrieved, normalization was needed to mitigate the influence of extreme outliers and account for the inherent physiological differences between individuals.
To do this, normalization was performed within-patient, with a robust Z-score approach and the formula:

$$Z = \frac{0.6745 \cdot (X_i - \text{Median})}{\text{MAD}} \quad where\ MAD = median\ (|X - median|) \tag{3.1}$$

where the 0.6745 multiplication factor makes it comparable to a standard Z-score, X_i represents the numerical values for each feature i and MAD corresponds to the Mean Absolute Deviation of each single feature. The usage of the median and MAD, instead of the usual mean and standard deviation values, makes this normalization approach more robust and less sensitive to extreme anomalies in the data [71].
After this step, the resulting normalized features become similar in scale, allowing for direct comparisons between them and higher stability.

3.2.3 Mass Univariate Analysis

Following the within-patient normalization performed with a robust Z-score approach, an exploratory mass univariate analysis was conducted before creating the morphometric similarity networks, to assess regional group differences between the PD and HC cohorts. This analysis independently evaluated the five selected morphometric features across all 68 cortical regions. Group comparisons were performed using a General Linear Model (GLM), adopting the model design outlined in Section 3.2.5, with the exception that each feature served as the dependent variable instead of node strength. To control for Type I errors across the multiple tests, statistical significance was evaluated after applying False Discovery Rate (FDR) correction. Of the model results, t-values were extracted to allow visualization of deviation patterns within the cohorts.

3.2.4 Morphometric Similarity Networks

Proceeding from the exploratory analysis, morphometric similarity networks were created for each individual to quantify the complex structural similarity relationships between different anatomical regions. For a given patient, each predefined ROI was characterized by a specific 5-feature vector, which included the normalized values for each of the five chosen metrics.
To evaluate the divergence between any given pair of regions x and y, the Multivariate Euclidean Distance, $d(x, y)$, was calculated across the feature space to obtain distance information between the 68 regions making up the Desikan Killiany Atlas. The following formula was used:

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{3.2}$$

where x_i and y_i represent the values of morphometric feature i at regions x and y, and n corresponds to the total number of areas included in the parcellation atlas, which in the Desikan-Killiany case is equal to 68.
The use of multivariate Euclidean distance was motivated by its greater information retention compared to other commonly employed measures like correlations[36]. Additionally, Euclidean

distance was considered more suitable and robust given the limited dimensionality of the feature space (n=5), which would otherwise lead to statistical instability and insufficient degrees of freedom for correlation-based measures. Since the Euclidean distance acts as a measure of structural dissimilarity, where larger values denote highly divergent morphological profiles, the resulting distances were systematically converted into similarity scores, $S(x, y)$, before constructing the MSNs.

This was achieved through the application of a standard inverse transformation, where the distance $d(x, y)$ was also normalized by the number of features, through the formula:

$$S(x, y) = \frac{1}{1 + \frac{d(x,y)}{\#metrics}} \tag{3.3}$$

The result of this transformation is a similarity weight bounded between 0 and 1, where values approaching 1 indicate that two regions possess matching morphometric profiles and values closer to 0 point to two highly dissimilar regions in morphometric terms. Finally, the resulting pairwise similarity scores were aggregated into a symmetric 68×68 similarity matrix, which maps effectively the patient-specific MSN, yielding a weigthed, undirected graph that serves as the foundation for the following network analysis.

To characterize the topological properties of the generated networks, graph analysis was applied to the MSNs. The metric used in this work was node strength, which serves as the weighted equivalent of degree centrality in complex networks and was already employed in other studies using MSNs[39][38]. While unweighted degree centrality merely counts the number of connections a node possesses, node strength provides a more nuanced measure by accounting for the magnitude of those connections.

For each specific ROI x, which represents a node of the graph, its node strength $Strength(x)$ was mathematically defined as the sum of all its incident edge weights, which correspond to morphometric similarity scores in this work. This was computed through the following equation:

$$Strength(x) = \sum_{x \neq y} S(x, y) \tag{3.4}$$

where $S(x,y)$ corresponds to the similarity score between regions x and y. Self similarity weights are excluded from the sum as they all present a similarity score of 1 and are thus irrelevant in the overall calculation.

Node strength in the context of MSNs holds biological relevance, as a region exhibiting a high node strength shares a highly concordant morphometric profile with a large number of other anatomical areas. This similarity pattern allows the identification of highly integrated structural hubs within the patient's overall cortical structure[2]. On the other hand, the regions exhibiting

a low node strength score usually equate to areas with more morphologically distinct or isolated profiles compared to the rest of the structural connectome.

3.2.5 Group Differences Analysis

To evaluate alterations in the morphometric network architecture associated with the disease cohort, their regional node strength values were statistically compared with the healthy controls. As demographic factors have a strong influence on structural brain metrics [72] and are also sometimes correlated to neurodegenerative diseases' progression or outcome[73], a General Linear Model was employed to assess group-level differences while systematically controlling for these potential confounders. For each region of interest, the node strength of a subject was modeled as the dependent variable. The clinical diagnosis (PD compared to HC) was entered as the primary categorical independent variable of interest, alongside age, sex, eTIV and scanner model as covariates. The scanner model in particular was included through one-hot encoding, in order to reduce the potential noise and variance added by the multiple scanners[74].
The GLM was formulated as follows:

$$Strength_i = \beta_0 + \beta_1 \text{Group}_i + \beta_2 \text{Age}_i + \beta_3 \text{Sex}_i + \beta_4 \text{eTIV}_i + \beta_5 \text{ScannerModel}_i + \epsilon_i \quad (3.5)$$

where β_0 represents the intercept, and the coefficients $\beta_1, \beta_2, \beta_3, \beta_4$ and β_5 quantify the independent effects of diagnostic group, age, sex, eTIV and scanner model, respectively. The residual error term is denoted by ϵ_i.
In this framework, the primary coefficient of clinical interest is β_1. This parameter captures the main effect of the pathology, representing the magnitude and direction of the change in regional node strength driven specifically by the PD diagnosis, independent of the variance attributable to the subject's included confounder variables.
To rigorously account for the inflated risk of Type I errors arising from simultaneous statistical testing across all 68 regions in the network, the resulting p-values were corrected for multiple comparison using the Benjamini-Hochberg False Discovery Rate (FDR) procedure. Group-level differences in regional node strength were considered statistically significant at an FDR-adjusted threshold of $p < 0.05$, although some considerations could still be made about uncorrected p-values < 0.05, as it is common practice. Of model results, beta coefficients, t-statistics and corrected p-values were kept for the following analyses and visualizations, obtaining for each of these variables a vector of 68 values. The results were visualized as bar plots of t-values, divided by left and right hemisphere, in order to see at a quick glance the amount of changes in nodal strength.
Of all the retained results, in particular, the vector of beta coefficients, as it represented the ef-

fect size of the node strength change between cohorts, was used as the phenotype map for the subsequent imaging transcriptomics analysis.

3.2.6 Imaging Transcriptomics

To investigate the potential molecular and cellular correlates underlying the observed macroscale network alterations, an imaging transcriptomics step was implemented. To do this, two main actions are needed: a first step where a phenotype map is cross-referenced with the spatial information coming from post-mortem microarray gene expression data, and a second step where the just obtained gene rankings serve as input for a gene set enrichment analysis to uncover molecular pathways potentially related to the disease cohort morphometric alterations.
The primary input for this analysis was the 68-dimensional vector of unstandardized β_1 coefficients derived from the group-level GLM, which served as the phenotypic input map. This vector effectively acts as a spatial map, representing the regional magnitude and direction of PD-related changes in node strength across all the anatomical regions. The microarray gene expression data used in this work is derived from the Allen Human Brain Atlas[49], and by using the Imaging Transcriptomics Toolbox[75] to perform the first step, it is also implicitly preprocessed in the same way as in Martins et al. (2021)[47]. The final number of genes from the AHBA that passed the preprocessing and filtering steps with abagen[53], and were compared to the strength-differences spatial map was 15,633.

To identify specific genes whose spatial expression patterns covaried with the PD-related network topological changes, the first half of the vector of regional β_1 coefficients was statistically associated with the regional gene expression matrix. In this case, the first half of the array corresponded to the left hemisphere data. This was done because of AHBA's inherent limitation: cortical data from the right hemisphere derives from just 2 post-mortem subjects, thus it is preferable to use only the left hemisphere information for this type of analyses[75].
Instead of the widely used simple spatial correlations, a Partial Least Squares (PLS) regression was employed, consistent with recent methodological approaches[10][70], to identify a weighted linear combination of genes that maximally covaried with the regional β map. Pearson or Spearman correlations, on the other hand, only test gene indipendently, not accounting for inter-gene relationships.
Only results for the first PLS component, the one explaining the maximum covariance between the two maps, were retained for subsequent use. With the PLS analysis, a value of cumulative variance was provided by the Imaging Transcriptomics Toolbox, together with a permutation based p-value that had the goal of assessing how significant the original spatial phenotype association with gene expression was compared to null maps. From PLS regression, a ranked list of all the genes was obtained, with each gene possessing a Z-Score value and a p-value. These in-

dicate the magnitude of spatial covariance between the phenotype and individual transcriptomic maps and their statistical significance. Because adjacent brain regions inherently share similar structural and transcriptomic profiles, the statistical significance of the imaging-transcriptomic associations was evaluated using spatial permutation testing, more specifically using the Váša algorithm[76] for spin testing, which generates null models by randomly rotating cortical maps on a sphere while preserving their spatial autocorrelation.

To determine the biological significance of the genes spatially associated with PD-related network alterations, functional and cell-type enrichment analyses were then performed on the whole ranked gene list. Rather than analyzing individual genes in isolation, this approach identifies whether specific biological pathways or cellular populations are overrepresented within the transcriptomic findings.

Different gene libraries were tested in this work: GO Biological Process[77], to observe which biological goals or outcomes a gene contributes to, KEGG Pathways (Kyoto Encyclopedia of Genes and Genomes)[78], to look into metabolic reactions and interaction networks, DisGeNET [79], for extensive gene-disease associations, and the neuronal Cell-Type library[80], to match gene signatures to specific adult human brain cell data obtained from single-cell RNA sequencing (scRNA-seq).

The statistical significance estimation of all the enrichment analyses was evaluated automatically by the *gseapy* python package using permutation testing and then multiple comparisons correction, keeping the threshold at FDR $p < 0.05$. Each enrichment analysis yielded a finalized list of terms, each with their own enrichment score, normalized enrichment score, FDR p-val and leading genes. These results were subsequently used to identify and interpret the molecular pathways and biological processes underlying the observed imaging–transcriptomic associations.

Alongside the aforementioned steps, the pipeline generates a series of graphical outputs. Specifically, it generates plots detailing the morphometric matrices, strength vectors, and maps of strength differences both as bar plots and overlapped on cortical surfaces. It also outputs tables consisting of the top covarying positive and negative PLS1 genes, as well as the most significantly enriched functional terms based on their normalized enrichment scores. Finally, a PDF report is generated, with the results from all the previous phases and the figures included.

3.2.7 Code and Implementation

The analysis pipeline was developed and used in a Python 3.9 environment. The Imaging Transcriptomics Toolbox[75] was employed to perform the Partial Least Squares (PLS) regression and its associated spatial permutations to generate gene rankings. Subsequently, the gseapy package[81] was utilized to conduct gene set enrichment analysis on the resulting ranked lists. To visualize the neuroimaging maps on cortical surfaces, the Nilearn package was

used[82]. All custom code written and used in this study is publicly available on GitHub at `https://github.com/tfraa`.

Chapter 4

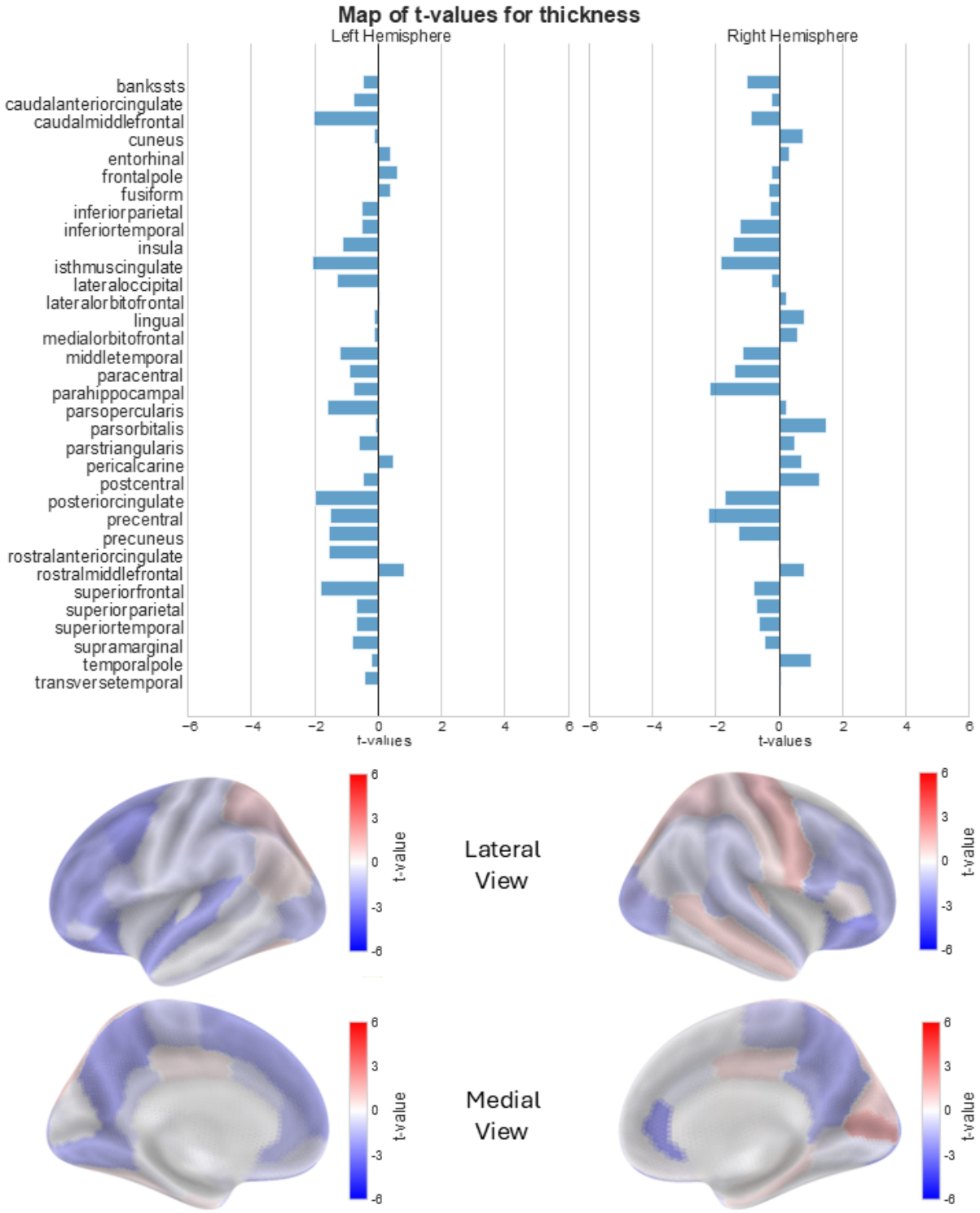
Results

4.1 Morphometric Similarity Networks

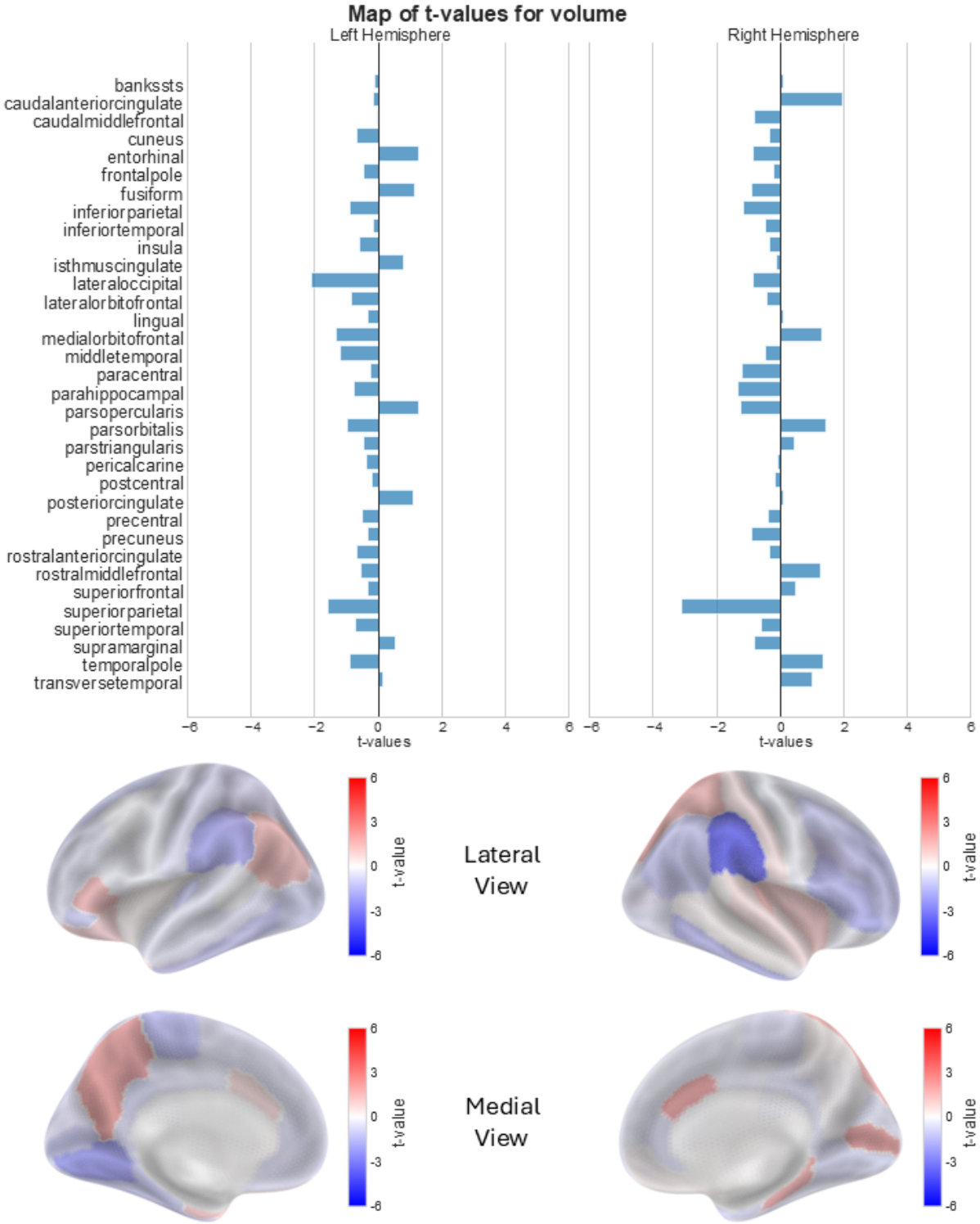
Once the preprocessing step was performed, before the computation of the morphometric similarity networks, an explorative mass univariate group difference analysis was conducted between the PD cohort and the HC group for each of the 5 selected morphometric features. None of the features in none of the 68 regions showed significant change after FDR correction. However, the deviation patterns, obtained with the same GLM used in the main steps of the pipeline and represented by t-values, were consistent with neurodegeneration, as the PD group brought general decreases in surface area, cortical thickness and gray matter volume, and an increase in the two metrics related to curvature. In Figure 4.1 the t-value maps can be visualized both as bar plots and projected on cortical surfaces.

After this first step, the morphometric similarity matrices were created. As shown in Figure 4.2, which represents the average per-group morphometric similarity matrices, upon a first visual inspection they look almost identical, with few small, almost imperceptible and localized changes. However, at the individual level, certain MSNs exhibit areas of reduced similarity compared to the average healthy individual.

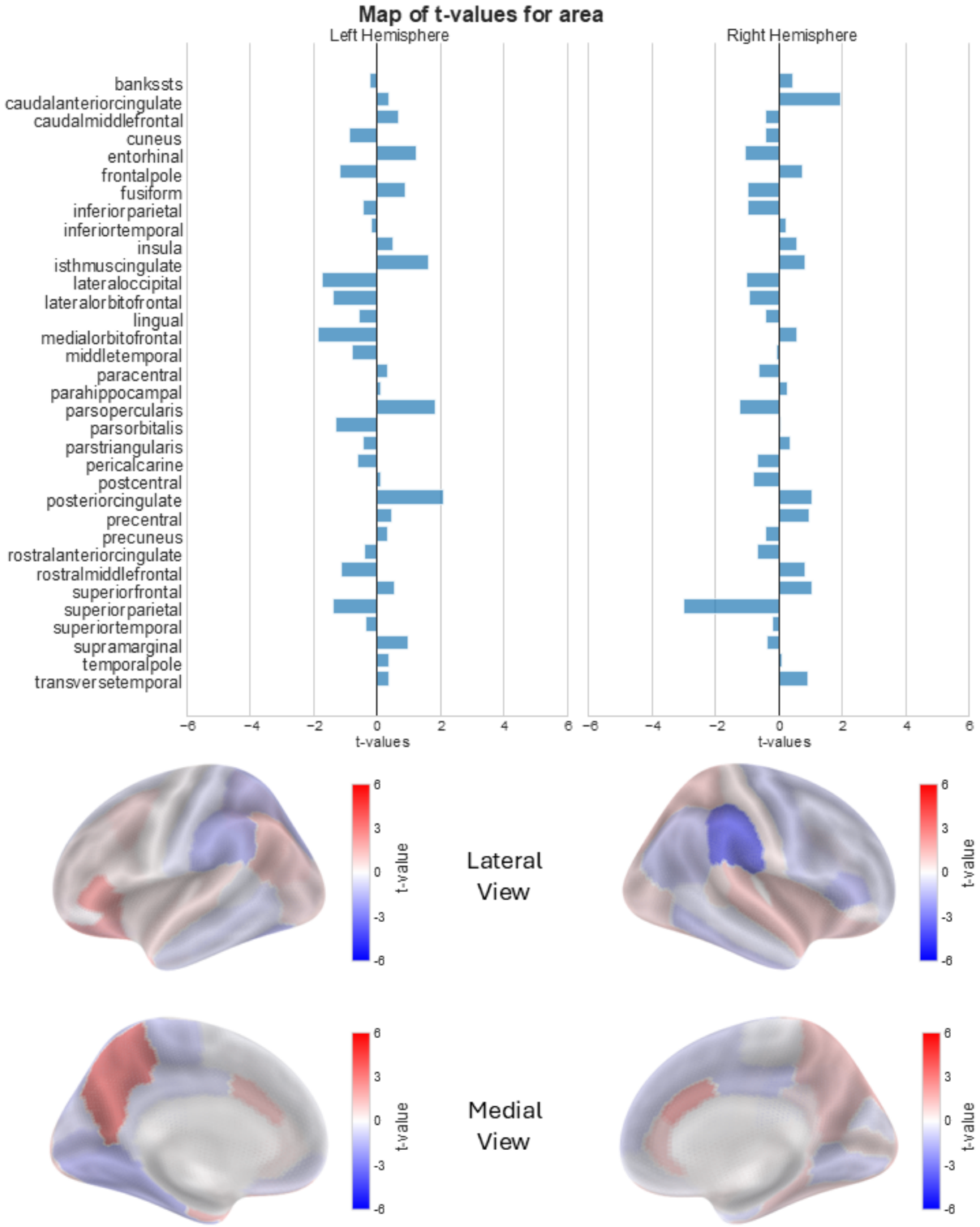
The same fate applied to the topological analysis with node strength, where on average virtually no differences were present, as can also be seen from Figure 4.3, which illustrates the strength values for all the 68 regions for each cohort, averaged across all subjects. In this case, just like before, when looking at individual node strength maps, the differences were minimal, with few localized and subtle changes for some of the subjects.



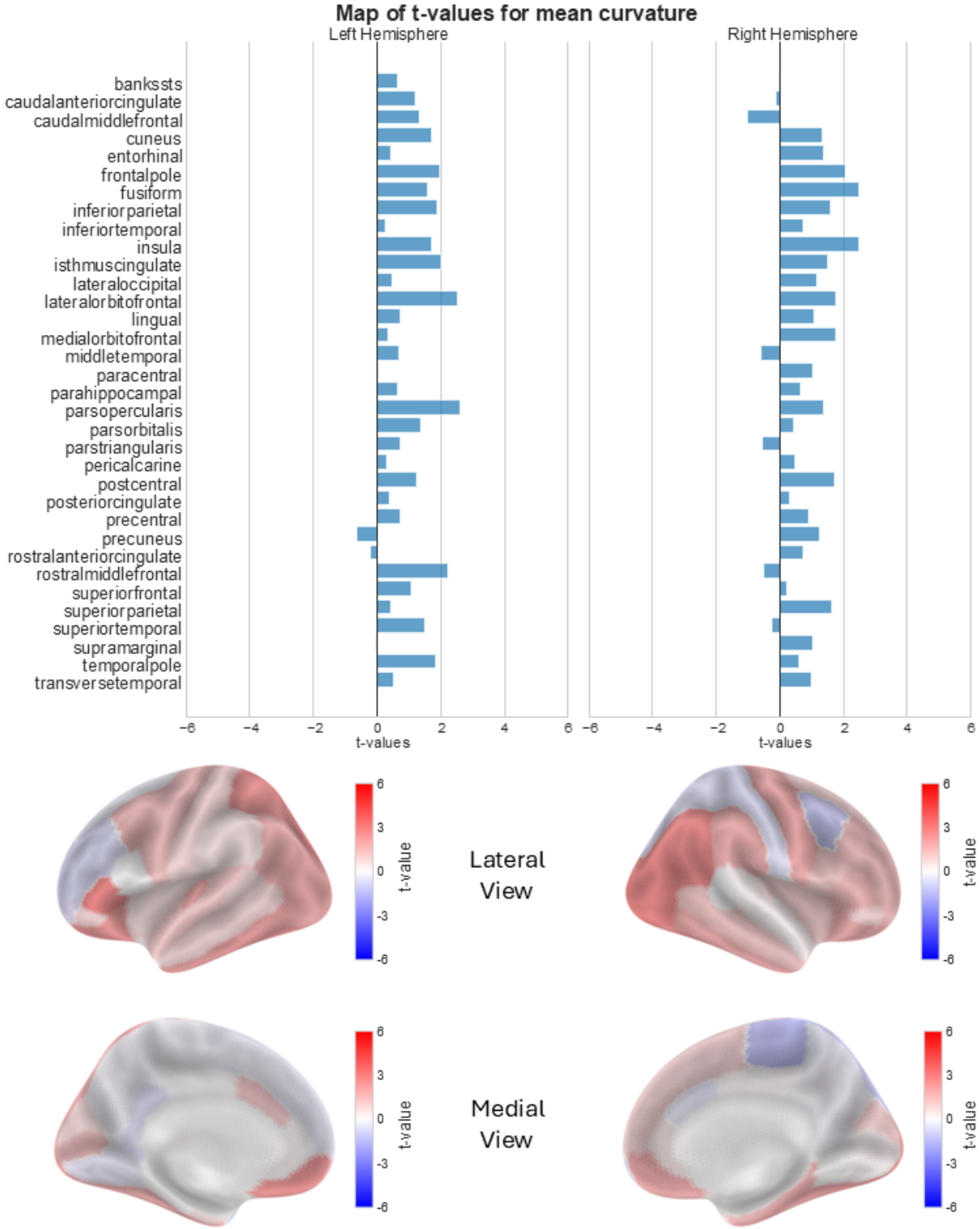
(a) Average cortical thickness.



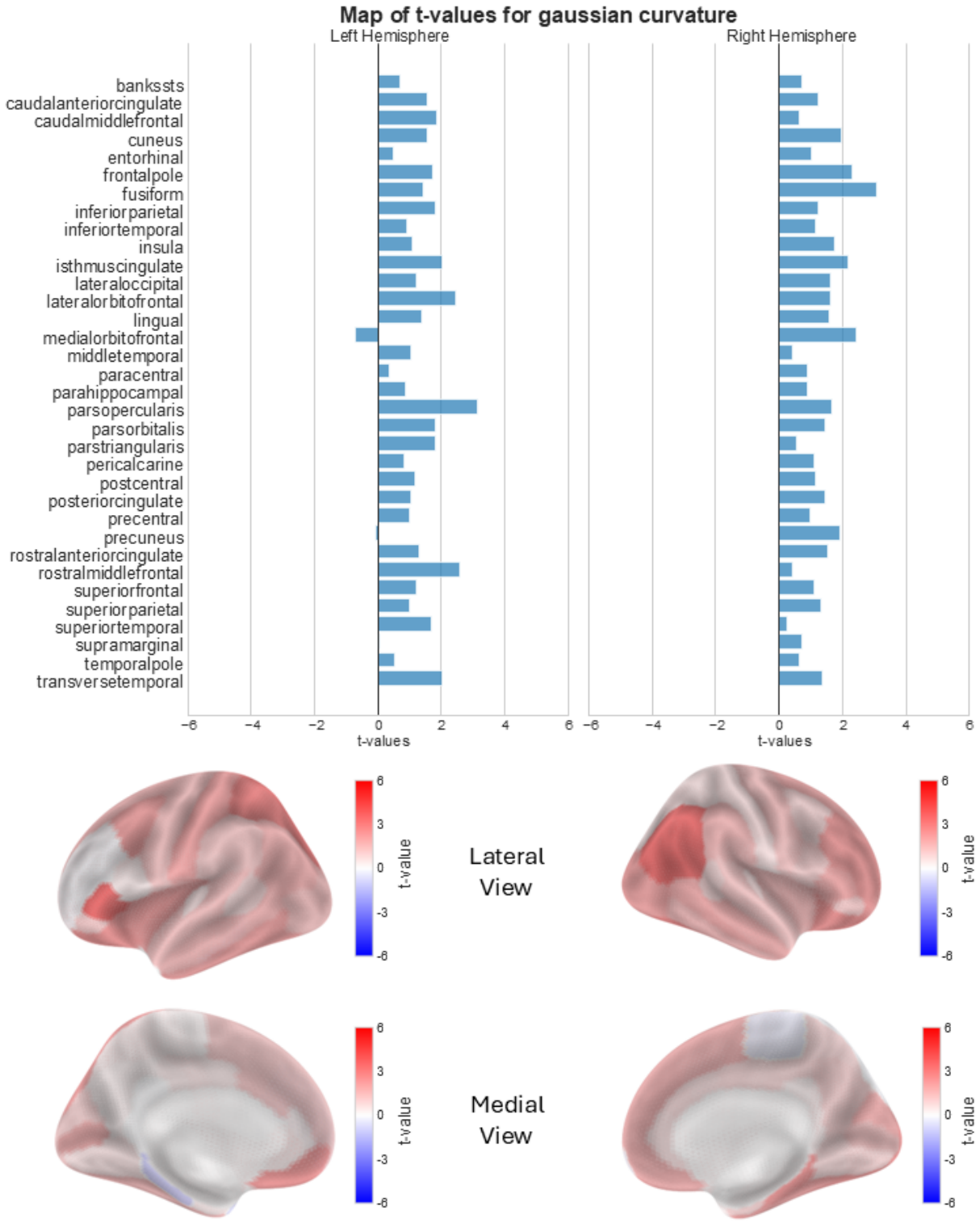
(b) Cortical volume.



(c) Cortical surface area.



(d) Mean Curvature.



(e) Gaussian Curvature.

Figure 4.1: Univariate analysis for the 5 morphometric features between PD and HC.

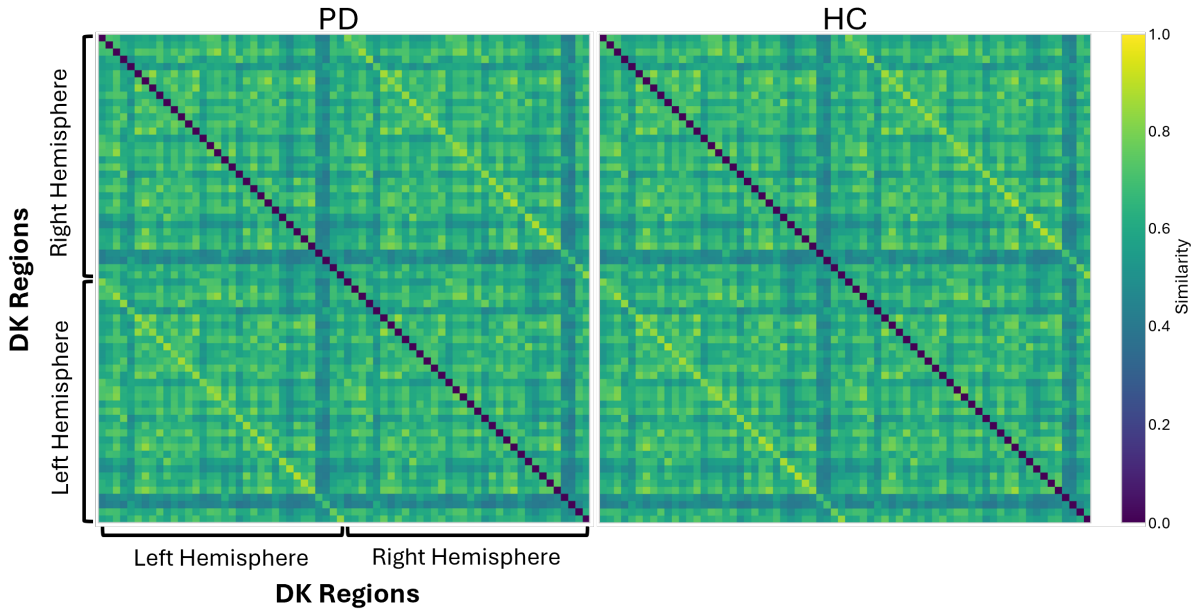


Figure 4.2: Averaged morphometric similarity matrices for each cohort, PD and HC, showing no visually distinguishable differences. Regions are ordered according to FreeSurfer's default regional ordering.

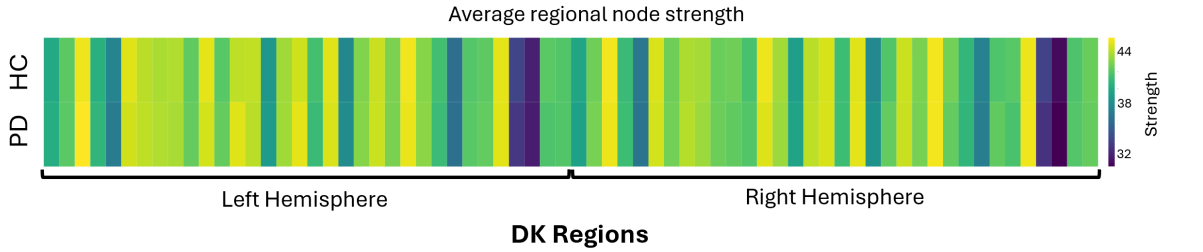


Figure 4.3: Averaged nodal strength across the 68 brain regions for each of the two clinical groups, PD and HC. As for the morphometric similarity matrices, the differences are here also imperceptible. Regions are ordered according to FreeSurfer's default regional ordering.

4.2 Group-Level differences

To quantify the extent of the differences in node strength between groups, the GLM was employed, systematically controlling for age, sex, estimated total intracranial volume and scanner model, as described in the Materials and Methods section of this work. From the GLM results, the beta values for the scanner type were considered to investigate the amount of influence of this variable on the results, as well as beta coefficients, t-values, and p-values from the Group variable, to evaluate potential group differences. The beta coefficients for the scanner model predictor variable resulted in extremely low values, meaning the scanner type had little impact on the overall results and explained only a small amount of variance in the model.

Hemi	Region	Beta coefficient	t-value	p-value	FDR p-value
L	Caudal middle frontal gyrus	0.498	2.621	0.009	0.219
L	Medial orbito frontal cortex	1.019	2.772	0.006	0.219
L	Pars opercularis	0.528	2.612	0.010	0.219
R	Superior parietal	0.434	2.195	0.029	0.498

Table 4.1: Results of the generalized linear model for the brain regions exhibiting a significant difference in node strength, before FDR correction, with $p_{unc} < 0.05$.

Figure 4.4 illustrates the t-values extracted from the group difference analysis and the projections of this map on cortical surfaces, both on the lateral and medial view of the brain, with the use of the inflated representation of the cortex instead of the pial one to better distinguish the gyral and sulcal surfaces.

As can be seen from table 4.1, four main regions were indentified as the ones presenting a significant regional shift in nodal strength: the left caudal middle frontal gyrus, left medial orbitofrontal cortex, left pars opercularis, and the right superior parietal lobule. These results were achieved prior to multiple comparison correction ($p < 0.05$). Notably, all the significant areas belonging to the left hemisphere appear to belong to the frontal lobe.
After the computation of the node strength difference patterns, we visualized the behavior of the structural hubs, in order to investigate the potential presence of patterns like dedifferentiation, hypercoupling, hyperdifferentiation and hub decoupling. This analysis was conducted similarly to previous investigations[47], correlating MS group differences with mean MS in healthy controls.

As shown in Figure 4.5, patterns are uniform for most of the regions except the ones showing an already low morphometric similarity (MS) in healthy controls. The lower the MS are in controls, the more the distribution is skewed towards negative beta coefficients and thus hyperdifferentiation. The most frequent pattern was the hub decoupling, linked to a decrease of nodal strength in the PD cohort where it is usually high on average in HC, with a total of 32.4% of regions falling in this category. Following this, the second most frequent pattern was the hypercoupling, with 27.9% of regions included. Then differentiation with 20.6% and last hyperdifferentiation with only 19.1% of regions.

The unstandardized beta coefficients (β_1), representing the magnitude and direction of these disease-driven topological changes, were extracted from the model's group variable to finally serve as the spatial phenotype map for the subsequent transcriptomic profiling.

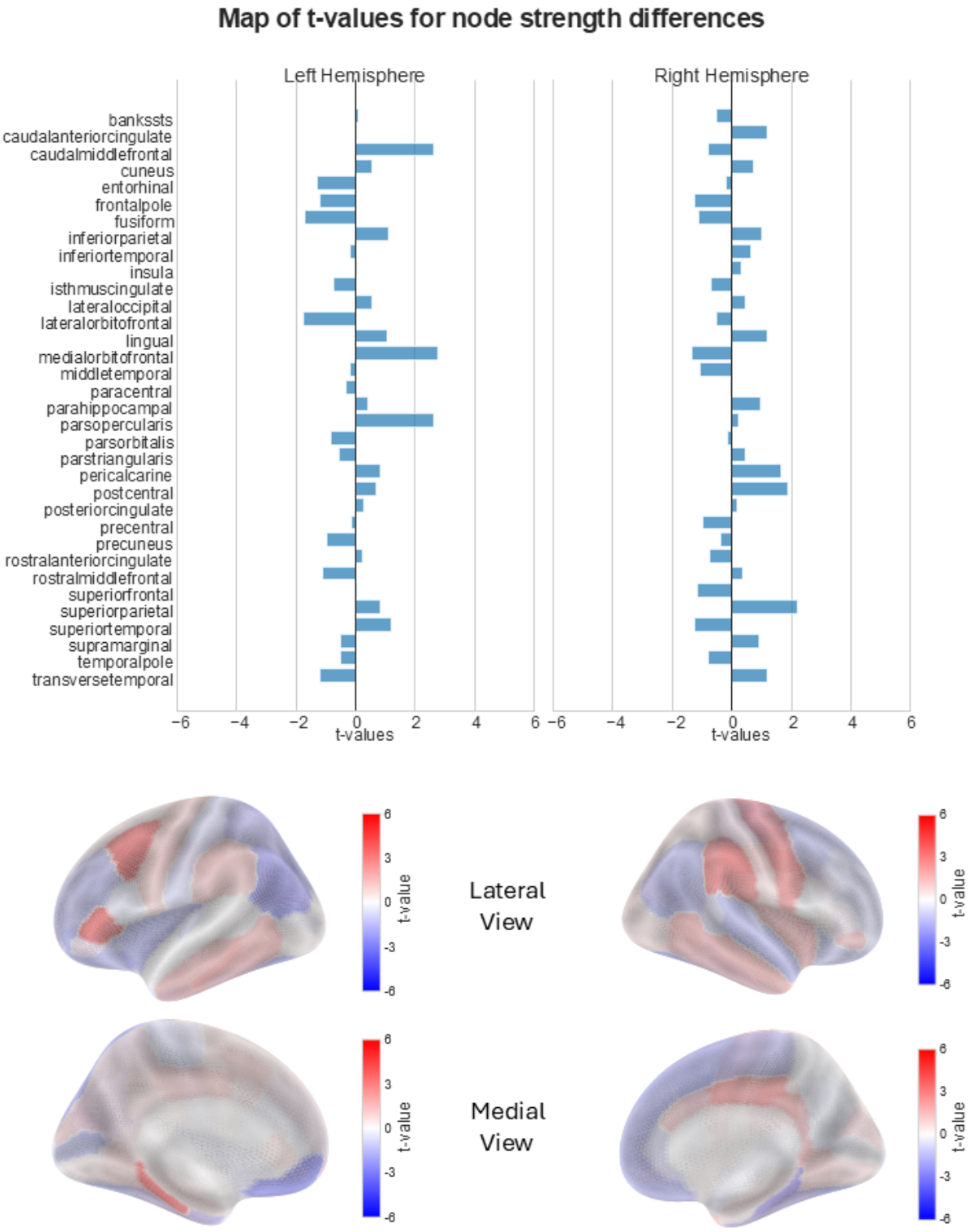


Figure 4.4: T-statistics values of the differences in nodal strength for each of the 68 brain regions, FDR corrected, after the application of the generalized linear model.

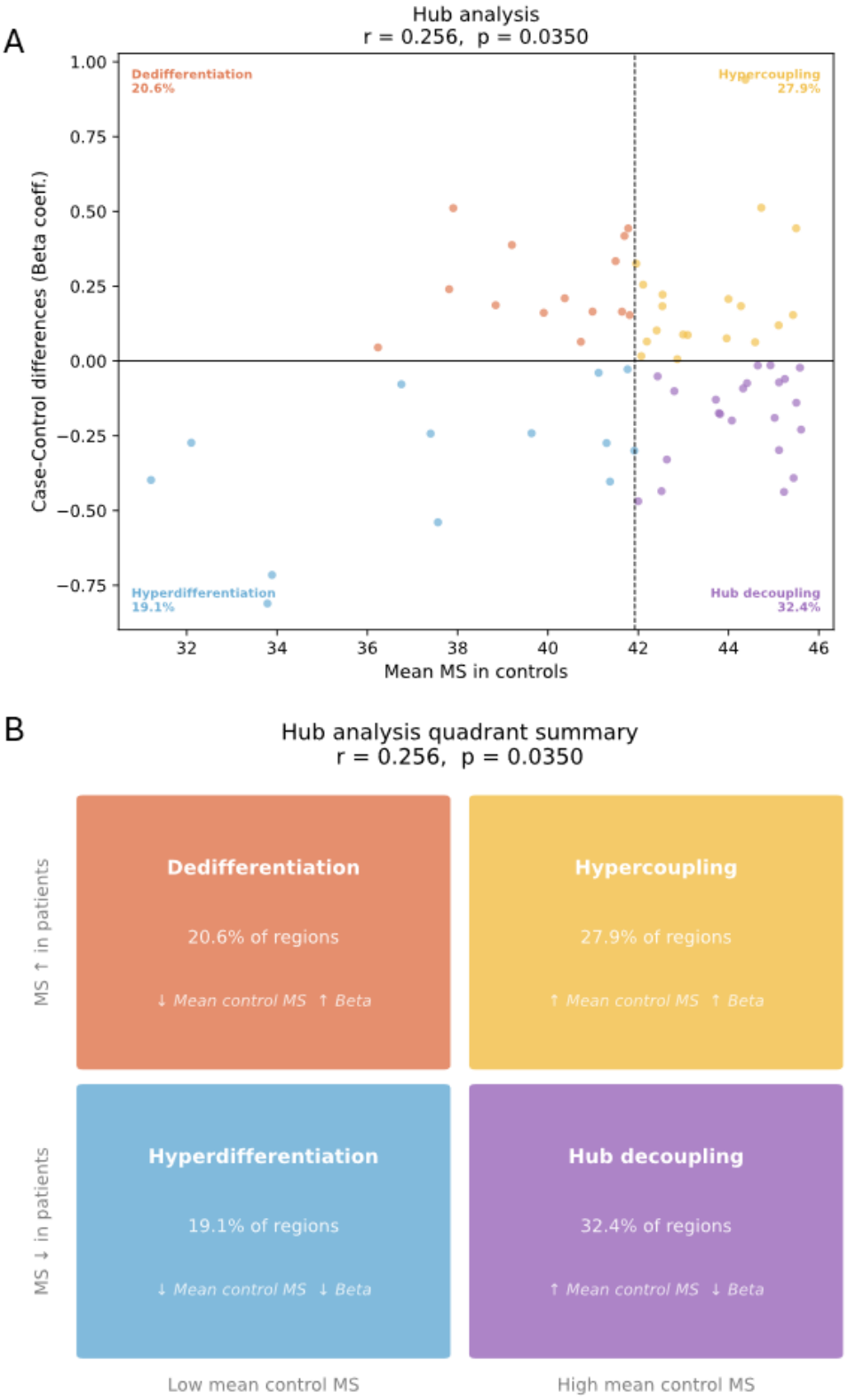


Figure 4.5: Hub susceptibility analysis: (A) correlation between mean values in healthy subjects and morphometric similarity is presented, together with the rate (B) for each of the possible result options.

4.3 Imaging Transcriptomics

To bridge these macroscale morphological alterations with their potential underlying molecular mechanisms, the spatial map of left-hemisphere beta coefficients was cross-referenced with microarray gene expression data from the Allen Human Brain Atlas (AHBA), where the 15633 genes that survived preprocessing and filtering had their expression spatially assigned to their respective brain regions. Partial Least Squares (PLS) regression was used to identify the first principal component of this weighted linear combination of genes that maximally covaried with the PD-related changes in node strength. Spatial permutation testing (spin tests) utilizing the Váša algorithm [83] was also used, in order to account for spatial relationships between close regions. As a result, the statistical significance of the first PLS component was p-value = 0.937, and it explained 0.114 of its covariance with the spatial phenotype map. These results, while not ideal, appear consistent with the previous steps' results, as no significant deviations from the healthy controls leads to spatial covariance maps that are very likely to be coming from chance or due to the spatial structure itself, rather than from a disease-related pattern.

Still, with the use of the PLS1 component, the PLS step finally generated a ranked list of all the 15633 genes each with a Z-score and a p-value corrected for multiple comparisons and subjected to spin testing spatial permutations, resulting in a total of 891 genes presenting an FDR corrected p-value lower than 0.05 and thus considered significant. The genes that showed the most positive and negative spatial covariance with the node strength map are presented in Table 4.2.

For each one of these top genes, their downstream function was investigated, and while none of them was specifically related to PD, they had a consistent connection to either system immunity, inflammation or cancer.

Inside this ranked list, we explored whether genes previously linked to Parkinson's disease were present and exhibited a high Z-score and thus spatial covariance with the phenotype map composed of nodal strength differences. The genes were obtained from previous studies that showed strong association between them and PD and included: SNCA, PRKN, PARK7, LRRK2, PINK1 and ATP13A2 [84][46][85]. Of these, only LRRK2 appeared to show a significant spatial covariance with our map, placing at position 381/15633 in the PLS1 ranked list, with a Z-Score = 1.702 and a FDR p-val = 0.025. Its high location in the gene list and thus positive Z-Score indicates that LRRK2 presents higher expression in areas where the node strength becomes higher, so where regions show a pattern of becoming more similar to each other.

The full resulting PLS1 gene ranking was subsequently subjected to functional and cell-type enrichment analyses to uncover the biological relevance of the transcriptomic signature. Gene Set Enrichment Analysis (GSEA) results for the GO Biological Process, KEGG, DisGeNET and Cell-Type libraries are presented in Figures 4.64.74.84.9, each with their top positively and neg-

Top 10 Positive Z-Scores		
Gene	**Z-Score**	**FDR p-value**
AAR2	2.711	0.0011
LOC494141	2.686	0.0011
TRAF7	2.643	0.0014
RGS11	2.598	0.0020
IQSEC3	2.527	0.0023
SOCS2	2.523	0.0024
VASH2	2.458	0.0027
GADD45A	2.438	0.0028
SUPT16H	2.401	0.0028
C3orf35	2.395	0.0035
Bottom 10 Negative Z-Scores		
Gene	**Z-Score**	**FDR p-value**
MPEG1	-2.904	0.0014
CYSLTR1	-2.713	0.0026
LOC145783	-2.709	0.0029
C1orf56	-2.694	0.0030
STAG3	-2.559	0.0035
OSTC	-2.520	0.0040
DPF3	-2.458	0.0044
CYB5A	-2.456	0.0045
NFYC-AS1	-2.431	0.0046
ADARB2	-2.413	0.0047

Table 4.2: Top positive and negative Z-Scores with their corrected p-value obtained from the Partial Least Squares regression first component.

atively enriched terms. Reported in the enrichment plots are the normalized enrichment scores together with the significance of each term. Only significant terms ($p_{FDR} < 0.05$) are shown.
For GO Biological Process, linked to biological and molecular processes, 30 terms mostly related to immunity and with predominantly negative normalized enrichment scores (NES) were present, meaning that the expression of this genes was higher were the nodal strength decreased.
For KeGG, related to biological and metabolic pathways, 42 total terms reached the enrichment significance threshold. Here, just like before, the results showed potential connections with immunity alterations and metabolism disruptions, and almost all of the normalized enrichment scores were negative.
For DisGeNET, used to link genes with specific diseases, 61 terms in total were enriched at the end of the analysis. As the previous gene libraries results already showed, here they were also mainly related to immune terms, with mostly negative NES apart from Epilepsy, which pre-

sented a high and distinctive positive NES.
Finally, for the neuronal cell-type gene library, 24 terms in total were recongized as singificantly enriched. They mostly were excitatory neurons with positive NES, while astrocytes, oligodendrocytes and oligodendrocyte progenitor cells (OPC), appeared with a strongly negative normalized enrichment score.

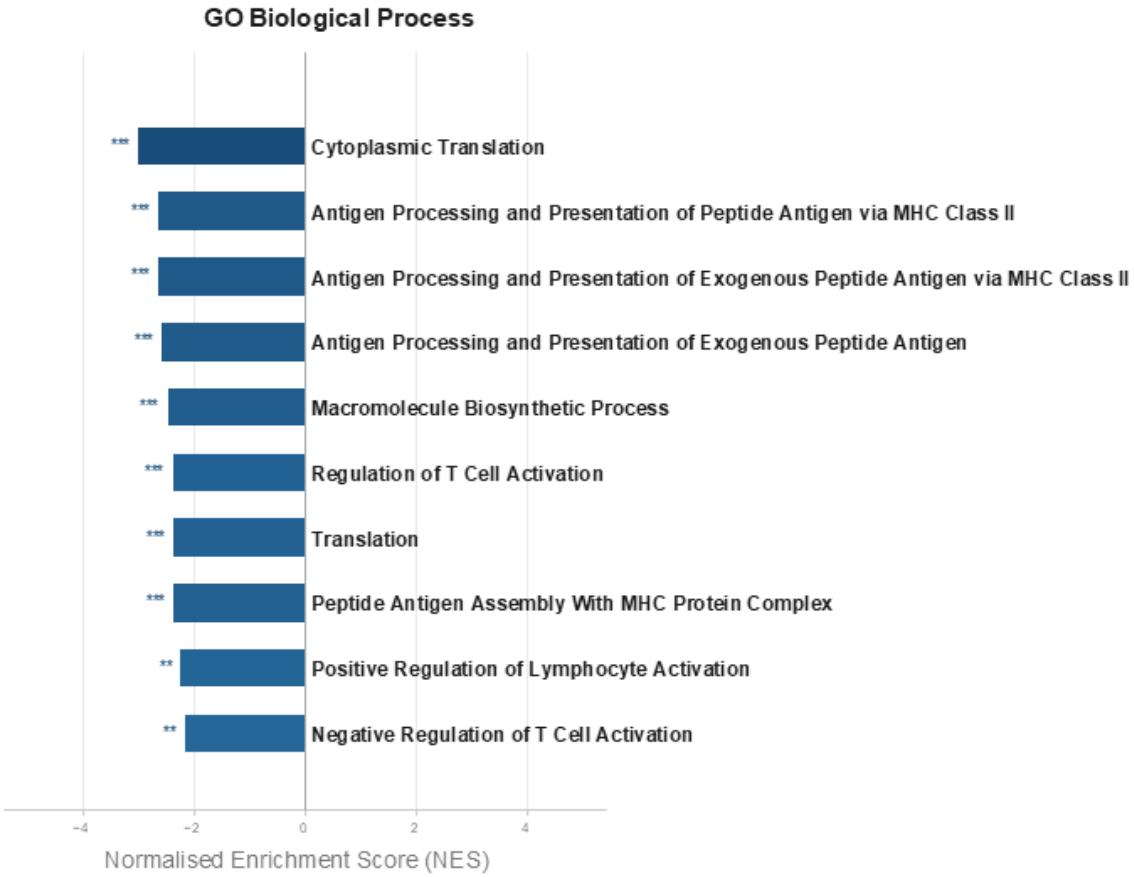


Figure 4.6: GO Biological Process 2025 gene library.

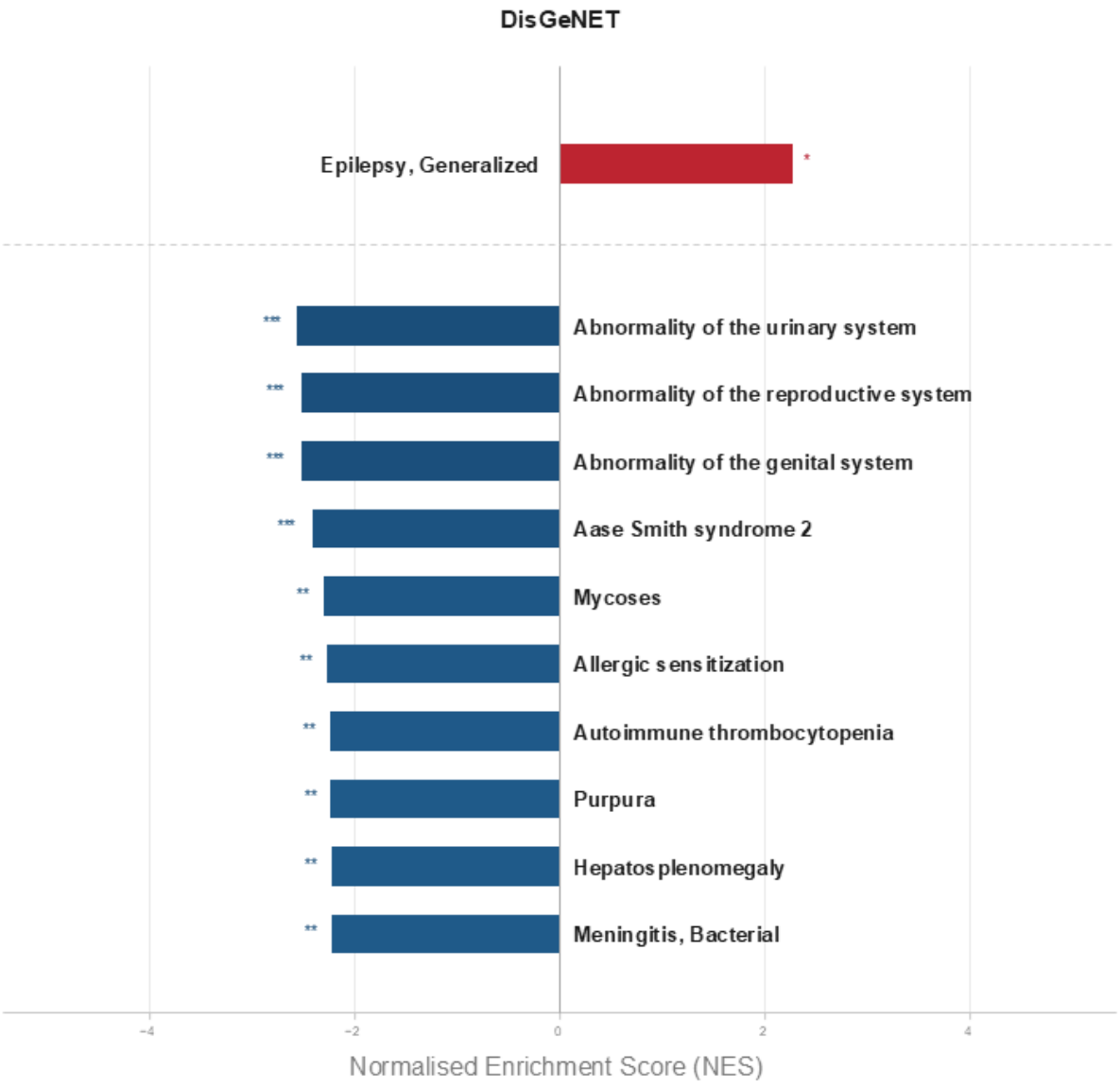


Figure 4.7: DisGeNET gene library.

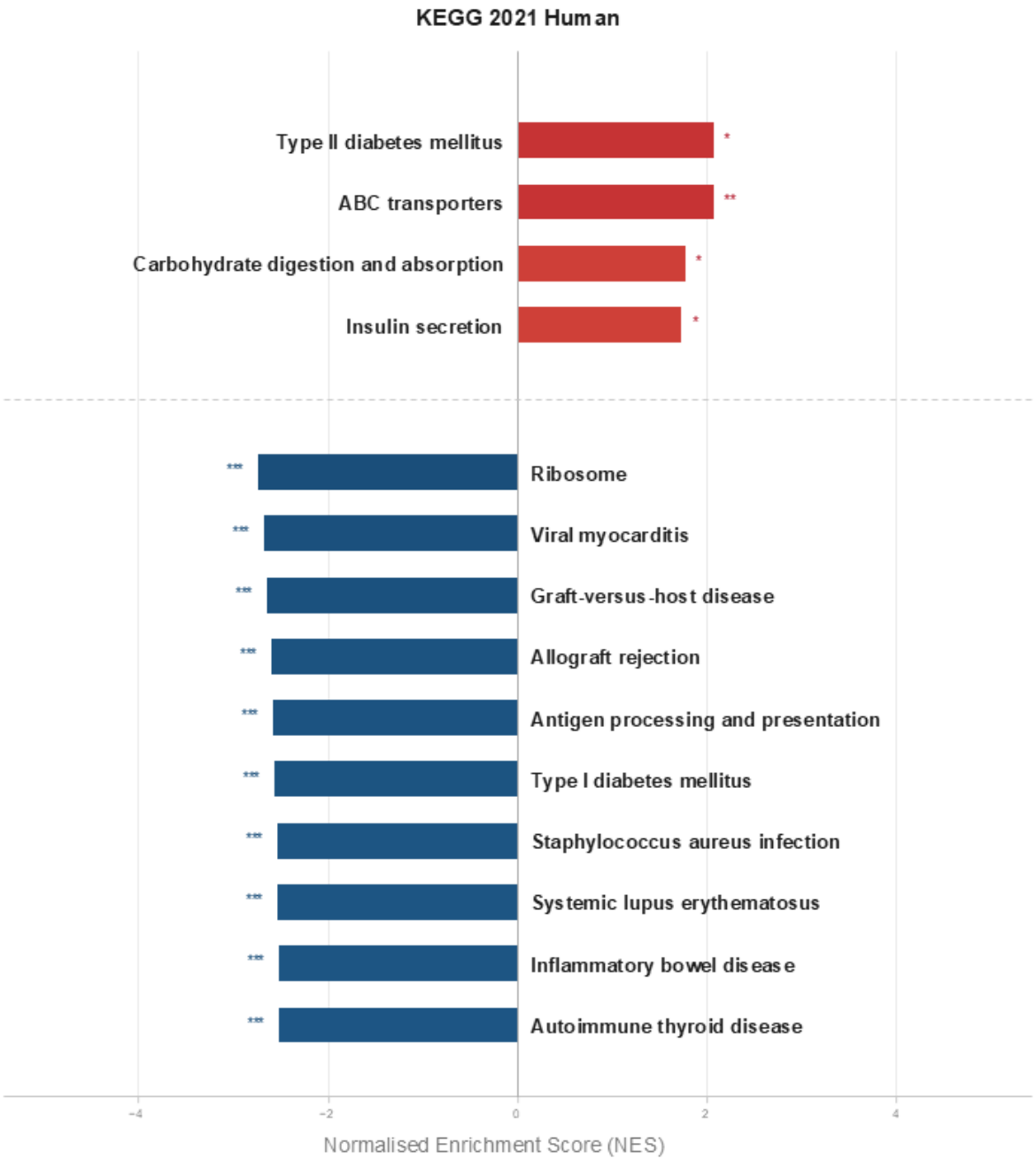


Figure 4.8: KEGG 2021 (Human) gene library.

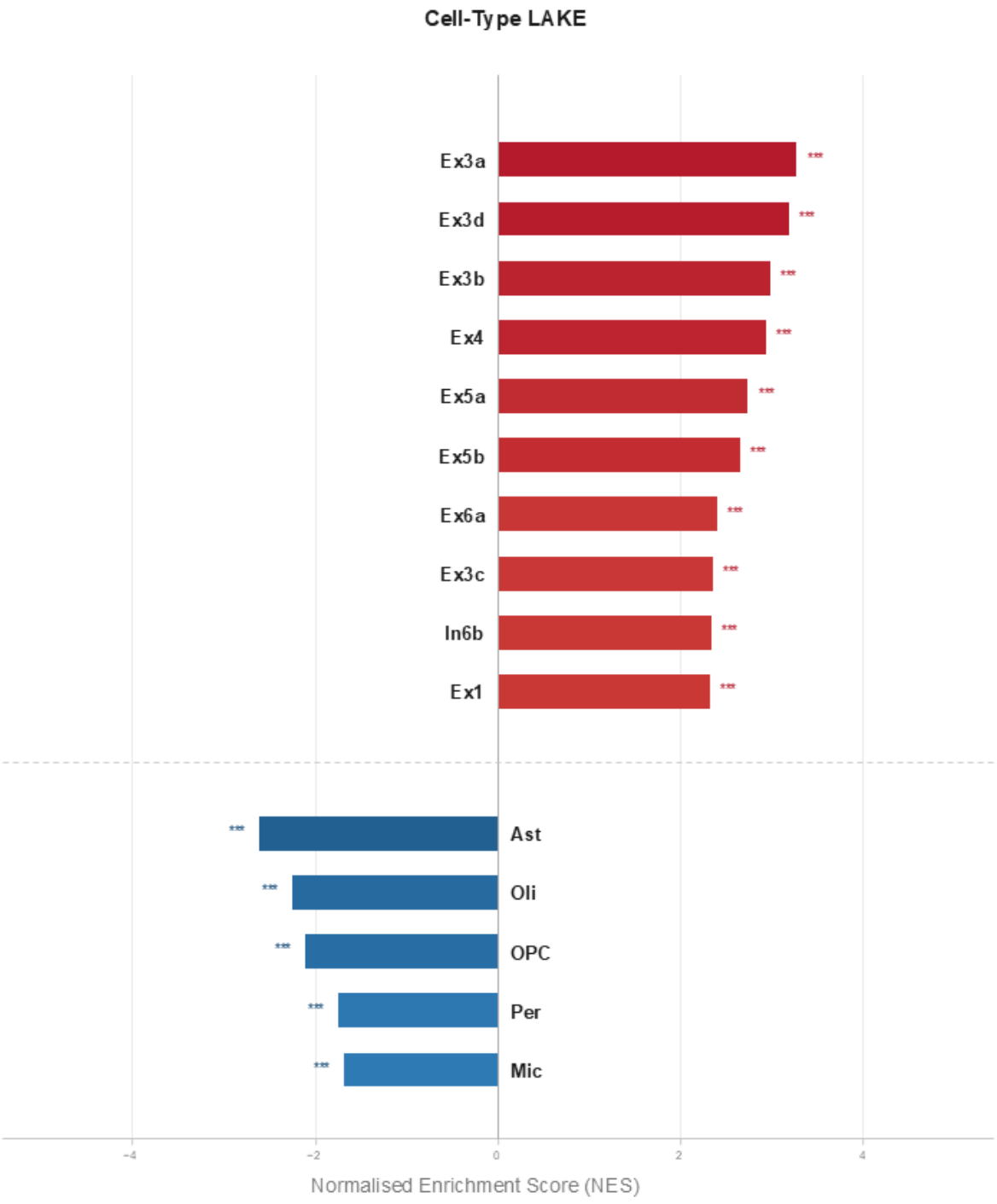


Figure 4.9: Neuronal cell-type gene library.

4.4 Further validation: Frontotemporal dementia dataset

To further investigate our framework's validity, the full pipeline performed on the PPMI dataset was also applied to to a frontotemporal dementia (FTD) dataset. Through the analysis of the morphometric differences in the genetically stratified cohorts comprising the dataset, containing groups HC and prodromal patients with the FTD-linked mutations C9orf72, GRN and MAPT, more significant differences were found, compared to the overall results coming from the PPMI dataset.

This analysis served as a further test of the framework developed in this work, providing a dataset related to a disease specifically more involved in cortical atrophy. This way, the assessment of the current framework was possible in different clinical cases, which enabled a better explainability of the final results and the development of biologically relevant insights about the disease and the structural connectome.

Chapter 5

Discussion

The main objective of this study was to create a framework to track and investigate topological alterations of Morphometric Similarity Networks (MSN) between clinical cohorts, to then check for spatial covariance with brain transcriptomic data through a step of imaging transcriptomics. To test this methodology, a dataset consisting of Parkinson's disease patients and healthy controls was analyzed using the developed pipeline.

Overall, **our findings suggest that the large-scale macroscopic structural covariance of the brain remains highly preserved in the first stages of PD, consistent with existing literature [86]** and confirm the pipeline's ability to correctly detect expected null-like results. Specifically, the results of the starting explorative mass univariate analysis yielded low-magnitude but consistent patterns of structural alteration. The PD cohort, compared to the healthy controls, exhibited trends of decreased surface area, cortical thickness and gray matter volume, accompanied by an increase in mean and gaussian curvature. Despite being non-significant changes, an inspection of the direction of these changes can still be done. In fact, while the decreased area, thickness and volume are characteristic findings linked with neurodegeneration, the increases of curvature metrics are less straightforward to interpret. An increased curvature usually means sharper folds in the cortex, with gyral curves being sharper due to tissue collapse and the subsequent bending of the remaining cortex. Current literature usually associates atrophy of the cerebral cortex to a widening of sulcal regions and a decrease in the width of the sulci, conditions that commonly lead to decreases in the values of curvature indexes[87]. While some studies report localized decreases in curvature tracking with volume loss, the generalized increase in curvature observed in this work for the PD cohort may reflect morphological tension shifts or compensatory folding alterations that are secondary to gray matter thinning. This is also supported by the way the results of this work match with previous studies showing that increases in curvature actually relate to white matter losses in volume[88], which, despite not being tracked in this analysis, are characteristic of PD and come before the disruptions in the morphometry of the cortex[89]. For

this reason, new approaches that integrate white matter morphometry should also be explored. The lack of significance in any of the regions after FDR correction also suggests that, unlike the severe macroscopic atrophy seen in later-stage Alzheimer's, cortical structural degradation in the evaluated PD cohort is either highly heterogeneous across subjects or insufficiently advanced to cross stringent statistical thresholds, a result that actually reflects their recent diagnosis and early disease stage[66]. This is consistent to previous studies showing no difference in selected metrics between PD and HC cohorts[90].

Visual inspection on average MSNs confirmed that the global similarity profiles of PD patients and healthy controls are virtually indistinguishable, despite some subjects showing slight individual variance. Without a proper statistical analysis, though, this visual inspection doesn't provide conclusive evidence.
The same applied to the node strength vectors across all the 68 regions of the Desikan-Killiany atlas, with both average and single-subject maps not showing remarkable differences compared to the healthy controls maps. Node strength was utilized as the main topological metric in line with previous studies [47], and for of its biological relevance, as it provides a measure of regional centrality and reflects the degree of morphometric integration of a node within the overall structural network. The use of other topological metrics such as betweenness centrality, which is more closely related to the path of information flows, could also be explored to investigate different aspects of the structural connectome.

When quantifying differences between the PD and HC cohorts via topological node strength, the GLM controlling for age, sex, eTIV, and scanner type, identified four distinct cortical regions exhibiting altered structural integration: the left caudal middle frontal gyrus, left medial orbitofrontal cortex, left pars opercularis, and the right superior parietal lobule. Notably, these findings did not survive FDR correction and should be therefore considered exploratory.
Interestingly, the left hemisphere frontal regions demonstrated positive beta coefficients, hinting at a localized shift in regional network embedding rather than patterns of global disconnection. Though these regions did not survive FDR correction, their spatial distribution is clinically relevant. The frontal and orbitofrontal cortices are integral to the frontostriatal networks, which are known to be heavily impacted by dopaminergic depletion in PD. Alterations in these networks are frequently associated with some of the symptoms of PD [91]. The middle frontal gyrus, in particular, has already been linked to cognitive impairment[92]. Similarly, involvement of the pars opercularis (part of Broca's area) may relate to speech manifestations, while right superior parietal alterations could be linked to deficits often observed in PD patients [93].
It needs to be recalled that the changes found within the GLM framework are not directly translatable to cortical atrophy, but to structural similarity, where a higher value means that the corresponding area is showing an increase in likeness to other areas compared to healthy controls,

with a contribution of all the different regions. Still, changes in some areas could indicate structural deviations in cortical regions related to them. For example, the orbitofrontal cortex is strongly connected to the piriform cortex, which in turn is the main location of the olfactory cortex. Notably, one of the first non-motor symptoms in Parkinson's disease is usually olfactory dysfunction[94][95]. While olfactory dysfunction in PD is still not clearly understood, new analyses focusing on structural deviations in both gray matter and white matter structures may help elucidate undelying relationships that have thus far remained elusive.

To further investigate the network-level impact of Parkinson's disease and the ability of the developed pipeline to detect it, we evaluated the behavior of structural hubs by correlating group differences with mean MS in healthy controls. This approach, already introduced in another study[47], allowed us to determine whether network reorganization in PD follows specific topological trajectories, which can be divided into 4 groups: dedifferentiation, hypercoupling, hyperdifferentiation, and hub decoupling. What was found in this work is that structural reorganization in PD is weakly dependent on the baseline topological role of the regions. In particular, the most prominent shift was hub decoupling, which accounted for approximately one third of the analyzed regions (32.4%). This pattern reflects a decrease in nodal strength in the PD cohort within regions that are normally connected hubs (and thus have high MS) in healthy controls. The prevalence of hub decoupling aligns with the hypothesis of hub vulnerability frequently observed in PD and other neurodegenerative disorders[18][96][97]. The rationale is that the inherent physiological stress of being topologically central and metabolically demanding renders the structural hubs more susceptible to neurodegeneration and pathological protein accumulation. Interestingly, the second most frequent pattern observed was hypercoupling (27.9%), followed by dedifferentiation (20.6%), both related to regions where an increase in similarity is exhibited compared to healthy individuals. This may mark an attempt of the brain network to maintain communication efficiency and reroute connections as a compensatory mechanism[98]. Lastly, 19.1% of regions present a hyperdifferentiation pattern, which highlights regions that were already isolated (low MS in HC) becoming even more morphologically distinct from the others with the disease.
Looking at the node strength deviation map globally, the present alterations, due to the lack of significance of the comparison, support an hypothesis of network resilience. This could also be attributed to the subjects present in the used PD cohort, as it is heterogeneous, small in size and not guaranteed to present any cortical anomaly, as it consists of early-stage patients. Compared to other neurodegenerative diseases, in fact, the cortical element of the illness in PD doesn't reach the same severity until the later stages.

To investigate the molecular mechanisms potentially driving these small localized topological shifts, the left-hemisphere spatial phenotype map was cross-referenced with normative mi-

croarray gene expression data from the Allen Human Brain Atlas (AHBA). Utilizing Partial Least Squares (PLS) regression, the first principal component (PLS1) was isolated to identify the weighted linear combination of genes that maximally covaried with the PD-related changes in node strength.

It is crucial to interpret the resulting transcriptomic signature with methodological caution. While the PLS1 component explained a certain amount of the covariance between the two maps, spatial permutation testing performed using the Váša algorithm[83], yielded a non-significant and extremely high p-value of 0.937. This indicates that the spatial alignment between the overall transcriptomic profile and our morphological map does not definitively exceed the boundaries of spatial autocorrelation. In other words, the global spatial correspondence could be driven by the brain's inherent spatial network structure rather than a highly specific, disease-driven molecular gradient. Consequently, the downstream gene rankings and enrichment analyses must be considered as exploratory insights rather than definitive mechanistic links.

Despite the global model's limitations, the PLS1 component successfully generated a ranked list of individual genes, yielding 891 genes with an FDR-corrected p-value < 0.05. When looking through this list for established PD-associated genes (including SNCA, PRKN, PARK7, LRRK2, PINK1, and ATP13A2), LRRK2 emerged as the sole significant correlate.

Ranking at 381 out of 15,633 genes with a positive Z-Score ($Z = 1.702$), LRRK2 expression demonstrated a positive spatial covariance with the node strength map. This suggests that cortical regions with higher expression of LRRK2 correspond to areas in our PD cohort that exhibit increases of structural similarity (higher node strength). LRRK2 mutations are one of the most common genetic causes of familial PD, heavily implicated in kinase pathway dysfunction, vesicular trafficking, and autophagy[99][100]. The positive covariance observed may reflect a dynamic where disruptions in LRRK2 signaling might drive compensatory coupling in these specific areas, as other works have already suggested[101].

A new narrative, despite its uncertain overall significance, emerges from the Gene Set Enrichment Analysis (GSEA) performed on the PLS1 ranked list: there is a functional difference between regions exhibiting increasing and decreasing node strength.

The negative pole of PLS1 represents genes highly expressed in regions where PD patients display decreased node strength, and was dominated by terms related to immunity and metabolism across the GO Biological Process, KeGG and DisGeNET gene libraries. Furthermore, cell-type enrichment analysis revealed that these negatively covarying regions are heavily enriched for glial cells, specifically astrocytes, oligodendrocytes, and oligodendrocyte precursor cells (OPCs). This convergence of immune pathways and glial cell types aligns strongly with the neuroinflammatory hypothesis behind Parkinson's disease, with phenomena like microglial activation usually preceding neuronal loss. This could support a model of neuroinflammation and

glial dysfunction that serve as precursors to the macroscopic loss of similarity observed in the data [102][103].

In contrast, the positive pole of PLS1 was enriched mainly for excitatory neuronal cell types. In addition, the DisGeNET library highlghted a strong positive association with epilepsy, and the KeGG library showed positive enrichment scores for terms related to metabolism and especially to glucose and insulin.
Research on Parkinson's disease so far has proved the presence of altered excitatory balance in the brain[104], with ties to excitotoxicity and cell death. The enrichment of excitatory neuronal cells and epilepsy related terms in regions of increased node strength suggests that these areas may be undergoing a compensatory structural integration: as some network degrade, other cortical networks rich in excitatory neurons might become hyper-coupled, resulting in potential excitotoxicity issues.

The final picture painted by this developed pipeline presents, for this dataset, results coherent with biology, despite its limitations. What emerges is its ability to correctly capture the expected topological signatures of early-stage neurodegeneration, and linking them to plausible molecular drivers. To further validate this framework, though, results coming from frontotemporal dementia patients were used. The study focused on prodromal FTD patients stratified by genetic mutation, namely C9orf72, GRN and MAPT. The test with this new dataset allowed to also inspect a neurodegenerative disease that presents characteristic and more prominent cortical degeneration patterns compared to PD, which in turn is more heavily associated with white matter until the later stages of the condition. This, together with the fact that the framework only analyzes cortical data, may help explain the substantial differences observed between the two analyses, since the one on FTD patients ended with more significant and numerically greater results.
The PPMI dataset, in fact, was not restricted to late-stage PD, PD-MCI, or PDD patients, but instead included a broader population of subjects with Parkinson's disease, focusing primarily on early-stage patients. The cohort had a mean age of 61.2 years ($\pm$ 5.9), further reducing the likelihood that these participants were in advanced stages of the disease, where cognitive decline is more prevalent. Although the relationship between disease progression, age, and cognitive impairment is complex, and age alone cannot fully account for cognitive decline, it has been established as a primary risk factor for Parkinson's disease[105].

5.1 Study Limitations

This study presents several methodological and conceptual limitations that must be considered when interpreting its findings.

First, neither the mass univariate analysis nor the topological node strength alterations survived stringent FDR correction. While the unthresholded spatial maps provide valuable phenotypes of disease vulnerability, the lack of statistical significance mainly indicates that macroscopic structural covariance is largely preserved in this specific PD cohort. This may stem from the cohort's disease stage, sample size limitations or from the inherent clinical and spatial heterogeneity of Parkinson's disease, which can dilute group-level statistical significance. Therefore, this reflects more the intrinsic nature of the data rather than a limitation of the methodological pipeline.
Second, while the PLS regression yielded a biologically plausible gene ranking, the global spatial permutation testing for the first PLS component did not reach statistical significance. As the covariance between the neuroimaging map and the transcriptomic map cannot be separated from the brain's own spatial autocorrelation, the subsequent gene-level findings and enrichment analyses must be strictly interpreted as exploratory, rather than definitive evidence.
A third limitation comes from the data sourced from the Allen Human Brain Atlas: the limited availability to left hemisphere data and lack of gene expression data in diseased patients appear as a fundamental limit to investigative work, where lateralized changes or asymmetric molecular vulnerabilities could be present.
Finally, the cross-sectional design of this study precludes any causal inferences regarding disease progression, as the patients data was all collected at baseline.

Chapter 6

Conclusion

This work proposed and implemented a comprehensive methodological framework designed to bridge macroscale brain topology and microscale molecular signatures. By integrating Morphometric Similarity Networks (MSNs) with post-mortem microarray gene expression data from the Allen Human Brain Atlas, the developed pipeline successfully established a systematic approach allowing to track structural reorganization and exploring its underlying transcriptomic correlates in neurodegenerative diseases.

By applying this framework to a cohort of Parkinson's disease patients and healthy controls from the PPMI dataset, key findings revealed that the macroscopic structural covariance of the cortex remains largely preserved in earlier stages of the disease, as expected. The structural connectome in PD exhibits subtle signs of topological shifting, most notably characterized by hub decoupling. This pattern aligns with the established hypothesis of hub vulnerability, wherein highly connected regions succumb to neurodegeneration earlier than isolated areas. Simultaneously, results also suggest the presence of compensatory mechanisms attempting to maintain network communication efficiency in the face of structural decline.

The exploratory imaging transcriptomics step of this study further enriched these macroscopic findings by proposing potential molecular drivers of the newfound structural reorganization. Although global spatial permutation testing indicated that the alignment between the transcriptomic profile and morphological maps did not significantly exceed inherent spatial autocorrelation, the resulting genetic rankings offered compelling avenues for future research, supporting the hypothesis that neuroinflammation, glial dysfunction and excitotoxicity are involved in macroscopic structural degradation. Overall, the biological consistency between the MSN-derived structural patterns and the transcriptomic signatures recovered by the pipeline itself constitutes evidence of the framework's internal validity and correctness.

Crucially, the successful application of this pipeline to a frontotemporal dementia (FTD) dataset validated the framework's broader utility and sensitivity, with the more robust results

observed in the FTD cohort underscoring that the framework, in its current iteration, is optimized for neurodegenerative conditions characterized by prominent early cortical atrophy.

Ultimately, despite the inherent challenges, this work successfully delivers a scalable, automated and biologically grounded pipeline. It provides a vital step for future research aiming to decode the complex relationship between genetic vulnerability and whole-brain structural reorganization in a standardized way.

Bibliography

[1] D. E. Katharina A Schindlbeck, "Network imaging biomarkers: Insights and clinical applications in parkinson's disease," *The Lancet Neurology*, vol. 17, no. 7, pp. 629–640, 2018, ISSN: 1474-4422. DOI: 10.1016/s1474-4422(18)30169-8.

[2] J. Seidlitz et al., "Morphometric similarity networks detect microscale cortical organization and predict inter-individual cognitive variation," *Neuron*, vol. 97, no. 1, 231–247.e7, Jan. 2018, ISSN: 0896-6273. DOI: 10.1016/j.neuron.2017.11.039.

[3] J. Liu et al., "Recent developments in imaging transcriptomics for psychiatric disorders," *Psychoradiology*, Mar. 2026, ISSN: 2634-4416. DOI: 10.1093/psyrad/kkag010.

[4] O. Hansson, "Biomarkers for neurodegenerative diseases," *Nature Medicine*, vol. 27, no. 6, pp. 954–963, Jun. 2021, ISSN: 1546-170X. DOI: https://doi.org/10.1038/s41591-021-01382-x.

[5] E. A. Ayeni et al., "Neurodegenerative diseases: Implications of environmental and climatic influences on neurotransmitters and neuronal hormones activities," *International Journal of Environmental Research and Public Health*, vol. 19, no. 19, p. 12 495, Sep. 2022, ISSN: 1660-4601. DOI: https://doi.org/10.3390/ijerph191912495.

[6] D. M. Wilson, M. R. Cookson, L. Van Den Bosch, H. Zetterberg, D. M. Holtzman, and I. Dewachter, "Hallmarks of neurodegenerative diseases," *Cell*, vol. 186, no. 4, pp. 693–714, Feb. 2023, ISSN: 0092-8674. DOI: https://doi.org/10.1016/j.cell.2022.12.032.

[7] S. Wang, Y. Jiang, A. Yang, F. Meng, and J. Zhang, "The expanding burden of neurodegenerative diseases: An unmet medical and social need," *Aging and disease*, p. 0, 2024, ISSN: 2152-5250. DOI: https://doi.org/10.14336/AD.2024.1071.

[8] Y. Zeighami et al., "Network structure of brain atrophy in de novo parkinson's disease," *eLife*, vol. 4, Sep. 2015, ISSN: 2050-084X. DOI: https://doi.org/10.7554/eLife.08440.

[9] V. L. Feigin et al., "Global, regional, and national burden of neurological disorders, 1990–2016: A systematic analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 18, no. 5, pp. 459–480, May 2019, ISSN: 1474-4422. DOI: 10.1016/s1474-4422(18)30499-x.

[10] Y. Li et al., "Imaging transcriptomics of brain functional alterations in ms and neuromyelitis optica spectrum disorder," *American Journal of Neuroradiology*, vol. 45, no. 12, pp. 1901–1909, Nov. 2024, ISSN: 1936-959X. DOI: 10.3174/ajnr.a8480.

[11] B. R. Bloem, M. S. Okun, and C. Klein, "Parkinson's disease," *The Lancet*, vol. 397, no. 10291, pp. 2284–2303, Jun. 2021, ISSN: 0140-6736. DOI: 10.1016/s0140-6736(21)00218-x.

[12] S. Chandrasekaran and D. Bonchev, "A network view on parkinson's disease," *Computational and Structural Biotechnology Journal*, vol. 7, no. 8, e201304004, Apr. 2013, ISSN: 2001-0370. DOI: 10.5936/csbj.201304004.

[13] A. Recchia, P. Debetto, A. Negro, D. Guidolin, S. D. Skaper, and P. Giusti, "A-synuclein and parkinson's disease," *The FASEB Journal*, vol. 18, no. 6, pp. 617–626, Apr. 2004, ISSN: 1530-6860. DOI: https://doi.org/10.1096/fj.03-0338rev.

[14] D. Aarsland et al., "Cognitive decline in parkinson disease," *Nature Reviews Neurology*, vol. 13, no. 4, pp. 217–231, Mar. 2017, ISSN: 1759-4766. DOI: 10.1038/nrneurol.2017.27.

[15] Y. Yau et al., "Network connectivity determines cortical thinning in early parkinson's disease progression," *Nature Communications*, vol. 9, no. 1, Jan. 2018, ISSN: 2041-1723. DOI: 10.1038/s41467-017-02416-0.

[16] H. R. Morris, M. G. Spillantini, C. M. Sue, and C. H. Williams-Gray, "The pathogenesis of parkinson's disease," *The Lancet*, vol. 403, no. 10423, pp. 293–304, Jan. 2024, ISSN: 0140-6736. DOI: 10.1016/s0140-6736(23)01478-2.

[17] H. Braak, K. D. Tredici, U. Rüb, R. A. de Vos, E. N. Jansen Steur, and E. Braak, "Staging of brain pathology related to sporadic parkinson's disease," *Neurobiology of Aging*, vol. 24, no. 2, pp. 197–211, Mar. 2003, ISSN: 0197-4580. DOI: 10.1016/s0197-4580(02)00065-9.

[18] S. Nigro et al., "Characterizing structural neural networks in de novo parkinson disease patients using diffusion tensor imaging," *Human Brain Mapping*, vol. 37, no. 12, pp. 4500–4510, Jul. 2016, ISSN: 1097-0193. DOI: 10.1002/hbm.23324.

[19] X. Xu et al., "Brain atrophy and reorganization of structural network in parkinson's disease with hemiparkinsonism," *Frontiers in Human Neuroscience*, vol. 12, Mar. 2018, ISSN: 1662-5161. DOI: 10.3389/fnhum.2018.00117.

[20] M. Bergamino, E. G. Keeling, N. J. Ray, A. Macerollo, M. Silverdale, and A. M. Stokes, "Structural connectivity and brain network analyses in parkinson's disease: A cross-sectional and longitudinal study," *Frontiers in Neurology*, vol. 14, Mar. 2023, ISSN: 1664-2295. DOI: 10.3389/fneur.2023.1137780.

[21] C. Huang et al., "Changes in network activity with the progression of parkinson's disease," *Brain*, vol. 130, no. 7, pp. 1834–1846, May 2007, ISSN: 1460-2156. DOI: 10.1093/brain/awm086.

[22] J. H. Ko, P. G. Spetsieris, and D. Eidelberg, "Network structure and function in parkinson's disease," *Cerebral Cortex*, pp. 1–15, Oct. 2017, ISSN: 1460-2199. DOI: 10.1093/cercor/bhx267.

[23] A. Alexander-Bloch, J. N. Giedd, and E. Bullmore, "Imaging structural co-variance between human brain regions," *Nature Reviews Neuroscience*, vol. 14, no. 5, pp. 322–336, Mar. 2013, ISSN: 1471-0048. DOI: 10.1038/nrn3465.

[24] K. Xue et al., "Transcriptional signatures of the cortical morphometric similarity network gradient in first-episode, treatment-naive major depressive disorder," *Neuropsychopharmacology*, vol. 48, no. 3, pp. 518–528, 2023, ISSN: 1740-634X. DOI: 10.1038/s41386-022-01474-3. [Online]. Available: https://doi.org/10.1038/s41386-022-01474-3.

[25] D. J. King and A. G. Wood, "Clinically feasible brain morphometric similarity network construction approaches with restricted magnetic resonance imaging acquisitions," *Network Neuroscience*, vol. 4, no. 1, pp. 274–291, Mar. 2020, ISSN: 2472-1751. DOI: 10.1162/netn_a_00123.

[26] Y. Wang et al., "Individualized morphometric similarity predicts body mass index and food approach behavior in school-age children," *Cerebral Cortex*, vol. 33, no. 8, pp. 4794–4805, Oct. 2022, ISSN: 1460-2199. DOI: 10.1093/cercor/bhac380.

[27] A. Goulas et al., "Cytoarchitectonic similarity is a wiring principle of the human connectome," Aug. 2016. DOI: 10.1101/068254.

[28] K. Wagstyl, L. Ronan, I. M. Goodyer, and P. C. Fletcher, "Cortical thickness gradients in structural hierarchies," *NeuroImage*, vol. 111, pp. 241–250, May 2015, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2015.02.036.

[29] S. E. Morgan et al., "Cortical patterning of abnormal morphometric similarity in psychosis is associated with brain expression of schizophrenia-related genes," *Proceedings of the National Academy of Sciences*, vol. 116, no. 19, pp. 9604–9609, Apr. 2019, ISSN: 1091-6490. DOI: 10.1073/pnas.1820754116.

[30] Y. Zhang, M. Ma, Z. Xie, H. Wu, N. Zhang, and J. Shen, "Bridging the gap between morphometric similarity mapping and gene transcription in alzheimer's disease," *Frontiers in Neuroscience*, vol. 15, Sep. 2021, ISSN: 1662-453X. DOI: 10.3389/fnins.2021.731292.

[31] R. S. Desikan et al., "An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest," *NeuroImage*, vol. 31, no. 3, pp. 968–980, Jul. 2006, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2006.01.021.

[32] C. Destrieux, B. Fischl, A. Dale, and E. Halgren, "Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature," *NeuroImage*, vol. 53, no. 1, pp. 1–15, Oct. 2010, ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2010.06.010.

[33] V. Vuksanović, "Brain morphometric similarity and flexibility," *Cerebral Cortex Communications*, vol. 3, no. 3, Jun. 2022, ISSN: 2632-7376. DOI: 10.1093/texcom/tgac024.

[34] A. Schaefer et al., "Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri," *Cerebral Cortex*, vol. 28, no. 9, pp. 3095–3114, Jul. 2017, ISSN: 1460-2199. DOI: 10.1093/cercor/bhx179.

[35] L. H. Scholtens, M. A. de Reus, S. C. de Lange, R. Schmidt, and M. P. van den Heuvel, "An mri von economo – koskinas atlas," *NeuroImage*, vol. 170, pp. 249–256, Apr. 2018, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2016.12.069.

[36] K. Yu, X. Wang, Q. Li, X. Zhang, X. Li, and S. Li, "Individual morphological brain network construction based on multivariate euclidean distances between brain regions," *Frontiers in Human Neuroscience*, vol. 12, May 2018, ISSN: 1662-5161. DOI: 10.3389/fnhum.2018.00204.

[37] I. Sebenius et al., "Robust estimation of cortical similarity networks from brain mri," *Nature Neuroscience*, vol. 26, no. 8, pp. 1461–1471, Jul. 2023, ISSN: 1546-1726. DOI: 10.1038/s41593-023-01376-7.

[38] S. Han et al., "Revealing structural brain-cognition relationships in children: A comparison of morphometric similarity and inverse divergence networks," *Neuroinformatics*, vol. 24, no. 1, Jan. 2026, ISSN: 1559-0089. DOI: 10.1007/s12021-025-09764-z.

[39] A. Del Casale et al., "Multimodal morphometric similarity network analysis of autism spectrum disorder," *Brain Sciences*, vol. 15, no. 3, p. 247, Feb. 2025, ISSN: 2076-3425. DOI: https://doi.org/10.1073/pnas.1820754116.

[40] J. Seidlitz et al., "Transcriptomic and cellular decoding of regional brain vulnerability to neurodevelopmental disorders," Mar. 2019. DOI: 10.1101/573279.

[41] J. Janssen et al., "Heterogeneity of morphometric similarity networks in health and schizophrenia," *Schizophrenia*, vol. 11, no. 1, Apr. 2025, ISSN: 2754-6993. DOI: 10.1038/s41537-025-00612-2.

[42] G. Yao et al., "Cortical structural changes of morphometric similarity network in early-onset schizophrenia correlate with specific transcriptional expression patterns," *BMC Medicine*, vol. 21, no. 1, Dec. 2023, ISSN: 1741-7015. DOI: 10.1186/s12916-023-03201-1.

[43] Y. Wang et al., "Altered morphometric similarity networks in insomnia disorder," *Brain Structure and Function*, vol. 229, no. 6, pp. 1433–1445, May 2024, ISSN: 1863-2661. DOI: 10.1007/s00429-024-02809-0.

[44] P. Moghimi, A. T. Dang, Q. Do, T. I. Netoff, K. O. Lim, and G. Atluri, "Evaluation of functional mri-based human brain parcellation: A review," *Journal of Neurophysiology*, vol. 128, no. 1, pp. 197–217, Jul. 2022, ISSN: 1522-1598. DOI: 10.1152/jn.00411.2021.

[45] FreeSurfer, *Freesurfer parcellation*, [https://surfer.nmr.mgh.harvard.edu/].

[46] C. Klein and A. Westenberger, "Genetics of parkinson's disease," *Cold Spring Harbor Perspectives in Medicine*, vol. 2, no. 1, a008888–a008888, Jan. 2012, ISSN: 2157-1422. DOI: 10.1101/cshperspect.a008888.

[47] D. Martins et al., "Transcriptional and cellular signatures of cortical morphometric remodelling in chronic pain," *Pain*, vol. 163, no. 6, e759–e773, Sep. 2021, ISSN: 1872-6623. DOI: 10.1097/j.pain.0000000000002480.

[48] A. Arnatkeviciute, B. D. Fulcher, M. A. Bellgrove, and A. Fornito, "Imaging transcriptomics of brain disorders," *Biological Psychiatry Global Open Science*, vol. 2, no. 4, pp. 319–331, Oct. 2022, ISSN: 2667-1743. DOI: 10.1016/j.bpsgos.2021.10.002.

[49] E. H. Shen, C. C. Overly, and A. R. Jones, "The allen human brain atlas," *Trends in Neurosciences*, vol. 35, no. 12, pp. 711–714, Dec. 2012, ISSN: 0166-2236. DOI: 10.1016/j.tins.2012.09.005.

[50] M. Li et al., "Integrative functional genomic analysis of human brain development and neuropsychiatric risks," *Science*, vol. 362, no. 6420, Dec. 2018, ISSN: 1095-9203. DOI: 10.1126/science.aat7615.

[51] S. Akbarian et al., "The psychencode project," *Nature Neuroscience*, vol. 18, no. 12, pp. 1707–1712, Nov. 2015, ISSN: 1546-1726. DOI: 10.1038/nn.4156.

[52] B. B. Lake et al., "Neuronal subtypes and diversity revealed by single-nucleus rna sequencing of the human brain," *Science*, vol. 352, no. 6293, pp. 1586–1590, Jun. 2016, ISSN: 1095-9203. DOI: 10.1126/science.aaf1204.

[53] R. D. Markello, A. Arnatkeviciute, J.-B. Poline, B. D. Fulcher, A. Fornito, and B. Misic, "Standardizing workflows in imaging transcriptomics with the abagen toolbox," *eLife*, vol. 10, Nov. 2021, ISSN: 2050-084X. DOI: 10.7554/elife.72129.

[54] A. Arnatkeviciute, R. D. Markello, B. D. Fulcher, B. Misic, and A. Fornito, "Toward best practices for imaging transcriptomics of the human brain," *Biological Psychiatry*, vol. 93, no. 5, pp. 391–404, Mar. 2023, ISSN: 0006-3223. DOI: 10.1016/j.biopsych.2022.10.016.

[55] R. Dear et al., "Cortical gene expression architecture links healthy neurodevelopment to the imaging, transcriptomics and genetics of autism and schizophrenia," *Nature Neuroscience*, vol. 27, no. 6, pp. 1075–1086, Apr. 2024, ISSN: 1546-1726. DOI: 10.1038/s41593-024-01624-4.

[56] Y. Ji et al., "Genes associated with gray matter volume alterations in schizophrenia," *NeuroImage*, vol. 225, p. 117 526, Jan. 2021, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2020.117526.

[57] Y. Xie et al., "Brain mrna expression associated with cortical volume alterations in autism spectrum disorder," *Cell Reports*, vol. 32, no. 11, p. 108 137, Sep. 2020, ISSN: 2211-1247. DOI: 10.1016/j.celrep.2020.108137.

[58] R. Romero-Garcia, V. Warrier, E. T. Bullmore, S. Baron-Cohen, and R. A. I. Bethlehem, "Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism," *Molecular Psychiatry*, vol. 24, no. 7, pp. 1053–1064, Feb. 2018, ISSN: 1476-5578. DOI: 10.1038/s41380-018-0023-7.

[59] B. Freeze, D. Acosta, S. Pandya, Y. Zhao, and A. Raj, "Regional expression of genes mediating trans-synaptic alpha-synuclein transfer predicts regional atrophy in parkinson disease," *NeuroImage: Clinical*, vol. 18, pp. 456–466, 2018, ISSN: 2213-1582. DOI: 10.1016/j.nicl.2018.01.009.

[60] G. E. C. Thomas et al., "Regional brain iron and gene expression provide insights into neurodegeneration in parkinson's disease," *Brain*, vol. 144, no. 6, pp. 1787–1798, Mar. 2021, ISSN: 1460-2156. DOI: 10.1093/brain/awab084.

[61] T. Rittman et al., "Regional expression of the mapt gene is associated with loss of hubs in brain networks and cognitive impairment in parkinson disease and progressive supranuclear palsy," *Neurobiology of Aging*, vol. 48, pp. 153–160, Dec. 2016, ISSN: 0197-4580. DOI: 10.1016/j.neurobiolaging.2016.09.001.

[62] M. J. Grothe et al., "Molecular properties underlying regional vulnerability to alzheimer's disease pathology," *Brain*, Jul. 2018, ISSN: 1460-2156. DOI: 10.1093/brain/awy189.

[63] A. S. Mandal, M. Gandal, J. Seidlitz, and A. Alexander-Bloch, "A critical appraisal of imaging transcriptomics," *Biological Psychiatry Global Open Science*, vol. 2, no. 4, pp. 311–313, Oct. 2022, ISSN: 2667-1743. DOI: 10.1016/j.bpsgos.2022.08.001.

[64] Y. Shen et al., "A spatial imaging-transcriptomics paradigm for deciphering the molecular basis of microscopic mri in the normal brain and alzheimer's disease," *Cell Reports*, vol. 44, no. 8, p. 116 073, Aug. 2025, ISSN: 2211-1247. DOI: 10.1016/j.celrep.2025.116073.

[65] T. M. J. F. F. for Parkinson's Research, *Parkinson's progression markers initiative*, [https://www.ppmi-info.org/].

[66] K. Marek et al., "The parkinson's progression markers initiative (ppmi) – establishing a pd biomarker cohort," *Annals of Clinical and Translational Neurology*, vol. 5, no. 12, pp. 1460–1477, Oct. 2018, ISSN: 2328-9503. DOI: https://doi.org/10.1002/acn3.644.

[67] K. Marek et al., "The parkinson progression marker initiative (ppmi)," *Progress in Neurobiology*, vol. 95, no. 4, pp. 629–635, Dec. 2011, ISSN: 0301-0082. DOI: https://doi.org/10.1016/j.pneurobio.2011.09.005.

[68] C. M. Stonnington et al., "Interpreting scan data acquired from multiple scanners: A study with alzheimer's disease," *NeuroImage*, vol. 39, no. 3, pp. 1180–1185, Feb. 2008, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2007.09.066.

[69] B. Fischl, "Freesurfer," *NeuroImage*, vol. 62, no. 2, pp. 774–781, Aug. 2012, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2012.01.021.

[70] D. Martins et al., "Transcriptional and cellular signatures of cortical morphometric remodelling in chronic pain," *PAIN*, vol. 163, no. 6, 2022, ISSN: 0304-3959. DOI: 10.1097/j.pain.0000000000002480. [Online]. Available: https://journals.lww.com/pain/fulltext/2022/06000/transcriptional_and_cellular_signatures_of.22.aspx.

[71] K. S. Kannan, K Manoj, and S Arumugam, "Labeling methods for identifying outliers," *International Journal of Statistics and Systems*, vol. 10, no. 2, pp. 231–238, 2015.

[72] B. Pakkenberg and H. J. G. Gundersen, "Neocortical neuron number in humans: Effect of sex and age," *The Journal of Comparative Neurology*, vol. 384, no. 2, pp. 312–320, Jul. 1997, ISSN: 1096-9861. DOI: 10.1002/(sici)1096-9861(19970728)384:2<312::aid-cne10>3.0.co;2-k.

[73] N. A. Mushkudiani et al., "Prognostic value of demographic characteristics in traumatic brain injury: Results from the impact study," *Journal of Neurotrauma*, vol. 24, no. 2, pp. 259–269, Feb. 2007, ISSN: 1557-9042. DOI: 10.1089/neu.2006.0028.

[74] D. Kostro et al., "Correction of inter-scanner and within-subject variance in structural mri based automated diagnosing," *NeuroImage*, vol. 98, pp. 405–415, Sep. 2014, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2014.04.057.

[75] A. Giacomel et al., "Integrating neuroimaging and gene expression data using the imaging transcriptomics toolbox," *STAR Protocols*, vol. 3, no. 2, p. 101 315, Jun. 2022, ISSN: 2666-1667. DOI: 10.1016/j.xpro.2022.101315.

[76] F. Váša et al., "Adolescent tuning of association cortex in human structural brain networks," *Cerebral Cortex*, vol. 28, no. 1, pp. 281–294, Oct. 2017, ISSN: 1460-2199. DOI: 10.1093/cercor/bhx249.

[77] M. Ashburner et al., "Gene ontology: Tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, May 2000, ISSN: 1546-1718. DOI: 10.1038/75556.

[78] M. Kanehisa, "Kegg: Kyoto encyclopedia of genes and genomes," *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, Jan. 2000, ISSN: 1362-4962. DOI: 10.1093/nar/28.1.27.

[79] J. Piñero et al., "Disgenet: A comprehensive platform integrating information on human disease-associated genes and variants," *Nucleic Acids Research*, vol. 45, no. D1, pp. D833–D839, Oct. 2016, ISSN: 1362-4962. DOI: 10.1093/nar/gkw943.

[80] B. B. Lake et al., "Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain," *Nature Biotechnology*, vol. 36, no. 1, pp. 70–80, Dec. 2017, ISSN: 1546-1696. DOI: 10.1038/nbt.4038.

[81] Z. Fang, X. Liu, and G. Peltz, "Gseapy: A comprehensive package for performing gene set enrichment analysis in python," *Bioinformatics*, vol. 39, no. 1, Z. Lu, Ed., Nov. 2022, ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btac757.

[82] Nilearn, *Nilearn package*, [https://nilearn.github.io/].

[83] F. Váša and B. Mišić, "Null models in network neuroscience," *Nature Reviews Neuroscience*, vol. 23, no. 8, pp. 493–504, May 2022, ISSN: 1471-0048. DOI: 10.1038/s41583-022-00601-9.

[84] C. Blauwendraat, M. A. Nalls, and A. B. Singleton, "The genetic architecture of parkinson's disease," *The Lancet Neurology*, vol. 19, no. 2, pp. 170–178, Feb. 2020, ISSN: 1474-4422. DOI: 10.1016/s1474-4422(19)30287-x.

[85] C. M. Lill, "Genetics of parkinson's disease," *Molecular and Cellular Probes*, vol. 30, no. 6, pp. 386–396, Dec. 2016, ISSN: 0890-8508. DOI: 10.1016/j.mcp.2016.11.001.

[86] M. Zarei et al., "Cortical thinning is associated with disease stages and dementia in parkinson's disease," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 84, no. 8, pp. 875–882, 2013, ISSN: 0022-3050. DOI: 10.1136/jnnp-2012-304126. eprint: https://jnnp.bmj.com/content/84/8/875.full.pdf. [Online]. Available: https://jnnp.bmj.com/content/84/8/875.

[87] V. A. Magnotta, "Quantitative in vivo measurement of gyrification in the human brain: Changes associated with aging," *Cerebral Cortex*, vol. 9, no. 2, pp. 151–160, Mar. 1999, ISSN: 1460-2199. DOI: 10.1093/cercor/9.2.151.

[88] M. Deppe et al., "Increased cortical curvature reflects white matter atrophy in individual patients with early multiple sclerosis," *NeuroImage: Clinical*, vol. 6, pp. 475–487, 2014, ISSN: 2213-1582. DOI: 10.1016/j.nicl.2014.02.012.

[89] S. Lee et al., "Regional volume analysis of the parkinson disease brain in early disease stage: Gray matter, white matter, striatum, and thalamus," *American Journal of Neuroradiology*, vol. 32, no. 4, pp. 682–687, Feb. 2011, ISSN: 1936-959X. DOI: 10.3174/ajnr.a2372.

[90] T. Jubault et al., "Patterns of cortical thickness and surface area in early parkinson's disease," *NeuroImage*, vol. 55, no. 2, pp. 462–467, Mar. 2011, ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2010.12.043.

[91] R. de la Fuente-Fernández, "Frontostriatal cognitive staging in parkinson's disease," *Parkinson's Disease*, vol. 2012, pp. 1–8, 2012, ISSN: 2042-0080. DOI: 10.1155/2012/561046.

[92] L. Li, B. Ji, T. Zhao, X. Cui, J. Chen, and Z. Wang, "The structural changes of gray matter in parkinson disease patients with mild cognitive impairments," *PLOS ONE*, vol. 17, no. 7, E. Bartolini, Ed., e0269787, Jul. 2022, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0269787.

[93] L. Pelizzari et al., "Parietal perfusion alterations in parkinson's disease patients without dementia," *Frontiers in Neurology*, vol. 11, Jun. 2020, ISSN: 1664-2295. DOI: 10.3389/fneur.2020.00562.

[94] M. Barresi et al., "Evaluation of olfactory dysfunction in neurodegenerative diseases," *Journal of the Neurological Sciences*, vol. 323, no. 1–2, pp. 16–24, Dec. 2012, ISSN: 0022-510X. DOI: 10.1016/j.jns.2012.08.028.

[95] R. L. Doty, "Olfactory dysfunction in parkinson disease," *Nature Reviews Neurology*, vol. 8, no. 6, pp. 329–339, May 2012, ISSN: 1759-4766. DOI: 10.1038/nrneurol.2012.80.

[96] K. T. E. Olde Dubbelink et al., "Disrupted brain network topology in parkinson's disease: A longitudinal magnetoencephalography study," *Brain*, vol. 137, no. 1, pp. 197–207, Nov. 2013, ISSN: 0006-8950. DOI: 10.1093/brain/awt316.

[97] J. Carp, J. Park, A. Hebrank, D. C. Park, and T. A. Polk, "Age-related neural dedifferentiation in the motor system," *PLoS ONE*, vol. 6, no. 12, N. Wenderoth, Ed., e29411, Dec. 2011, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0029411.

[98] W. de Haan, K. Mott, E. C. W. van Straaten, P. Scheltens, and C. J. Stam, "Activity dependent degeneration explains hub vulnerability in alzheimer's disease," *PLoS Computational Biology*, vol. 8, no. 8, O. Sporns, Ed., e1002582, Aug. 2012, ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1002582.

[99] E. Tolosa, M. Vila, C. Klein, and O. Rascol, "Lrrk2 in parkinson disease: Challenges of clinical trials," *Nature Reviews Neurology*, vol. 16, no. 2, pp. 97–107, Jan. 2020, ISSN: 1759-4766. DOI: 10.1038/s41582-019-0301-2.

[100] D. R. Alessi and E. Sammler, "Lrrk2 kinase in parkinson's disease," *Science*, vol. 360, no. 6384, pp. 36–37, Apr. 2018, ISSN: 1095-9203. DOI: 10.1126/science.aar5683.

[101] R. C. Helmich et al., "Reorganization of corticostriatal circuits in healthy g2019s lrrk2 carriers," *Neurology*, vol. 84, no. 4, pp. 399–406, Jan. 2015, ISSN: 1526-632X. DOI: 10.1212/wnl.0000000000001189.

[102] Y. S. Kim and T. H. Joh, "Microglia, major player in the brain inflammation: Their roles in the pathogenesis of parkinson's disease," *Experimental and Molecular Medicine*, vol. 38, no. 4, pp. 333–347, Aug. 2006, ISSN: 2092-6413. DOI: 10.1038/emm.2006.40.

[103] C. M. Long-Smith, A. M. Sullivan, and Y. M. Nolan, "The influence of microglia on the pathogenesis of parkinson's disease," *Progress in Neurobiology*, vol. 89, no. 3, pp. 277–287, Nov. 2009, ISSN: 0301-0082. DOI: 10.1016/j.pneurobio.2009.08.001.

[104] W. M. Caudle and J. Zhang, "Glutamate, excitotoxicity, and programmed cell death in parkinson disease," *Experimental Neurology*, vol. 220, no. 2, pp. 230–233, Dec. 2009, ISSN: 0014-4886. DOI: 10.1016/j.expneurol.2009.09.027.

[105] A. Reeve, E. Simcox, and D. Turnbull, "Ageing and parkinson's disease: Why is advancing age the biggest risk factor?" *Ageing Research Reviews*, vol. 14, pp. 19–30, Mar. 2014, ISSN: 1568-1637. DOI: 10.1016/j.arr.2014.01.004.