6-20-2023 3:00 PM

Longitudinal dynamics of cerebrospinal fluid Aβ, pTau and sTREM2 reveal predictive preclinical trajectories of Alzheimer's pathology

Bahaaldin Helal, *Western University*

Supervisor: Schmitz, Taylor W., *The University of Western Ontario*
Co-Supervisor: Mur, Marieke, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Engineering Science degree in Biomedical Engineering

Abstract

Alzheimer's disease (AD) is a prevalent neurodegenerative disorder with limited early-stage treatment options. There is an urgent and unmet need for accurate biomarkers which can identify patients at risk for AD before cognitive symptoms emerge. Here we compared the performance of two analytical methods, univariate and multivariate classification, for identifying cognitively normal (CN) and mild cognitive impairment (MCI) patients based on their cerebrospinal fluid (CSF) biomarkers of Aβ42, pTau-181, sTREM2. Post-hoc analyses were then employed to assess patient progression in each of the SNF clusters. We found that SNF identified subgroups within the CN and MCI cohorts, based solely on conjoint patterns of CSF, uncaptured by univariate strategies. In both CN and MCI, a fast progressor patient cluster was identified. Our findings suggest that multivariate modeling of CSF data can uncover predictive patterns of AD progression which may help to stratify patients in clinical trials of preventative therapeutics.

Keywords

Alzheimer’s disease, human, multivariate classification, similarity network fusion, longitudinal analysis, cerebrospinal fluid, biomarkers

Summary for Lay Audience

Alzheimer's disease (AD) affects over 50 million people globally, yet reliable pre-symptomatic diagnosis is lacking. Current treatments focus on symptom management during advanced stages of the disease. To address this, low-cost sensitive biomarkers are needed to identify individuals at risk before symptoms appear, akin to cholesterol levels and heart disease. Such proactive care strategies are currently absent in AD. The objective of my thesis is to evaluate whether warning signals in another type of bodily tissue, known as cerebrospinal fluid (CSF), can alert clinicians to risk for future AD, even when the patient has no signs of dementia. CSF is a liquid that bathes the brain and spinal cord and can be reliably and safely collected from the spine. It has proven to be one of the most effective tools for early detection, because it is in direct contact with the brain making it highly sensitive to abnormal biological processes in the brain. Biological markers, known as biomarkers, can be extracted from the CSF and measured to give us insight into the biological processes occurring in the brain. Over time, these biomarkers interact in ways that can be compared to the performance of an orchestra. In an orchestra, each musician (CSF biomarker) contributes to the overall performance (disease), and the complexity of the music comes from their coordinated interaction. Similarly, understanding the dynamics of AD involves considering each biomarker not just individually, but also in their interplay over time.

Previous research analyzed these biomarkers individually and at one point in time. In this thesis, we employed a new technique to understand this 'biomarker orchestra' by tracking multiple CSF biomarkers together over the course of multiple years. With this approach,

we identified a pattern of biomarkers that can determine patients who are at high risk of developing AD, even before clinical symptoms emerge. We hope these analysis techniques can be used in combination with CSF biomarkers to identify high risk cognitively normal older adults who can then be enrolled in clinical trials to evaluate drugs for slowing or even preventing AD progression, akin to strategies in the fields of cardiovascular medicine.

Acknowledgments

I want to express my utmost gratitude to my supervisors, Dr. Taylor Schmitz and Dr. Marieke Mur, for shaping my research interest, continued mentorship, and guiding me to work diligently throughout the past two years. Without their patience, consideration, guidance, and support, this work would not have been possible.

I would like to extend my thanks to my committee members, Dr. Ali Khan, Dr. Elizabeth Finger, and Dr. Stefan Everling, whose feedback and support was valuable.

I am grateful for all my colleagues that I shared lab meetings and discussions with. Special thanks to Hayley Shanks and Kate Onuska for always making time to encourage, support, and assist me throughout this work.

To my colleague and friend, Alaa Taha, thank you for your unwavering support and company over the years. I want to express my deepest appreciation to my friends, Hamza and Mohammad, for always being there for me.

Last but not least, I want to dedicate this work to my parents, Mohammed Noor and Halima, who despite being thousands of miles away have never failed to support me and instilled in me the resilience to overcome the many obstacles I have faced in my life.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations and Nomenclature

AD	Alzheimer’s disease
CN	cognitively normal
MCI	mild cognitive impairment
CSF	cerebrospinal fluid
Aβ	amyloid-beta
pTau	phosphorylated tau
TREM2	triggering receptor expressed on myeloid cells-2
APOE	apolipoprotein E
MRI	magnetic resonance imaging
PET	positron emission tomography
SNF	similarity network fusion
APC	annual percent change
ADNI	Alzheimer’s Disease Neuroimaging Initiative

1 Introduction

1.1 Biomarkers for preclinical Alzheimer’s disease: an urgent and unmet need

Alzheimer's disease (AD) is the most common neurodegenerative disorder and form of dementia in Canada, accounting for 60-80% of all cases parahippocamp. It is estimated that the global prevalence of dementia is 57.4 million cases in 2019 and expected to increase to 152.8 million cases by 2050 (Nichols et al. 2022). AD is not only a significant health concern but also has significant economic implications. In 2010, the global economic impact of AD was estimated to be around 604 billion US dollars (Wimo et al. 2013). Due to the considerable global and societal burden of AD, an estimated $42.5 billion US dollars have been invested into the research and development of AD drugs (Cummings et al. 2022).

While multiple therapies have been developed to manage symptoms and temporarily slow the progression of AD, there is a lack of efficacious drugs that can prevent or cure the underlying neurodegenerative disease process (Livingston et al. 2017; Atri 2019; Cooper et al. 2013). One potential explanation for the limited effectiveness of drugs for AD is the design of clinical trials. Typically, these trials assess drugs in advanced-stage AD patients where irreversible brain damage has occurred (Figure 1.1). By this stage, the neurodegenerative processes have already caused substantial harm, making it difficult for drugs to have a significant impact.

A promising strategy for overcoming the inefficacy of current AD therapeutics is to develop biomarkers which capture AD pathology in its earliest stages, before the onset of clinical symptoms. These biomarkers can then be used to stratify patients in clinical trials according to their AD risk (e.g., low versus high risk). Drugs which target the preclinical stages of pathology are more likely to be effective at slowing or even preventing AD progression. In turn, clinical trials which target patients in the preclinical stages of AD are more likely to observe treatment effects.

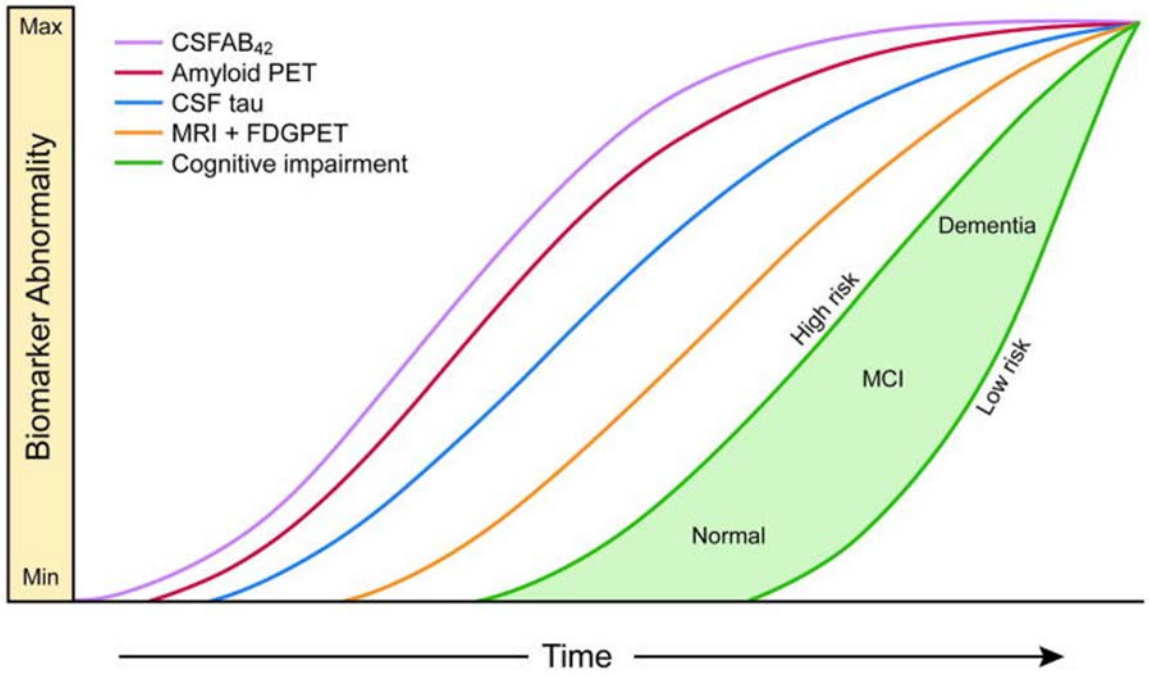

Figure 1.1: A model showing the temporal changes in biomarkers associated with Alzheimer's disease. Biomarkers increasingly become abnormal in the progression towards Alzheimer's disease (AD), following a certain temporal pattern before cognitive impairment appears. Cerebrospinal fluid (CSF) Aβ42 shown in purple and positron emission tomography (PET) amyloid imaging shown in red represent amyloid-beta (Aβ). CSF tau is shown in blue. 18F-fluorodeoxyglucose (FDG) PET paired with structural magnetic resonance imaging (MRI) measures neurodegeneration and is shown in orange. Cognitive function is shown as a green area with borders denoting low and high risk, as function deteriorates from normal to mild cognitive impairment (MCI) to dementia. Figure adapted from Jack et al. 2013

To test the sensitivity and specificity of potential preclinical AD biomarkers, we need (1) large sample datasets where multiple different biomarkers are collected in parallel across multiple different points in time, and (2) data sampled from cognitively normal (CN) older adults who may differ in their risk status but are otherwise clinically indistinguishable from one another. Large longitudinal multimodal open-access datasets,

specifically the Alzheimer's Disease Neuroimaging Initiative (ADNI), have recruited hundreds of CN (i.e., preclinical) and symptomatic older adults and collected years of multimodal longitudinal data, such as cerebrospinal fluid (CSF) proteins, Magnetic Resonance Imaging (MRI) markers of structural degeneration, and Positron Emission Tomography (PET) markers of amyloid-β (Aβ) and tau (R. C. Petersen et al. 2010). These extensive data have made it possible to study these biomarkers, and literature has shown that many of these biomarkers become abnormal long before clinical symptoms emerge (Fagan et al. 2014; Gustafson et al. 2007; Saykin et al. 2006; Yang Wang et al. 2013; Wolfsgruber et al. 2019; Pascoal et al. 2018).

Of these multimodal biomarkers, CSF has been shown to exhibit the highest sensitivity and specificity to preclinical AD pathology (Blennow et al. 2015; Blennow and Hampel 2003; Palmqvist, Mattsson, and Hansson 2016; Schmand, Huizenga, and van Gool 2010; Hansson et al. 2006). The CSF is in direct contact with the extracellular space of the brain and provides a fluid medium for the clearance of cellular waste products, including misfolded proteins. The CSF concentration of proteins has been shown to reflect multiple different types of pathology related to AD, and to dynamically change with disease progression (Brydon et al. 1995; 1996; Blennow et al. 2010). The core CSF biomarkers associated with AD include amyloid-β-42 (Aβ42) and phosphorylated-tau-181 (pTau-181) (Blennow et al. 2010). CSF Aβ42 reflects Aβ plaques in the brain, which are a hallmark of AD. Amyloid plaques are aggregates of beta-amyloid protein that accumulate extracellularly and interfere with normal brain function. CSF pTau-181 reflects the accumulation of intracellular neurofibrillary tangles in the brain, another hallmark pathology of AD. These two biomarkers are the best characterized biological fluid

markers for AD (Zou, Abdullah, and Michikawa 2020). In addition to the two core pathologies of extracellular Aβ and intracellular phosphorylated-tau, emerging evidence is pointing towards chronic neuroinflammation as a third core pathology of AD (Zotova et al. 2010). This inflammatory response is effectuated by the immune system, with the major player being microglia, the resident immune cells in the central nervous system. In early AD, neurotoxic Aβ plaques are thought to drive the activation of microglia, which subsequently migrate and phagocytose Aβ plaques, resulting in increased clearance of Aβ (Kinney et al. 2018; Baik et al. 2016; Simard et al. 2006; Yuyama et al. 2012; Chakrabarty et al. 2010; Shaftel et al. 2007). As the inflammatory response is sustained however, the microglia become enlarged and less able to clear Aβ (Chakrabarty et al. 2010; Shaftel et al. 2007; Hickman, Allison, and El Khoury 2008; Bard et al. 2000). The action of Triggering Receptor Expressed on Myeloid Cells-2 (TREM2), a receptor protein on microglia, is vital for the many functions of the microglia (Takahashi, Rochford, and Neumann 2005; N'Diaye et al. 2009; Atagi et al. 2015; Yeh et al. 2016; Kleinberger et al. 2014).

Collectively, CSF Aβ, tau and TREM2 proteins exhibit strong potential as biomarkers for preclinical AD. The collection of CSF samples also gives researchers the ability to (1) assay multiple diverse types of proteins and (2) perform within-subject longitudinal assessments. However, these advantages have created two major technical obstacles. The first obstacle concerns how the multiple CSF proteins are statistically analyzed relative to one another. Most studies typically use univariate strategies, where each protein is treated as a separate variable (Blennow et al. 2019; Salvadó et al. 2023; Hansson et al. 2007; Snider et al. 2009; Adamczuk et al. 2015; Shoji et al. 1998). The second obstacle

concerns whether and how the CSF proteins are analyzed relative to their own longitudinal changes over time. Most studies typically use cross-sectional experimental designs, where each protein is measured at only a single timepoint (Blennow et al. 2019; Salvadó et al. 2023; Hansson et al. 2007; Snider et al. 2009; Adamczuk et al. 2015; Shoji et al. 1998). To provide a categorical boundary for statistical comparison of CSF pathology in a cross-sectional design, patients must be grouped according to an independent risk variable, such as cognitive status, apolipoprotein E (APOE) genotype or amyloid PET abnormality. Hence, univariate statistical analysis and cross-sectional designs necessitate multiple parallel statistical comparisons to identify how the central tendency, variability, and distribution of each individual protein may differ between groups of patients who may also differ on many other characteristics, which are difficult to counterbalance. These strategies also do not provide insight into the dynamic interrelationships among AD pathologies, which are known to reflect a ‘cascade’ model of progression (Hadjichrysanthou et al. 2020; Jack et al. 2013; Palmqvist et al. 2019). Under the cascade model, CSF biomarkers do not all become abnormal simultaneously, but rather change in a predictive multivariate pattern as the brain’s compensatory mechanisms systematically give way to pathology and neurodegeneration. A univariate cross-sectional comparison of a single CSF protein would obscure these dynamic multivariate patterns. The pervasive use of univariate statistical analysis and cross-sectional experimental designs may account for the substantial variability in the research literature as to which CSF biomarker, Aβ, tau or TREM2, exhibits superior sensitivity and specificity to pathology in preclinical stages. Even in longitudinal studies, univariate

strategies are typically implemented to assess each CSF protein in parallel (Brys et al. 2009; Hansson et al. 2007; Mattsson et al. 2009).

To overcome these obstacles, there is a need for analytical strategies which capture both the dynamic (biomarker levels changing over time) and multifactor complexity (biomarkers interacting with each other) of CSF protein biomarkers. In this work, we therefore explored a novel multivariate strategy for analyzing dynamic longitudinal patterns in CSF Aβ, tau and TREM2 data collected in a large, well-characterized sample of older adults who were CN at baseline. A key feature of this analytical strategy is that it obviates any need for arbitrary selection of a separate grouping factor, such as *APOE-ε4* or cognitive status. Rather, the groups emerge from the multivariate longitudinal patterns detected in the CSF, providing an unbiased stratification of preclinical AD risk.

1.2 Background

1.2.1 Preclinical Alzheimer's disease pathophysiology

The pathology of preclinical AD is typically characterized by the accumulation of abnormal proteins in the brain, including Aβ plaques and tau tangles. Aβ is a peptide that forms from the breakdown of a larger protein called amyloid precursor protein (APP), a single transmembrane protein that is highly concentrated in the synapses of neurons in the brain (Haass and Selkoe 2007). Once produced, Aβ can form soluble oligomers which can aggregate to form larger insoluble Aβ plaques (Golde, Eckman, and Younkin 2000). These Aβ plaques are dense clumps of protein that accumulate extracellularly in the brain. In healthy individuals, Aβ is cleared from the brain via several mechanisms,

namely enzymatic degradation, cellular uptake, and drainage outside the brain (Iwata et al. 2001; Eckman et al. 2006; Mawuenyega et al. 2010b; Rasmussen, Mestre, and Nedergaard 2018). However, in AD patients, an imbalance between Aβ production and clearance results in larger amounts of Aβ plaques, ultimately leading to the activation of neuroinflammatory response mechanisms and damage to nearby cells (Mawuenyega et al. 2010a; Selkoe 2003; Zlokovic et al. 2000; Hardy and Selkoe 2002). Aβ deposition is observed in both sporadic and familial AD (FAD) (O'Brien and Wong 2011). FAD is associated with mutations in three genes: APP, Presenilin 1 (PSEN1) and Presenilin 2 (PSEN2) (Duff et al. 1996, 1; Levy-Lahad et al. 1995; Sherrington et al. 1995). The presenilin's are catalytic components of the γ-secretase enzyme complex, which cleaves APP to produce Aβ (Karch and Goate 2015). Mutations in these genes leads to a pathogenic increase in the production of the longer and more amyloidogenic form of Aβ called Aβ42, typically resulting in earlier onset (<65 years of age) of AD in individuals carrying these genes (Bettens, Sleegers, and Van Broeckhoven 2010). Early onset AD, comprising both familial and sporadic AD subtypes, is estimated to represent 1-5% of all AD cases (Zhu et al. 2015). The remaining number of cases are sporadic late onset, which affect individuals aged 65 or older and have age as the single most important risk factor (Van Der Flier 2005). Outside of subtype specific AD risk factors, the APOE-ε4 allele is the most important genetic risk factor for late onset AD, where heterozygotes and homozygotes have an increased risk of 2-4 times and 8-12 times, revealing a dose-dependent effect (Lindsay A. Farrer et al. 1997). APOE-ε4 is thought to impair processing and clearance of amyloid, differentially impacting Aβ clearance from the brain (Verghese et al. 2013; Huynh et al. 2017; Kim, Basak, and Holtzman 2009).

Tau is an essential protein encoded by the microtubule associated protein tau (MAPT) gene that normally helps to stabilize microtubules within neurons to facilitate the transportation of nutrients and other materials across the cell. In AD, tau becomes abnormally phosphorylated, leading to the misfolding and aggregation of phosphorylated tau (pTau) to form intracellular neurofibrillary tangles, which can disrupt the normal functioning of neurons.

In AD, tau follows a characteristic spatiotemporal pattern when spreading through the brain. This stereotyped pattern of progression can be subdivided into six Braak stages. Stages I-II represent preclinical AD where tangles are present in the transentorhinal cortex. Stages III-IV correspond to incipient AD where tangles are present in the limbic regions. Late AD where tau pathology has spread to the neocortex is represented by staged V-VI (Braak and Braak 1991).

This pathogenic cascade of tau hyperphosphorylation followed by misfolding and spreading throughout the cortex is thought to be triggered by Aβ (Musiek and Holtzman 2015). Several studies provide evidence for the complex interplay between Aβ and tau (Bennett et al. 2017; Israel et al. 2012; Oddo et al. 2003; Lee et al. 2016; Choi et al. 2014). Tau pathology is more closely associated with cognitive decline and disease severity compared to amyloid burden (Spires-Jones and Hyman 2014; Nelson et al. 2012). Unlike Aβ, no mutations in MAPT are considered causative of familial AD. However, different MAPT gene variants could confer increased or decreased risk of AD (Schraen-Maschke et al. 2004; Gerrish et al. 2012; Allen et al. 2014).

In addition to the two core pathologies of extracellular Aβ and intracellular phosphorylated-tau, emerging evidence is pointing towards chronic neuroinflammation as a third core pathology of AD (Zotova et al. 2010). This inflammatory response is effectuated by the immune system, with the major player being microglia, the resident immune cells in the central nervous system. In their inactive state, microglia are responsible for surveying the brain and communicating with other cells. Microglia become activated when a threat is detected, such as pathogen or injury. This activation process comprises morphological changes including cell enlargement and migration (Bolmont et al. 2008; Graeber et al. 1988; Mrak 2012). In early AD, neurotoxic Aβ plaques are thought to drive the activation of microglia, which subsequently migrate and phagocytose Aβ plaques, resulting in increased clearance of Aβ (Kinney et al. 2018; Baik et al. 2016; Simard et al. 2006; Yuyama et al. 2012; Chakrabarty et al. 2010; Shaftel et al. 2007). As the inflammatory response is sustained, the microglia become enlarged and less able to clear Aβ (Chakrabarty et al. 2010; Shaftel et al. 2007; Hickman, Allison, and El Khoury 2008; Bard et al. 2000). This results in exacerbation of AD pathology due to the compromised clearance of Aβ, coupled with the continued release of pro-inflammatory cytokines, which acts to aggravate neurodegeneration (Hickman, Allison, and El Khoury 2008; Meda et al. 1995; Sheng et al. 1998; Krabbe et al. 2013; Michelucci et al. 2009).

Emerging lines of evidence that highlight the importance of the immune response carried out by microglia in AD pathogenesis come from genetic studies that reported the R47H variant of the TREM2 gene to be the second strongest genetic risk factor, conferring approximately a threefold increase risk of AD (Guerreiro et al. 2013; Jonsson et al.

2013). TREM2 is a receptor protein expressed on myeloid lineage cells, including macrophages, dendritic cells, osteoclasts, and most importantly for AD, microglia (Colonna 2003). TREM2 is a single transmembrane glycoprotein that consists of an extracellular immunoglobulin-like domain, a transmembrane domain, and a short, cytoplasmic tail (Forabosco et al. 2013). Soluble TREM2 (sTREM2), which is present in the blood and CSF, is produced from proteolytic shedding of the receptor ectodomain or translation of an alternative spliced TREM2 transcript that lacks the transmembrane domain (Del-Aguila et al. 2019; Piccio et al. 2008; Wunderlich et al. 2013).

Microglia are responsible for engulfing and clearing debris to maintain homeostasis. TREM2 is critical for microglial phagocytosis of apoptotic neurons, cellular debris, bacteria, lipoproteins, and Aβ (Takahashi, Rochford, and Neumann 2005; N'Diaye et al. 2009; Atagi et al. 2015; Yeh et al. 2016; Kleinberger et al. 2014). When Aβ plaques are complexed with APOE, uptake by microglia is enhanced through TREM2 interaction with the lipoproteins (Yeh et al. 2016). Furthermore, TREM2 plays a role in modulating the proliferation and survival of microglia. When TREM2 is reduced or absent, microglial proliferation appears to decrease (Yaming Wang et al. 2016). Additionally, TREM2 is a key factor in promoting survival of microglia (Zheng et al. 2017).

1.2.2 Biomarkers of AD pathology

The intricate network of pathological events and molecular alterations manifesting in preclinical AD, involving the accumulation of Aβ plaques and tau tangles as well as chronic neuroinflammation, plays a critical role in the disease's progression.

Biomarkers provide measurable and objective characteristics that indicate the presence or progress of pathological processes, such as Aβ, or a response to a therapeutic intervention (Biomarkers Definitions Working Group. 2001). For AD, biomarkers may be used to improve the accuracy of diagnosis, predict conversion between clinical disease states, and monitor the effectiveness of treatment (Blennow et al. 2010; Mahaman et al. 2022).

1.2.3 Imaging biomarkers of AD pathology

Imaging biomarkers have played a significant role in elucidating the pathology of early AD. The most widely used imaging techniques include MRI and PET, which can quantify structural, functional, and molecular measures of the disease's pathology.

MRI enables the quantification of structural alterations, like the shrinkage in cortical volume or thickness in specific brain regions. In AD, MRI studies reveal a pattern of brain atrophy, beginning in the entorhinal cortex, followed by the hippocampus, amygdala, and parahippocampal gyrus (Lehéricy et al. 1994; Chan et al. 2001; Dickerson et al. 2001; Killiany et al. 2002). As such, atrophy of the medial temporal lobe structures is considered to be a well-established MRI marker of AD (Juottonen et al. 1998; Du et al. 2001; Duara et al. 2008; Frisoni et al. 2010). It also possesses predictive potential, indicating the possible transition from mild cognitive impairment (MCI) to AD (DeCarli et al. 2007; Korf et al. 2004).

PET imaging leverages specific radioactive tracers to generate images of different molecular and pathological processes in the brain. Three main types of PET have become a mainstay in AD research: [^{18}F]Fluorodeoxyglucose (FDG-PET), amyloid PET, and tau

PET. FDG-PET is a measure of glucose metabolism and has demonstrated prognostic value in identifying individuals at risk of progression from MCI to AD (Teipel et al. 2015). Amyloid and tau PET have emerged with the advent of new PET tracers and have allowed us to visualize in vivo the fibrillar Aβ plaques and tau neurofibrillary tangles, respectively. Amyloid PET serves to confirm AD diagnosis, assists in the early detection of MCI patients, and track changes in brain amyloid load due to anti-amyloid therapy (Suppiah, Didier, and Vinjamuri 2019). Tau PET, while still in relatively early stages, has shown utility in the diagnosis and staging of AD diagnosis (G. C. Petersen et al. 2022).

Imaging biomarkers, such as MRI and PET have played a crucial role in the detection and assessment of AD. However, they also have certain limitations. First, imaging biomarkers may not be sensitive enough to detect early pathological changes in the brain during the preclinical stages of AD. Instead, these imaging biomarkers may primarily capture the effects of disease progression at later stages, wherein symptoms are already evident. Second, the interpretation of MRI and PET images requires expertise and trained personnel. The identification and quantification of specific biomarkers or pathological features requires many analytical steps which may involve subjective judgments. Standardization and consensus on imaging protocols and interpretation criteria are crucial to ensure reliability and consistency across different centers, but remain lacking for MRI and PET. Lastly, MRI and PET imaging techniques are also expensive and may not be readily available in all healthcare settings. Limited access to these imaging technologies can restrict their widespread use, especially in resource-limited or remote areas.

1.2.4 Fluid biomarkers of AD pathology

Fluid biomarkers comprise CSF and plasma. CSF has been studied extensively and is regarded as the most optimal tool for detecting preclinical AD pathology. First, since the CSF is in direct contact with the extracellular space of the brain, its concentration of proteins has been shown to reflect multiple different types of pathology *in the brain* that are related to AD and to dynamically change with disease progression (Brydon et al. 1995; 1996; Blennow et al. 2010). Second, changes in CSF biomarkers are observed earlier than in MRI or PET (Schmand, Huizenga, and van Gool 2010; Palmqvist, Mattsson, and Hansson 2016). Third, a variety of proteins, each of which reflects a unique pathology, can be measured from a single CSF acquisition, including the core pathologies: Aβ42, pTau-181, and sTREM2.

CSF is typically extracted through a procedure called lumbar puncture (LP) or spinal tap. During this procedure, a healthcare provider will insert a thin, hollow needle into the lower back, between the vertebrae in the lumbar region. The needle is inserted into the subarachnoid space, which is filled with CSF, and a small amount of the fluid is collected for testing. LPs are considered safe, while serious post-LP complications are rare. More common complications following LPs include headache and back pain (Duits et al. 2016). Unlike the expensive infrastructure required for PET and MRI, CSF acquisition requires only a trained nurse and inpatient clinical care. CSF is therefore considerably more accessible than imaging biomarkers in real world clinical settings.

After the CSF aliquot is obtained, it can do undergo different measurement techniques to quantify the biomarkers of interest. Most commonly, an antibody-based approach is

implemented, such as the enzyme-linked immunosorbent assay (ELISA) or electro-chemiluminescence immunoassays (ECLIA).

ELISA is a biochemical technique that uses an enzyme-linked antibody to detect the presence of a specific protein or antigen in a sample. In ELISA, a specific antigen is immobilized on a surface (such as a microplate well) and a corresponding antibody linked to an enzyme (such as alkaline phosphatase or horseradish peroxidase) is added to the sample. If the antigen is present in the sample, it will bind to the antibody, forming an antigen-antibody complex. After washing away any unbound substances, a substrate specific to the enzyme is added to the well. If the enzyme is present, it will catalyze a colorimetric reaction of the substrate, producing a detectable signal such as a color change, which indicates the presence and amount of the antigen in the sample.

One significant early challenge for CSF biomarkers of AD pathology was their variability in the measured protein concentrations, which was observed to differ between different labs and even between repeated analyses of the same samples (Blennow and Hampel 2003; Olsson et al. 2005; Hort et al. 2010). This variability may be due to differences in pre-analytical factors such as sample handling and storage, analytical factors such as between/within differences in laboratory procedures, and assay-related factors such as variations in kit manufacturing (Hok-A-Hin et al. 2019; Teunissen et al. 2009; Andreasson et al. 2018; Teunissen et al. 2010; Fourier et al. 2015). This variability in pre-analytical techniques was identified in multicenter studies such as ADNI and the Swedish BIOFINDER study, leading to international efforts to standardize CSF procedures (Hort et al. 2010). To address the issues, Roche Diagnostics developed the CSF Elecsys assays, which are fully automated ECLIAs that have been shown to have high analytical

performance, reliability, and sensitivity (Bittner et al. 2016; Shaw et al. 2019; Doecke et al. 2020).

In ECLIA, instead of using an enzyme to generate a colorimetric signal, a proprietary Sulfo-Tag reporter molecule is used (Bolton et al. 2020). The Sulfo-Tag is a compound that produces a chemiluminescent signal when an electrical current is applied. A sample is added to the assay plate, and if it contains the antigen of interest, it will bind to the antigen coated on the plate. A secondary antibody that recognizes the first antibody is added, which is linked to the Sulfo-Tag. The assay plate is then placed in a reader, which applies an electric pulse to the plate, triggering a reaction that produces a chemiluminescent signal. The signal is detected and measured by a camera in the reader, and the amount of signal is proportional to the amount of antigen-antibody complex in the sample.

ECLIA has several advantages over ELISA, including a wider dynamic range, higher sensitivity, and better precision (Bolton et al. 2020). In addition, the ECLIA method eliminates the potential for variability associated with timing of substrate addition, as the substrate is activated by the reader (Bolton et al. 2020). These rigorous methodological features have led to standardization of CSF cutpoints on multiple AD pathological markers, including Aβ42 and pTau-181. Several studies have evaluated the concordance between CSF Aβ42 and pTau-181 as measured by Elecsys assays in different populations, including ADNI and the Swedish BIOFINDER studies and converged on a common cut-off (Hansson et al. 2018; Shaw et al. 2018; Doecke et al. 2020). The Elecsys assay has since been adopted by leading studies, including ADNI, the Swedish BioFINDER Study, and the Australian Imaging, Biomarkers and Lifestyle (AIBL) study

(Doecke et al. 2020; Lifke et al. 2019; Shaw et al. 2018). Therefore, evidence using these new methods supports the reliability issues of CSF have been largely addressed, opening the door to its utility as a preclinical biomarker.

1.2.5 Cross-sectional univariate analysis of CSF biomarkers

The vast majority of research using CSF to study AD pathology has employed cross-sectional univariate comparisons, where single timepoint measurements of a CSF biomarker are used to classify a group of control and AD patients. A quintessential example of this approach is the usage of receiver operating characteristic (ROC) analysis to assess the diagnostic accuracy of CSF biomarkers for AD. An ROC curve is a graphical representation of the performance of a binary classifier as the discrimination threshold is varied. In the context of disease diagnosis, the classifier assigns a patient to either a positive or negative category based on a biomarker value. Consequently, the ROC curve charts the true positive rate (sensitivity) against the false positive rate (1-specificity) at varying thresholds.

In the paper by Hansson et al., the authors used ROC analysis to evaluate the diagnostic accuracy of CSF biomarkers (Aβ42, t-tau and pTau-181) for identifying AD pathology in two independent cohorts of patients (Hansson et al. 2018). The binary classifier was set to be amyloid-β PET scans of patients, which were labelled to be positive or negative based on blinded visual assessment. The sensitivity and specificity of the different CSF biomarkers at different cut-off points would be plotted on the ROC curve, and the

optimal cut-off point would be chosen based on the point on the curve that maximizes the sum of sensitivity and specificity.

The univariate approach to biomarker analysis, which typically involves the establishment of a cut-off point for a single biomarker or the ratio of two biomarkers, offers an array of benefits. Its ease of computation and interpretation has undoubtedly contributed to its widespread usage. This approach has seen successes in increasing the diagnostic accuracy of AD and, on occasion, predicting progression from MCI to AD (Lewczuk et al. 2004; Maddalena et al. 2003; Koopman et al. 2009).

However, the inherent limitations of this univariate, cut-off-based strategy necessitate caution in its application and interpretation. It assumes a linear relationship between biomarkers and disease outcomes, whereby higher or lower biomarker values correspond linearly to specific outcomes. Yet, this assumption overlooks the potential for non-linearity in these relationships and fails to account for possible interactions among different biomarkers (Toledo et al. 2015; Popescu et al. 2020; de Leon et al. 2018).

Consider an analogy of an orchestra to illustrate these interactions and their significance. In an orchestra, each musician contributes to the overall performance. Listening to individual musicians or instruments may provide some insight, but the true beauty and complexity of the performance come from the coordinated and harmonious interaction of all the musicians. Similarly, understanding the 'orchestra' of biomarkers involved in AD necessitates not only looking at each biomarker individually (univariate) but also considering their interactions (multivariate) and monitoring them over time (longitudinal).

Another fundamental limitation of the univariate approach is its attempt to apply a 'one size fits all' model in a disease context that is far from homogenous. Establishing a cut-off point from a general cohort of patients and then applying this globally overlooks the substantial variability in the manifestation and progression of AD (Jellinger 2022; Birkenbihl et al. 2022; Ferrari and Sorbi 2021; Badhwar et al. 2020). Patients may vary in their rate of progression, the severity of symptoms, and the specific cognitive domains affected. Additionally, factors like age, sex, genetics, and comorbidities can significantly influence disease progression and the associated biomarker profiles.

Ignoring this heterogeneity can lead to misclassifications and reduced diagnostic accuracy. It can also stifle the development of personalized therapeutic approaches and complicate drug discovery and screening processes (Ferrari and Sorbi 2021). It might contribute to the mixed results of clinical trials, where the potential benefits of a treatment could be obscured when tested in a diverse group of patients, as the treatment might only be effective for a specific subpopulation.

Additionally, the cross-sectional nature of this analysis provides a snapshot in time, thereby failing to capture the temporal evolution of a disease like AD, characterized by its lengthy preclinical phase and gradual changes over time. This limitation may lead to missing crucial alterations in biomarker levels that occur over years.

1.2.6 Longitudinal univariate analysis of CSF biomarkers

Since AD is a progressive disease, studying its pathologies longitudinally provides great insight into the dynamics of the pathophysiology of the different stages of the disease as

they manifest. However, like cross-sectional CSF studies, most longitudinal CSF studies use univariate analysis strategies (Hansson et al. 2007; Brys et al. 2009; Mattsson et al. 2009; Salvadó et al. 2023). This approach necessitates the concatenation of different univariate analyses of Aβ and tau, followed by a qualitative interpretation of their interrelationships. These studies also typically use linear analytical strategies, such as linear mixed models, linear logistic regression, and K-means cluster analysis (Vlies et al. 2009; Choo et al. 2013; Anders M. Fjell et al. 2010; Fagan et al. 2014; Alcolea et al. 2014). While these linear methods have proven useful in many contexts, they fall short in modeling the nonlinear patterns and relationships between CSF biomarkers, a critical consideration given the intricate pathophysiology of AD (de Leon et al. 2018; Popescu et al. 2020; Williams et al. 2011). The linearity assumption these methods rely upon oversimplifies the relationships among biomarkers and can lead to less accurate, incomplete representations of the disease processes.

1.2.7 Similarity Network Fusion: A multivariate strategy for examining the preclinical utility of AD CSF biomarkers

AD is a multifaceted neurodegenerative disorder, with intricate biological processes and heterogeneous manifestations that defy simple explanations or straightforward diagnostic paradigms. Given its complex nature, the study of AD calls for robust multivariate analytical methods capable of integrating multiple data types to provide comprehensive phenotyping and unravel the underlying pathophysiological processes. This is particularly pressing in the current era of big data and international collaborative data-sharing initiatives, which provide extensive and diverse datasets for exploration.

The simplest analytical strategy for integrating multi-view data is concatenation, which combines different types of data into a single data structure. This approach has its appeal, given its straightforwardness and relatively low computational requirements. However, concatenation carries notable limitations, one of which is the potential reduction in signal-to-noise ratio, leading to difficulties in identifying true patterns amidst the increased noise. Additionally, it overlooks the unique inherent structure within each data modality, effectively treating each data type as if it were the same (B. Wang et al. 2014; Markello et al. 2021).

To overcome these challenges, a range of advanced techniques have been developed to handle multimodal data. One of these techniques is known as Similarity Network Fusion (SNF) (B. Wang et al. 2014; Markello et al. 2021). In the original paper, SNF was employed to integrate multiple types of genomic data—mRNA expression, DNA methylation, and miRNA expression—obtained from the same group of cancer patients (B. Wang et al. 2014). When clustering patients on the multivariate data, SNF revealed a previously hidden patient subgroup that was not identifiable when data types were considered separately. This subtype had distinctive survival outcomes and therapeutic responses. Furthermore, SNF was demonstrated to outperform data concatenation and a method based on joint latent variable models in identifying clinically distinct clusters in all cases (Cancer Genome Atlas Research Network 2008; B. Wang et al. 2014; Markello et al. 2021). Since its inception, SNF has been employed in influential research papers with substantial impact. In another study, SNF led to the discovery of a concealed subtype in medulloblastoma cancer and stratifying pancreatic ductal adenocarcinoma tumors, further reinforcing its usefulness and effectiveness in decoding complex diseases

(Cavalli et al. 2017; Raphael et al. 2017). These examples collectively highlight the potential of SNF for revealing clinically relevant disease subgroups that might be missed by other techniques.

SNF operates by constructing networks for samples and encompasses two main steps. Initially, it establishes individual patient similarity networks for each type of data. These networks represent pairwise similarities between patients, with the similarity metrics potentially varying based on the data type being integrated. Measures such as Euclidean distance, Pearson correlation, or mutual information can serve this purpose.

Subsequently, SNF fuses these individual patient similarity networks into a single integrated data matrix through an iterative, non-linear process. This process systematically incorporates the information from each data source, ensuring that the incremental value from each additional data source builds upon the previous one. This approach allows SNF to capture both shared and complementary information from various data types.

A particularly remarkable feature of SNF is its robustness, even with a limited number of samples. It also shows high resilience to varying noise levels and data heterogeneity, which is an essential characteristic for handling the complex, multi-layered data involved in AD research. By effectively integrating multi-view data and capturing non-linear relationships among AD biomarkers, such as those derived from CSF samples, SNF can provide a more nuanced understanding of the pathophysiological processes underlying AD. Hence, its adoption could be a significant step forward in enhancing our understanding of AD and potentially advancing diagnostic and therapeutic strategies.

1.2.8 A novel implementation of SNF: Longitudinal CSF biomarkers

Thus far, patient clustering with SNF has been evaluated primarily in the context of integrating cross-sectional data. There has been less attention on evaluating patient clustering with SNF in the context of longitudinal data. One particularly interesting and unexplored area is using SNF to evaluate clustering of baseline and follow up CSF biomarkers in patients at risk for AD. That is, rather than fusing patient networks from single snapshot features collected across different data modalities, one would fuse patient networks from multiple different temporal features (e.g., baseline, relative change, etc.) collected in a single data modality. Multimodal data are expensive and technically challenging to acquire in patients. This latter 'multitemporal' strategy has the advantage of requiring fewer types of different data modalities for classifying patients, making it more feasible for patient stratification in clinical trials.

Before we can implement our multitemporal strategy of modeling CSF with SNF, we must first tackle two methodological obstacles to modelling longitudinal biomarker data. First, one must interpret relative changes in CSF protein concentrations over time, given a certain starting point in baseline pathology. Longitudinal studies often do not report whether or how participants differ at baseline CSF values, nor do they explicitly account for these potential baseline differences. Rather they analyze absolute differences (time2 – time1). This can lead to ambiguous longitudinal differences if there is a systematic bias in baseline CSF in one group compared to another. One cannot determine whether the difference is due to difference in baseline, follow-up, or both.

A second obstacle for longitudinal studies concerns missing data, which can lead to variable number of timepoints per subject, variable inter-timepoint intervals, and variation in the number among subjects in terms of the intervals between baseline and subsequent visits, as well as inconsistencies in baseline levels. These are typically not accounted for in longitudinal studies of CSF.

To overcome these two obstacles, we and others have developed the annual percent change (APC) measure, which explicitly controls for variation in baseline measures and timepoint intervals (Rechberger et al. 2022).

The APC serves as a summary measure of relative change, providing a standardized way to examine changes over time. The APC formula is shown below (x_1 and x_2 as measurements obtained at timepoints t_1 and t_2)

$$APC = \left(\left(\frac{x_2}{x_1} \right)^{\frac{365}{t_2 - t_1}} - 1 \right) \cdot 100\% \quad \textbf{(1)}$$

Typically, x_1 and x_2 correspond to the participant's baseline value and final value for a variable of interest, respectively. Similarly, timepoints t_1 and t_2 correspond to the dates of the visits during which x_1 and x_2 were measured (i.e., baseline/first visit and final visit), respectively. For example, to compute the APC of CSF Aβ42 of patient Y, we extract the CSF Aβ42 values at the first and the last visits, where the CSF Aβ42 values correspond to x_1 and x_2 and the visit dates correspond to t_1 and t_2.

Because the APC calculates the ratio of the two measurements from a given subject, with the baseline/starting value in the denominator, the APC accounts for variations in

baseline measurements between subjects. In a hypothetical scenario, a researcher intends to investigate the longitudinal CSF Aβ42 reduction in two AD patients over a 1-year study. At baseline, participant A has a CSF Aβ42 level of 1000 pg/mL, while participant B has a CSF Aβ42 level of 400 pg/mL (Figure 1.2). If their final CSF Aβ42 levels are 800 pg/mL and 200 pg/mL, respectively, both individuals have the same reduction of 200 pg/mL. However, because participant B started with a lower CSF Aβ42 level, they have experienced a 50% reduction per year compared to the 20% reduction per year experienced by participant A.

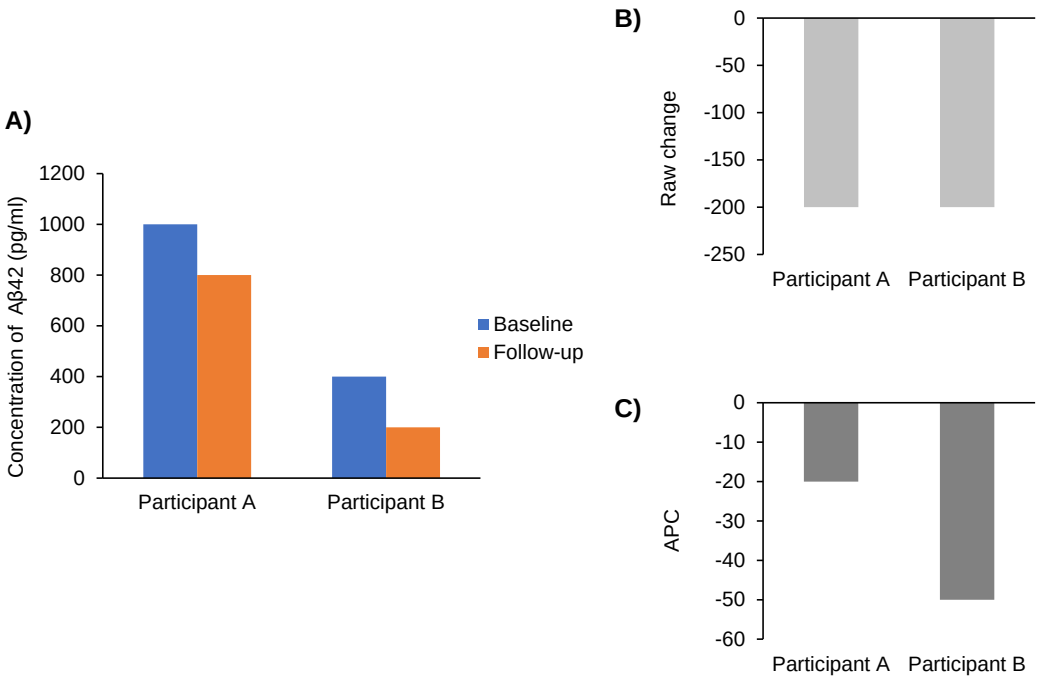

Figure 1.2: A hypothetical model demonstrating the utility of APC in controlling for baseline variations. A) Two participants with different baseline and follow-up levels of CSF Aβ42 exhibit B) equivalent reductions in the raw values, but C) APC provides a more accurate measure of relative change.

A useful mathematical property of the APC is that by removing the baseline, one obtains a pure metric of relative change in the observation window. Baseline values can (and should) then be compared as a separate complementary comparison between groups.

Furthermore, the APC includes and accounts for the time interval between the visits at which the measurements were obtained for each subject. In another hypothetical scenario, both participants A and B have a CSF Aβ42 level of 1000 pg/mL at baseline (Figure 1.3). On their final visits, both participants have a CSF Aβ42 level of 500 pg/ml. However, the time interval between the baseline and final visits for participant A is 5 years, while that of participant B is 1 year. The reduction for both participants is 500 pg/ml, however due to the longer time interval between visits for participant A, the reduction per year that participant A exhibits is 13%, while that of participant B is 50%. Additionally, because the APC computes relative change (i.e., ratio of two measurements, the APC standardizes the unit space.

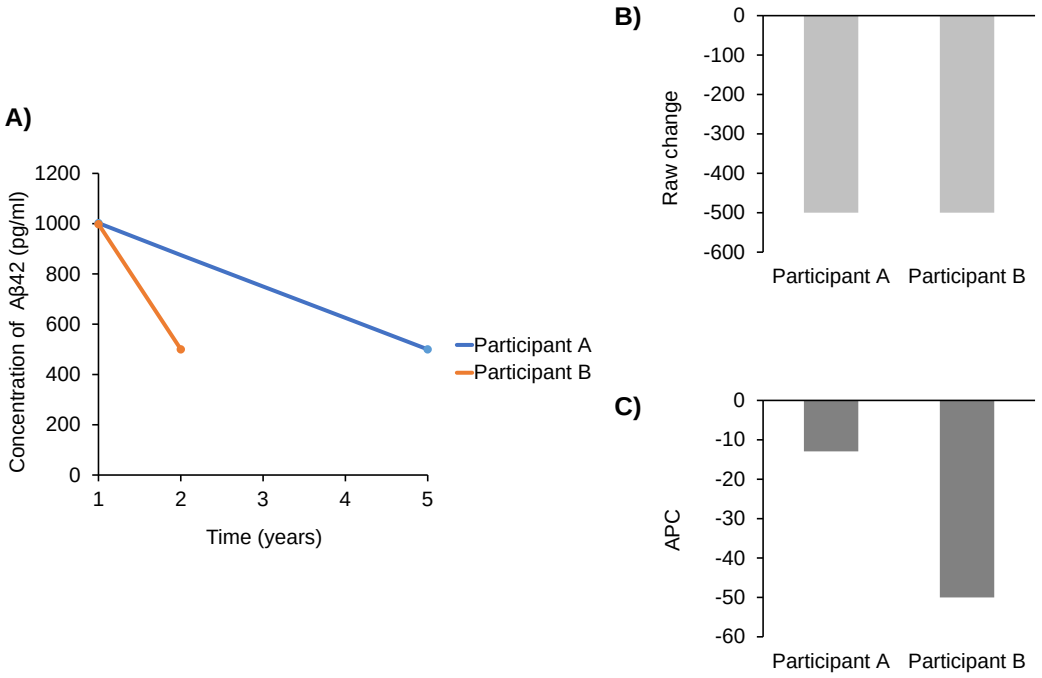

Figure 1.3: A hypothetical model demonstrating the utility of APC in controlling for differences in time intervals. A) Two participants with the same baseline and follow-up levels of CSF Aβ42 over different time periods exhibit B) equivalent reductions in the raw values, but C) APC provides a more accurate measure of relative change.

These features make the APC the ideal tool for comparing longitudinal change between subjects, where subjects typically have different baseline levels and varying time intervals, as well as between data types in a heterogenous population.

1.3 Rationale

To our knowledge, SNF has not been used to explore multitemporal measures of CSF pathology to cluster patients. In this thesis, we will therefore examine the potential for SNF to cluster CN older adults at risk for AD according to their baseline and APC

measure of CSF pathology. Given that CSF Aβ42, pTtau-181 and sTREM2 capture multidomain pathology in preclinical AD, we will focus on multitemporal modelling of all three of these CSF biomarkers in parallel. A secondary goal of this work is to demonstrate the feasibility and potential for using SNF to cluster preclinical AD patients on CSF alone, without the addition of MRI or PET imaging data, or cognitive tests. In doing so, we hope to demonstrate that SNF, derived from two timepoints of CSF, would be sufficient for a widely accessible preclinical diagnostic standard.

The open-access data in ADNI provides access to hundreds of CN, MCI, and AD participants with longitudinal measures of CSF Aβ42, pTau-181 and sTREM2. Therefore, to address the limitations of univariate and cross-sectional strategies, we will use SNF in combination with baseline and APC measures of multifactor CSF pathology (pTau-181, Aβ42 and sTREM2) to better model how these biomarkers move in tandem with longitudinal AD progression.

1.4 Objective and hypothesis

Our study encompasses two main objectives that will be investigated within two distinct cohorts: 1) CN older adults and 2) older adults with MCI symptoms using data from three CSF biomarkers (Aβ42, pTau-181, and sTREM2).

Aim 1: Compare the sensitivity and specificity of cross-sectional multivariate SNF to standard cross-sectional univariate classification approaches.

Aim 2: Evaluate SNF using a novel multitemporal strategy, which includes both baseline and longitudinal CSF data.

We hypothesize that SNF will reveal a high-risk subgroup of cognitively normal older adults based on their baseline and/or multivariate longitudinal pattern of CSF biomarkers. Additionally, we expect SNF to identify distinct subgroups within the MCI cohort depending on their risk level for AD progression. These hypotheses will be validated by post-hoc assessments to investigate potential differences in patient clusters, including the impact of primary AD risk factors such as APOE-ε4 genotype, age, sex, and multidomain markers of cognitive function.

2 Methods

2.1 Data collection

All data used in the preparation of this study were obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information, see www.adni-info.org. The study received approval from the Institutional Review Boards of all participating institutions, and informed written consent was obtained from all participants at each site R. C. Petersen et al. 2010).

Older adults were eligible to participate in ADNI if they met the following criteria: 55 to 90 years old (inclusive), Hachinski Ischemic Score ≤ 4, use of permitted medications stable for four weeks prior to screening, Geriatric Depression Scale ≤ 6, study partner with 10 hours/week of contact who can accompany the participant to the clinic, reasonable visual and auditory acuity to allow for neuropsychological assessment, good general health with no diseases precluding enrollment, six grades of education or work history equivalent, and fluent in English or Spanish. Additionally, women had to be 2 years past childbearing potential or sterile. Exclusion criteria included any other significant neurologic disease, use of specific psychoactive medications, and history of significant head trauma. More details can be found on the ADNI website (www.adni-info.org).

Participants in this study were classified as CN, MCI, or AD based on scores from a neuropsychological battery of tests including the Mini-Mental State Examination (MMSE), Clinical Dementia Rating (CDR), and Wechsler Memory Scale (WMS-IV). The participants underwent baseline and periodic standard neuropsychological assessments and provided biological samples (blood, urine, and CSF) at various visits.

CSF samples were collected after an overnight fast by lumbar puncture into collection tubes. Within an hour, they were transferred into polypropylene transfer tubes for freezing on dry ice to be shipped overnight to the ADNI Biomarker Core laboratory at the University of Pennsylvania Medical Center. Samples were thawed for an hour at room temperature and gently mixed for aliquot preparation, which were stored in vials at −80°C (Shaw et al. 2009). Roche Elecsys® assays were used to measure Aβ42 and pTau-181. sTREM2 levels were measured using ELISA based on the MSD platform (Kleinberger et al. 2014; Suárez-Calvet et al. 2016; Suárez-Calvet et al. 2019).

2.2 Data selection and curation

In this study, we investigated the baseline and longitudinal trajectories of three CSF biomarkers, namely Aβ42, pTau-181, and sTREM2, in individuals classified as CN and MCI. ADNI participants who were categorized as CN or MCI at their baseline visit and possessed a minimum of two visits of each of the aforementioned CSF biomarkers were selected.

2.3 Cognitive assessments

The ADNI neuropsychological test battery comprises a diverse array of cognitive tests, with several versions for each, such as three for the AD Assessment Scale - Cognition (ADAS-Cog) and two for the Rey Auditory Verbal Learning Test (RAVLT). This variety poses analytical challenges for integrating and interpreting these data. One solution is to create composite scores from these neuropsychological tests (Crane et al. 2012; Gibbons et al. 2012). There are several advantages to the composite score strategy. By combining elements of all individual test types, a composite score can provide a summary measure of cognitive functioning in specific domains. Individual neuropsychological tests can have inherent variability and measurement error. By aggregating multiple measures into a composite score, the impact of random measurement error can be reduced, leading to a more reliable estimate of cognitive functioning. Composite scores can also increase the sensitivity to detect cognitive impairment or changes over time. By pooling information from various tests, the composite score can capture broader aspects of cognitive functioning and potentially detect subtle changes that may not be apparent with individual tests alone. Prior work in the ADNI neuropsychological working group used an item response theory method for deriving four composite scores from the ADNI battery: (1) memory, (2) executive functioning, (3) language, and (4) visuospatial functioning. These composite scores were demonstrated to outperform their constituents in detecting change and predicting conversion from MCI to dementia. Furthermore, they were shown to correlate with a priori relevant neuroimaging markers previously found to relate to memory, such as hippocampal volume and thickness of the parahippocampal

gyrus, entorhinal cortex, and fusiform gyrus (Yonelinas et al. 2007; Walhovd et al. 2009; A. M. Fjell et al. 2008; Murphy et al. 2010; Van Petten et al. 2004).

In addition to increasing sensitivity and reducing redundancy in analyses of ADNI neuropsychological test data, a pertinent property of the composite scores is their linear scaling properties (Mungas and Reed 2000; Crane et al. 2008). This means that these composite scores maintain an interval level of measurement where a given degree of change in score represents the same amount of change in the underlying cognitive ability, regardless of the part of the ability continuum in which this change occurs. Thus, a specific score change at a lower level of cognitive functioning is equivalent to the same score change at a higher level of cognitive functioning. Such linear measurement properties are crucial in longitudinal analyses to ensure comparability of changes over time and across different cognitive ability levels.

Therefore, given the numerous advantages of these composite scores, including their ability to increase sensitivity, reduce redundancy in analyses, and provide linear scaling properties, we opted to use them for our longitudinal analyses. This approach is intended to optimize the accuracy of our evaluations and provide the most valuable insights into the progression of cognitive impairment and dementia.

2.4 Similarity Network Fusion

SNF is a data integration method used in biomedical research to integrate multiple data modalities for a group of patients or samples. The goal of SNF is to create a single

integrated network that captures the shared patterns of similarity between patients across multiple data modalities.

For Aim 1 (comparison of cross-sectional SNF versus cross-sectional univariate strategies), we constructed three distinct feature vectors corresponding to the baseline measurements of Aβ42, pTau-181, and sTREM2. These were entered into SNF (see below). As our benchmark, we used a cutpoint for CSF pTau-181/Aβ42 derived from a prior univariate analysis (Hansson et al. 2018). Briefly, univariate ROC analysis was employed to determine an optimal cutoff for CSF pTau-181/Aβ42 for concordance with PET visual read. The derived cutoff of 0.025 was then validated to predict future progression in MCI patients.

For Aim 2 (evaluation of multitemporal data with SNF), we used both the three baseline vectors from Aim1 and added the APC values calculated for each of the CSF biomarkers. For participants with n>2 CSF samples, we chose the follow-up visit separated farthest in time from the baseline visit, which we refer to as the final visit. This strategy was chosen to capture maximal changes in CSF. The APC values of the three CSF biomarkers were extracted and used to generate a matrix consisting of three columns. It is important to highlight that while the baseline CSF biomarker data have varying ranges for each biomarker, the APC shares a common unit space. Consequently, the APC data were amalgamated into a single matrix. SNF first constructs separate similarity networks for each data modality (three baseline vectors and one APC matrix of Aβ42, pTau-181, sTREM2) (Figure 2.1). Then SNF applies a graph fusion technique to these individual similarity networks to integrate them into a single network (Figure 2.1). This fusion technique involves iteratively computing the weighted average of the adjacency matrices

of the individual networks, where the weights are determined based on the similarity between nodes (i.e., patients) in the different networks. This process is repeated until convergence is achieved, resulting in a final integrated network that captures the shared patterns of similarity between patients across all data modalities. The resulting integrated network can then be analyzed using graph-based methods to identify patient clusters, subgroups, or other patterns of interest.

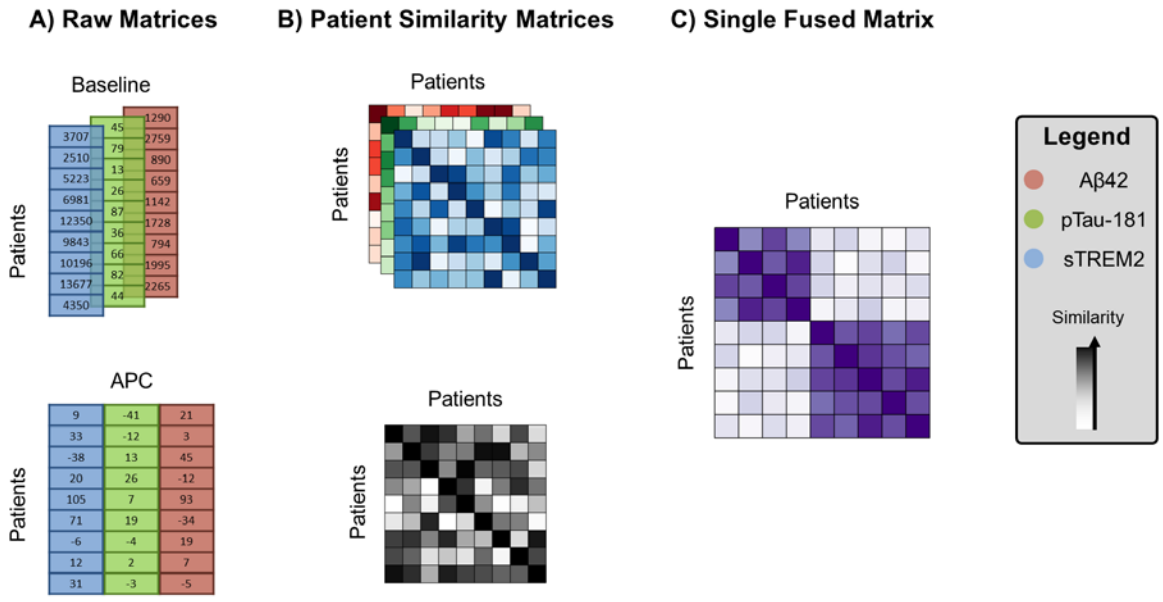

Figure 2.1: Visualization example of Similarity Network Fusion (SNF). A) Raw data representations of cerebrospinal fluid biomarkers (Aβ42, pTau-181, and sTREM2) corresponding to baseline (three individual vectors) and to annual percent change (APC). B) Patient similarity matrices constructed independently for each data type using Euclidean distance from the raw matrices. Each patient similarity matrix corresponds to one raw matrix/vector. C) Single fused matrix constructed by iterative fusion of the patient similarity matrices showing two clusters. Matrix illustration ordered based on obtained spectral clustering solution. In each matrix, darker colors correspond to greater similarity between patients.

The main steps of SNF are described as follows. Suppose we start with *n* samples (e.g., patients). First, four raw matrices—three *n* × *1* baseline vectors and one *n* × *3* APC matrix—were created. Then, the four raw matrices were converted into four *n* × *n* patient similarity matrices $\mathbf{W}$, using a scaled exponential similarity kernel. Here, $\mathbf{W}(i,j)$ denotes the similarity between patients *i* and *j*:

$$\mathbf{W}(i,j) = \exp\left(-\frac{\rho^2(x_i, x_j)}{\mu\varepsilon_{i,j}}\right) \quad \textbf{(2)}$$

where $\rho(x_i, x_j)$ represents the Euclidean distance between patients *i* and *j*, where x_i and x_j are of sizes *1* × *m* and correspond to the raw data of patients *i* and *j* obtained from the *n* × *m* raw matrix (in our case, these could be the baseline vectors or the APC matrix). The variable μ is a hyperparameter that acts as a scaling factor.

$\varepsilon_{i,j}$ is used to eliminate the scaling problem and is defined as follows:

$$\varepsilon_{i,j} = \frac{mean(\rho(x_i, N_i)) + mean\left(\rho(x_j, N_j)\right) + \rho(x_i, x_j)}{3} \quad \textbf{(3)}$$

where $mean(\rho(x_i, N_i))$ is the average Euclidean distance between x_i, the raw data of patient *i*, and each of its *K* most similar neighbors, N_i, where *K* is a hyperparameter controlling the number of neighbors.

For the fusion process, a global and a local kernel are computed from each of the individual similarity matrices, $\mathbf{W}$. The global kernel is an *n* × *n* normalized similarity matrix $\mathbf{P}$. One way to perform the normalization is $\mathbf{P} = \mathbf{D}^{-1}\mathbf{W}$, where $\mathbf{D}$ represents the

diagonal matrix with $\mathbf{D}(i,i) = \sum_j \mathbf{W}(i,j)$, such that $\sum_j \mathbf{P}(i,j) = 1$. This normalization process may encounter numerical instability due to the inclusion of self-similarities in the diagonal entries of matrix, **W**. To address this concern and achieve improved normalization, an alternative approach is proposed as follows:

$$\mathbf{P}(i,j) = \begin{cases} \dfrac{\mathbf{W}_{(i,j)}}{2\sum_{k \neq i} \mathbf{W}_{(i,k)}}, & j \neq i \\ \dfrac{1}{2}, & j = i \end{cases} \quad \textbf{(4)}$$

The matrix **P** represents information about the similarity between a given patient and all other patients. This normalization is stable because it is free of the scale of self-similarity in the diagonal entries and $\sum_j \mathbf{P}(i,j) = \mathbf{1}$ remains true**.**

Additionally, the $n \times n$ local kernel **S** was computed to measure local affinity as shown below:

$$\mathbf{S}(i,j) = \begin{cases} \dfrac{\mathbf{W}_{(i,j)}}{\sum_{k \in N_i} \mathbf{W}_{(i,k)}}, & j \in N_i \\ 0, & otherwise \end{cases} \quad \textbf{(5)}$$

By performing this operation, the similarity values between a patient and its non-neighboring patients (i.e., patients that are not within the most similar *K* neighbors) are set to zero. Therefore, the matrix **S** represents information about the similarity between a given patient and N_i, the patient's *K* most similar neighbors. This is based on the assumption that local similarities (larger similarity values) bear greater importance or

reliability compared to those that are distant—an assumption commonly adopted by various manifold learning algorithms. Importantly, matrix **P** comprehensively encapsulates the similarity information of each patient to all other patients, whereas matrix **S** solely captures the similarity values pertaining to the *K* most similar patients for each individual. The algorithm consistently initializes matrix **P** as the initial state and integrates matrix **S** as the kernel matrix during the fusion process. This approach ensures the capacity to effectively capture local graph structures while maintaining computational efficiency.

To explain the fusion process, consider the case where we have two data types (i.e., two initial raw matrices). First, from two raw matrices, similarity matrices $\mathbf{W^{(1)}}$ and $\mathbf{W^{(2)}}$, corresponding to data types 1 and 2, respectively. Subsequently, from the similarity matrices, the global matrices $\mathbf{P^{(1)}}$ and $\mathbf{P^{(2)}}$ are computed, followed by the kernel matrices $\mathbf{S^{(1)}}$ and $\mathbf{S^{(2)}}$.

At iteration t = 0, consider that $\mathbf{P}^{(\mathbf{1})}_{t=0} = \mathbf{P^{(1)}}$ and $\mathbf{P}^{(\mathbf{2})}_{t=0} = \mathbf{P^{(2)}}$. SNF performs the fusion through iteratively updating the global matrices of each data type as follows:

$$\mathbf{P}^{(\mathbf{1})}_{t=0} = \mathbf{S}^{(1)} \times \mathbf{P}^{(2)}_{t} \times (\mathbf{S^1})^T \quad \textbf{(6)}$$

$$\mathbf{P}^{(\mathbf{2})}_{t=0} = \mathbf{S}^{(2)} \times \mathbf{P}^{(1)}_{t} \times (\mathbf{S^2})^T \quad \textbf{(7)}$$

where $\mathbf{P}^{(\mathbf{1})}_{t+1}$ **and** $\mathbf{P}^{(\mathbf{2})}_{t+1}$ represent the global matrices after *t* iterations corresponding to data types 1 and 2, respectively. This procedure facilitates the simultaneous updating of the global matrices through two parallel and interchanging diffusion processes. At each

iteration, $\mathbf{P}_{t+1}^{(1)}$ **and** $\mathbf{P}_{t+1}^{(2)}$ underwent normalization as in equation (4). This ensures two key properties throughout SNF iterations: maintaining the highest similarity of each patient to themselves compared to others and generating a full-rank final network, essential for classification and clustering tasks. This normalization also facilitates faster convergence of SNF. The final matrix, after *t* iterations, is computed as follows:

$$\mathbf{P}^{(c)} = \frac{\mathbf{P}_t^{(1)} + \mathbf{P}_t^{(2)}}{2} \quad \textbf{(8)}$$

An alternative representation of equation 6 is

$$\mathbf{P}_{t+1}^{(1)}(i,j) = \sum_{k \in N_i} \sum_{l \in N_j} S^{(1)}(i,k) \times S^{(1)}(j,l) \times P_t^2(k,l) \quad \textbf{(9)}$$

and similarly for $\mathbf{P}_{t+1}^{(2)}$. Remember that N_i and N_j are the *K* most similar neighbors of patients *i* and *j*, respectively. This fusion formula highlights that the propagation of similarity information occurs exclusively within the common neighborhood. For example, this formula will consider patients *i* and *j* to likely be from the same cluster if they share the same neighbors.

To extend this beyond two data types, consider q distinct data types, where similarity matrices $\mathbf{W}^{(v)}$ were constructed using equation (2), corresponding to the v^{th} data type, $v = 1,2,\cdots,\mathrm{q}$. Next, $\mathbf{P}^{(v)}$ and $\mathbf{S}^{(v)}$ were computed using equations (4) and (5), respectively. The equation below represents equations (7) and (8), but for more than two data types (i.e., q > 2):

$$\mathbf{P}^{(v)} = \mathbf{S}^{(v)} \times \left(\frac{\sum_{k \neq v} \mathbf{P}^{(k)}}{q - 1} \right) \times \left(\mathbf{S}^{(v)} \right)^{T}, \; v = 1,2,\cdots,q \quad \textbf{(10)}$$

To ensure optimal data fusion in SNF, we did not pre-select a single pair of parameters for fusion. Instead, we conducted an exhaustive parameter search, where we used 100 values for each of *K* and μ, of recommended ranges [10, 30] and [0.3, 0.8], respectively (B. Wang et al. 2014). Spectral clustering was applied on each fusion. From each of the resulting 10,000 fusions, the ideal number of clusters, ranging from 2 to 5, was determined using the rotation cost and eigen-gap methods (J. Huang, Nie, and Huang 2013; S. Park and Zhao 2018). The clustering solution detected with the highest frequency in this parameter sweep was then selected (S. Park and Zhao 2018; B. Wang et al. 2014). In sum, this unsupervised, soft clustering method segregates the study participants into different subtypes according to the fully fused matrix.

2.5 Post-hoc statistical analyses

SNF produces clusters of patients derived from fused networks of multivariate CSF data only. To interpret whether and how these CSF derived clusters may differentiate patients based on concentrations of CSF proteins, longitudinal outcomes, and AD risk factors, we performed several types of post hoc analyses on data hidden from SNF.

CSF comparisons: Baseline, APC and time interval data were not normally distributed (Shapiro-Wilk $p<0.05$ for each) and therefore Mann-Whitney U tests were used to test whether each of these variables differed significantly in the clustering solutions.

Patient disease conversion: A Log-rank test was performed to evaluate and compare the risk of disease status conversion over time between the clustering solutions.

Patient demographics: Age was normally distributed and therefore two-tailed unpaired t-tests were used to test whether age differed significantly between the clustering solutions. For categorical outcome measures (APOE-ε4 status, sex, ethnicity), Chi-square tests of independence (with Monte Carlo simulation, if chi-square test conditions were not met) were performed to test for differences between clusters on each variable (Bradley and Cutcomb 1977).

Statistical significance was determined at an alpha<0.05.

3 Results

3.1 Validation of SNF and preclinical utility

3.1.1 Baseline clustering and conversion risk analysis in CN cohort

We clustered the CN cohort (n=122) based on the pTau-181/Aβ42 ratio, computed using baseline CSF pTau-181 and Aβ42. Using the standard univariate methods, we divided the participants into two groups (cluster 1, n=81; cluster 2, n=41), with participants in cluster 2 having a ratio of ≥0.025 (Hansson et al. 2018). After following each of the participants longitudinally, their respective diagnoses at their last visit were extracted and binned into either CN, MCI, or AD (Table 3.1). Log-rank tests comparing the risk of conversion to MCI or AD between the two clusters yielded no significant differences (p=0.9).

We performed SNF on the same cohort using only baseline CSF Aβ, pTau-181, and sTREM2, which yielded two clusters (cluster 1, n=64; cluster 2, n=58) after the parameter sweep (Appendix A). Log-rank tests showed significant differences between the two groups in the risk of conversion to MCI or AD, with cluster 1 exhibiting an increased likelihood of disease progression ($p<0.05$).

Additionally, individual matrices (prior to SNF fusion) were computed for each of the CSF biomarkers. Then, spectral clustering was performed (same as SNF). Log-rank tests were performed which showed that individual data types did not lead to significantly different clusters in terms of disease progression, whereas SNF did (Appendix B).

Table 3.1: Number of participants with a diagnosis of cognitively normal (CN), mild cognitive impairment (MCI), or Alzheimer's disease (AD) of the clusters of the CN cohort at the last visit available.

Clustering Method	Number of Clusters	Cluster 1	Cluster 2	*p* Value
pTau-181/Aβ42 ratio	2	**CN:** 59 (73%) **MCI:** 13 (16%) **AD:** 9 (11%)	27 (66%) 9 (22%) 5 (12%)	0.9
Baseline SNF	2	**CN:** 51 (80%) **MCI:** 9 (14%) **AD:** 4 (6%)	35 (60%) 13 (22%) 10 (17%)	*<0.05*

Presented values are counts (%). *p* value shown as obtained from Log-rank test comparing risk of conversion to MCI or AD.

3.1.2 Baseline clustering and conversion risk analysis in MCI cohort

We divided the MCI (N = 194) cohort into two groups based on the pTau-181/Aβ42 ratio as discussed in the previous section (cluster 1, n=56; cluster 2, n=138). Log-rank tests performed to compare the risk of conversion to AD between the two clusters yielded no significant differences (p=0.87) (Table 3.2).

The baseline SNF model identified three clusters as the most stable solution (cluster 1, n=57; cluster 2, n=104; cluster 3, n=33) (Appendix C). Subsequent analysis using log-rank tests showed that the three clusters have significantly different risks of conversion to AD over time (*p<0.001*).

Except for pTau-181, single data type analysis did not lead to clusters with significantly different risks of conversion, unlike SNF (Appendix D).

Table 3.2: Number of participants with a diagnosis of cognitively normal (CN), mild cognitive impairment (MCI), or Alzheimer's disease (AD) of the clusters of the MCI cohort at the last visit available.

Clustering Method	Number of Clusters	Cluster 1	Cluster 2	Cluster 3	*p* Value
pTau-181/Aβ42 ratio	2	**CN:** 5 (9%) **MCI:** 32 (57%) **AD:** 19 (34%)	11 (8%) 67 (49%) 60 (43%)	N/A	0.87
Baseline SNF	3	**CN:** 9 (16%) **MCI:** 37 (65%) **AD:** 11 (19%)	4 (4%) 44 (42%) 56 (54%)	3 (9%) 18 (54%) 12 (36%)	*<0.001*

Presented values are counts (%). *p* value shown as obtained from Log-rank test comparing risk of conversion (AD).

3.2 Employing SNF Longitudinally

From the prior analyses, we confirmed that SNF identified clusters that are statistically significant with regards to risk of conversion whereas traditional cross-sectional univariate models did not produce significant clusters. In this next section, we examined whether and how multitemporal clustering of baseline and APC measures of the CSF data may further improve characterization of disease trajectories in the cognitively normal and MCI patients.

3.2.1 Cognitively normal cohort

From the 10,000-parameter sweep, the two-cluster solution was found to be the most optimal (Appendix E). Subsequently, spectral clustering of the SNF-derived fused matrix of the cognitively normal cohort (n=122) identified two clusters (n=59; n=63) (Figure

3.1). There were no significant inter-cluster differences for age (p=0.07), education (p=0.14), study duration (p=0.21), sex (p=0.21), *APOE-ε4* (p=0.08), and ethnicity (p=0.30) (Table 3.3).

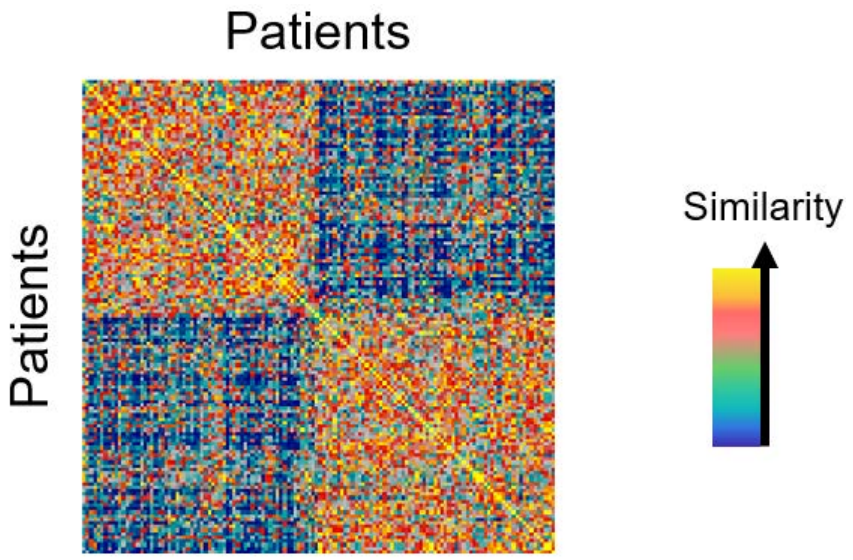

Figure 3.1: Single fused matrix constructed by similarity network fusion of the cognitively normal cohort. From the Alzheimer's Disease Neuroimaging Initiative study, a cohort of participants who were cognitively normal participants at baseline (n=122) were identified. Baseline and annual percent change data of cerebrospinal fluid Aβ42, pTau-181, and sTREM2 were integrated by similarity network fusion. Spectral clustering on the single fused matrix identified two clusters (n=59; n=63). Warmer colors correspond to greater similarity between patients.

Table 3.3: Demographic characteristics of clusters of the cognitively normal cohort

Characteristic	Cluster 1 (n = 59)	Cluster 2 (n = 63)	*p* Value
Age, mean ± SEM, yr	75.4 ± 0.7	73.4 ± 0.8	0.0725
Education, mean ± SEM, yr	16.1 ± 0.4	16.9 ± 04	0.1415
Study duration, mean ± SEM, yr	3.1 ± 0.1	3.0 ± 0.2	0.2084
% Female	50.8	38.1	0.2053
***APOE-ε4*, %**			0.0768
Carriers	28.8	14.3	
Noncarriers	71.2	85.7	
Ethnicity, %			0.3042
Asian	1.7	1.6	
African American	3.4	11.1	
White	94.9	85.7	
Other	0	1.6	

Two-tailed unpaired t-tests or Mann-Whitney U tests were performed for mean ± SEM, based on the distribution of the variable and Chi-squared tests of independence for %.

3.2.2 Mild cognitive impairment

From the 10,000-parameter sweep, the two-cluster solution was found to be the most optimal (Appendix F). Therefore, spectral clustering of the SNF-derived fused matrix of the MCI cohort (n=194) identified two clusters (n=116; n=78) (Figure 3.2). There were no significant inter-cluster differences for age (p=0.22), education (p=0.23), study duration (p=0.99), sex (p=1.00), and ethnicity (p=0.26) (Table 3.4). However, APOE-ε4 status was significantly different between the two clusters (p<0.001). Post-hoc analysis

revealed that cluster 1 had a significantly higher number of *APOE-ε4* carriers than cluster 2 ($p<0.001$).

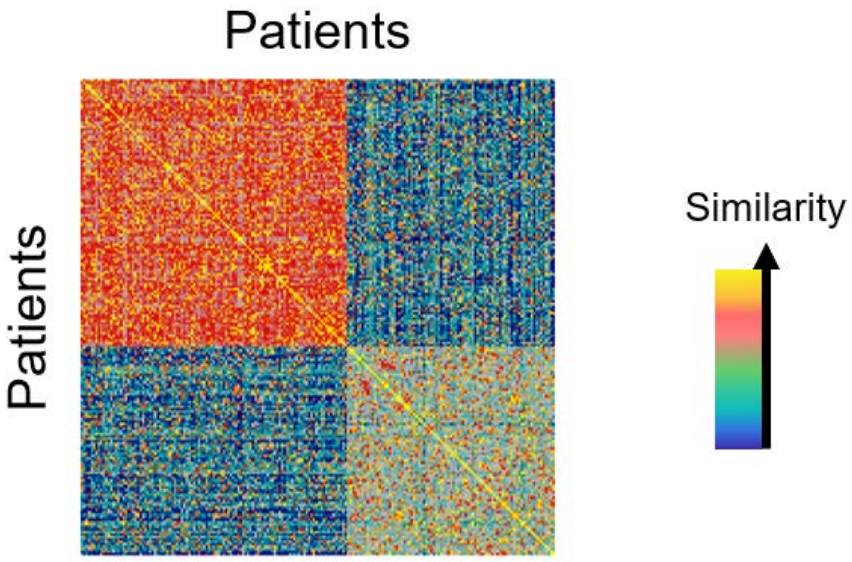

Figure 3.2: Single fused matrix constructed by similarity network fusion of the mild cognitive impairment cohort. From the Alzheimer's Disease Neuroimaging Initiative study, a cohort of participants with mild cognitive impairment at baseline (n=194) were identified. Baseline and annual percent change data of cerebrospinal fluid Aβ42, pTau-181, and sTREM2 were integrated by similarity network fusion. Spectral clustering on the single fused matrix identified two clusters (n=116; n=78). Warmer colors correspond to greater similarity between patients.

Table 3.4: Demographic characteristics of clusters of the mild cognitive impairment cohort.

Characteristic	Cluster 1 (n = 116)	Cluster 2 (n = 78)	*p* Value
Age, mean ± SEM, yr	73.1 ± 0.6	71.8 ± 0.9	0.2232
Education, mean ± SEM, yr	16.5 ± 0.3	16.1 ± 0.3	0.2340
Study duration, mean ± SEM, yr	2.9 ± 0.1	2.8 ± 0.1	0.9886
% Female	39.8	38.5	1.00
***APOE-ε4*, %**			*<0.001*
Carriers	67.2	34.6	
Noncarriers	32.8	65.4	
Ethnicity, %			0.2615
Asian	0	2.6	
African American	0.9	2.6	
White	97.4	92.2	
Other	1.7	2.6	

Two-tailed unpaired t-tests or Mann-Whitney U test were performed for mean ± SEM, based on the distribution of the variable and Chi-squared tests of independence for %.

3.3 Summary measures of CSF biomarkers

3.3.1 Cognitively normal cohort

The mean baseline levels of all three CSF biomarkers (Aβ42, pTau-181, and sTREM2) were higher in cluster 1 than cluster 2; albeit, only significantly higher for pTau-181 ($p<0.001$) and sTREM2 ($p<0.001$) (Figure 3.3A). In terms of longitudinal patterns, cluster 1 demonstrated a significantly greater negative annual percent change of Aβ42 ($p<0.001$). Cluster 1 also demonstrated a significantly greater positive annual percent

change of pTau-181 (p<0.01), reflecting greater longitudinal increases in p-Tau-181 compared to cluster 2 (Figure 3.3B). Both clusters showed a positive annual rate of change of sTREM2, with cluster 2 being insignificantly greater than cluster 1 (p=0.68). Overall, we see significantly increased longitudinal accumulation of the neurodegenerative biomarker pTau-181 and decreasing Aβ42—likely reflecting Aβ plaque accumulation in the brain in cluster 1, suggestive of a heightened risk for AD.

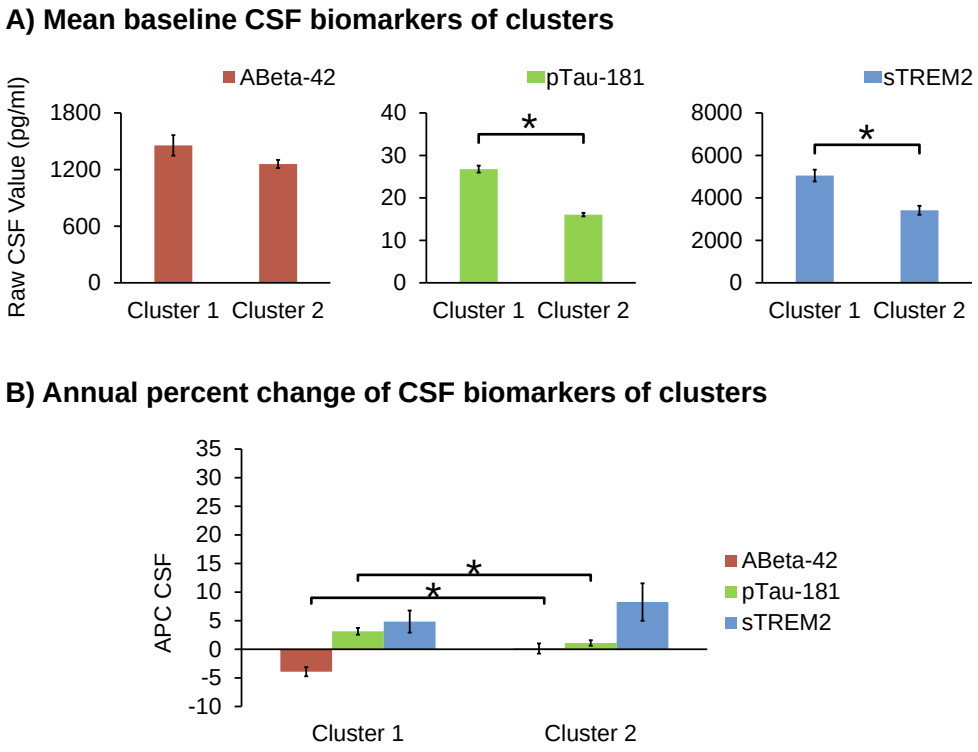

Figure 3.3: Comparison of the mean A) baseline and B) annual percent change levels of the CSF Aβ42, pTau-181, and sTREM2 between clusters of the cognitively normal cohort. Baseline CSF pTau-181 and sTREM2 were significantly higher in cluster 1 than cluster 2. The rates of change of CSF Aβ42 (negative) and CSF pTau-181 (positive) for cluster 1 were significantly different from cluster 2. In both clusters, sTREM2 increased, with no significance between clusters. Bars represent mean ± standard error of mean. Asterisks indicate statistically significant differences (*p<0.05), as determined by Mann-Whitney U-tests.

3.3.2 Mild cognitive impairment cohort

The mean baseline level of CSF Aβ42 was significantly lower in cluster 1 than cluster 2 ($p<0.01$), while CSF pTau-181 ($p<0.001$) and sTREM2 ($p<0.001$) were significantly higher in cluster 1 than cluster 2 (Figure 3.4A). Furthermore, cluster 1 demonstrated a significantly greater negative APC of Aβ42 ($p<0.01$) and a significantly lower positive APC of sTREM2 ($p<0.001$), compared to the pattern of change in cluster 2 (Figure 3.4B). Both clusters showed a positive annual percent change of pTau-181, with cluster 2 being insignificantly greater than cluster 1 ($p=0.18$). Overall, we see low baseline Aβ42, which is further dropping and high starting pTau-181 levels, indicating potentially higher risk of AD.

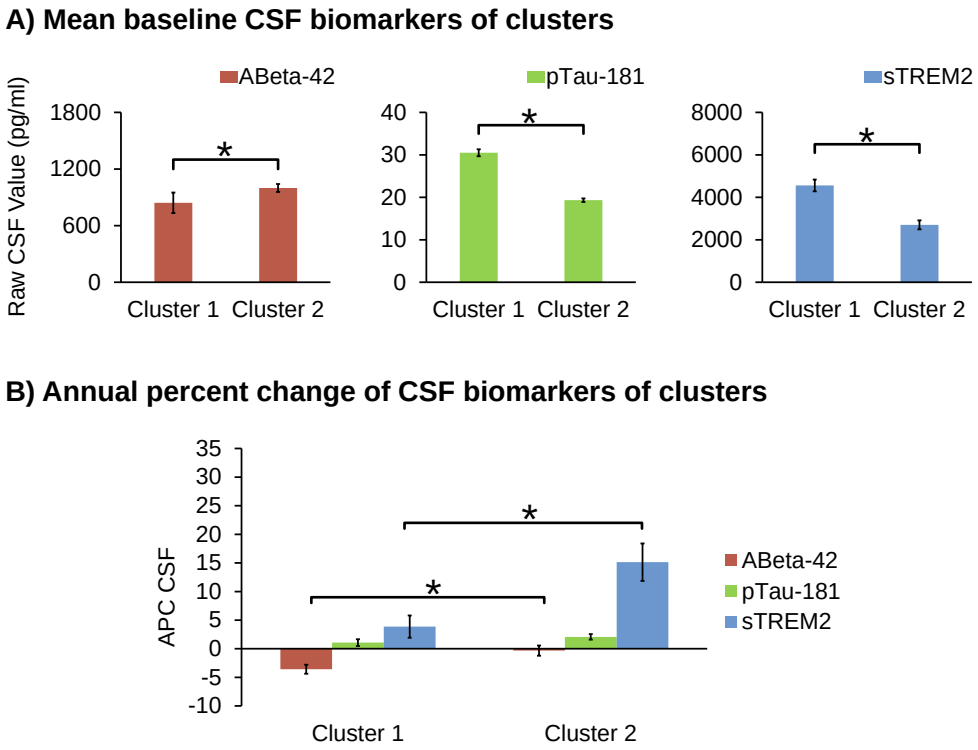

Figure 3.4: Comparison of the mean A) baseline and B) annual percent change levels of the CSF Aβ42, pTau-181, and sTREM2 between clusters of the mild cognitive impairment cohort. Baseline CSF Aβ42 was significantly lower while baseline CSF pTau-181 and sTREM2 were significantly higher in cluster 1 than cluster 2. The rates of change of CSF Aβ42 (negative) and CSF sTREM2 (positive) for cluster 1 were significantly different from cluster 2. In both clusters, pTau-181 increased, with no significance between clusters. Bars represent mean ± standard error of mean. Asterisks indicate statistically significant differences (*$p<0.05$), as determined by Mann-Whitney U-tests.

3.4 Post-hoc summary measures of cognition composite scores

For both the cognitively normal and MCI patients, SNF yielded two clusters, and in both cases cluster 1 exhibited a pattern of baseline and longitudinal trajectory consistent with

more aggressive AD progression. We next examined with post-hoc tests whether these differences in CSF translated to differences in cognitive progression.

3.4.1 Cognitively normal cohort

At baseline, the memory composite score of cluster 1 was significantly lower than cluster 2 ($p<0.05$), while the mean composite scores of executive function and language were not significantly different between the two clusters ($p=0.13$; $p=0.14$, respectively) (Figure 3.5A). Similarly, the memory composite score of cluster 1 decreased at a significantly greater rate than that of cluster 2 ($p<0.01$), while the rate of change of the executive function and language composite scores for both clusters did not significantly differ ($p=0.88$; $p=0.33$, respectively) (Figure 3.5B). In summary, even though SNF was blinded to measures of cognition, the SNF-derived clusters corresponded to a higher risk group, as evidenced by lower baseline and decreasing memory scores, as well as a lower risk group.

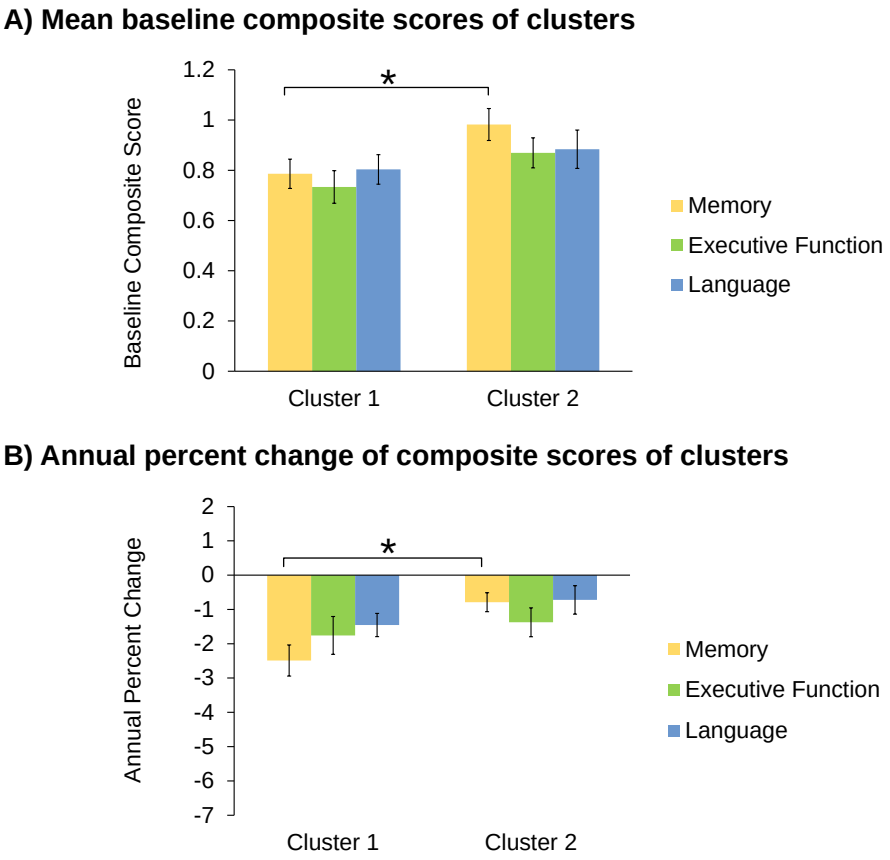

Figure 3.5: Comparison of the mean A) baseline and B) annual percent change composite score values of memory, executive function, and language between clusters of the cognitively normal cohort. Baseline memory composite score was significantly lower in cluster 1 than cluster 2, while baseline executive function and language composite scores showed no difference between clusters. Memory composite score decreased significantly in cluster 1 compared to cluster 2. Executive function and language composite scores decreased in both clusters but were not significantly different. Bars represent mean ± standard error of mean. Asterisks indicate statistically significant differences (*p<0.05), as determined by unpaired t-tests or Mann-Whitney U-tests.

3.4.2 Mild cognitive impairment cohort

At baseline, the memory composite score of cluster 1 was significantly lower than cluster 2 (p<0.05), while the composite scores of executive function and language did not

significantly differ between clusters (p=0.33; p=0.82, respectively) (Figure 3.6A). Similarly, the memory, executive function, and language composite scores of cluster 1 decreased at a significantly greater rate than cluster 2 (p<0.05) (Figure 3.6B). In summary, cluster 1 could indicate a subgroup within the MCI cohort that is at higher risk or faster rate of progression to AD as compared to cluster 2.

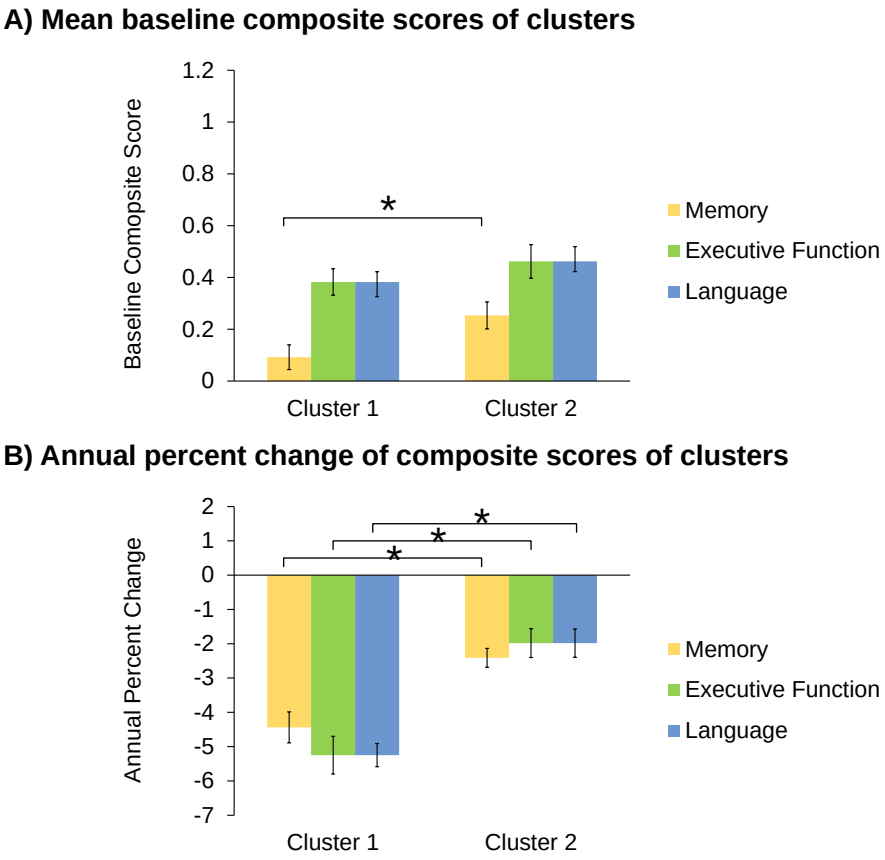

Figure 3.6: Comparison of the mean A) baseline and B) annual percent change composite score values of memory, executive function, and language between clusters of the mild cognitive impairment cohort. Baseline memory composite score was significantly lower in cluster 1 than cluster 2, while baseline executive function and language composite scores showed no significant difference between clusters. Memory, executive function, and language composite scores decreased significantly more in cluster 1 compared to cluster 2. Bars represent mean ± standard error of mean. Asterisks indicate statistically significant differences (*p<0.05), as determined by unpaired t-tests or Mann-Whitney U-tests.

3.5 Post-hoc clinical diagnosis at last follow-up

3.5.1 Cognitively normal cohort

Similar to Section 3.1.1, diagnosis at last follow-up was extracted for the CN cohort (Table 3.5). Log-rank tests were performed to compare the risk of conversion to MCI or AD between the two clusters. Cluster 1 exhibited a significantly higher risk of conversion compared to cluster 2 ($p<0.01$). In other words, participants in cluster 1 are significantly more likely than participants in cluster 2 to progress towards symptomatic disease stages within the same time frame. This result further supports the finding that cluster 1 is a high-risk group.

Table 3.5: Number of participants with a diagnosis of cognitively normal, mild cognitive impairment, or Alzheimer's disease of the clusters of the cognitively normal cohort at the last visit available.

Final diagnosis	Cluster 1	Cluster 2	p value
Cognitively normal	34 (58%)	52 (83%)	
Mild cognitive impairment	13 (22%)	9 (14%)	
Alzheimer's disease	12 (20%)	2 (3%)	
			<0.01

Presented values are counts (%). *p* value shown as obtained from log-rank tests comparing the risk of conversion to MCI or AD.

3.5.2 Mild cognitive impairment cohort

The process of extracting and categorizing the final diagnosis was replicated in the MCI cohort. Similar to the CN cohort, log-rank tests showed that cluster 1 had a significantly

higher risk of conversion to AD compared to cluster 2 (p<0.01) (Table 3.6). In other words, during the same time period, MCI patients in cluster 1 are converting to AD at a more rapid rate than patients in cluster 2, reinforcing the interpretation that cluster 1 might comprise a set of MCI patients who progress rapidly.

Table 3.6: Number of participants with a diagnosis of cognitively normal, mild cognitive impairment, or Alzheimer's disease of the clusters of the mild cognitive impairment cohort at the last visit available.

Final diagnosis	**Cluster 1**	**Cluster 2**	**p value**
Cognitively normal	8 (7%)	8 (10%)	
Mild cognitive impairment	52 (45%)	47 (60%)	
Alzheimer's disease	56 (48%)	23 (30%)	
			<0.01

Presented values are counts (%). *p* value shown as obtained from log-rank tests comparing the risk of conversion to AD.

4 Discussion

In this study, we aimed to explore the utility of a novel non-linear data integration technique, SNF, on baseline and APC CSF Aβ42, pTau-181, sTREM2 in two cohorts: CN and MCI. Our first objective was to validate SNF by comparing it to traditional cross-sectional univariate models. To that end, we performed SNF on baseline CSF Aβ42, pTau-181, sTREM2. Our second objective was to test if SNF can identify distinct subgroups within each cohort driven by the longitudinal dynamic patterns exhibited by CSF biomarkers. To investigate that, we performed SNF using both baseline and longitudinal data encoded in APC metrics.

Consistent with increased sensitivity of multivariate methods for classifying presymptomatic patients, the SNF method identified distinct subgroups both in the cognitively normal and mild cognitive impairment cohorts, based solely on conjoint patterns of CSF pathology, which were not detected by standard univariate strategies. Post-hoc validation analyses confirmed that a cluster of cognitively normal and mild cognitive impairment patients exhibited a higher risk of accelerated cognitive decline and disease progression. In the mild cognitive impairment high risk patient cluster, these individuals tended to have higher probability for the APOE4 genotype, which is also a higher risk for AD. Our findings suggest that the fusion analysis of CSF data can uncover patterns that serve as effective predictors of AD progression, offering a promising avenue for stratifying patients in future clinical trials evaluating early-stage preventative therapeutics.

4.1 Validating SNF

We conducted a comprehensive comparison on the risk of disease status conversion between clustering performed by SNF and univariate statistical methods (Hansson et al. 2018; Blennow et al. 2019). SNF effectively stratified the CN participants at baseline into a low-risk and a high-risk group, demonstrating a greater likelihood of progressing to MCI or AD. On the other hand, the two groups identified by the conventional univariate statistics showed comparable conversion rates, indicating limited discriminatory power. Similarly, for the MCI cohort, SNF identified three unique clusters that exhibited different risks of converting to AD, while the univariate method failed to capture this information. In a similar fashion to prior studies employing SNF, we also compared clustering performance using data prior to SNF fusion (i.e., matrices of individual data types; see Appendix A) (B. Wang et al. 2014; Markello et al. 2021). Clusters generated based on these data did not significantly differ in their risk of conversion (except for clusters based on pTau-181).

These findings underscore the clear advantage of SNF over traditional univariate statistical methods in revealing and capturing complex relationships and patterns within the same data and further validate the effectiveness of data fusion. By non-linearly integrating multiple variables and their interactions, SNF offers a more comprehensive and nuanced analysis, enabling the identification of individuals at greater risk of disease progression. As such, we highlight the potential for SNF as a superior tool for predictive modeling and risk stratification in the field of neurodegenerative diseases.

4.2 Longitudinal SNF

After validating SNF's performance relative to univariate approaches, we employed SNF longitudinally by clustering on both baseline and longitudinal CSF data, to better understand the multivariate profile of disease progression in CSF Aβ42, pTau-181, sTREM2. Consistent with our hypothesis, longitudinal SNF identified two subgroups within the CN cohort: a high-risk subgroup and a low-risk subgroup. The high-risk subgroup displayed lower initial memory scores, a decline in memory composite score over time, and a higher likelihood of progressing to MCI or AD. Conversely, the low-risk subgroup showed more favorable cognitive outcomes and a lower risk of disease progression. Notably, these subgroups exhibited distinct patterns in the levels of CSF biomarkers, indicating different underlying disease processes at play. Similar findings were observed in the MCI cohort, wherein high-risk and low-risk subgroups were discerned. The high-risk subgroup exhibited lower memory scores at baseline, more pronounced decline on all composite scores, and an increased probability of converting to AD—possibly representing a subgroup of rapid progressors. In contrast, the low-risk or slow progressors subgroup showed relatively slower decline of all composite scores and a lower risk of disease progression. The key finding of this analysis is that SNF was able to effectively differentiate meaningful patient subgroups solely using CSF biomarker data. This outcome underscores the predictive potential of these dynamic CSF biomarker patterns in forecasting disease progression, thereby providing critical insights for early intervention and disease management strategies.

4.2.1 SNF clusters

The two clusters within the CN cohort displayed no significant differences in demographic characteristics, such as age, education, study duration, sex, APOE-ε4 status, and ethnicity. Similarly, in the MCI cohort, no differences were observed except for APOE-ε4 status, where cluster 1 (high-risk) displayed a larger proportion of *APOE-ε4* carriers. This is in line with previous research, as *APOE-ε4* has been identified to be the strongest genetic risk factor for sporadic AD (L. A. Farrer et al. 1997; Strittmatter et al. 1993; Corder et al. 1993). Moreover, studies have previously reported an association between the presence of *APOE-ε4* and increased risk of progression from MCI to AD (Fleisher et al. 2007; Elias-Sonnenschein et al. 2011; R. C. Petersen et al. 1995). This increased risk is thought to be due to the interaction of APOE-ε4 genotype with AD pathologies, including Aβ, tau, and neuroinflammation (Vemuri et al. 2010; Morris et al. 2010; Leoni 2011; Ghisays et al. 2021; Benson et al. 2022). Our results suggest that SNF was successful in capturing the complex interaction between the APOE-ε4 genotype and AD-related pathologies, thereby confirming its association with disease risk. These findings motivated a more thorough investigation between clusters.

4.2.2 CSF patterns

Consistent with previous findings, our data support the notion that decreasing CSF Aβ42 and increasing pTau-181 are predictive of cognitive decline in CN individuals and progression to MCI or AD (Fagan et al. 2007; Li et al. 2007). Interestingly, our study did not find a significantly lower baseline CSF Aβ42 in the high-risk cluster within the CN

cohort, a finding that appears to diverge from previous literature which suggested that lower baseline CSF Aβ42 levels can predict subsequent cognitive decline in healthy older adults (Stomrud et al. 2007; Roe et al. 2013; Gustafson et al. 2007). Previous investigations that supported this observation utilized univariate models. However, recent work posited that baseline CSF Aβ42 retains its predictive utility for CN to MCI conversion when evaluated independently but loses this predictive capability in a multivariable regression model (Prosser et al. 2023). Prosser et al. further emphasized the superior predictive performance of their multivariate model over individual models for CN to MCI conversion, suggesting that other CSF patterns may supersede the importance of reduced baseline Aβ42 and thus reinforcing the relevance of multivariate techniques.

The pattern of increased pTau-181 and decreased Aβ42 at baseline has been previously identified as predictive of conversion from MCI to AD (Hansson et al. 2006; Andreasen et al. 2003; Buchhave et al. 2012). Additionally, in the presence of the APOE-ε4 allele, elevated pTau-181 has been linked with more rapid disease progression (Blom et al. 2009; Herukka et al. 2007). This is in line with our observation that cluster 1 of the MCI cohort, which exhibited a more rapid decline in composite scores and greater risk of conversion, also had higher pTau-181 and lower Aβ42 as well as a greater proportion of *APOE-ε4* carriers. In contrast to the CN cohort, we did not observe a significant difference in the longitudinal rate of change between the two MCI clusters. This finding resonates with previous reports that indicated a positive rate of change of pTau-181 during the preclinical stage, which diminishes at the onset of cognitive impairment (Morar et al. 2022; Llibre-Guerra et al. 2019).

Baseline sTREM2 was significantly higher in cluster 1 (higher-risk subgroup) of both the CN and MCI cohorts, aligning with prior research indicating that CSF sTREM2 increases in a disease-stage-dependent fashion along the AD continuum, potentially following a non-linear pattern, with peak levels in the early stages (Suárez-Calvet et al. 2016; Suárez-Calvet et al. 2016; Biel et al. 2023; Falcon et al. 2019; Pillai et al. 2021). As elevated CSF sTREM2 concentrations have been correlated with markers of neuronal damage and tau pathology (total-Tau and p-Tau), an early sTREM2 increase is hypothesized to reflect an enhanced inflammatory response and reactive microgliosis triggered by tau-induced neurodegeneration (Suárez-Calvet et al. 2019; Suárez-Calvet et al. 2016; Heslegrave et al. 2016; S.-H. Park et al. 2021; Knapskog et al. 2020; Piccio et al. 2016). In line with these studies, our data indicate that the higher-risk subgroups from both cohorts, characterized by elevated sTREM2, also exhibited higher levels of pTau. Furthermore, we observe that a more rapid increase in sTREM2 is associated with a smaller decline in Aβ42, possibly evidencing the neuroprotective role of sTREM2, as observed in animal studies, in mitigating plaque formation and toxicity in the brain (Y. Huang et al. 2021; Parhizkar et al. 2019). To our knowledge, our study is the first to replicate this finding, recently found in a longitudinal study of autosomal-dominant AD, in sporadic AD (Morenas-Rodríguez et al. 2022).

Overall, longitudinal SNF revealed multivariate patterns that are predictive of a higher risk of conversion in CN and faster progression in MCI.

4.3 Limitations and future improvements

While our study provides valuable insights into the utility of SNF analysis in analyzing longitudinal CSF biomarkers and identifying subgroups within seemingly homogenous cohorts, there are several limitations that should be acknowledged.

Our sample sizes are relatively small due to the strict filtering criteria necessitating at least two time points for each CSF biomarker. Therefore, a future direction of this work is replicating our findings in a larger independent sample to enhance the validity and generalizability of our results.

Even though we carefully selected a panel of biomarkers that is sensitive to preclinical AD and facilitated the uncovering of unique subgroups, the complexity of AD suggests that a more diverse set of biomarkers might yield additional insights. Longitudinal SNF analysis incorporating this broader range of biomarkers may offer deeper insights into concealed patterns, potentially revealing distinct subgroups, thereby providing novel insights into the pathophysiology of AD.

Utilizing two timepoints to capture the longitudinal trajectory of the CSF biomarkers proved useful allowing SNF to discern unique dynamic trajectories. However, future endeavors should aim to delve deeper into the patterns by collecting data from three visits to encapsulate non-linearity more accurately. For instance, employing a dynamic method to compute APC as opposed to the current static method could account for the dynamic changes across each visit, thereby providing a more detailed summary metric for each time point rather than a ‘lump sum’.

Additionally, forthcoming studies should focus on extracting the features and patterns that SNF used for clustering and encoding these into traditional machine learning classifiers. Given the high discriminatory power of SNF observed in the present study, such machine learning models could potentially demonstrate high efficacy in classifying disease trajectories and distinguishing them from normal aging.

4.4 Implications

This study underscores the utility of SNF, an advanced multivariate and non-linear data fusion method, in illuminating dynamic patterns of CSF biomarkers across the continuum of AD. By applying a multivariate longitudinal modeling approach to these biomarkers, it is possible to uncover distinct patterns indicative of disease progression during the preclinical stage. The identified biomarker patterns could potentially be incorporated into the design of clinical trials that assess preclinical treatments, thus enhancing participant stratification and recruitment processes. This could substantially enhance the precision of these trials to ultimately lead to development of treatments that can slow down or halt disease progression at the preclinical stage.

Identifying subgroups with varying disease risks and progression patterns holds substantial prognostic implications. Such information allows healthcare providers to make more accurate predictions regarding disease outcomes and tailor management strategies to the individual patient's needs. The early identification of individuals at higher risk for disease progression could facilitate timely interventions, potentially delaying or even preventing the onset of severe cognitive decline.

References

Adamczuk, Katarzyna, Jolien Schaeverbeke, Hugo M. J. Vanderstichele, Johan Lilja, Natalie Nelissen, Koen Van Laere, Patrick Dupont, Kelly Hilven, Koen Poesen, and Rik Vandenberghe. 2015. “Diagnostic Value of Cerebrospinal Fluid Aβ Ratios in Preclinical Alzheimer’s Disease.” *Alzheimer’s Research & Therapy* 7 (1): 75. https://doi.org/10.1186/s13195-015-0159-5.

Ahmed, Olfa Ben, Jenny Benois-Pineau, Michelle Allard, Gwénaëlle Catheline, and Chokri Ben Amar. 2017. “Recognition of Alzheimer’s Disease and Mild Cognitive Impairment with Multimodal Image-Derived Biomarkers and Multiple Kernel Learning.” *Neurocomputing*, Recent Research in Medical Technology Based on Multimedia and Pattern Recognition, 220 (January): 98–110. https://doi.org/10.1016/j.neucom.2016.08.041.

Alcolea, Daniel, María Carmona-Iragui, Marc Suárez-Calvet, M. Belén Sánchez-Saudinós, Isabel Sala, Sofía Antón-Aguirre, Rafael Blesa, Jordi Clarimón, Juan Fortea, and Alberto Lleó. 2014. “Relationship Between β-Secretase, Inflammation and Core Cerebrospinal Fluid Biomarkers for Alzheimer’s Disease.” *Journal of Alzheimer’s Disease* 42 (1): 157–67. https://doi.org/10.3233/JAD-140240.

Andreasen, Niels, Eugeen Vanmechelen, Hugo Vanderstichele, Pia Davidsson, and Kaj Blennow. 2003. “Cerebrospinal Fluid Levels of Total-Tau, Phospho-Tau and A Beta 42 Predicts Development of Alzheimer’s Disease in Patients with Mild Cognitive Impairment.” *Acta Neurologica Scandinavica. Supplementum* 179: 47–51. https://doi.org/10.1034/j.1600-0404.107.s179.9.x.

Andreasson, Ulf, Julia Kuhlmann, Josef Pannee, Robert M. Umek, Erik Stoops, Hugo Vanderstichele, Anja Matzen, et al. 2018. “Commutability of the Certified Reference Materials for the Standardization of β-Amyloid 1-42 Assay in Human Cerebrospinal Fluid: Lessons for Tau and β-Amyloid 1-40 Measurements.” *Clinical Chemistry and Laboratory Medicine* 56 (12): 2058–66. https://doi.org/10.1515/cclm-2018-0147.

Atagi, Yuka, Chia-Chen Liu, Meghan M. Painter, Xiao-Fen Chen, Christophe Verbeeck, Honghua Zheng, Xia Li, et al. 2015. “Apolipoprotein E Is a Ligand for Triggering Receptor Expressed on Myeloid Cells 2 (TREM2).” *The Journal of Biological Chemistry* 290 (43): 26043–50. https://doi.org/10.1074/jbc.M115.679043.

Atri, Alireza. 2019. “Current and Future Treatments in Alzheimer’s Disease.” *Seminars in Neurology* 39 (2): 227–40. https://doi.org/10.1055/s-0039-1678581.

Badhwar, AmanPreet, G Peggy McFall, Shraddha Sapkota, Sandra E Black, Howard Chertkow, Simon Duchesne, Mario Masellis, Liang Li, Roger A Dixon, and Pierre Bellec. 2020. “A Multiomics Approach to Heterogeneity in Alzheimer’s Disease: Focused Review and Roadmap.” *Brain* 143 (5): 1315–31. https://doi.org/10.1093/brain/awz384.

Baik, Sung Hoon, Seokjo Kang, Sung Min Son, and Inhee Mook-Jung. 2016. “Microglia Contributes to Plaque Growth by Cell Death Due to Uptake of Amyloid β in the Brain of Alzheimer’s Disease Mouse Model.” *Glia* 64 (12): 2274–90. https://doi.org/10.1002/glia.23074.

Bard, F., C. Cannon, R. Barbour, R. L. Burke, D. Games, H. Grajeda, T. Guido, et al. 2000. “Peripherally Administered Antibodies against Amyloid Beta-Peptide Enter the Central Nervous System and Reduce Pathology in a Mouse Model of Alzheimer Disease.” *Nature Medicine* 6 (8): 916–19. https://doi.org/10.1038/78682.

Benson, Gloria S., Chris Bauer, Lucrezia Hausner, Samuel Couturier, Piotr Lewczuk, Oliver Peters, Michael Hüll, et al. 2022. “Don’t Forget about Tau: The Effects of ApoE4 Genotype on Alzheimer’s Disease Cerebrospinal Fluid Biomarkers in Subjects with Mild Cognitive Impairment—Data from the Dementia Competence Network.” *Journal of Neural Transmission* 129 (5–6): 477–86. https://doi.org/10.1007/s00702-022-02461-0.

Bettens, Karolien, Kristel Sleegers, and Christine Van Broeckhoven. 2010. “Current Status on Alzheimer Disease Molecular Genetics: From Past, to Present, to Future.” *Human Molecular Genetics* 19 (R1): R4–11. https://doi.org/10.1093/hmg/ddq142.

Biel, Davina, Marc Suárez-Calvet, Paul Hager, Anna Rubinski, Anna Dewenter, Anna Steward, Sebastian Roemer, et al. 2023. “sTREM2 Is Associated with Amyloid-related P-tau Increases and Glucose Hypermetabolism in Alzheimer’s Disease.” *EMBO Molecular Medicine* 15 (2): e16987. https://doi.org/10.15252/emmm.202216987.

Biomarkers Definitions Working Group. 2001. “Biomarkers and Surrogate Endpoints: Preferred Definitions and Conceptual Framework.” *Clinical Pharmacology and Therapeutics* 69 (3): 89–95. https://doi.org/10.1067/mcp.2001.113989.

Birkenbihl, Colin, Yasamin Salimi, Holger Fröhlich, for the Japanese Alzheimer’s Disease Neuroimaging Initiative, and the Alzheimer’s Disease Neuroimaging Initiative. 2022. “Unraveling the Heterogeneity in Alzheimer’s Disease Progression across Multiple Cohorts and the Implications for Data-Driven Disease Modeling.” *Alzheimer’s & Dementia* 18 (2): 251–61. https://doi.org/10.1002/alz.12387.

Bittner, Tobias, Henrik Zetterberg, Charlotte E. Teunissen, Richard E. Ostlund, Michael Militello, Ulf Andreasson, Isabelle Hubeek, et al. 2016. “Technical Performance of a Novel, Fully Automated Electrochemiluminescence Immunoassay for the Quantitation of β-Amyloid (1–42) in Human Cerebrospinal Fluid.” *Alzheimer’s & Dementia* 12 (5): 517–26. https://doi.org/10.1016/j.jalz.2015.09.009.

Blennow, Kaj, Bruno Dubois, Anne M. Fagan, Piotr Lewczuk, Mony J. de Leon, and Harald Hampel. 2015. “Clinical Utility of Cerebrospinal Fluid Biomarkers in the Diagnosis of Early Alzheimer’s Disease.” *Alzheimer’s & Dementia: The Journal of the Alzheimer’s Association* 11 (1): 58–69. https://doi.org/10.1016/j.jalz.2014.02.004.

Blennow, Kaj, and Harald Hampel. 2003. “CSF Markers for Incipient Alzheimer’s Disease.” *The Lancet. Neurology* 2 (10): 605–13. https://doi.org/10.1016/s1474-4422(03)00530-1.

Blennow, Kaj, Harald Hampel, Michael Weiner, and Henrik Zetterberg. 2010. “Cerebrospinal Fluid and Plasma Biomarkers in Alzheimer Disease.” *Nature Reviews Neurology* 6 (3): 131–44. https://doi.org/10.1038/nrneurol.2010.4.

Blennow, Kaj, Leslie M. Shaw, Erik Stomrud, Niklas Mattsson, Jon B. Toledo, Katharina Buck, Simone Wahl, et al. 2019. “Predicting Clinical Decline and Conversion to Alzheimer’s Disease or Dementia Using Novel Elecsys Aβ(1–42), PTau and TTau CSF Immunoassays.” *Scientific Reports* 9 (1): 19024. https://doi.org/10.1038/s41598-019-54204-z.

Blom, Elin S., Vilmantas Giedraitis, Henrik Zetterberg, Hiroaki Fukumoto, Kaj Blennow, Bradley T. Hyman, Michael C. Irizarry, Lars-Olof Wahlund, Lars Lannfelt, and Martin Ingelsson. 2009. “Rapid Progression from Mild Cognitive Impairment to Alzheimer’s Disease in Subjects with Elevated Levels of Tau in Cerebrospinal Fluid and the APOE E4/E4 Genotype.” *Dementia and Geriatric Cognitive Disorders* 27 (5): 458–64. https://doi.org/10.1159/000216841.

Bolmont, Tristan, Florent Haiss, Daniel Eicke, Rebecca Radde, Chester A. Mathis, William E. Klunk, Shinichi Kohsaka, Mathias Jucker, and Michael E. Calhoun. 2008. “Dynamics of the Microglial/Amyloid Interaction Indicate a Role in Plaque Maintenance.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 28 (16): 4283–92. https://doi.org/10.1523/JNEUROSCI.4814-07.2008.

Bolton, Jessica S., Sidhartha Chaudhury, Sheetij Dutta, Scott Gregory, Emily Locke, Tony Pierson, and Elke S. Bergmann-Leitner. 2020. “Comparison of ELISA with Electro-Chemiluminescence Technology for the Qualitative and Quantitative Assessment of Serological Responses to Vaccination.” *Malaria Journal* 19 (1): 159. https://doi.org/10.1186/s12936-020-03225-5.

Braak, H., and E. Braak. 1991. “Neuropathological Stageing of Alzheimer-Related Changes.” *Acta Neuropathologica* 82 (4): 239–59. https://doi.org/10.1007/BF00308809.

Bradley, Drake R., and Steven Cutcomb. 1977. “Monte Carlo Simulations and the Chi-Square Test of Independence.” *Behavior Research Methods & Instrumentation* 9 (2): 193–201. https://doi.org/10.3758/BF03214499.

Brydon, H. L., R. Hayward, W. Harkness, and R. Bayston. 1995. “Physical Properties of Cerebrospinal Fluid of Relevance to Shunt Function. 2: The Effect of Protein upon CSF Surface Tension and Contact Angle.” *British Journal of Neurosurgery* 9 (5): 645–51. https://doi.org/10.1080/02688699550040936.

———. 1996. “Does the Cerebrospinal Fluid Protein Concentration Increase the Risk of Shunt Complications?” *British Journal of Neurosurgery* 10 (3): 267–73. https://doi.org/10.1080/02688699650040124.

Brys, Miroslaw, Elizabeth Pirraglia, Kenneth Rich, Sindre Rolstad, Lisa Mosconi, Remigiusz Switalski, Lidia Glodzik-Sobanska, et al. 2009. “Prediction and Longitudinal Study of CSF Biomarkers in Mild Cognitive Impairment.” *Neurobiology of Aging* 30 (5): 682–90. https://doi.org/10.1016/j.neurobiolaging.2007.08.010.

Buchhave, Peder, Lennart Minthon, Henrik Zetterberg, Åsa K. Wallin, Kaj Blennow, and Oskar Hansson. 2012. “Cerebrospinal Fluid Levels Ofβ-Amyloid 1-42, but Not of Tau, Are Fully Changed Already 5 to 10 Years Before the Onset of Alzheimer Dementia.” *Archives of General Psychiatry* 69 (1): 98–106. https://doi.org/10.1001/archgenpsychiatry.2011.155.

Cancer Genome Atlas Research Network. 2008. “Comprehensive Genomic Characterization Defines Human Glioblastoma Genes and Core Pathways.” *Nature* 455 (7216): 1061–68. https://doi.org/10.1038/nature07385.

Cavalli, Florence M. G., Marc Remke, Ladislav Rampasek, John Peacock, David J. H. Shih, Betty Luu, Livia Garzia, et al. 2017. “Intertumoral Heterogeneity within Medulloblastoma Subgroups.” *Cancer Cell* 31 (6): 737-754.e6. https://doi.org/10.1016/j.ccell.2017.05.005.

Chakrabarty, Paramita, Karen Jansen-West, Amanda Beccard, Carolina Ceballos-Diaz, Yona Levites, Christophe Verbeeck, Abba C. Zubair, Dennis Dickson, Todd E. Golde, and Pritam Das. 2010. “Massive Gliosis Induced by Interleukin-6 Suppresses Abeta Deposition in Vivo: Evidence against Inflammation as a Driving Force for Amyloid Deposition.” *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 24 (2): 548–59. https://doi.org/10.1096/fj.09-141754.

Chan, D., N. C. Fox, R. I. Scahill, W. R. Crum, J. L. Whitwell, G. Leschziner, A. M. Rossor, J. M. Stevens, L. Cipolotti, and M. N. Rossor. 2001. “Patterns of Temporal Lobe Atrophy in Semantic Dementia and Alzheimer’s Disease.” *Annals of Neurology* 49 (4): 433–42.

Choo, IL Han, Ruiqing Ni, Michael Schöll, Anders Wall, Ove Almkvist, and Agneta Nordberg. 2013. “Combination of 18F-FDG PET and Cerebrospinal Fluid Biomarkers as a Better Predictor of the Progression to Alzheimer’s Disease in Mild Cognitive Impairment Patients.” *Journal of Alzheimer’s Disease* 33 (4): 929–39. https://doi.org/10.3233/JAD-2012-121489.

Colonna, Marco. 2003. “TREMs in the Immune System and Beyond.” *Nature Reviews. Immunology* 3 (6): 445–53. https://doi.org/10.1038/nri1106.

Cooper, Claudia, Ryan Li, Constantine Lyketsos, and Gill Livingston. 2013. “Treatment for Mild Cognitive Impairment: Systematic Review.” *British Journal of Psychiatry* 203 (4): 255–64. https://doi.org/10.1192/bjp.bp.113.127811.

Corder, E. H., A. M. Saunders, W. J. Strittmatter, D. E. Schmechel, P. C. Gaskell, G. W. Small, A. D. Roses, J. L. Haines, and M. A. Pericak-Vance. 1993. “Gene Dose of Apolipoprotein E Type 4 Allele and the Risk of Alzheimer’s Disease in Late Onset

Families." *Science (New York, N.Y.)* 261 (5123): 921–23. https://doi.org/10.1126/science.8346443.

Crane, Paul K., Adam Carle, Laura E. Gibbons, Philip Insel, R. Scott Mackin, Alden Gross, Richard N. Jones, et al. 2012. "Development and Assessment of a Composite Score for Memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI)." *Brain Imaging and Behavior* 6 (4): 502–16. https://doi.org/10.1007/s11682-012-9186-z.

Crane, Paul K., Kaavya Narasimhalu, Laura E. Gibbons, Dan M. Mungas, Sebastien Haneuse, Eric B. Larson, Lewis Kuller, Kathleen Hall, and Gerald van Belle. 2008. "Item Response Theory Facilitated Co-Calibrating Cognitive Tests and Reduced Bias in Estimated Rates of Decline." *Journal of Clinical Epidemiology* 61 (10): 1018-27.e9. https://doi.org/10.1016/j.jclinepi.2007.11.011.

Cummings, Jeffrey, Dana P. Goldman, Nicholas R. Simmons-Stern, and Eric Ponton. 2022. "The Costs of Developing Treatments for Alzheimer's Disease: A Retrospective Exploration." *Alzheimer's & Dementia* 18 (3): 469–77. https://doi.org/10.1002/alz.12450.

DeCarli, Charles, Giovanni B. Frisoni, Christopher M. Clark, Danielle Harvey, Michael Grundman, Ronald C. Petersen, Leon J. Thal, et al. 2007. "Qualitative Estimates of Medial Temporal Atrophy as a Predictor of Progression from Mild Cognitive Impairment to Dementia." *Archives of Neurology* 64 (1): 108–15. https://doi.org/10.1001/archneur.64.1.108.

Del-Aguila, Jorge L., Bruno A. Benitez, Zeran Li, Umber Dube, Kathie A. Mihindukulasuriya, John P. Budde, Fabiana H. G. Farias, et al. 2019. "TREM2 Brain Transcript-Specific Studies in AD and TREM2 Mutation Carriers." *Molecular Neurodegeneration* 14 (1): 18. https://doi.org/10.1186/s13024-019-0319-3.

Dickerson, B. C., I. Goncharova, M. P. Sullivan, C. Forchetti, R. S. Wilson, D. A. Bennett, L. A. Beckett, and L. deToledo-Morrell. 2001. "MRI-Derived Entorhinal and Hippocampal Atrophy in Incipient and Very Mild Alzheimer's Disease☆☆This Research Was Supported by Grants P01 AG09466 and P30 AG10161 from the National Institute on Aging, National Institutes of Health." *Neurobiology of Aging* 22 (5): 747–54. https://doi.org/10.1016/S0197-4580(01)00271-8.

Doecke, James D., Larry Ward, Samantha C. Burnham, Victor L. Villemagne, Qiao-Xin Li, Steven Collins, Christopher J. Fowler, et al. 2020. "Elecsys CSF Biomarker Immunoassays Demonstrate Concordance with Amyloid-PET Imaging." *Alzheimer's Research & Therapy* 12 (1): 36. https://doi.org/10.1186/s13195-020-00595-5.

Du, A. T., N. Schuff, D. Amend, M. P. Laakso, Y. Y. Hsu, W. J. Jagust, K. Yaffe, et al. 2001. "Magnetic Resonance Imaging of the Entorhinal Cortex and Hippocampus in Mild Cognitive Impairment and Alzheimer's Disease." *Journal of Neurology, Neurosurgery, and Psychiatry* 71 (4): 441–47. https://doi.org/10.1136/jnnp.71.4.441.

Duara, R., D. A. Loewenstein, E. Potter, J. Appel, M. T. Greig, R. Urs, Q. Shen, et al. 2008. “Medial Temporal Lobe Atrophy on MRI Scans and the Diagnosis of Alzheimer Disease.” *Neurology* 71 (24): 1986–92. https://doi.org/10.1212/01.wnl.0000336925.79704.9f.

Duff, K., C. Eckman, C. Zehr, X. Yu, C. M. Prada, J. Perez-tur, M. Hutton, et al. 1996. “Increased Amyloid-Beta42(43) in Brains of Mice Expressing Mutant Presenilin 1.” *Nature* 383 (6602): 710–13. https://doi.org/10.1038/383710a0.

Duits, Flora H., Pablo Martinez-Lage, Claire Paquet, Sebastiaan Engelborghs, Alberto Lleó, Lucrezia Hausner, José L. Molinuevo, et al. 2016. “Performance and Complications of Lumbar Puncture in Memory Clinics: Results of the Multicenter Lumbar Puncture Feasibility Study.” *Alzheimer’s & Dementia* 12 (2): 154–63. https://doi.org/10.1016/j.jalz.2015.08.003.

Eckman, Elizabeth A., Stephanie K. Adams, Frederick J. Troendle, Becky A. Stodola, Murad A. Kahn, Abdul H. Fauq, Hong D. Xiao, Kenneth E. Bernstein, and Christopher B. Eckman. 2006. “Regulation of Steady-State Beta-Amyloid Levels in the Brain by Neprilysin and Endothelin-Converting Enzyme but Not Angiotensin-Converting Enzyme.” *The Journal of Biological Chemistry* 281 (41): 30471–78. https://doi.org/10.1074/jbc.M605827200.

Elias-Sonnenschein, Lyzel S., Wolfgang Viechtbauer, Inez H. G. B. Ramakers, Frans R. J. Verhey, and Pieter Jelle Visser. 2011. “Predictive Value of APOE-E4 Allele for Progression from MCI to AD-Type Dementia: A Meta-Analysis.” *Journal of Neurology, Neurosurgery, and Psychiatry* 82 (10): 1149–56. https://doi.org/10.1136/jnnp.2010.231555.

Fagan, Anne M., Catherine M. Roe, Chengjie Xiong, Mark A. Mintun, John C. Morris, and David M. Holtzman. 2007. “Cerebrospinal Fluid Tau/β-Amyloid42 Ratio as a Prediction of Cognitive Decline in Nondemented Older Adults.” *Archives of Neurology* 64 (3): 343–49. https://doi.org/10.1001/archneur.64.3.noc60123.

Fagan, Anne M., Chengjie Xiong, Mateusz S. Jasielec, Randall J. Bateman, Alison M. Goate, Tammie L. S. Benzinger, Bernardino Ghetti, et al. 2014. “Longitudinal Change in CSF Biomarkers in Autosomal-Dominant Alzheimer’s Disease.” *Science Translational Medicine* 6 (226): 226ra30-226ra30. https://doi.org/10.1126/scitranslmed.3007901.

Falahati, Farshad, Eric Westman, and Andrew Simmons. 2014. “Multivariate Data Analysis and Machine Learning in Alzheimer’s Disease with a Focus on Structural Magnetic Resonance Imaging.” *Journal of Alzheimer’s Disease* 41 (3): 685–708. https://doi.org/10.3233/JAD-131928.

Falcon, Carles, Gemma C. Monté-Rubio, Oriol Grau-Rivera, Marc Suárez-Calvet, Raquel Sánchez-Valle, Lorena Rami, Beatriz Bosch, Christian Haass, Juan Domingo Gispert, and José Luis Molinuevo. 2019. “CSF Glial Biomarkers YKL40 and STREM2 Are Associated with Longitudinal Volume and Diffusivity Changes in Cognitively

Unimpaired Individuals." *NeuroImage: Clinical* 23: 101801. https://doi.org/10.1016/j.nicl.2019.101801.

Farrer, L. A., L. A. Cupples, J. L. Haines, B. Hyman, W. A. Kukull, R. Mayeux, R. H. Myers, M. A. Pericak-Vance, N. Risch, and C. M. van Duijn. 1997. "Effects of Age, Sex, and Ethnicity on the Association between Apolipoprotein E Genotype and Alzheimer Disease. A Meta-Analysis. APOE and Alzheimer Disease Meta Analysis Consortium." *JAMA* 278 (16): 1349–56.

Farrer, Lindsay A., L. Adrienne Cupples, Jonathan L. Haines, Bradley Hyman, Walter A. Kukull, Richard Mayeux, Richard H. Myers, Margaret A. Pericak-Vance, Neil Risch, and Cornelia M. van Duijn. 1997. "Effects of Age, Sex, and Ethnicity on the Association Between Apolipoprotein E Genotype and Alzheimer Disease: A Meta-Analysis." *JAMA* 278 (16): 1349–56. https://doi.org/10.1001/jama.1997.03550160069041.

Ferrari, Camilla, and Sandro Sorbi. 2021. "The Complexity of Alzheimer's Disease: An Evolving Puzzle." *Physiological Reviews* 101 (3): 1047–81. https://doi.org/10.1152/physrev.00015.2020.

Fjell, A. M., K. B. Walhovd, I. Amlien, A. Bjørnerud, I. Reinvang, L. Gjerstad, T. Cappelen, et al. 2008. "Morphometric Changes in the Episodic Memory Network and Tau Pathologic Features Correlate with Memory Performance in Patients with Mild Cognitive Impairment." *AJNR. American Journal of Neuroradiology* 29 (6): 1183–89. https://doi.org/10.3174/ajnr.A1059.

Fjell, Anders M., Kristine B. Walhovd, Christine Fennema-Notestine, Linda K. McEvoy, Donald J. Hagler, Dominic Holland, James B. Brewer, and Anders M. Dale. 2010. "CSF Biomarkers in Prediction of Cerebral and Clinical Change in Mild Cognitive Impairment and Alzheimer's Disease." *The Journal of Neuroscience* 30 (6): 2088–2101. https://doi.org/10.1523/JNEUROSCI.3785-09.2010.

Fleisher, A. S., B. B. Sowell, C. Taylor, A. C. Gamst, R. C. Petersen, L. J. Thal, and Alzheimer's Disease Cooperative Study. 2007. "Clinical Predictors of Progression to Alzheimer Disease in Amnestic Mild Cognitive Impairment." *Neurology* 68 (19): 1588–95. https://doi.org/10.1212/01.wnl.0000258542.58725.4c.

Forabosco, Paola, Adaikalavan Ramasamy, Daniah Trabzuni, Robert Walker, Colin Smith, Jose Bras, Adam P. Levine, et al. 2013. "Insights into TREM2 Biology by Network Analysis of Human Brain Gene Expression Data." *Neurobiology of Aging* 34 (12): 2699–2714. https://doi.org/10.1016/j.neurobiolaging.2013.05.001.

Fourier, Anthony, Erik Portelius, Henrik Zetterberg, Kaj Blennow, Isabelle Quadrio, and Armand Perret-Liaudet. 2015. "Pre-Analytical and Analytical Factors Influencing Alzheimer's Disease Cerebrospinal Fluid Biomarker Variability." *Clinica Chimica Acta*, Biomarkers in Neuropsychiatric Disorders, 449 (September): 9–15. https://doi.org/10.1016/j.cca.2015.05.024.

Frisoni, Giovanni B., Nick C. Fox, Clifford R. Jack, Philip Scheltens, and Paul M. Thompson. 2010. “The Clinical Use of Structural MRI in Alzheimer Disease.” *Nature Reviews. Neurology* 6 (2): 67–77. https://doi.org/10.1038/nrneurol.2009.215.

Ghisays, Valentina, Willemijn J. Jansen, Yinghua Chen, Hillary D. Protas, Michael H. Malek-Ahmadi, Ji Luo, Wendy Lee, et al. 2021. “CSF Amyloid-beta, Tau, and Neurodegenerative and Inflammatory Biomarkers in Cognitively Unimpaired Late Middle-aged and Older Adult *APOE E4* Homozygotes, Heterozygotes, and Non-carriers from the Arizona APOE Cohort.” *Alzheimer’s & Dementia* 17 (S5). https://doi.org/10.1002/alz.057242.

Gibbons, Laura E., Adam C. Carle, R. Scott Mackin, Danielle Harvey, Shubhabrata Mukherjee, Philip Insel, S. McKay Curtis, Dan Mungas, Paul K. Crane, and for the Alzheimer’s Disease Neuroimaging Initiative. 2012. “A Composite Score for Executive Functioning, Validated in Alzheimer’s Disease Neuroimaging Initiative (ADNI) Participants with Baseline Mild Cognitive Impairment.” *Brain Imaging and Behavior* 6 (4): 517–27. https://doi.org/10.1007/s11682-012-9176-1.

Golde, Todd E, Christopher B Eckman, and Steven G Younkin. 2000. “Biochemical Detection of Aβ Isoforms: Implications for Pathogenesis, Diagnosis, and Treatment of Alzheimer’s Disease.” *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1502 (1): 172–87. https://doi.org/10.1016/S0925-4439(00)00043-0.

Graeber, M. B., W. Tetzlaff, W. J. Streit, and G. W. Kreutzberg. 1988. “Microglial Cells but Not Astrocytes Undergo Mitosis Following Rat Facial Nerve Axotomy.” *Neuroscience Letters* 85 (3): 317–21. https://doi.org/10.1016/0304-3940(88)90585-x.

Guerreiro, Rita, Aleksandra Wojtas, Jose Bras, Minerva Carrasquillo, Ekaterina Rogaeva, Elisa Majounie, Carlos Cruchaga, et al. 2013. “TREM2 Variants in Alzheimer’s Disease.” *The New England Journal of Medicine* 368 (2): 117–27. https://doi.org/10.1056/NEJMoa1211851.

Gustafson, Deborah R., Ingmar Skoog, Lars Rosengren, Henrik Zetterberg, and Kaj Blennow. 2007. “Cerebrospinal Fluid β-Amyloid 1–42 Concentration May Predict Cognitive Decline in Older Women.” *Journal of Neurology, Neurosurgery & Psychiatry* 78 (5): 461–64. https://doi.org/10.1136/jnnp.2006.100529.

Haass, Christian, and Dennis J. Selkoe. 2007. “Soluble Protein Oligomers in Neurodegeneration: Lessons from the Alzheimer’s Amyloid β-Peptide.” *Nature Reviews Molecular Cell Biology* 8 (2): 101–12. https://doi.org/10.1038/nrm2101.

Hadjichrysanthou, Christoforos, Stephanie Evans, Sumali Bajaj, Loizos C. Siakallis, Kevin McRae-McKee, Frank de Wolf, Roy M. Anderson, and the Alzheimer’s Disease Neuroimaging Initiative. 2020. “The Dynamics of Biomarkers across the Clinical Spectrum of Alzheimer’s Disease.” *Alzheimer’s Research & Therapy* 12 (1): 74. https://doi.org/10.1186/s13195-020-00636-z.

Hansson, Oskar, John Seibyl, Erik Stomrud, Henrik Zetterberg, John Q. Trojanowski, Tobias Bittner, Valeria Lifke, et al. 2018. “CSF Biomarkers of Alzheimer’s Disease Concord with Amyloid-β PETand Predict Clinical Progression: A Study of Fully Automated Immunoassays in BioFINDER and ADNI Cohorts.” *Alzheimer’s & Dementia : The Journal of the Alzheimer’s Association* 14 (11): 1470–81. https://doi.org/10.1016/j.jalz.2018.01.010.

Hansson, Oskar, Henrik Zetterberg, Peder Buchhave, Ulf Andreasson, Elisabet Londos, Lennart Minthon, and Kaj Blennow. 2007. “Prediction of Alzheimer’s Disease Using the CSF Aβ42/Aβ40 Ratio in Patients with Mild Cognitive Impairment.” *Dementia and Geriatric Cognitive Disorders* 23 (5): 316–20. https://doi.org/10.1159/000100926.

Hansson, Oskar, Henrik Zetterberg, Peder Buchhave, Elisabet Londos, Kaj Blennow, and Lennart Minthon. 2006. “Association between CSF Biomarkers and Incipient Alzheimer’s Disease in Patients with Mild Cognitive Impairment: A Follow-up Study.” *The Lancet Neurology* 5 (3): 228–34. https://doi.org/10.1016/S1474-4422(06)70355-6.

Hardy, John, and Dennis J. Selkoe. 2002. “The Amyloid Hypothesis of Alzheimer’s Disease: Progress and Problems on the Road to Therapeutics.” *Science (New York, N.Y.)* 297 (5580): 353–56. https://doi.org/10.1126/science.1072994.

Herukka, Sanna-Kaisa, Seppo Helisalmi, Merja Hallikainen, Susanna Tervo, Hilkka Soininen, and Tuula Pirttilä. 2007. “CSF Abeta42, Tau and Phosphorylated Tau, APOE Epsilon4 Allele and MCI Type in Progressive MCI.” *Neurobiology of Aging* 28 (4): 507–14. https://doi.org/10.1016/j.neurobiolaging.2006.02.001.

Heslegrave, Amanda, Wendy Heywood, Ross Paterson, Nadia Magdalinou, Johan Svensson, Per Johansson, Annika Öhrfelt, et al. 2016. “Increased Cerebrospinal Fluid Soluble TREM2 Concentration in Alzheimer’s Disease.” *Molecular Neurodegeneration* 11 (1): 3. https://doi.org/10.1186/s13024-016-0071-x.

Hickman, Suzanne E., Elizabeth K. Allison, and Joseph El Khoury. 2008. “Microglial Dysfunction and Defective Beta-Amyloid Clearance Pathways in Aging Alzheimer’s Disease Mice.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 28 (33): 8354–60. https://doi.org/10.1523/JNEUROSCI.0616-08.2008.

Hok-A-Hin, Yanaika S., Eline A. J. Willemse, Charlotte E. Teunissen, and Marta Del Campo. 2019. “Guidelines for CSF Processing and Biobanking: Impact on the Identification and Development of Optimal CSF Protein Biomarkers.” *Methods in Molecular Biology (Clifton, N.J.)* 2044: 27–50. https://doi.org/10.1007/978-1-4939-9706-0_2.

Hort, J., A. Bartos, T. Pirttilä, and P. Scheltens. 2010. “Use of Cerebrospinal Fluid Biomarkers in Diagnosis of Dementia across Europe.” *European Journal of Neurology* 17 (1): 90–96. https://doi.org/10.1111/j.1468-1331.2009.02753.x.

Huang, Jin, Feiping Nie, and Heng Huang. 2013. “Spectral Rotation versus K-Means in Spectral Clustering.” *Proceedings of the AAAI Conference on Artificial Intelligence* 27 (1): 431–37. https://doi.org/10.1609/aaai.v27i1.8683.

Huang, Youtong, Kaisa E. Happonen, Patrick G. Burrola, Carolyn O’Connor, Nasun Hah, Ling Huang, Axel Nimmerjahn, and Greg Lemke. 2021. “Microglia Use TAM Receptors to Detect and Engulf Amyloid β Plaques.” *Nature Immunology* 22 (5): 586–94. https://doi.org/10.1038/s41590-021-00913-5.

Huynh, Tien-Phat V., Albert A. Davis, Jason D. Ulrich, and David M. Holtzman. 2017. “Apolipoprotein E and Alzheimer’s Disease: The Influence of Apolipoprotein E on Amyloid-β and Other Amyloidogenic Proteins.” *Journal of Lipid Research* 58 (5): 824–36. https://doi.org/10.1194/jlr.R075481.

Iwata, N., S. Tsubuki, Y. Takaki, K. Shirotani, B. Lu, N. P. Gerard, C. Gerard, E. Hama, H. J. Lee, and T. C. Saido. 2001. “Metabolic Regulation of Brain Abeta by Neprilysin.” *Science (New York, N.Y.)* 292 (5521): 1550–52. https://doi.org/10.1126/science.1059946.

Jack, Clifford R., David S. Knopman, William J. Jagust, Ronald C. Petersen, Michael W. Weiner, Paul S. Aisen, Leslie M. Shaw, et al. 2013. “Update on Hypothetical Model of Alzheimer’s Disease Biomarkers.” *Lancet Neurology* 12 (2): 207–16. https://doi.org/10.1016/S1474-4422(12)70291-0.

Jellinger, Kurt A. 2022. “Recent Update on the Heterogeneity of the Alzheimer’s Disease Spectrum.” *Journal of Neural Transmission* 129 (1): 1–24. https://doi.org/10.1007/s00702-021-02449-2.

Jonsson, Thorlakur, Hreinn Stefansson, Stacy Steinberg, Ingileif Jonsdottir, Palmi V. Jonsson, Jon Snaedal, Sigurbjorn Bjornsson, et al. 2013. “Variant of TREM2 Associated with the Risk of Alzheimer’s Disease.” *The New England Journal of Medicine* 368 (2): 107–16. https://doi.org/10.1056/NEJMoa1211103.

Juottonen, K., M. P. Laakso, R. Insausti, M. Lehtovirta, A. Pitkänen, K. Partanen, and H. Soininen. 1998. “Volumes of the Entorhinal and Perirhinal Cortices in Alzheimer’s Disease.” *Neurobiology of Aging* 19 (1): 15–22. https://doi.org/10.1016/s0197-4580(98)00007-4.

Karch, Celeste M., and Alison M. Goate. 2015. “Alzheimer’s Disease Risk Genes and Mechanisms of Disease Pathogenesis.” *Biological Psychiatry* 77 (1): 43–51. https://doi.org/10.1016/j.biopsych.2014.05.006.

Killiany, R. J., B. T. Hyman, T. Gomez-Isla, M. B. Moss, R. Kikinis, F. Jolesz, R. Tanzi, K. Jones, and M. S. Albert. 2002. “MRI Measures of Entorhinal Cortex vs Hippocampus in Preclinical AD.” *Neurology* 58 (8): 1188–96. https://doi.org/10.1212/WNL.58.8.1188.

Kim, Jungsu, Jacob M. Basak, and David M. Holtzman. 2009. “The Role of Apolipoprotein E in Alzheimer’s Disease.” *Neuron* 63 (3): 287–303. https://doi.org/10.1016/j.neuron.2009.06.026.

Kinney, Jefferson W., Shane M. Bemiller, Andrew S. Murtishaw, Amanda M. Leisgang, Arnold M. Salazar, and Bruce T. Lamb. 2018. “Inflammation as a Central Mechanism in Alzheimer’s Disease.” *Alzheimer’s & Dementia : Translational Research & Clinical Interventions* 4 (September): 575–90. https://doi.org/10.1016/j.trci.2018.06.014.

Kleinberger, Gernot, Yoshinori Yamanishi, Marc Suárez-Calvet, Eva Czirr, Ebba Lohmann, Elise Cuyvers, Hanne Struyfs, et al. 2014. “TREM2 Mutations Implicated in Neurodegeneration Impair Cell Surface Transport and Phagocytosis.” *Science Translational Medicine* 6 (243): 243ra86-243ra86. https://doi.org/pet.

Knapskog, Anne-Brita, Kristi Henjum, Ane-Victoria Idland, Rannveig Sakshaug Eldholm, Karin Persson, Ingvild Saltvedt, Leiv Otto Watne, Knut Engedal, and Lars N. G. Nilsson. 2020. “Cerebrospinal Fluid STREM2 in Alzheimer’s Disease: Comparisons between Clinical Presentation and AT Classification.” *Scientific Reports* 10 (1): 15886. https://doi.org/10.1038/s41598-020-72878-8.

Koopman, Karen, Nathalie Le Bastard, Jean-Jacques Martin, Guy Nagels, Peter P. De Deyn, and Sebastiaan Engelborghs. 2009. “Improved Discrimination of Autopsy-Confirmed Alzheimer’s Disease (AD) from Non-AD Dementias Using CSF P-Tau(181P).” *Neurochemistry International* 55 (4): 214–18. https://doi.org/10.1016/j.neuint.2009.02.017.

Korf, Esther S. C., Lars-Olof Wahlund, Pieter Jelle Visser, and Philip Scheltens. 2004. “Medial Temporal Lobe Atrophy on MRI Predicts Dementia in Patients with Mild Cognitive Impairment.” *Neurology* 63 (1): 94–100. https://doi.org/10.1212/01.wnl.0000133114.92694.93.

Krabbe, Grietje, Annett Halle, Vitali Matyash, Jan L. Rinnenthal, Gina D. Eom, Ulrike Bernhardt, Kelly R. Miller, Stefan Prokop, Helmut Kettenmann, and Frank L. Heppner. 2013. “Functional Impairment of Microglia Coincides with Beta-Amyloid Deposition in Mice with Alzheimer-like Pathology.” *PloS One* 8 (4): e60921. https://doi.org/10.1371/journal.pone.0060921.

Lehéricy, S., M. Baulac, J. Chiras, L. Piérot, N. Martin, B. Pillon, B. Deweer, B. Dubois, and C. Marsault. 1994. “Amygdalohippocampal MR Volume Measurements in the Early Stages of Alzheimer Disease.” *AJNR. American Journal of Neuroradiology* 15 (5): 929–37.

Leon, Mony J. de, Elizabeth Pirraglia, Ricardo S. Osorio, Lidia Glodzik, Les Saint-Louis, Hee-Jin Kim, Juan Fortea, et al. 2018. “The Nonlinear Relationship between Cerebrospinal Fluid Aβ42 and Tau in Preclinical Alzheimer’s Disease.” *PLoS ONE* 13 (2): e0191240. https://doi.org/10.1371/journal.pone.0191240.

Leoni, Valerio. 2011. “The Effect of Apolipoprotein E (ApoE) Genotype on Biomarkers of Amyloidogenesis, Tau Pathology and Neurodegeneration in Alzheimer’s Disease.” *Clinical Chemistry and Laboratory Medicine* 49 (3): 375–83. https://doi.org/10.1515/CCLM.2011.088.

Levy-Lahad, E., W. Wasco, P. Poorkaj, D. M. Romano, J. Oshima, W. H. Pettingell, C. E. Yu, P. D. Jondro, S. D. Schmidt, and K. Wang. 1995. “Candidate Gene for the Chromosome 1 Familial Alzheimer’s Disease Locus.” *Science (New York, N.Y.)* 269 (5226): 973–77. https://doi.org/10.1126/science.7638622.

Lewczuk, Piotr, Hermann Esselmann, Markus Otto, Juan Manuel Maler, Andreas Wolfram Henkel, Maria Kerstin Henkel, Oliver Eikenberg, et al. 2004. “Neurochemical Diagnosis of Alzheimer’s Dementia by CSF Abeta42, Abeta42/Abeta40 Ratio and Total Tau.” *Neurobiology of Aging* 25 (3): 273–81. https://doi.org/10.1016/S0197-4580(03)00086-1.

Li, G., I. Sokal, J. F. Quinn, J. B. Leverenz, M. Brodey, G. D. Schellenberg, J. A. Kaye, et al. 2007. “CSF Tau/Aβ42 Ratio for Increased Risk of Mild Cognitive Impairment: A Follow-up Study.” *Neurology* 69 (7): 631–39. https://doi.org/10.1212/01.wnl.0000267428.62582.aa.

Lifke, Valeria, Gwendlyn Kollmorgen, Ekaterina Manuilova, Tobias Oelschlaegel, Lars Hillringhaus, Monika Widmann, Christine A. F. von Arnim, et al. 2019. “Elecsys® Total-Tau and Phospho-Tau (181P) CSF Assays: Analytical Performance of the Novel, Fully Automated Immunoassays for Quantification of Tau Proteins in Human Cerebrospinal Fluid.” *Clinical Biochemistry*, Alzheimer’s Disease: Making the Point, 72 (October): 30–38. https://doi.org/10.1016/j.clinbiochem.2019.05.005.

Liu, Fayao, Luping Zhou, Chunhua Shen, and Jianping Yin. 2014. “Multiple Kernel Learning in the Primal for Multimodal Alzheimer’s Disease Classification.” *IEEE Journal of Biomedical and Health Informatics* 18 (3): 984–90. https://doi.org/10.1109/JBHI.2013.2285378.

Livingston, Gill, Andrew Sommerlad, Vasiliki Orgeta, Sergi G. Costafreda, Jonathan Huntley, David Ames, Clive Ballard, et al. 2017. “Dementia Prevention, Intervention, and Care.” *Lancet (London, England)* 390 (10113): 2673–2734. https://doi.org/10.1016/S0140-6736(17)31363-6.

Llibre-Guerra, Jorge J., Yan Li, Suzanne E. Schindler, Brian A. Gordon, Anne M. Fagan, John C. Morris, Tammie L. S. Benzinger, et al. 2019. “Association of Longitudinal Changes in Cerebrospinal Fluid Total Tau and Phosphorylated Tau 181 and Brain Atrophy With Disease Progression in Patients With Alzheimer Disease.” *JAMA Network Open* 2 (12): e1917126. https://doi.org/10.1001/jamanetworkopen.2019.17126.

Maddalena, Alessia, Andreas Papassotiropoulos, Britta Müller-Tillmanns, Hans H. Jung, Thomas Hegi, Roger M. Nitsch, and Christoph Hock. 2003. “Biochemical Diagnosis of Alzheimer Disease by Measuring the Cerebrospinal Fluid Ratio of Phosphorylated Tau Protein to β-Amyloid Peptide42.” *Archives of Neurology* 60 (9): 1202–6. https://doi.org/10.1001/archneur.60.9.1202.

Mahaman, Yacoubou Abdoul Razak, Kidane Siele Embaye, Fang Huang, Longfei Li, Feiqi Zhu, Jian-Zhi Wang, Rong Liu, Jun Feng, and Xiaochuan Wang. 2022. “Biomarkers Used in Alzheimer’s Disease Diagnosis, Treatment, and Prevention.”

Ageing Research Reviews 74 (February): 101544. https://doi.org/10.1016/j.arr.2021.101544.

Markello, Ross D., Golia Shafiei, Christina Tremblay, Ronald B. Postuma, Alain Dagher, and Bratislav Misic. 2021. “Multimodal Phenotypic Axes of Parkinson’s Disease.” *Npj Parkinson’s Disease* 7 (1): 1–12. https://doi.org/10.1038/s41531-020-00144-9.

Mattsson, Niklas, Henrik Zetterberg, Oskar Hansson, Niels Andreasen, Lucilla Parnetti, Michael Jonsson, Sanna-Kaisa Herukka, et al. 2009. “CSF Biomarkers and Incipient Alzheimer Disease in Patients With Mild Cognitive Impairment.” *JAMA* 302 (4): 385–93. https://doi.org/10.1001/jama.2009.1064.

Mawuenyega, Kwasi G., Wendy Sigurdson, Vitaliy Ovod, Ling Munsell, Tom Kasten, John C. Morris, Kevin E. Yarasheski, and Randall J. Bateman. 2010a. “Decreased Clearance of CNS Amyloid-β in Alzheimer’s Disease.” *Science (New York, N.Y.)* 330 (6012): 1774. https://doi.org/10.1126/science.1197623.

———. 2010b. “Decreased Clearance of CNS Beta-Amyloid in Alzheimer’s Disease.” *Science (New York, N.Y.)* 330 (6012): 1774. https://doi.org/10.1126/science.1197623.

Meda, L., M. A. Cassatella, G. I. Szendrei, L. Otvos, P. Baron, M. Villalba, D. Ferrari, and F. Rossi. 1995. “Activation of Microglial Cells by Beta-Amyloid Protein and Interferon-Gamma.” *Nature* 374 (6523): 647–50. https://doi.org/10.1038/374647a0.

Michelucci, Alessandro, Tony Heurtaux, Luc Grandbarbe, Eleonora Morga, and Paul Heuschling. 2009. “Characterization of the Microglial Phenotype under Specific Pro-Inflammatory and Anti-Inflammatory Conditions: Effects of Oligomeric and Fibrillar Amyloid-Beta.” *Journal of Neuroimmunology* 210 (1–2): 3–12. https://doi.org/10.1016/j.jneuroim.2009.02.003.

Morar, Ulyana, Walter Izquierdo, Harold Martin, Parisa Forouzannezhad, Elaheh Zarafshan, Elona Unger, Zoran Bursac, et al. 2022. “A Study of the Longitudinal Changes in Multiple Cerebrospinal Fluid and Volumetric Magnetic Resonance Imaging Biomarkers on Converter and Non-converter Alzheimer’s Disease Subjects with Consideration for Their Amyloid Beta Status.” *Alzheimer’s & Dementia: Diagnosis, Assessment & Disease Monitoring* 14 (1). https://doi.org/10.1002/dad2.12258.

Morenas-Rodríguez, Estrella, Yan Li, Brigitte Nuscher, Nicolai Franzmeier, Chengjie Xiong, Marc Suárez-Calvet, Anne M. Fagan, et al. 2022. “Soluble TREM2 in CSF and Its Association with Other Biomarkers and Cognition in Autosomal-Dominant Alzheimer’s Disease: A Longitudinal Observational Study.” *The Lancet. Neurology* 21 (4): 329–41. https://doi.org/10.1016/S1474-4422(22)00027-8.

Morris, John C., Catherine M. Roe, Chengjie Xiong, Anne M. Fagan, Alison M. Goate, David M. Holtzman, and Mark A. Mintun. 2010. “APOE Predicts Amyloid-Beta but Not Tau Alzheimer Pathology in Cognitively Normal Aging.” *Annals of Neurology* 67 (1): 122–31. https://doi.org/10.1002/ana.21843.

Mrak, Robert E. 2012. “Microglia in Alzheimer Brain: A Neuropathological Perspective.” *International Journal of Alzheimer’s Disease* 2012: 165021. https://doi.org/10.1155/2012/165021.

Mungas, D., and B. R. Reed. 2000. “Application of Item Response Theory for Development of a Global Functioning Measure of Dementia with Linear Measurement Properties.” *Statistics in Medicine* 19 (11–12): 1631–44. https://doi.org/Does the Cerebrospinal Fluid Protein Concentration Increase the Risk of Shunt Complications.

Murphy, E. A., D. Holland, M. Donohue, L. K. McEvoy, D. J. Hagler, A. M. Dale, J. B. Brewer, and Alzheimer’s Disease Neuroimaging Initiative. 2010. “Six-Month Atrophy in MTL Structures Is Associated with Subsequent Memory Decline in Elderly Controls.” *NeuroImage* 53 (4): 1310–17. https://doi.org/10.1016/j.neuroimage.2010.07.016.

N’Diaye, Elsa-Noah, Catherine S. Branda, Steven S. Branda, Lisette Nevarez, Marco Colonna, Clifford Lowell, Jessica A. Hamerman, and William E. Seaman. 2009. “TREM-2 (Triggering Receptor Expressed on Myeloid Cells 2) Is a Phagocytic Receptor for Bacteria.” *The Journal of Cell Biology* 184 (2): 215–23. https://doi.org/10.1083/jcb.200808080.

Nelson, Peter T., Irina Alafuzoff, Eileen H. Bigio, Constantin Bouras, Heiko Braak, Nigel J. Cairns, Rudolph J. Castellani, et al. 2012. “Correlation of Alzheimer Disease Neuropathologic Changes with Cognitive Status: A Review of the Literature.” *Journal of Neuropathology and Experimental Neurology* 71 (5): 362–81. https://doi.org/10.1097/NEN.0b013e31825018f7.

Nichols, Emma, Jaimie D Steinmetz, Stein Emil Vollset, Kai Fukutaki, Julian Chalek, Foad Abd-Allah, Amir Abdoli, et al. 2022. “Estimation of the Global Prevalence of Dementia in 2019 and Forecasted Prevalence in 2050: An Analysis for the Global Burden of Disease Study 2019.” *The Lancet Public Health* 7 (2): e105–25. https://doi.org/10.1016/S2468-2667(21)00249-8.

O’Brien, Richard J., and Philip C. Wong. 2011. “Amyloid Precursor Protein Processing and Alzheimer’s Disease.” *Annual Review of Neuroscience* 34: 185–204. https://doi.org/10.1146/annurev-neuro-061010-113613.

Olsson, Annika, Hugo Vanderstichele, Niels Andreasen, Geert De Meyer, Anders Wallin, Björn Holmberg, Lars Rosengren, Eugeen Vanmechelen, and Kaj Blennow. 2005. “Simultaneous Measurement of Beta-Amyloid(1-42), Total Tau, and Phosphorylated Tau (Thr181) in Cerebrospinal Fluid by the XMAP Technology.” *Clinical Chemistry* 51 (2): 336–45. https://doi.org/10.1373/clinchem.2004.039347.

Palmqvist, Sebastian, Philip S. Insel, Erik Stomrud, Shorena Janelidze, Henrik Zetterberg, Britta Brix, Udo Eichenlaub, et al. 2019. “Cerebrospinal Fluid and Plasma Biomarker Trajectories with Increasing Amyloid Deposition in Alzheimer’s Disease.” *EMBO Molecular Medicine* 11 (12): e11170. https://doi.org/10.15252/emmm.201911170.

Palmqvist, Sebastian, Niklas Mattsson, and Oskar Hansson. 2016. “Cerebrospinal Fluid Analysis Detects Cerebral Amyloid-β Accumulation Earlier than Positron Emission Tomography.” *Brain* 139 (4): 1226–36. https://doi.org/10.1093/brain/aww015.

Parhizkar, Samira, Thomas Arzberger, Matthias Brendel, Gernot Kleinberger, Maximilian Deussing, Carola Focke, Brigitte Nuscher, et al. 2019. “Loss of TREM2 Function Increases Amyloid Seeding but Reduces Plaque-Associated ApoE.” *Nature Neuroscience* 22 (2): 191–204. https://doi.org/10.1038/s41593-018-0296-9.

Park, Seyoung, and Hongyu Zhao. 2018. “Spectral Clustering Based on Learning Similarity Matrix.” *Bioinformatics (Oxford, England)* 34 (12): 2069–76. https://doi.org/10.1093/bioinformatics/bty050.

Park, So-Hee, Eun-Hye Lee, Hyung-Ji Kim, Sungyang Jo, Sunju Lee, Sang Won Seo, Hyun-Hee Park, Seong-Ho Koh, and Jae-Hong Lee. 2021. “The Relationship of Soluble TREM2 to Other Biomarkers of Sporadic Alzheimer’s Disease.” *Scientific Reports* 11 (1): 13050. https://doi.org/10.1038/s41598-021-92101-6.

Pascoal, Tharick A., Sulantha Mathotaarachchi, Monica Shin, Ah Yeon Park, Sara Mohades, Andrea L. Benedet, Min Su Kang, et al. 2018. “Amyloid and Tau Signatures of Brain Metabolic Decline in Preclinical Alzheimer’s Disease.” *European Journal of Nuclear Medicine and Molecular Imaging* 45 (6): 1021–30. https://doi.org/10.1007/s00259-018-3933-3.

Petersen, Gabriel Cassinelli, Michelle Roytman, Gloria C. Chiang, Yi Li, Marc L. Gordon, and Ana M. Franceschi. 2022. “Overview of Tau PET Molecular Imaging.” *Current Opinion in Neurology* 35 (2): 230–39. https://doi.org/10.1097/WCO.0000000000001035.

Petersen, R. C., P. S. Aisen, L. A. Beckett, M. C. Donohue, A. C. Gamst, D. J. Harvey, C. R. Jack, et al. 2010. “Alzheimer’s Disease Neuroimaging Initiative (ADNI): Clinical Characterization.” *Neurology* 74 (3): 201–9. https://doi.org/10.1212/WNL.0b013e3181cb3e25.

Petersen, R. C., G. E. Smith, R. J. Ivnik, E. G. Tangalos, D. J. Schaid, S. N. Thibodeau, E. Kokmen, S. C. Waring, and L. T. Kurland. 1995. “Apolipoprotein E Status as a Predictor of the Development of Alzheimer’s Disease in Memory-Impaired Individuals.” *JAMA* 273 (16): 1274–78.

Piccio, Laura, Cecilia Buonsanti, Marina Cella, Ilaria Tassi, Robert E. Schmidt, Chiara Fenoglio, John Rinker II, et al. 2008. “Identification of Soluble TREM-2 in the Cerebrospinal Fluid and Its Association with Multiple Sclerosis and CNS Inflammation.” *Brain* 131 (11): 3081–91. https://doi.org/10.1093/brain/awn217.

Piccio, Laura, Yuetiva Deming, Jorge L. Del-Águila, Laura Ghezzi, David M. Holtzman, Anne M. Fagan, Chiara Fenoglio, Daniela Galimberti, Barbara Borroni, and Carlos Cruchaga. 2016. “Cerebrospinal Fluid Soluble TREM2 Is Higher in Alzheimer Disease

and Associated with Mutation Status." *Acta Neuropathologica* 131 (6): 925–33. https://doi.org/10.1007/s00401-016-1533-5.

Pillai, Jagan A., Maria Khrestian, James Bena, James B. Leverenz, and Lynn M. Bekris. 2021. "Temporal Ordering of Inflammatory Analytes STNFR2 and STREM2 in Relation to Alzheimer's Disease Biomarkers and Clinical Outcomes." *Frontiers in Aging Neuroscience* 13 (June): 676744. https://doi.org/10.3389/fnagi.2021.676744.

Popescu, Sebastian G., Alex Whittington, Roger N. Gunn, Paul M. Matthews, Ben Glocker, David J Sharp, James H Cole, and for the Alzheimer's Disease Neuroimaging Initiative. 2020. "Nonlinear Biomarker Interactions in Conversion from Mild Cognitive Impairment to Alzheimer's Disease." *Human Brain Mapping* 41 (15): 4406–18. https://doi.org/10.1002/hbm.25133.

Prince, Martin, Anders Wimo, Maelenn Guerchet, Gemma Claire Ali, Yu-Tzu Wu, Matthew Prina, and others. 2015. "The Global Impact of Dementia: An Analysis of Prevalence, Incidence, Cost and Trends." *World Alzheimer Report* 2015.

Prosser, Lloyd, Amy Macdougall, Carole H. Sudre, Emily N. Manning, Ian B. Malone, Phoebe Walsh, Olivia Goodkin, et al. 2023. "Predicting Cognitive Decline in Older Adults Using Baseline Metrics of AD Pathologies, Cerebrovascular Disease, and Neurodegeneration." *Neurology* 100 (8): e834–45. https://doi.org/10.1212/WNL.0000000000201572.

Raphael, Benjamin J., Ralph H. Hruban, Andrew J. Aguirre, Richard A. Moffitt, Jen Jen Yeh, Chip Stewart, A. Gordon Robertson, et al. 2017. "Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma." *Cancer Cell* 32 (2): 185-203.e13. https://doi.org/10.1016/j.ccell.2017.07.007.

Rasmussen, Martin Kaag, Humberto Mestre, and Maiken Nedergaard. 2018. "The Glymphatic Pathway in Neurological Disorders." *The Lancet. Neurology* 17 (11): 1016–24. https://doi.org/10.1016/S1474-4422(18)30318-1.

Rechberger, Simon, Yong Li, Sebastian J. Kopetzky, and Markus Butz-Ostendorf. 2022. "Automated High-Definition MRI Processing Routine Robustly Detects Longitudinal Morphometry Changes in Alzheimer's Disease Patients." *Frontiers in Aging Neuroscience* 14 (June): 832828. https://doi.org/10.3389/fnagi.2022.832828.

Roe, Catherine M., Anne M. Fagan, Elizabeth A. Grant, Jason Hassenstab, Krista L. Moulder, Denise Maue Dreyfus, Courtney L. Sutphen, et al. 2013. "Amyloid Imaging and CSF Biomarkers in Predicting Cognitive Impairment up to 7.5 Years Later." *Neurology* 80 (19): 1784–91. https://doi.org/10.1212/WNL.0b013e3182918ca6.

Salvadó, Gemma, Victoria Larsson, Karly A Cody, Nicholas C Cullen, Erin M Jonaitis, Erik Stomrud, Gwendlyn Kollmorgen, et al. 2023. "Optimal Combinations of CSF Biomarkers for Predicting Cognitive Decline and Clinical Conversion in Cognitively Unimpaired Participants and Mild Cognitive Impairment Patients: A Multi-Cohort Study." *Alzheimer's & Dementia* n/a (n/a). https://doi.org/10.1002/alz.12907.

Saykin, A. J., H. A. Wishart, L. A. Rabin, R. B. Santulli, L. A. Flashman, J. D. West, T. L. McHugh, and A. C. Mamourian. 2006. “Older Adults with Cognitive Complaints Show Brain Atrophy Similar to That of Amnestic MCI.” *Neurology* 67 (5): 834–42. https://doi.org/10.1212/01.wnl.0000234032.77541.a2.

Schmand, B., H. M. Huizenga, and W. A. van Gool. 2010. “Meta-Analysis of CSF and MRI Biomarkers for Detecting Preclinical Alzheimer’s Disease.” *Psychological Medicine* 40 (1): 135–45. https://doi.org/10.1017/S0033291709991516.

Selkoe, Dennis J. 2003. “Aging, Amyloid, and Alzheimer’s Disease: A Perspective in Honor of Carl Cotman.” *Neurochemical Research* 28 (11): 1705–13. https://doi.org/10.1023/a:1026065122854.

Shaftel, Solomon S., Stephanos Kyrkanides, John A. Olschowka, Jen-nie H. Miller, Renee E. Johnson, and M. Kerry O’Banion. 2007. “Sustained Hippocampal IL-1 Beta Overexpression Mediates Chronic Neuroinflammation and Ameliorates Alzheimer Plaque Pathology.” *The Journal of Clinical Investigation* 117 (6): 1595–1604. https://doi.org/10.1172/JCI31450.

Shaw, Leslie M., Oskar Hansson, Ekaterina Manuilova, Colin L. Masters, James D. Doecke, Qiao-Xin Li, Sandra Rutz, Monika Widmann, Andreas Leinenbach, and Kaj Blennow. 2019. “Method Comparison Study of the Elecsys® β-Amyloid (1–42) CSF Assay versus Comparator Assays and LC-MS/MS.” *Clinical Biochemistry*, Alzheimer’s Disease: Making the Point, 72 (October): 7–14. https://doi.org/10.1016/j.clinbiochem.2019.05.006.

Shaw, Leslie M., Hugo Vanderstichele, Malgorzata Knapik-Czajka, Christopher M. Clark, Paul S. Aisen, Ronald C. Petersen, Kaj Blennow, et al. 2009. “Cerebrospinal Fluid Biomarker Signature in Alzheimer’s Disease Neuroimaging Initiative Subjects.” *Annals of Neurology* 65 (4): 403–13. https://doi.org/10.1002/ana.21610.

Shaw, Leslie M., Teresa Waligorska, Leona Fields, Magdalena Korecka, Michal Figurski, John Q. Trojanowski, Udo Eichenlaub, et al. 2018. “Derivation of Cutoffs for the Elecsys® Amyloid β (1–42) Assay in Alzheimer’s Disease.” *Alzheimer’s & Dementia: Diagnosis, Assessment & Disease Monitoring* 10 (1): 698–705. https://doi.org/10.1016/j.dadm.2018.07.002.

Sheng, J. G., X. Q. Zhou, R. E. Mrak, and W. S. Griffin. 1998. “Progressive Neuronal Injury Associated with Amyloid Plaque Formation in Alzheimer Disease.” *Journal of Neuropathology and Experimental Neurology* 57 (7): 714–17. https://doi.org/10.1097/00005072-199807000-00008.

Sherrington, R., E. I. Rogaev, Y. Liang, E. A. Rogaeva, G. Levesque, M. Ikeda, H. Chi, et al. 1995. “Cloning of a Gene Bearing Missense Mutations in Early-Onset Familial Alzheimer’s Disease.” *Nature* 375 (6534): 754–60. https://doi.org/10.1038/375754a0.

Shoji, M., E. Matsubara, M. Kanai, M. Watanabe, T. Nakamura, Y. Tomidokoro, M. Shizuka, et al. 1998. “Combination Assay of CSF Tau, A Beta 1-40 and A Beta 1-42(43)

as a Biochemical Marker of Alzheimer's Disease." *Journal of the Neurological Sciences* 158 (2): 134–40. https://doi.org/10.1016/s0022-510x(98)00122-1.

Simard, Alain R., Denis Soulet, Genevieve Gowing, Jean-Pierre Julien, and Serge Rivest. 2006. "Bone Marrow-Derived Microglia Play a Critical Role in Restricting Senile Plaque Formation in Alzheimer's Disease." *Neuron* 49 (4): 489–502. https://doi.org/10.1016/j.neuron.2006.01.022.

Snider, Barbara J., Anne M. Fagan, Catherine Roe, Aarti R. Shah, Elizabeth A. Grant, Chengjie Xiong, John C. Morris, and David M. Holtzman. 2009. "Cerebrospinal Fluid Biomarkers and Rate of Cognitive Decline in Very Mild Dementia of the Alzheimer Type." *Archives of Neurology* 66 (5): 638–45. https://doi.org/10.1001/archneurol.2009.55.

Spires-Jones, Tara L., and Bradley T. Hyman. 2014. "The Intersection of Amyloid Beta and Tau at Synapses in Alzheimer's Disease." *Neuron* 82 (4): 756–71. https://doi.org/10.1016/j.neuron.2014.05.004.

Stomrud, Erik, Oskar Hansson, Kaj Blennow, Lennart Minthon, and Elisabet Londos. 2007. "Cerebrospinal Fluid Biomarkers Predict Decline in Subjective Cognitive Function over 3 Years in Healthy Elderly." *Dementia and Geriatric Cognitive Disorders* 24 (2): 118–24. https://doi.org/10.1159/000105017.

Strittmatter, W. J., A. M. Saunders, D. Schmechel, M. Pericak-Vance, J. Enghild, G. S. Salvesen, and A. D. Roses. 1993. "Apolipoprotein E: High-Avidity Binding to Beta-Amyloid and Increased Frequency of Type 4 Allele in Late-Onset Familial Alzheimer Disease." *Proceedings of the National Academy of Sciences of the United States of America* 90 (5): 1977–81. https://doi.org/10.1073/pnas.90.5.1977.

Suárez-Calvet, Marc, Miguel Ángel Araque Caballero, Gernot Kleinberger, Randall J. Bateman, Anne M. Fagan, John C. Morris, Johannes Levin, et al. 2016. "Early Changes in CSF STREM2 in Dominantly Inherited Alzheimer's Disease Occur after Amyloid Deposition and Neuronal Injury." *Science Translational Medicine* 8 (369): 369ra178-369ra178. https://doi.org/10.1126/scitranslmed.aag1767.

Suárez-Calvet, Marc, Gernot Kleinberger, Miguel Ángel Araque Caballero, Matthias Brendel, Axel Rominger, Daniel Alcolea, Juan Fortea, et al. 2016. "STREM2 Cerebrospinal Fluid Levels Are a Potential Biomarker for Microglia Activity in Early-stage Alzheimer's Disease and Associate with Neuronal Injury Markers." *EMBO Molecular Medicine* 8 (5): 466–76. https://doi.org/10.15252/emmm.201506123.

Suárez-Calvet, Marc, Estrella Morenas-Rodríguez, Gernot Kleinberger, Kai Schlepckow, Miguel Ángel Araque Caballero, Nicolai Franzmeier, Anja Capell, et al. 2019. "Early Increase of CSF STREM2 in Alzheimer's Disease Is Associated with Tau Related-Neurodegeneration but Not with Amyloid-β Pathology." *Molecular Neurodegeneration* 14 (January): 1. https://doi.org/10.1186/s13024-018-0301-5.

Suppiah, Subapriya, Mellanie-Anne Didier, and Sobhan Vinjamuri. 2019. “The Who, When, Why, and How of PET Amyloid Imaging in Management of Alzheimer’s Disease—Review of Literature and Interesting Images.” *Diagnostics* 9 (2): 65. https://doi.org/10.3390/diagnostics9020065.

Takahashi, Kazuya, Christian D. P. Rochford, and Harald Neumann. 2005. “Clearance of Apoptotic Neurons without Inflammation by Microglial Triggering Receptor Expressed on Myeloid Cells-2.” *The Journal of Experimental Medicine* 201 (4): 647–57. https://doi.org/10.1084/jem.20041611.

Teipel, Stefan, Alexander Drzezga, Michel J. Grothe, Henryk Barthel, Gaël Chételat, Norbert Schuff, Pawel Skudlarski, et al. 2015. “Multimodal Imaging in Alzheimer’s Disease: Validity and Usefulness for Early Detection.” *The Lancet. Neurology* 14 (10): 1037–53. https://doi.org/10.1016/S1474-4422(15)00093-9.

Teunissen, C. E., A. Petzold, J. L. Bennett, F. S. Berven, L. Brundin, M. Comabella, D. Franciotta, et al. 2009. “A Consensus Protocol for the Standardization of Cerebrospinal Fluid Collection and Biobanking.” *Neurology* 73 (22): 1914–22. https://doi.org/10.1212/WNL.0b013e3181c47cc2.

Teunissen, C. E., Niek A. Verwey, Maartje I. Kester, Kees van Uffelen, and Marinus A. Blankenstein. 2010. “Standardization of Assay Procedures for Analysis of the CSF Biomarkers Amyloid β (1-42), Tau, and Phosphorylated Tau in Alzheimer’s Disease: Report of an International Workshop.” *International Journal of Alzheimer’s Disease* 2010 (September): 635053. https://doi.org/10.4061/2010/635053.

Toledo, Jon B., Maria Bjerke, Xiao Da, Susan M. Landau, Norman L Foster, William Jagust, Clifford Jack Jr, et al. 2015. “Nonlinear Association Between Cerebrospinal Fluid and Florbetapir F-18 β-Amyloid Measures Across the Spectrum of Alzheimer Disease.” *JAMA Neurology* 72 (5): 571–81. https://doi.org/10.1001/jamaneurol.2014.4829.

Van Der Flier, W M. 2005. “Epidemiology and Risk Factors of Dementia.” *Journal of Neurology, Neurosurgery & Psychiatry* 76 (suppl_5): v2–7. https://doi.org/10.1136/jnnp.2005.082867.

Van Petten, Cyma, Elena Plante, Patrick S. R. Davidson, Trudy Y. Kuo, Leslie Bajuscak, and Elizabeth L. Glisky. 2004. “Memory and Executive Function in Older Adults: Relationships with Temporal and Prefrontal Gray Matter Volumes and White Matter Hyperintensities.” *Neuropsychologia* 42 (10): 1313–35. https://doi.org/10.1016/j.neuropsychologia.2004.02.009.

Vemuri, Prashanthi, Heather J. Wiste, Stephen D. Weigand, David S. Knopman, Leslie M. Shaw, John Q. Trojanowski, Paul S. Aisen, et al. 2010. “Effect of Apolipoprotein E on Biomarkers of Amyloid Load and Neuronal Pathology in Alzheimer Disease.” *Annals of Neurology* 67 (3): 308–16. https://doi.org/10.1002/ana.21953.

Venugopalan, Janani, Li Tong, Hamid Reza Hassanzadeh, and May D. Wang. 2021. "Multimodal Deep Learning Models for Early Detection of Alzheimer's Disease Stage." *Scientific Reports* 11 (1): 3254. https://doi.org/10.1038/s41598-020-74399-w.

Verghese, Philip B., Joseph M. Castellano, Kanchan Garai, Yinong Wang, Hong Jiang, Aarti Shah, Guojun Bu, Carl Frieden, and David M. Holtzman. 2013. "ApoE Influences Amyloid-β (Aβ) Clearance despite Minimal ApoE/Aβ Association in Physiological Conditions." *Proceedings of the National Academy of Sciences of the United States of America* 110 (19): E1807-1816. https://doi.org/10.1073/pnas.1220484110.

Vlies, A. E. van der, N. A. Verwey, F. H. Bouwman, M. A. Blankenstein, M. Klein, P. Scheltens, and W. M. van der Flier. 2009. "CSF Biomarkers in Relationship to Cognitive Profiles in Alzheimer Disease." *Neurology* 72 (12): 1056–61. https://doi.org/10.1212/01.wnl.0000345014.48839.71.

Walhovd, K. B., A. M. Fjell, I. Amlien, R. Grambaite, V. Stenset, A. Bjørnerud, I. Reinvang, et al. 2009. "Multimodal Imaging in Mild Cognitive Impairment: Metabolism, Morphometry and Diffusion of the Temporal-Parietal Memory Network." *NeuroImage* 45 (1): 215–23. https://doi.org/10.1016/j.neuroimage.2008.10.053.

Wang, Bo, Aziz M. Mezlini, Feyyaz Demir, Marc Fiume, Zhuowen Tu, Michael Brudno, Benjamin Haibe-Kains, and Anna Goldenberg. 2014. "Similarity Network Fusion for Aggregating Data Types on a Genomic Scale." *Nature Methods* 11 (3): 333–37. https://doi.org/10.1038/nmeth.2810.

Wang, Yaming, Tyler K. Ulland, Jason D. Ulrich, Wilbur Song, John A. Tzaferis, Justin T. Hole, Peng Yuan, et al. 2016. "TREM2-Mediated Early Microglial Response Limits Diffusion and Toxicity of Amyloid Plaques." *The Journal of Experimental Medicine* 213 (5): 667–75. https://doi.org/10.1084/jem.20151948.

Wang, Yang, Shannon L. Risacher, John D. West, Brenna C. McDonald, Tamiko R. MaGee, Martin R. Farlow, Sujuan Gao, Darren P. O'Neill, and Andrew J. Saykin. 2013. "Altered Default Mode Network Connectivity in Older Adults with Cognitive Complaints and Amnestic Mild Cognitive Impairment." *Journal of Alzheimer's Disease* 35 (4): 751–60. https://doi.org/10.3233/JAD-130080.

Williams, Jonathan H., Gordon K. Wilcock, Jeffrey Seeburger, Aimee Dallob, Omar Laterza, William Potter, and A. David Smith. 2011. "Non-Linear Relationships of Cerebrospinal Fluid Biomarker Levels with Cognitive Function: An Observational Study." *Alzheimer's Research & Therapy* 3 (1): 5. https://doi.org/10.1186/alzrt64.

Wimo, Anders, Linus Jönsson, John Bond, Martin Prince, Bengt Winblad, and Alzheimer Disease International. 2013. "The Worldwide Economic Impact of Dementia 2010." *Alzheimer's & Dementia* 9 (1): 1. https://doi.org/10.1016/j.jalz.2012.11.006.

Wolfsgruber, Steffen, José Luis Molinuevo, Michael Wagner, Charlotte E. Teunissen, Lorena Rami, Nina Coll-Padrós, Femke H. Bouwman, et al. 2019. "Prevalence of Abnormal Alzheimer's Disease Biomarkers in Patients with Subjective Cognitive

Decline: Cross-Sectional Comparison of Three European Memory Clinic Samples." *Alzheimer's Research & Therapy* 11 (1): 8. https://doi.org/10.1186/s13195-018-0463-y.

Wunderlich, Patrick, Konstantin Glebov, Nadja Kemmerling, Nguyen T. Tien, Harald Neumann, and Jochen Walter. 2013. "Sequential Proteolytic Processing of the Triggering Receptor Expressed on Myeloid Cells-2 (TREM2) Protein by Ectodomain Shedding and γ-Secretase-Dependent Intramembranous Cleavage*." *Journal of Biological Chemistry* 288 (46): 33027–36. https://doi.org/10.1074/jbc.M113.517540.

Yeh, Felix L., Yuanyuan Wang, Irene Tom, Lino C. Gonzalez, and Morgan Sheng. 2016. "TREM2 Binds to Apolipoproteins, Including APOE and CLU/APOJ, and Thereby Facilitates Uptake of Amyloid-Beta by Microglia." *Neuron* 91 (2): 328–40. https://doi.org/10.1016/j.neuron.2016.06.015.

Yonelinas, Andrew P., Keith Widaman, Dan Mungas, Bruce Reed, Michael W. Weiner, and Helena C. Chui. 2007. "Memory in the Aging Brain: Doubly Dissociating the Contribution of the Hippocampus and Entorhinal Cortex." *Hippocampus* 17 (11): 1134–40. https://doi.org/10.1002/hipo.20341.

Yuyama, Kohei, Hui Sun, Susumu Mitsutake, and Yasuyuki Igarashi. 2012. "Sphingolipid-Modulated Exosome Secretion Promotes Clearance of Amyloid-β by Microglia." *The Journal of Biological Chemistry* 287 (14): 10977–89. https://doi.org/10.1074/jbc.M111.324616.

Zheng, Honghua, Lin Jia, Chia-Chen Liu, Zhouyi Rong, Li Zhong, Longyu Yang, Xiao-Fen Chen, et al. 2017. "TREM2 Promotes Microglial Survival by Activating Wnt/β-Catenin Pathway." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 37 (7): 1772–84. https://doi.org/10.1523/JNEUROSCI.2459-16.2017.

Zhu, Xi-Chen, Lan Tan, Hui-Fu Wang, Teng Jiang, Lei Cao, Chong Wang, Jun Wang, Chen-Chen Tan, Xiang-Fei Meng, and Jin-Tai Yu. 2015. "Rate of Early Onset Alzheimer's Disease: A Systematic Review and Meta-Analysis." *Annals of Translational Medicine* 3 (3): 38. https://doi.org/10.3978/j.issn.2305-5839.2015.01.19.

Zlokovic, B. V., S. Yamada, D. Holtzman, J. Ghiso, and B. Frangione. 2000. "Clearance of Amyloid Beta-Peptide from Brain: Transport or Metabolism?" *Nature Medicine* 6 (7): 718–19. https://doi.org/10.1038/77397.

Zotova, Elina, James AR Nicoll, Raj Kalaria, Clive Holmes, and Delphine Boche. 2010. "Inflammation in Alzheimer's Disease: Relevance to Pathogenesis and Therapy." *Alzheimers Res Ther* 2 (1): 1. https://doi.org/10.1186/alzrt24.

Zou, Kun, Mohammad Abdullah, and Makoto Michikawa. 2020. "Current Biomarkers for Alzheimer's Disease: From CSF to Blood." *Journal of Personalized Medicine* 10 (3): 85. https://doi.org/10.3390/jpm10030085.

Appendices

Appendix A: The proportions of the best clustering solutions out of the 10,000 parameter sweep for baseline similarity network fusion performed on the cognitively normal cohort

Best Clustering Solution	Proportion (%)
2	100

Appendix B: Benchmarking performance using individual data types before similarity network fusion (SNF) for the cognitively normal cohort

Comparison Metric	Number of clusters	*p* Value
Aβ42	4	0.2
pTau-181	3	0.5
sTREM2	2	0.8
SNF	2	0.026

Appendix C: The proportions of the best clustering solutions out of the 10,000 parameter sweep for baseline similarity network fusion on the mild cognitive impairment cohort.

Best Clustering Solution	Proportion (%)
2	17.57
3	67.87
4	11.40
5	2.25

Appendix D: Benchmarking performance using individual data types before similarity network fusion (SNF) for the mild cognitive impairment cohort

Comparison Metric	Number of clusters	*p* Value
Aβ42	3	0.5
pTau-181	3	0.02
sTREM2	2	0.5
SNF	3	1e-06

Appendix E: The proportions of the best clustering solutions out of the 10,000 parameter sweep for longitudinal similarity network fusion on the cognitively normal cohort.

Best Clustering Solution	Proportion (%)
2	100

Appendix F: The proportions of the best clustering solutions out of the 10,000 parameter sweep for longitudinal similarity network fusion on the mild cognitive impairment cohort.

Best Clustering Solution	Proportion (%)
2	68.35
3	24.46
4	4.50
5	2.69

Curriculum Vitae

Name:	Bahaaldin Helal
Post-secondary Education and Degrees:	Western University London, Ontario, Canada 2021-Present MESc, Biomedical Engineering
	Western University London, Ontario, Canada 2017-2021 BMSc, Medical Sciences
Honours and Awards:	Parkinson Society Graduate Award 2022
	Ontario Graduate Scholarships 2022
	Canada Graduate Scholarships – Master's 2022
	BrainsCAN Graduate Studentship Program 2021
	Dean's Undergraduate Research Opportunity Program 2021
	Undergraduate Summer Research Internship 2020