UNIVERSITY OF TURKU

Using 18F-fluorodeoxy-glucose positron emission tomography (FDG-PET) and Machine learning to Predict dementia in short term and long term in an ageing population

MDP Human Neuroscience
Master's thesis
Faculty of Medicine
University of Turku

Author:
Ayisha Mushtaq

30.07.2025
Turku

Master's thesis

Subject: Neuroscience
Author: Ayisha Mushtaq
Supervisor: Marco Bucci
Number of pages: 54
Date: 30.07.2025

Abstract

Dementia is a major health concern in aging populations. It requires early detection along with accurate prediction so that necessary steps for intervention can be taken. This is important for improving the patient as well as their loved one's quality of life. Brain imaging method such as 18F-fluorodeoxyglucose positron emission tomography (FDG-PET) can show that how the brains metabolism changes in dementia patients. At the same time machine learning methods gives us powerful tools to predict that who might develop the disease. In this study the potential of FDG-PET and machine learning to predict neurological disorders (including dementia) and to understand brain metabolism in aging is explored.

This study used brain imaging (18F-FDG PET scans) and machine learning to predict neurological diseases (including dementia) in the short term (1–5 years) and long term (>5 years) in an aging population. It also examined how age affects brain metabolism and whether brain metabolism patterns can point out differences between healthy people (Control), those with memory disorders (e.g., Alzheimer's disease), and others with different neurological conditions.

Using data from 1,447 participants in the AIVO database, it was found that older age is linked to lower brain metabolism across all 25 brain regions with the strongest effects in areas important for thinking and memory. Clear differences in brain metabolism between the Control and Memory Disorder groups could not be found possibly due to the small sample size (147 participants with complete data) and differences in age and gender. The machine learning models were built with all 25 brain regions. It showed that the long-term model worked well (AUC = 0.8718) and accurately predicting onset of neurological diseases after 5 years, with the bilateral inferior temporal gyri coming out as a key feature. The short-term model performed less well (AUC = 0.6571) to detect early onset of disease. This study highlights the role of brain metabolism in aging, memory disorders and neurological diseases (including dementia) and the value of using comprehensive brain data for prediction. This preliminary analysis of the dataset needs to be extended, the pre-existing image database needs to be further consolidated to obtain more valid cases for the analysis. Further research in the topic is required. Use of external validation cohorts might be needed.

Key words: Dementia, Alzheimer's disease, brain imaging, machine learning, neurodegeneration, [18F] FDG PET, dementia prediction

Table of contents

1 INTRODUCTION

Dementia is a multifactorial and common neurodegenerative disease with a growing global health load. Dementia is a general term for diseases associated with memory and Alzheimer's disease (AD) is the most prevalent form of dementia for which progressive cognitive impairment eventually causes moderate to complete disability and early death. The global prevalence of dementia has been estimated to exceed 50 million people (Livingston et al., 2020). The pharmaceutical and supportive aspects of treatments have advanced considerably; early diagnosis and prognosis remain to be a principal question, as current treatments only have a symptomatic approach and cannot stop the disease progress. New imaging tools and well-founded predictive models may offer a potential way to effectively deal with this disorder.

18F-FDG positron emission tomography (PET) is globally recognized as a noninvasive tool to assess cerebral glucose metabolism, which reflects the functional status of neurons and neuronal survival. Regional hypometabolism has also been demonstrated with FDG PET imaging, which has shown to have a consistent utility for differentiating normal aging and disease conditions including early AD asymptomatic stages (Mosconi et al., 2005). Reduced metabolism of the temporal-parietal and posterior cingulate cortices is a characteristic of AD and an indicator of progression from MCI to dementia (Knopman et al., 2014). Several epidemiological studies have linked lower cerebral metabolic capacity with future risk of premature death, and so FDG-PET may alternatively be an integrative biomarker of the combined impact of both neurodegenerative and systemic rates of change in brain metabolism.

More recently, the application of machine learning techniques has fundamentally changed the way data is analyzed in the biomedical field by allowing patterns and predictive characteristic features to be derived from complex multi-dimensional data sets. Applied in the field of neuroimaging, these computational tools permit an insightful analysis of the spatial homogeneity and temporal patterns of brain metabolism that often are masked in standard statistical analysis. More recently, machine learning algorithms (e.g., support vector machine, random forest, deep neural network) have been used to label the disease states and to predict the disease progression from the FDG-PET image data (Dukart et al., 2013; Shen Lu et al., 2017). Not only have these methods increased diagnostic accuracy; they have also facilitated comparisons of risk for cognitive decline over time.

Early identification of people at risk for developing the disease would make it possible to offer interventions in an early stage of the disease in the hope that its natural course would change. Reductions in brain metabolism measured by FDG-PET can be observed in the brain several years prior to the clinical appearance of dementia, suggesting a large pre-clinical window for interventions that would be most effective (Mosconi et al., 2008). Additionally, studies which demonstrate a relation between hypometabolism, and improved prediction of progression additionally support the appropriateness of considering PET as a diagnostic probe, but also as a prognostic marker with broader clinical applicability (Mosconi et al., 2008).

The merging of neuroimaging with machine learning for the diagnosis, management, and research of neurodegenerative diseases seems to be a very helpful strategy. Conventional MRI analysis focused primarily on region-of-interest measurements and involved linear statistics, whereas recent machine learning techniques enabled high-dimensional data processing and captured non-linear features representing the patterns of disease progression and recovery. These developments demonstrate a growing contribution of interdisciplinary approaches by combining advanced imaging, computational modeling, and clinical data to address the complex challenges posed by dementia and neurodegenerative diseases.

By integrating the imaging biomarkers into machine learning models, work is currently in progress to develop predictive models for the prediction of cognitive decline and also survival for longer life. Such dual prediction is expected to improve patient stratification, because it will help health care workers to discriminate between persons at a risk of developing a certain disease and those at the risk of bad health.

The study aims to utilize algorithms based on machine learning to predict onset of neurological diseases (including dementia) using FDG-PET data. Specifically dedicated to training and cross-validation of predictive models from longitudinal data to elucidate the metabolic (brain glucose metabolism) precursors to memory impairment and neurological diseases.. Computational methods of this sort with good imaging biomarkers may better predict accurate prognosis for each patient, be suggestive of personalized treatment options and cost-effective in clinic. If these prediction models are validated and successful, targeted interventions to reduce the societal and economic impact of dementia and its comorbidities may follow.

In summary, dementia represents significant health problems in the elderly and novel approaches of the early diagnosis and risk profiling are greatly needed. FDG-PET is a valuable

imaging technique to capture subtle metabolic brain shifts and is ideally suited for preclinical neurodegeneration. Alongside the potential to revolutionize predictive modelling with machine learning, FDG-PET presents an integrated biomarker that could double as an attractive candidate to predict long-term clinical outcomes. The present study seeks to connect MRI biomarkers in neuroimaging with predictive analytics for better prognosis and earlier intervention of dementia care.

2 REVIEW OF THE LITERATURE

2.1 Dementia

2.1.1 Epidemiology and Impact around the World

Globally, around 55 million people were suffering with dementia in 2020, and the figure was estimated to rise to 78 million by 2030 and 139 million by 2050, which creates challenging demands on the healthcare systems and the worldwide caregivers (World Health Organization, 2021). It is the seventh cause of death worldwide with costs upwards of US $1 trillion annual to include direct fee-for-service, medical costs as well as informal care. The cost will be double by 2030 (Alzheimer's Disease International, 2019). Low- and middle-income countries represent almost two-thirds of dementia cases, and are characterised by substantial disparities in diagnosis capacity, prevention initiatives, and access to treatment resources (Prince et al., 2015). The burden of dementia is also higher among women who make up the majority of both dementia patients and unpaid care givers (GBD 2019 Dementia Forecasting Collaborators, 2022). Latest analyses highlight the importance of modifiable risk factors, including hypertension, diabetes, low education and social isolation, and indicate that interventions that address these factors could possibly delay or prevent 40% of dementia cases worldwide (Livingston et al., 2020). The social and economic burden is further increased due to stigma, the under-diagnosis and a chronic deficit of specialized healthcare workers, especially in resource poor settings. These patterns reinforce the urgent requirement of scalable preventative and equitable health policies, and investment into dementia research and care infrastructure within diverse global settings.

2.1.2 Pathogenesis of Dementia

Proteinopathies: Aβ, Tau and α-Synuclein

Alzheimer's disease (AD), the most common type of dementia (60-70% according to WHO), is defined by extracellular β-amyloid (Aβ) plaques and tau, hyperphosphorylated, inside the cell neurofibrillary tangles with well-documented roles in synaptic impairment as well as correlations with neuronal loss and cognitive loss (Jack et al., 2018). The cascade theory of amyloid (Hardy et al., 1992) suggests Aβ aggregation to be an initial event that initiates downstream tauopathy, neuroinflammation, and neurodegeneration (Sperling et al., 2011), which, however, more closely associates with clinical symptoms. In Lewy body diseases

(LBDs), α-synuclein aggregates to intraneuronal inclusions (Lewy bodies), with variable superimposed AD pathology leading to fluctuating cognition.

Neuroinflammation and D innate immunity

Neuroinflammation is a characteristic of dementia pathophysiology, mostly mediated by activated microglia and astrocytes. These glia cells secrete pro-inflammatory cytokines such as IL-1β and TNF-α in addition to reactive oxygen species in response to misfolded proteins to further exacerbate synaptic dysfunction and neuronal damage signalling (Heneka et al., 2015). Genetic data also promotes the role of innate immunity, in particular the TREM2 R47H allele that disrupts microglial phagocytic ability and increases the risk for AD (Guerreiro et al., 2013). Dysregulation in microglial activity could in fact augment both impairment in clearance and neurotoxicity in a bidirectional manner, hence connecting immune dysfunction to protein accumulation and neurodegeneration.

Vascular Factors and Blood–Brain Barrier

Cerebrovascular disease plays a major role in the etiology of dementia. Ischemic injury, small-vessel disease, and cerebral amyloid angiopathy account for about 15% (VaD) and frequently coexist with AD pathology in "mixed dementia" (Iadecola, 2013). Vascular dysfunction disrupts the integrity of the blood–brain barrier, hinder clearing of neurotoxic proteins and enhance chronic neuroinflammation (Sweeney et al., 2019). Such dysregulations not only promote amyloid and tau pathologies but also, independently, cause neuronal dysfunction, emphasizing that it is important to incorporate vascular health in the preventive strategies of dementia. The clinical symptoms of dementia are thought to be the result of a progressive breakdown of connections between neurons and the loss of neurons themselves, in regions important for memory and thinking including the hippocampus and entorhinal cortex (Braak et al., 2015). Therefore, the treatment focussing on synaptic preservation and tau aggregation seems be the potential approach to modify the disease course.

2.1.3 Modifiable and Unmodifiable Dementia risk factors

Treatable Dementia risk factors account for a large proportion of dementia cases worldwide, with the 2020 Lancet Commission identifying nine key risks as low education, midlife hypertension, hearing impairment, smoking, obesity, depression, physical inactivity, diabetes, and low social contact and three additional risks as excessive alcohol consumption, traumatic

brain injury, and air pollution. Collectively explaining about 40% of the dementia burden; the 2024 update adds two more risks (high LDL cholesterol and untreated vision loss), leading to an estimated 45% of dementia cases being preventable or delayable (Livingston et al., 2020; The Lancet Commission, 2024). Dementia incidence increases exponentially with age, doubling approximately every 5 years after age 65 (Corrada et al., 2010). Early-onset familial Alzheimer's disease (<65 years) is driven by autosomal-dominant mutations in APP, PSEN1, and PSEN2 (Lanoiselée et al., 2017), while late-onset Alzheimer's disease is influenced by polygenic risk, notably APOE ε4 allele, which confers a 2–3-fold increased risk for heterozygotes and up to 10–15-fold risk for homozygotes (Kunkle et al., 2019). In contrast, factors related to protection (cognitive reserve with education and complex occupation, Mediterranean diet adherence, regular physical exercise, and active social lifestyle) can push back the clinical onset and modulate the risk as well (Stern et al., 2009; Scarmeas et al., 2006; Blondell et al., 2014; Kuiper et al., 2015).

2.1.4 Protective Factors and Resilience

Cognitive reserve, which is acquired through greater educational and occupational complexity, as well as more frequent mentally stimulating leisure activities, enables individuals to tolerate more neuropathology before any clinical manifestation; greater reserve associates with more efficient use of neural networks and compensatory recruitment (Stern et al., 2009). Consumption of a mediterranean diet that includes a variety of fruits, vegetables, whole grains, and healthy fats also has been linked to lower risk of Alzheimer's disease (Scarmeas et al., 2006). Physical activity attenuates dementia risk by 14% (Blondell et al., 2014), likely due to vascular, metabolic, and neurogenic mechanisms. Sound social ties and social participation are associated with reduced risk of incident dementia to a similar extent as other established risks (Kuiper et al., 2015). These protective factors may have an additive synergistic effect on delaying the onset of disease and its progression.

Contemporary clinical practice characterises Dementia according to syndromic or etiological classification in most current nomenclature. DSM-5 specifies the diagnosis of Major Neurocognitive Disorder based on the loss of at least one cognitive domain (i.e., complex attention, executive function, learning and memory, language, perceptual-motor, and social cognition) that interferes with activities of daily living (Sachdev et al., 2014). In ICD-11, neurocognitive disorders are reformulated into: block 6D8 (dementia) for which we can add post-coordination of etiology: 6D80 Alzheimer's dementia; 6D81 cerebrovascular dementia;

6D82 Lewy body dementia; 6D83 frontotemporal dementia; etc in order to get more detailed documentation and to reflect the multifaceted nature of dementia better (Word Health Organization, 2023). In addition to these syndromic criteria, the 2018 NIA-AA Research Framework (Jack et al., 2018) reconceptualizes AD biologically by the use of biomarkers of β-amyloid (A), pathological tau (T), and neurodegeneration (N) [the AT(N) scheme] for staging AD from preclinical to dementia stages irrespective of symptoms.

2.2 Biomarkers of machine learning

In the age of evidence-based decision-making, the timely and correct diagnosis of dementia has increasingly become dependent on biomarkers that are indicative of the underlying pathology. Fluid and imaging biomarkers permit the disease to be detected years before onset of clinical symptomatology and also allow for the disease to be monitored.

2.2.1 Fluid Biomarkers

Measurements of cerebrospinal fluid (CSF) reveal that low CSF Aβ42 combined with high levels of total tau (t-tau), or phosphorylated tau (p-tau) provide more than 85% sensitivity and specificity for diagnosing Alzheimer's disease (AD) compared to controls with normal cognition (Olsson et al., 2016; Hansson et al., 2018). CSF neurofilament light chain (NfL) levels represent axonal degeneration and are increased in neurodegenerative disorders, including FTLD (Benussi et al., 2020). YKL-40 as a marker of astroglial activation is associated with neuroinflammation and may differentiate between AD and other dementias (Craig-Schapiro et al., 2010). In more recent era, plasma biomarkers have transformed early diagnosis. Plasma p-tau217 as well as p-tau181 demonstrates robust correlation with CSF and PET results and also predicts the transition from MCI to dementia and thus would provide minimally invasive tools that are scalable for screening in large populations (Palmqvist et al., 2020).

2.2.2 Neuroimaging Biomarkers

Hippocampal volumetric losses measured through MRI and patterns of cortical thinning also predict MCI-to-AD conversion with approximately 70–80% accuracy (Dickerson et al., 2009). Fluorodeoxyglucose positron emission tomography (FDG-PET) shows typical hypometabolism in the temporoparietal areas of AD, and in the occipital lobes of DLB (Minoshima et al., 2001). With the advent of amyloid positron emission tomography (PET),

using ligands like [^{11}C] Pittsburgh Compound-B ([^{11}C] PiB), it has been possible to visualize amyloid plaques in vivo and prove the presence of early pathology already in asymptomatic individuals (Klunk et al., 2004; Villemagne et al., 2013). Recent advances in tau PET tracers, including [^{18}F] flortaucipir, offer more detailed in vivo mapping of neurofibrillary tangles and improve differential diagnosis of taupathies as well as treatment monitoring (Ossenkoppele et al. 2021).

Collectively, fluid and imaging biomarkers are dramatically changing dementia research and clinical care, leading to a paradigm shift toward the use of a precision medicine framework for diagnosing, staging and targeting intervention, for example of AD.

2.3 Classical classifiers for machine learning

Dementia research has been increasingly influenced by machine learning (ML) approaches that provide the ability to reveal insightful patterns buried in high-dimensional data and assist in early diagnosis, subtype discrimination and prognosis of cognitive impairment.

2.3.1 Support Vector Machines

Early explorations of support vector machines along whole-brain anatomical MRI have shown the potential of automated classification of Alzheimer's disease (AD). Magnin et al. (2009) trained an SVM classifier on parcellated grey matter features and obtained a robust mean correct classification of 94.5% (controls versus AD patients) using bootstrap resampling. The small sample size (n=38) demands the need for a large-scale replication. Beyond the global features, Chincarini et al. (2011) used local cortical thickness values as features in a SVM classifier, and they were able to reach high sensitivity for differentiating AD from healthy controls but in prodromal detection of AD the performance was more modest which highlights the challenges in early detection. More recently models that integrate MRI, CSF biomarkers and multivariate pattern classification have been used to predict conversion. Davatzikos et al. (2011) found that pattern-based SVM classifiers combining MRI and CSF data were able to successfully predict MCI converters, reporting very high baseline sensitivity.

2.3.2 Sparse Ensemble and Feature Fusion

Liu, Zhang, and Shen (2012) proposed the local patch-based subspace ensembles that integrated several sparse classifiers on MRI patches to enhance robustness of MCI conversion prediction (Liu et al., 2012). Later studies also built upon this approach by integrating MRI data of both

fMRI and structural MRI to refine MCI detection further. Gray et al. (2013) used random forest-based similarity measures to combine multimodal MRI markers and improve discrimination across the AD spectrum (Gray et al., 2013), this method achieved high accuracy of more than 85% in distinguishing AD from HC and around 75% for MCI vs. HC showing promising results of multi-modality approaches.

2.3.3 Advanced Fusion Strategies and Stability

The RF feature selection and fusion with majority voting were presented by Dimitriadis et al., 2018) for the classification of healthy controls, stable and progressive MCI, as well as AD (accuracy result in the blind validation: 61.9% for the four-class classification (Dimitriadis et al., 2018)). Lebedev et al. (2014) provided evidence on successful between-cohort transfer of RF ensembles, which outperformed the performance of support vector machines in AD detection and MCI-to-AD prediction (Lebedev et al., 2014). Combining morphometric measurements with ApoE genotype and demographics improved the sensitivity and specificity. Random forests have also been used in analysis of spontaneous speech, where Hason et al. (2022) demonstrated 82.2% accuracy in distinguishing AD from controls based on acoustic feature.

2.3.4 Gaussian Process Classifications

Gaussian process (GP) classifiers provide a principled, probabilistic approach for dementia prediction that are particularly attractive for clinical use as they can quantify uncertainty. Young et al. (2013) showed that GP models that incorporate both sMRI, FDG-PET, and CSF biomarkers outperforms support vector machines (SVMs) in the task of predicting MCI to AD conversion, particularly in the context of small data, where uncertainty modeling is crucial (Young et al., 2013). Unlike point-estimate methods, GPs provide posterior distributions over predictions, allowing for more conservative, interpretable diagnostics.

2.3.5 Convolutional Neural Networks

Due to their ability to extract hierarchical features of input data in multi-level scales, convolutional neural networks (CNN) have achieved prominent performances in automatic classification of the neuroimaging data for Alzheimer's disease (AD). Islam (2018) also designed a deep CNN ensemble for sMRI trained then tested AD classification, reporting higher performance when measured against traditional machine learning classifiers for early-stage AD

detection (Islam et al., 2018). The first studies that proposed 3D CNNs for neurodegenerative QS were introduced by Payan (2015), who used autoencoder-pretrained volumetric features to discriminate AD from controls on MRI scans (Payan et al., 2015). These progresses demonstrate the flexibility of CNN-based approaches to diverse imaging modalities and provide new perspective to their utility in improving the accuracy of diagnosis and prognosis in the strategy of the research of dementia.

2.3.6 Fusing Multimodalities with Deep Architectures

Deep learning architectures that integrate data from heterogenous sources (e.g., structural MRI, FDG-PET, CSF-biomarkers and clinical/genetic information) consistently improve upon single-modality models by leveraging complimentary disease signatures and refinement of high-level feature representations. Early deep fusion methods used hierarchical feature extraction to discover shared latent spaces between imaging modalities, and more recent attention- and graph-based networks seek to fine-tune modality weighting and inter-modality engagement. One of the first fully unified deep architectures for AD and MCI classification was presented by (Suk, Lee, & Shen, 2014), which initially applies independent autoencoder branches to learn high-level MRI and PET imaging features in separate, before combining these representations for co-prediction; the improvement of the joint decision over single-modality classifiers was pronounced (Suk et al., 2014).

2.3.7 Least absolute shrinkage and selection operator (LASSO)

Deep learning requires high dimensional and large datasets. An alternative approach that allows use of modest to small datasets is LASSO in association with cross-validation (optimally nested) with leave-one-out method. This has been already applied in dementia research in two studies that investigated the clinical value of plasma biomarkers in distinguishing diagnostic groups in the AD continuum and to predict amyloid PET positivity. (Bucci et al, 2024, 2025)FDG-PET and dementia studies

2.4 Clinical Accuracy in Differential Dementia

Early FDG-PET studies have established disease-specific hypometabolic patterns to enable differentiation of AD, frontal and anterior temporal in FTD and occipital in DLB. This provides a foundation for accurate diagnosis of subtypes of dementia. (Herholz et al., 2002) showed that automatic voxel-based analysis of FDG-PET scans allowed for differentiation of Alzheimer's

disease (AD) patients from control subjects with high precision. They reported a sensitivity and specificity of 93% for mild-to-moderate AD cases, and 84% sensitivity for very mild AD. Their multicenter investigation revealed typical temporoparietal and posterior cingulate hypometabolism in AD patients, which was detectable even at an early stage of the disease which highlights FDG PET`s importance for early diagnosis. Normalized disease-specific FDG-PET patterns have been consistent in providing high diagnostic accuracy for differentiation of dementia subtypes. In a multicenter study, the accuracy achieved was 95% for AD vs HC and 94% for FTD vs HC and 92% for DLB vs HC.

2.5 Quantitative analysis techniques

Quantitative methodology has increased the viability of FDG-PET for clinical application. Automatic comparison with normative databases (e.g., 3D-SSP) provides an objective assessment of small metabolic abnormalities like early hypometabolism (Minoshima et al., 1995). Among MCI cohorts, FDG-PET hypometabolism in the posterior cingulate, precuneus and medial temporal lobes are strong predictors of progression to AD with a positive predictive value higher than 80% (Mosconi et al., 2008). Longitudinal studies have shown that decline in metabolism typically occurs 1–3 years prior to MRI-detected atrophy although the medial temporal lobe atrophy occurs earlier than metabolic decline (de Leon et al., 2001).

FDG-PET is nonspecific in mixed or vascular diseases (Tarkin et al., 2014) due to uptake in both inflammatory and non-inflammatory conditions. Future studies combining FDG-PET with amyloid and tau PET, MRI and fluid biomarkers will yield multimodal signatures for individualized diagnosis with better specificity based on the stage of the disease and, eventually, for personalized prognosis (Sperling et al., 2011; Jack et al., 2010). 18F-FDG is a glucose analogue that is taken into neurons through glucose transporters (GLUT), and is phosphorylated by hexokinase, thereby becoming metabolically trapped and allowing for CMRglc calculation (Shivamurthy et al., 2015).

2.5.1 Quantitative Analysis and Machine Learning

Automated voxel-based z-scoring with respect to large normative samples such as three-dimensional stereotactic surface projections (3D-SSP) enables objective identification of regional deficits, with >90% sensitivity and specificity in AD versus controls (Minoshima et al., 1995). FDG-PET in AD has an established pattern of bilateral temporoparietal and posterior cingulate hypometabolism extending into the precuneus and medial temporal lobes,

with relative preservation of the sensorimotor and occipital cortices (Mosconi, 2008; Shivamurthy et al., 2015). Longitudinal cohorts with postmortem confirmation show that hypometabolism in these regions can predate clinical onset by more than a decade, highlighting its utility as an early imaging biomarker (Mosconi et al., 2009). In large multicentre populations, FDG-PET had an accuracy between 79 and 95% in distinguishing between probable AD and healthy participants, generating pooled sensitivity and specificity of 0.89 against healthy controls (Herholz et al., 2002; Silverman et al., 2001). Prospective FDG-PET studies show faster metabolic decline in relation to cognitive decline and time of conversion to dementia (de Leon et al., 2001; Caminiti et al., 2018).

2.5.2 Longitudinal and Predictive Utility

Longitudinal FDG-PET findings have shown that regional hypometabolism increases as MCI patients transition to AD dementia, and reduced glucose metabolism in the regions such as temporal neocortex is associated with future cognitive decline and brain volume loss (de Leon et al., 2001). Mosconi et al. (2009) have demonstrated that the posterior cingulate metabolic decrease can be detected years prior to the onset of symptoms in people who later develop Alzheimer's disease, emphasizing the sensitivity of this measure in the preclinical stage of AD (Mosconi et al., 2009).

2.5.3 FDG-PET Patterns in MCI

Among the MCI subtypes, amnestic MCI frequently displays posterior cingulate, precuneus, and medial temporal hypometabolism, which are consistent with the topography of early AD (Mosconi, 2005, Drzezga et al., 2003). The non-amnestic MCI subtypes present a more mixed picture of hypometabolism, with some cases showing frontal or parietal deficits in line with what would be expected in an FTD or DLB like progression (Anchisi et al., 2005). Automatic, voxel-wise, z-score comparison to huge normative datasets (e.g., 3D-SSP) provides sensitivity and specificity more than 85% in predicting conversion, which is superior to the visual assessment (Herholz et al., 2002).

2.5.4 Statistical Parametric Mapping

Statistical parametric maps for group comparisons and regression models were generated using a general linear model. SPM performs voxel-wise statistical contrasts with normative databases and produces z-score maps of spatially relative hypometabolism. The utility of the scoring

systems-based SPM z-scores considering diagnostic performance similar or better than original qualitative readings (SPM sensitivity ≈92%, specificity ≈90%) has also been proved for the diagnostic procedure in research samples (Herholz et al. 2002; Minoshima et al., 1995). SPM therefore allows the detection of early and subtle metabolic deficits, which are later associated with changes in glucose metabolism, ultimately turning PET signals into quantifiable values (Shivamurthy et al., 2015; Herholz et al., 2002).

3 AIMS OF THE STUDY

1. To refine and consolidate the metadata database for the 18F-FDG PET scans and to perform group classification, focusing on distinguishing control and memory disorders categories.

2. To establish the relationship between age and glucose uptake between different anatomical and functional brain regions.

3. To test the accuracy of statistical and machine learning methods using brain 18F-FDG PET in predicting occurrence of neurological diseases (including memory disorders) in short-term (<1y and 5y>) and long-term (>5y).

4 MATERIALS AND METHODS

4.1 Data Collection

The data which is used for this study was extracted from the Turku PET Center AIVO database (https://aivo.utu.fi/). It is a collection of a large brain imaging data collected and maintained by the Turku PET Centre (University of Turku and University Central Hospital, Turku, Finland). The aim of the AIVO project is to build a complete and comprehensive collection of brain PET and MRI imaging data by combining neuroimaging data with a pretty wide range of background data like psychological test scores, medical history, and social and demographic information which makes it a powerful tool for studying that how different aspects of a person's life like their social environment, family situation, or economic status can affect the brain and mental health. The AIVO database includes information from over 50,000 people. These people took part in different PET and MRI studies that were done for research into neurological conditions like Alzheimer's disease or memory disorders.

In the start 2,186 total subjects were chosen from the AIVO dataset, as they had to have in archive at least one positron emission tomography (PET) scan using 18F-fluorodeoxyglucose (FDG). Changes in glucose metabolism are associated with diseases that effects memory like Alzheimer's, this imaging method is particularly relevant and useful for dementia research.

To improve the quality of the data and to make sure that all the important variables for the study were present in the dataset, only those subjects who had an FDG-PET scan before the year 2016 and, had most of the values available for the important variables — such as age, weight, height, gender, scan timing, diagnosis, and group classification — were included. After meeting these requirements, the final dataset consisted of 1,447 total participants.

The dataset consists of different important types of information e.g. a unique identifier for the images produced because of the FDG-PET scans (image_id), along with demographic information of the subjects (age, gender, weight, height) and clinical diagnosis of the subjects (e.g., Mild Cognitive Impairment (MCI), Alzheimer's disease and other memory related and neurological conditions). The dataset also included technical details for the scans like the start time, dose of the radiotracer and project name. Most importantly, the dataset consisted of brain glucose uptake (BGU) values, calculated with Gjedde-Patlak, for different brain regions along with the whole brain, white matter, grey matter and also for anatomical and functional brain regions like the frontal lobe, temporal lobe, occipital lobe, and posterior cingulate cortex for a total of 25 region (listed in Table 13).

All the subjects were categorized into three main groups (control, memory disorder, other) by using the information from the dataset about the diagnosis group they were already assigned. The already present diagnoses for the subjects included MCI, AD, Dementia, Epilepsy, Parkinson, healthy controls and other disorders. The division of the groups into the 3 categories mentioned above were as follows. Subjects with no diagnoses (healthy people) were assigned the "control" group, subjects with Alzheimer's, dementia, MCI and memory problems were assigned the group "memory disorder", and in "Other" group were pooled other disorders, which included the rest of the non-healthy patients (including Parkinson, diabetic, epilepsy, anorexia and unknown). This group classification was done for the comparative analysis and helped to make meaningful distinctions between clinical populations. Categorization can be seen in Table 1.

To be able to perform the neurological prediction by using machine learning and statistical methods, two new outcome variables were extracted from available metadata database. The variable is_g_short_1y_5y which was used for short term prediction of dementia occurrence indicated if a participant was diagnosed with dementia within 1 to 5 years after they had their scan, while is_g_long_5yplus which was used for long term prediction of dementia occurrence indicated dementia diagnosis after 5 or more years of the time of the scan. These variables had binary values (1 = Yes and 0 = No). Not all participants had this follow-up information available so only those 41 subjects with known information were included in the machine learning analyses.

4.2 Data Preprocessing

All the preprocessing was done using **R Studio**. Firstly, it was checked that if the data consisted of duplicates of the image_id variable. The data had many duplicate image id`s so the one with the most complete information was retained and the others were removed. This step is important to avoid consider the same subject multiple times in the analysis. After visually analysis of the dataset in the excel worksheet, it was known that the demographic variables like gender, weight and height had a lot of missing values. These columns were merged with similar columns from other datasets, from the AIVO database, by image_id. In the dataset that was used for the study, the missing weight and height values were coded as 0, so these 0`s were re-coded as NA (not applicable values) to avoid misinterpretation of zeros as a valid data entry during the analysis.

The diagnostic group variable was classified into 3 categories as "control", "memory_disorder", and "other", where control included all the healthy controls, the memory disorder category included subjects with diagnosis for memory problems and the other category included

miscellaneous diagnoses. There were a lot of BGU values that were missing for certain brain regions and subjects. A total of 147 subjects had complete glucose uptake values for the selected anatomical and functional brain regions. No imputation or any other step was done for this missingness because these BGU values are required to be original. Outcome variables is_g_short_1y_5y and is_g_long_5yplus had also many missing values and only 41 rows had values for these columns. So only these 41 rows were used for prediction analysis. Gender was converted to binary column (0 for male, 1 for female).

4.3 Statistical Analysis

The demographic characteristics of subjects in each group (control, memory_disorder, and other) was summarized. For continuous demographic variables (age, weight, height) mean and standard deviation were calculated and for the gender column the proportions of males and females in each group were calculated. It was tested if these variables followed a normal distribution or not. For this, the Shapiro-Wilk test was used. Most variables were not normally distributed, so a non-parametric test was applied. To compare the three groups with the demographic variables, the Kruskal-Wallis test was applied. After the Kruskal-Wallis test showed significant differences between the groups across the demographic variables, the Dunn's test with Bonferroni correction was used to identify that which specific group pairs showed significant differences.

For gender, which was a categorical variable initially, the Pearson's chi-squared test was used to test for overall group differences. If the results of the chi-squared test were significant, the Fisher's exact test for pairwise comparisons was performed. This test helped to identify which specific group pairs had significantly different gender distributions.

4.3.1 Age and Glucose Uptake

To examine the relationship between age and the BGU across all brain regions (25) Spearman's rank correlation was calculated. Because many BGU variables (25 regions) were tested so the p-values were adjusted using the False Discovery Rate (FDR) correction method. This helped to reduce the chances of false positives due to testing multiple regions. The top 10 most significant and strongest correlations were selected for reporting and visualization. This helped to highlight the brain regions that show the most age-related changes in glucose uptake.

4.3.2 Relationship Between Glucose Uptake, and Group Differences

The dataset was filtered to include only participants belonging to either of the two categories (control and memory disorder). Out of the total 25 brain regions, the 15 selected brain regions were grouped into two categories: anatomical (8) and functional (7**).** The anatomical regions included the anterior cerebellum, posterior cerebellum, frontal lobe, fronto-temporal space, limbic lobe, occipital lobe, parietal lobe, and temporal lobe. The functional regions included the default mode network, dorsal attention network, frontoparietal network, limbic network, somatomotor network, ventral attention network, and visual network. The Kruskal-Wallis test was used to statistically compare glucose uptake values between the control and memory disorder groups because the data was not normally distributed. A custom R function was made that looped through each brain region in the list of selected brain regions and performed the Kruskal-Wallis test on that region and compared the glucose uptake values between the two groups. For each test the W statistic and the p-value was extracted and then adjusted all p-values using the False Discovery Rate (FDR) method to control for multiple testing.

4.4 Machine Learning Model Development

Nested cross-validation with Leave-One-Out Cross-Validation (LOOCV) was used to examine that how well the model performs at prediction and to avoid overfitting. The R-package "glmnet" was used to apply the LASSO method to do feature selection and estimation. In the outer loop the LOOCV performed for each iteration by using all the data except one observation as the training set, and the one left-out observation was then used for testing. This was repeated for every observation in the dataset which means that every data point got to be the test case at least once. Inside every outer loop fold the (k)-fold cross-validation was used, where k = n-1 (number of training observations) for the LASSO penalty parameter (λ), this was done for the inner loop and the cv.glmnet function was used for this. The best λ was chosen based on the lowest cross-validated error (Normalized Root Mean Square Error, NRMSE) also known as lambda.min.

4.4.1 Model Fitting

A LASSO regression model was trained for each outer loop by using the training data and the optimal lambda λ found in the inner loop. The binomial family was used because the outcome was binary and $\alpha = 1$ was used for the model. The predictions were made by including predicted probabilities (type = "response") and class predictions (type = "class"). The models were trained

using all 25 brain glucose uptake (BGU) regions along with demographic variables age and gender. Although the top 10 regions with the strongest age-BGU correlations were identified but still all 25 regions were included to capture a better representation of brain glucose metabolic activity. LASSO has a built-in feature selection which shrinks less predictive coefficients to zero which then allows the model to identify regions most important for prediction without pre-selection bias. This made sure that regions that are important for dementia were not excluded.

4.4.2 Feature Importance

To find out that which features were important, only the non-zero coefficients from each LASSO model across the folds were used. The absolute values of these coefficients were calculated and then they were averaged across the folds. This gave an idea of how much each feature contributed to the model.

4.4.3 Performance Metrics

The root mean squared error (RMSE) and normalized RMSE (NRMSE) were calculated for each fold by comparing the predicted probabilities to the actual outcomes. The fold with the lowest NRMSE was selected as the best model. The predictions, probabilities, coefficients, and feature importance were saved so that they could be further analyzed.

4.4.4 Model Evaluation

The final LASSO model was trained using the whole preprocessed dataset after doing a nested cross validation. The best λ for both the short-term and long-term models was used separately. The predicted probabilities were generated for the whole dataset. The cutpointr R-package was used to find the best cutpoint to turn the probabilities into binary predictions.

4.4.5 ROC Analysis

The pROC package was used to make Receiver Operating Characteristic curves. Area Under the Curve (AUC) for both the short-term and long-term models was calculated separately and the curves were plotted to show sensitivity versus 1-specificity.

4.5 Tools and Software

All analyses were conducted in RStudio, utilizing the standard R packages and the once mentioned earlier.

5 RESULTS

Table 1 Number of Subjects in Each Diagnostic Group

Group Category	Number of subjects
Control	331
Memory disorder	306
Other	728

Table 1 shows how many subjects fall into each category based on the AIVO database. 331 people were healthy with no diagnosed disorder, 306 people were diagnosed with a memory disorder and the largest group with 728 people includes subjects with disorders that do not fall under memory disorder category or healthy people. This information is also visualized below in Figure 1.

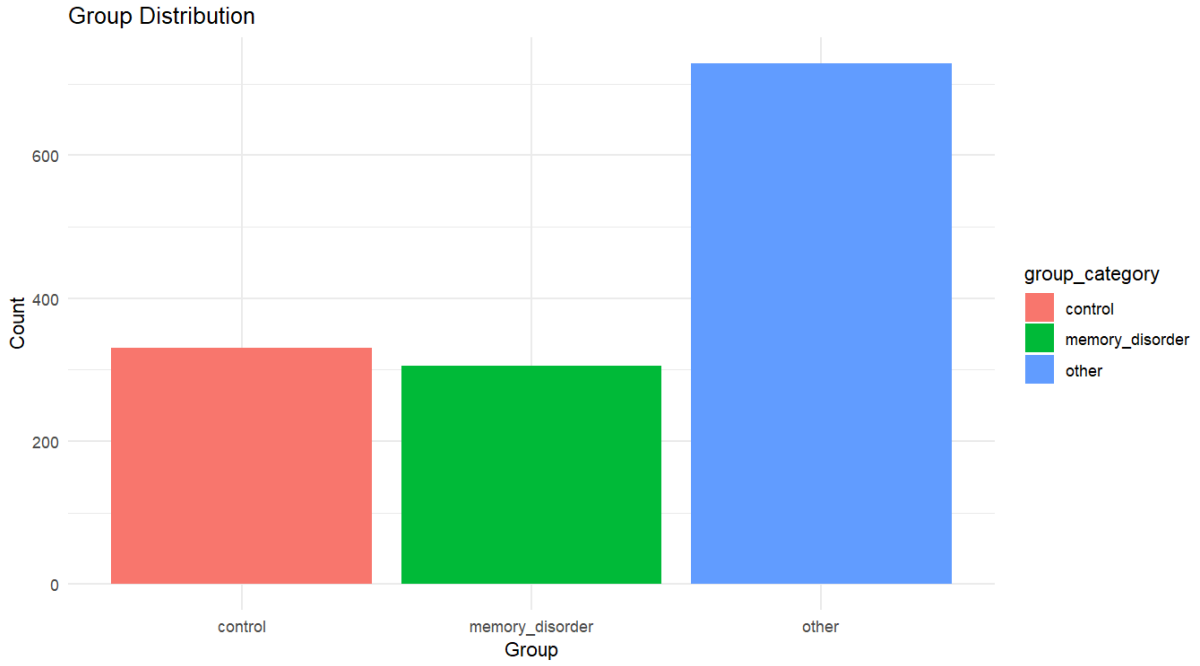

Figure 1 Distribution of Subjects Across Diagnostic Groups

The control and memory disorder groups are having almost same number of participants which is a good thing for statistical comparison of these groups.

Table 2 Demographic Characteristics by Diagnostic Group

Group	Mean Age	SD Age	Mean Weight	SD Weight	Mean Height	SD Height	Female (%)	Male (%)
Control	53.14	19.86	76.82	14.78	171.86	10.09	44.11	55.89
Memory disorder	63.19	14.97	73.70	14.59	166.95	20.06	58.17	41.83
other	55.15	19.34	78.44	17.51	167.42	21.08	46.57	53.43

From Table 2 we the memory disorder group having mean age of 63.19 years is the oldest which makes sense because memory disorders are more common in older people. The memory disorder group has a higher proportion of females (58.17%) compared to the control group (44.11%). This difference might be due to the longer life expectancy of females. The high value of SD for height is possibly due to missing data. Figure 2, Figure 3, Figure 4 and Figure 5 show visualizations for the data included in Table 2.

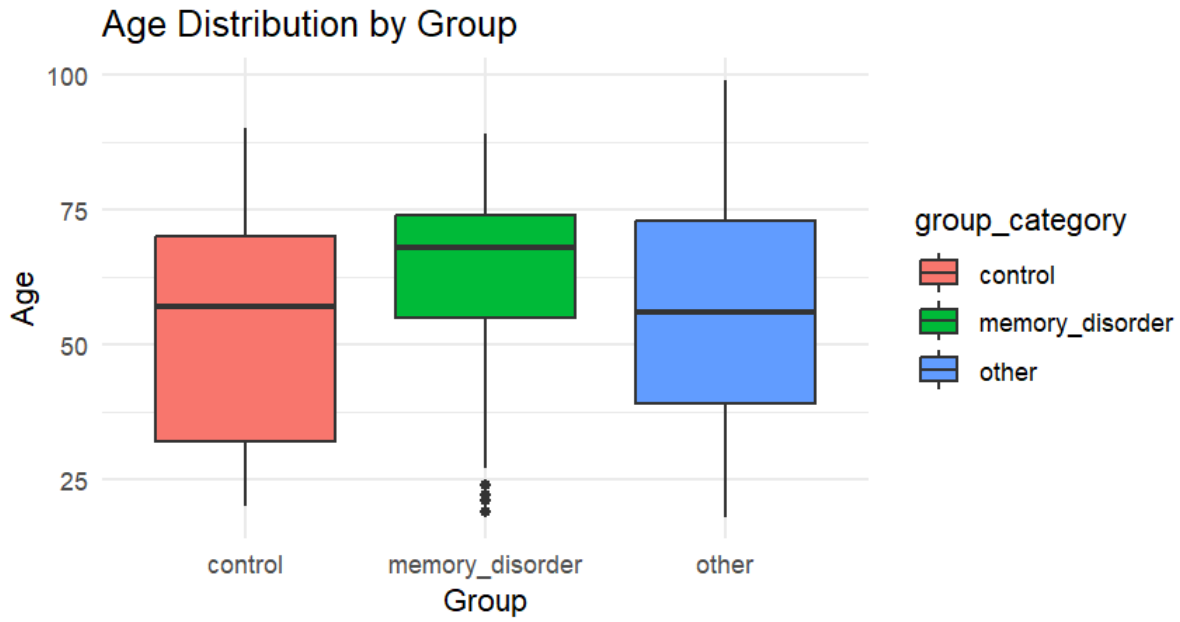

Figure 2 Distribution of Age by diagnostic groups

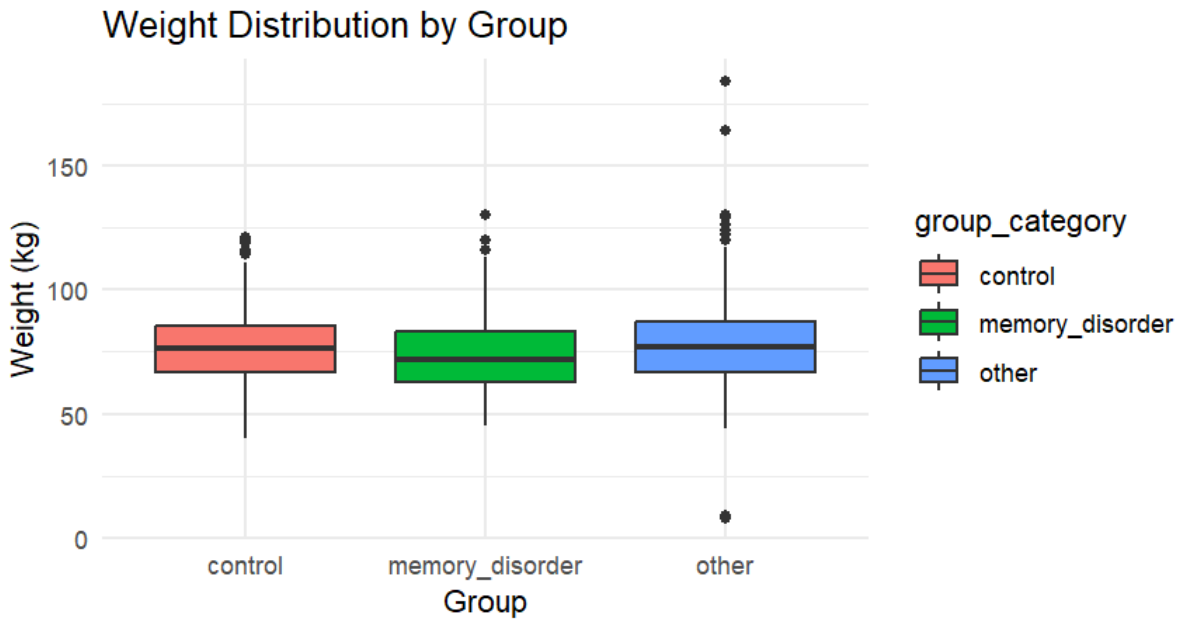

Figure 3 Distribution of Weight by diagnostic groups

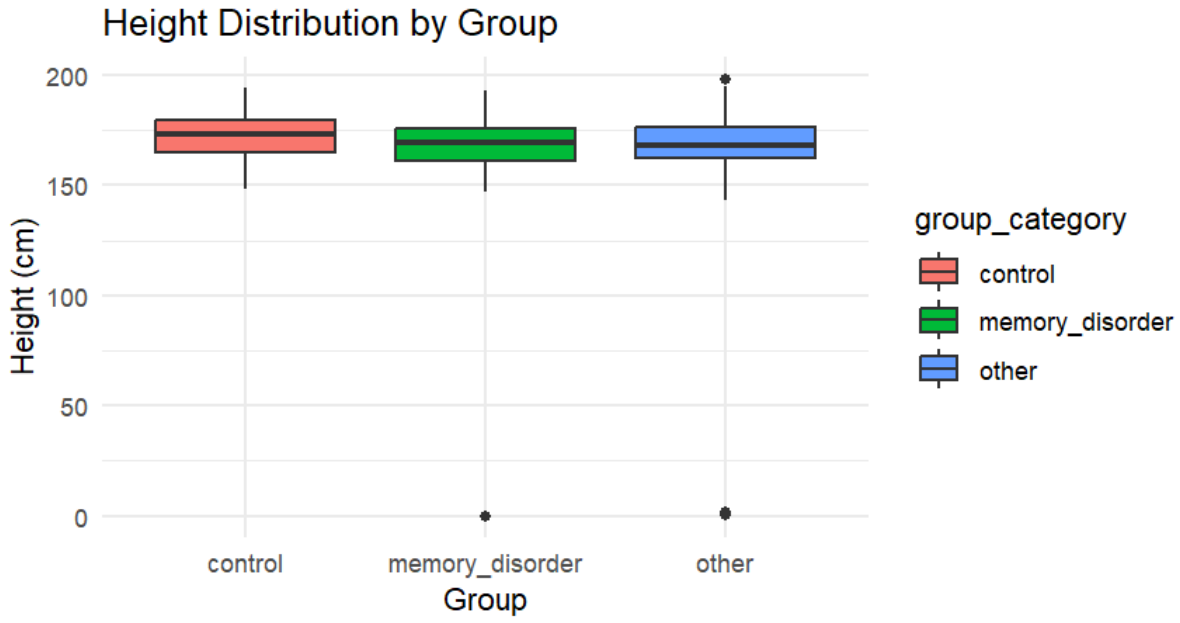

Figure 4 Distribution of Height by diagnostic groups

Table 3 Shapiro-Wilk Normality Tests and Kruskal-Wallis Test for Age Across Diagnostic Groups

Group	Shapiro-Wilk	p-value	Normality
Control	0.917	1.47×10^{-12}	No
Memory Disorder	0.915	3.75×10^{-12}	No
Other	0.939	$< 2.2 \times 10^{-16}$	No

Test	χ^2 (df)	p-value	Significant
Kruskal Wallis	44.80 (2)	1.88×10^{-10}	Yes

The Shapiro Wilk test results show that data for age for all the 3 groups is not normally distributed therefore, Kruskal Wallis test is applied to compare age distribution across the 3 groups. There are significant differences in age between the groups which means that the differences that were observed in Table 2 are not due to random chance.

Table 4 Shapiro-Wilk Normality Tests and Kruskal-Wallis Test for Weight Across Diagnostic Groups

Group	Shapiro-Wilk	p-value	Normality
Control	0.9765	0.00015	No
Memory Disorder	0.9723	0.00011	No
Other	0.9515	2.87×10^{-12}	No

Test	χ^2 (df)	p-value	Significant
Kruskal Wallis	14.55 (2)	0.00069	Yes

The Shapiro Wilk test results show that data for weight for all the 3 groups is not normally distributed therefore, Kruskal Wallis test is applied to compare weight distribution across the 3 groups. There are significant differences in weight between the groups which means that the differences that were observed in Table 2 are not due to random chance.

Table 5 Shapiro-Wilk Normality Tests and Kruskal-Wallis Test for Height Across Diagnostic Groups

Group	Shapiro-Wilk	p-value	Normality
Control	0.9869	0.02512	No
Memory Disorder	0.4846	$< 2.2 \times 10^{-16}$	No
Other	0.4935	$< 2.2 \times 10^{-16}$	No

Test	χ^2 (df)	p-value	Significant
Kruskal Wallis	13.75 (2)	0.00103	Yes

The Shapiro Wilk test results show that data for height for all the 3 groups is not normally distributed therefore, Kruskal Wallis test is applied to compare height distribution across the 3 groups. There are significant differences in height between the groups which means that the differences that were observed in Table 2 are not due to random chance.

Table 6 Gender Distribution Across Diagnostic Groups and Chi-Square Test of Independence

Groups	Female	Male
Control	146	185
Memory Disorder	178	128
Other	339	389

Test	χ^2 (df)	p-value	Significant
Pearson's Chi-squared test	15.10 (2)	0.00053	Yes

From Table 6 it is evident that the memory disorder group stands out with higher number of females compared to control. The significant Chi-Square test results prove that the gender distribution differs significantly across the 3 groups and that this difference is not due to a random chance.

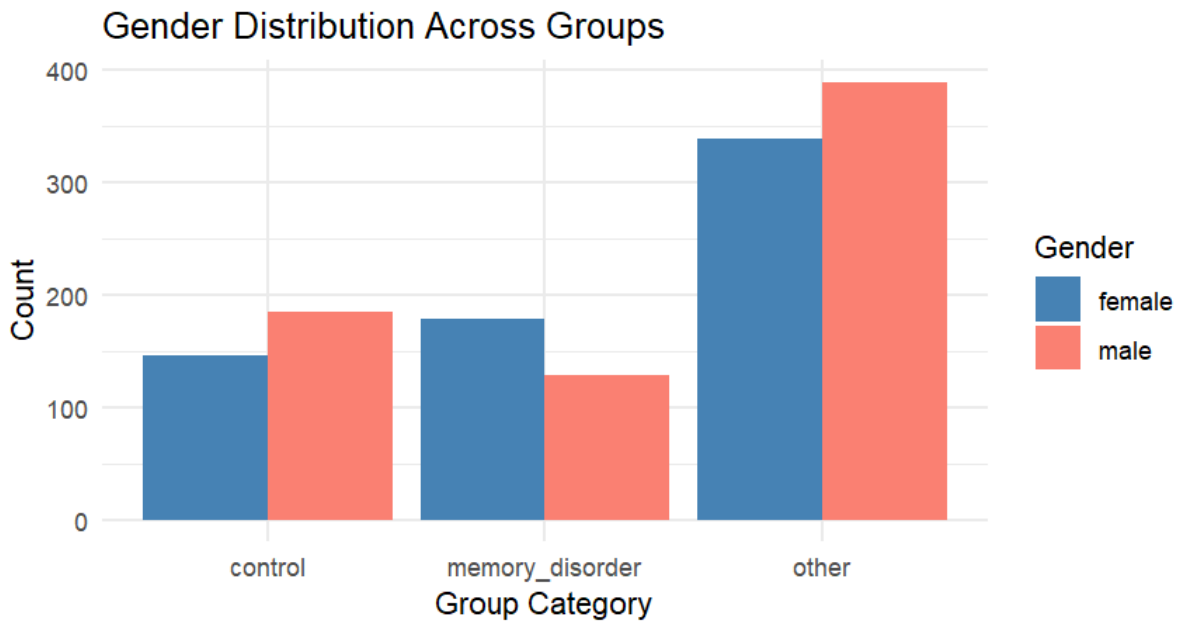

Figure 5 Distribution of Gender across groups

Table 7, Table 8 and Table 9 below, present the results of Dunn's Post Hoc Test with Bonferroni correction which was conducted to identify that which specific pairs of diagnostic groups (Control, Memory Disorder, Other) shows significant differences in age, weight, and height.

Table 7 Dunn's Post Hoc Test for Age Across Diagnostic Groups (Bonferroni Correction)

Comparison	Z-score	Unadjusted p-value	Adjusted p-value	Significant (α = 0.05)
Control vs Memory Disorder	-6.30	2.98×10^{-10}	8.93×10^{-10}	Yes
Control vs Other	-1.77	0.0765	0.2294	No
Memory Disorder vs Other	5.61	2.03×10^{-8}	6.08×10^{-8}	Yes

The Memory Disorder group (mean age 63.19, Table 2) is significantly older than the Control group (mean age 53.14). The large negative Z-score indicates a strong difference, with Memory Disorder participants being older. There is no significant difference in age between the Control group (mean age 53.14) and the Other group (mean age 55.15). The adjusted p-value is above 0.05, suggesting the small age difference is not statistically significant after correction. The Memory Disorder group is significantly older than the Other group. The positive Z-score confirms that the Memory Disorder group's higher mean age (63.19 vs. 55.15) is a significant difference.

Table 8 Dunn's Post Hoc Test for Weight Across Diagnostic Groups (Bonferroni Correction)

Comparison	Z-score	Unadjusted p-value	Adjusted p-value	Significant (α = 0.05)
Control vs Memory Disorder	2.50	0.0126	0.0378	Yes
Control vs Other	-1.02	0.3065	0.9195	No
Memory Disorder vs Other	-3.81	0.00014	0.00042	Yes

The Control group (mean weight 76.82 kg, Table 2) has significantly higher weight than the Memory Disorder group (73.70 kg). A positive Z-score indicates that the Control group has more weight. There is no significant difference in weight between the Control group (76.82 kg) and the Other group (78.44 kg). The adjusted p-value is above 0.05, suggesting that the small

weight difference is not statistically significant. The Other group (78.44 kg) has significantly higher weight than the Memory Disorder group (73.70 kg). The negative Z-score confirms that the Memory Disorder group has lower weight. The Memory Disorder group has significantly lower weight than both the Control and Other groups. This might point towards health-related factors e.g., weight loss in early dementia.

Table 9 Dunn's Post Hoc Test for Height Across Diagnostic Groups (Bonferroni Correction)

Comparison	Z-score	Unadjusted p-value	Adjusted p-value	Significant (α = 0.05)
Control vs Memory Disorder	3.38	0.00072	0.00217	Yes
Control vs Other	3.21	0.00132	0.00395	Yes
Memory Disorder vs Other	-0.66	0.50949	1.00000	No

The subjects in the Control group (mean height 171.86 cm, Table 2) are significantly taller than in the Memory Disorder group (166.95 cm). A positive Z-score indicates that the Control group has greater height. The Control group (171.86 cm) is also significantly taller than the Other group (167.42 cm). The positive Z-score confirms the Control group's greater height. There is no significant difference in height between the Memory Disorder group (166.95 cm) and the Other group (167.42 cm). The adjusted p-value of 1.0 indicates their heights are very similar.

Table 10 Pairwise Fisher's Exact Tests for Gender Distribution Between Diagnostic Groups

Comparison	Odds Ratio	95% CI (LCL-UCL)	p-value	Significant (α = 0.05)
Control vs Memory Disorder	0.57	0.41 – 0.79	0.00048	Yes
Control vs Other	0.91	0.69 – 1.19	0.465	No
Memory Disorder vs Other	1.59	1.21 – 2.11	0.00083	Yes

From Table 6, the Control group has 44.11% females (146/331), while the Memory Disorder group has 58.17% females (178/306). The odds ratio of 0.57 means the odds of being female in the Control group are 57% of the odds in the Memory Disorder group. There is no significant difference in gender distribution between the Control group and the Other group. The odds ratio of 0.91 indicates that the gender proportions are similar between these groups. The Memory Disorder group has 58.17% females (178/306), while the Other group has 46.57% females (339/728). The odds ratio of 1.59 means the odds of being female in the Memory Disorder group are 59% higher than in the Other group.

Table 11 Group Differences by Brain Glucose Uptake in Anatomical Brain Regions

Region	W-statistic	P-value	Adjusted P-value
r01_CerAnt_GU	4.07	0.0436	0.1706
r02_CerPost_GU	3.70	0.0544	0.1706
r08_OccipitalLobe_GU	3.43	0.0640	0.1706
r12_TempLobe_GU	1.79	0.1813	0.3627
r05_LimbicLobe_GU	1.18	0.2767	0.3690
r09_ParietalLobe_GU	1.20	0.2741	0.3690
r04_FrontTempSpace_GU	0.88	0.3475	0.3971
r03_FrontLobe_GU	0.66	0.4150	0.4150

Table 11 shows the results of the Kruskal Wallis test comparing brain glucose uptake (BGU) in eight anatomical brain regions between the Control and Memory Disorder groups. Unexpectedly no significant differences were found in BGU across any of the eight anatomical brain regions after FDR correction. This means that based on the available data (147 subjects with complete BGU values), there are no strong metabolic differences in these regions between the Control and Memory Disorder groups.

Table 12 Group Differences by Brain Glucose Uptake in Functional Brain Regions

Region	W-statistic	P-value	Adjusted P-value
r13_default_GU	0.54	0.4641	0.5415
r14_dorsal_attention_GU	0.85	0.3598	0.5415
r16_limbic_GU	0.86	0.3567	0.5415
r17_somatomotor_GU	1.24	0.2664	0.5415
r18_ventral_attention_GU	0.60	0.4427	0.5415
r19_visual_GU	2.73	0.0988	0.5415
r15_frontoparietal_GU	0.26	0.6123	0.6123

Table 12 shows the results of the Kruskal Wallis test comparing brain glucose uptake (BGU) in seven functional brain regions between the Control and Memory Disorder groups. No significant differences were found in BGU across any of the seven functional brain regions after FDR correction which is unexpected. This means that based on the available data (147 subjects with complete BGU values), there are no strong metabolic differences in these functional regions between the Control and Memory Disorder groups.

Table 13 Spearman Correlation Between Age and Regional Glucose Uptake (_GU)

Region	Spearman's ρ	p-value
r00a_WholeBrain_GU	-0.379	3.24e-06
r00b_GM_GU	-0.396	1.11e-06
r00c_WM_GU	-0.304	2.35e-04
r01_CerAnt_GU	-0.268	1.27e-03
r02_CerPost_GU	-0.239	4.20e-03
r03_FrontLobe_GU	-0.415	2.85e-07
r04_FrontTempSpace_GU	-0.428	1.05e-07
r05_LimbicLobe_GU	-0.416	2.57e-07
r06_Medulla_GU	-0.243	3.62e-03
r07_MidBrain_GU	-0.313	1.51e-04
r08_OccipitalLobe_GU	-0.306	2.08e-04
r09_ParietalLobe_GU	-0.375	4.30e-06
r10_Pons_GU	-0.166	4.79e-02
r11_Sub_lobar_GU	-0.411	3.84e-07
r12_TempLobe_GU	-0.361	9.89e-06
r13_default_GU	-0.439	4.64e-08

r14_dorsal_attention_GU	-0.404	6.11e-07
r15_frontoparietal_GU	-0.451	1.80e-08
r16_limbic_GU	-0.403	6.41e-07
r17_somatomotor_GU	-0.392	1.40e-06
r18_ventral_attention_GU	-0.427	1.15e-07
r19_visual_GU	-0.330	6.04e-05
r20_JNM_Bilateral_Inferior_Temporal_Gyri_GU	-0.309	1.87e-04
r21_JNM_Posterior_Cingulate_and_Precuneus_GU	-0.352	1.72e-05

Table 13 shows the results of Spearman's rank correlation between age and brain glucose uptake (BGU) across 25 brain regions. The Spearman's ρ shows the strength and direction of the relationship, and the p-value shows if the correlation is statistically significant or not. All the correlations are negative which means that older age is associated with lower glucose uptake in the brain. The strongest correlations ($\rho \geq 0.4$) are in functional regions e.g. frontoparietal, default mode, ventral attention. The temporal lobe, posterior cingulate, and inferior temporal gyri which are important areas in Alzheimer's and other dementias show significant but moderate correlations ($\rho = -0.309$ to -0.361).

Table 14 Spearman Correlation Between Age and Glucose Uptake (with FDR Correction) Top 10 regions

Rank	**Region**	**Spearman's ρ (r)**	**Raw p-value**	**FDR-adjusted p-value**	**Significant (α = 0.05)**
1	r15_frontoparietal_GU	-0.451	1.80×10^{-8}	4.50×10^{-7}	Yes
2	r13_default_GU	-0.439	4.64×10^{-8}	5.80×10^{-7}	Yes
3	r04_FrontTempSpace_GU	-0.428	1.05×10^{-7}	7.19×10^{-7}	Yes
4	r18_ventral_attention_GU	-0.427	1.15×10^{-7}	7.19×10^{-7}	Yes
5	r05_LimbicLobe_GU	-0.416	2.57×10^{-7}	1.19×10^{-6}	Yes
6	r03_FrontLobe_GU	-0.415	2.85×10^{-7}	1.19×10^{-6}	Yes
7	r11_Sub_lobar_GU	-0.411	3.84×10^{-7}	1.37×10^{-6}	Yes
8	r14_dorsal_attention_GU	-0.404	6.11×10^{-7}	1.78×10^{-6}	Yes
9	r16_limbic_GU	-0.403	6.41×10^{-7}	1.78×10^{-6}	Yes
10	r00b_GM_GU	-0.396	1.11×10^{-6}	2.76×10^{-6}	Yes

Table 14 lists the top 10 regions out of the 25 analyzed shown in Table 13 focusing on only those with the strongest negative correlations between age and BGU. All top 10 regions show significant negative correlations (ρ ranging from -0.396 to -0.451) after FDR correction which confirms that older age is associated with lower BGU across brain regions.

These results are visualized in Figure 6 while Figure 7 shows the scatter plot for age and brain glucose uptake values in the frontoparietal region.

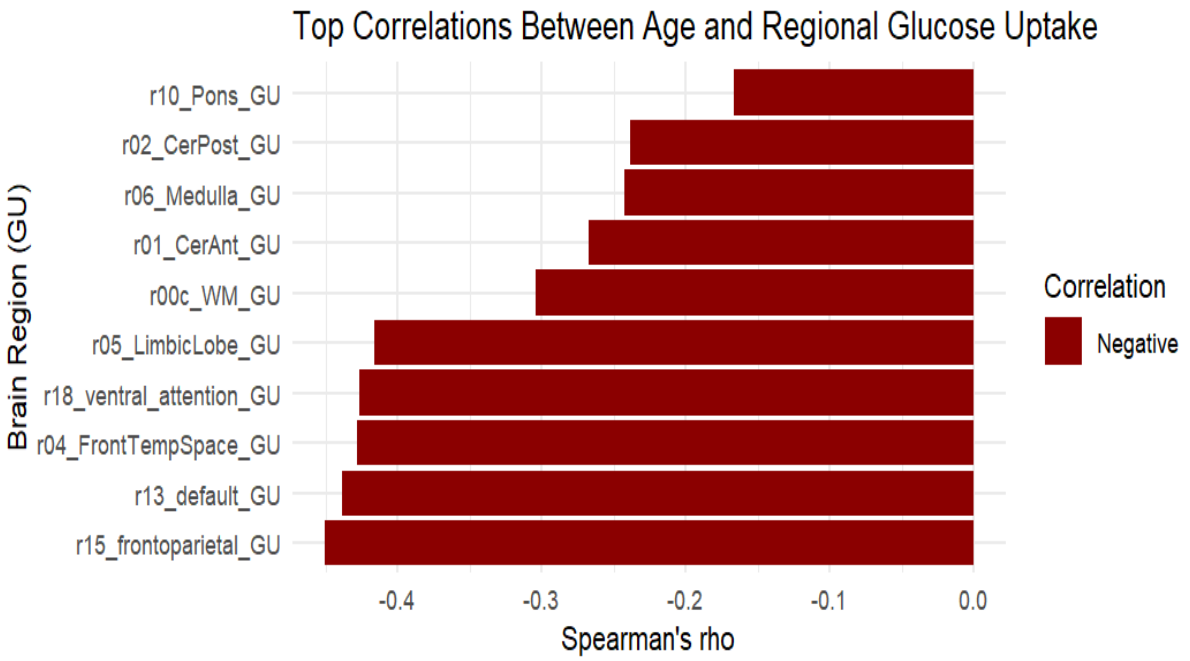

Figure 6 Distribution of correlations between age and brain glucose uptake values

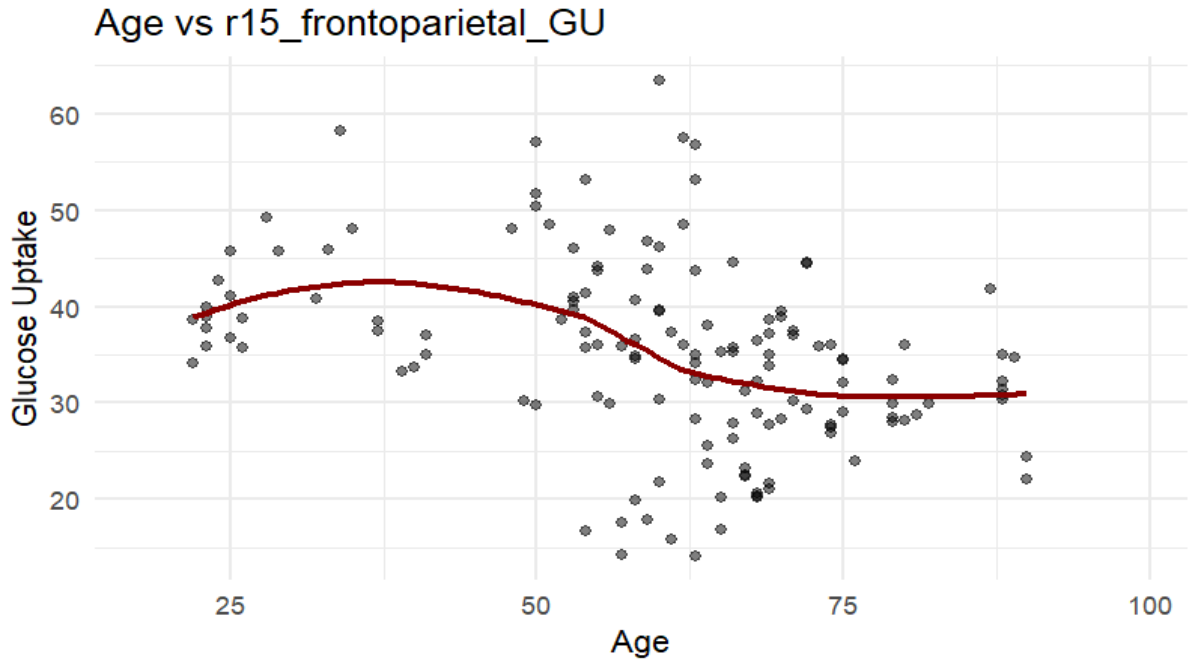

Figure 7 Distribution of age and brain glucose uptake values in the frontoparietal region

Table 15 Model Performance for Predicting Dementia Risk

Model	AUC	Optimal Cutpoint	Sensitivity	Specificity	Accuracy	Balanced Accuracy
Short-term (1–5y)	0.6571	0.275	0.40	0.93	0.79	0.66
Long-term (>5y)	0.8718	0.4236	0.67	1.00	0.89	0.83

Table 15 shows performance of the LASSO regression models for predicting in short-term (1–5 years) and long-term (>5 years). The short-term model (1–5y) has a moderate performance (AUC = 0.6571) visualized in Figure 8, with high specificity (93%) but low sensitivity (40%). This means that it is good at identifying participants who won't develop dementia but is not good to detect those who will develop dementia in the future and miss 60% of true dementia cases. The low sensitivity might be due to the small sample size (41 participants) and limited predictive features. The high specificity means that the model is conservative because it only predicts dementia when it is highly confident, but this happens because of missing many cases. The long-term model (>5y) shows good performance (AUC = 0.8718), with perfect specificity (100%) which is doubtful, and reasonable sensitivity (67%) visualized in Figure 8. It seems much better at differentiating long term dementia risk and correctly identifies all non-dementia cases while detecting two-thirds of dementia cases. The perfect specificity in the long-term model might indicate overfitting, especially given the small sample size (41 participants).

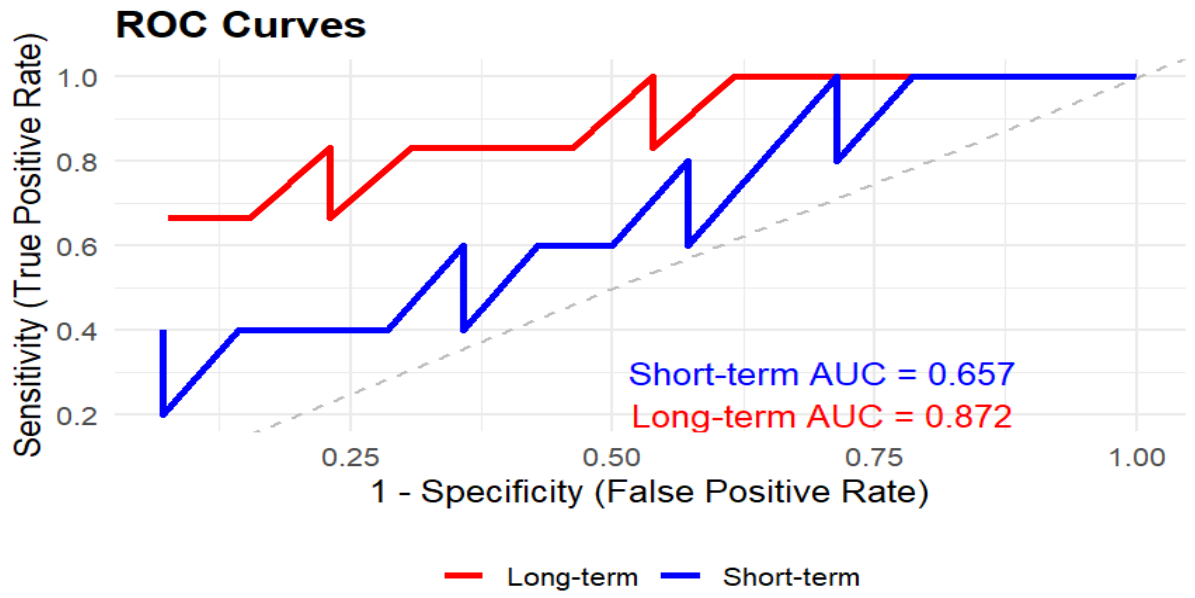

Figure 8 Distribution for receiver operating characteristic curve for both models (short and long)

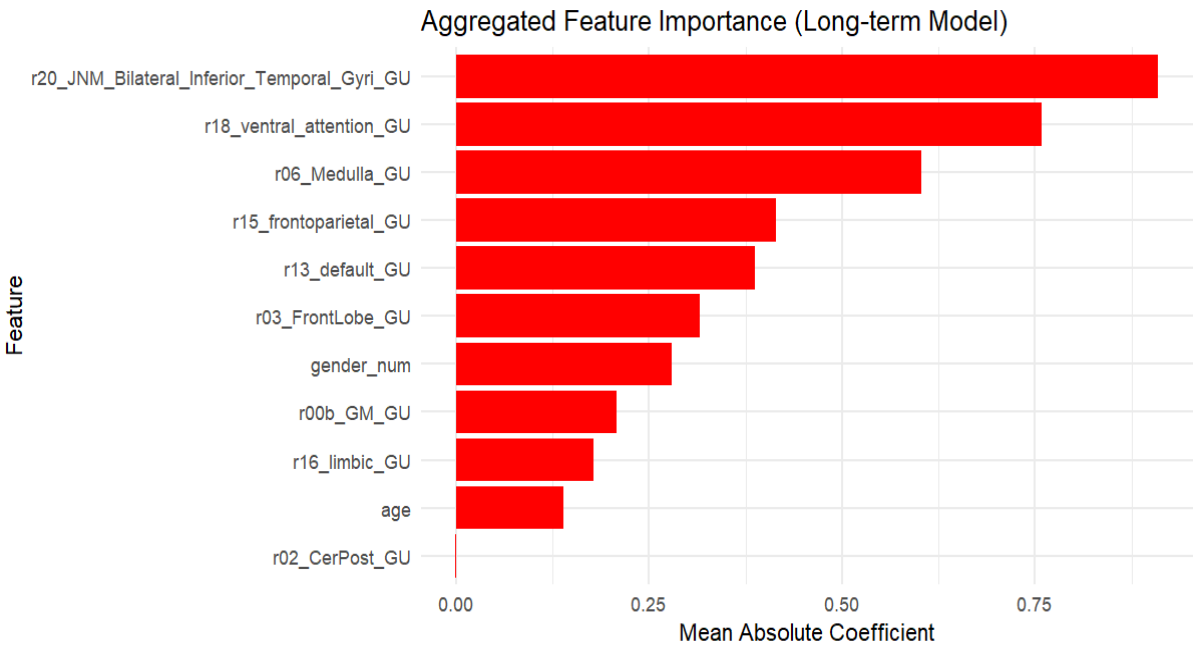

Figure 9 Distribution for feature importance for long term model

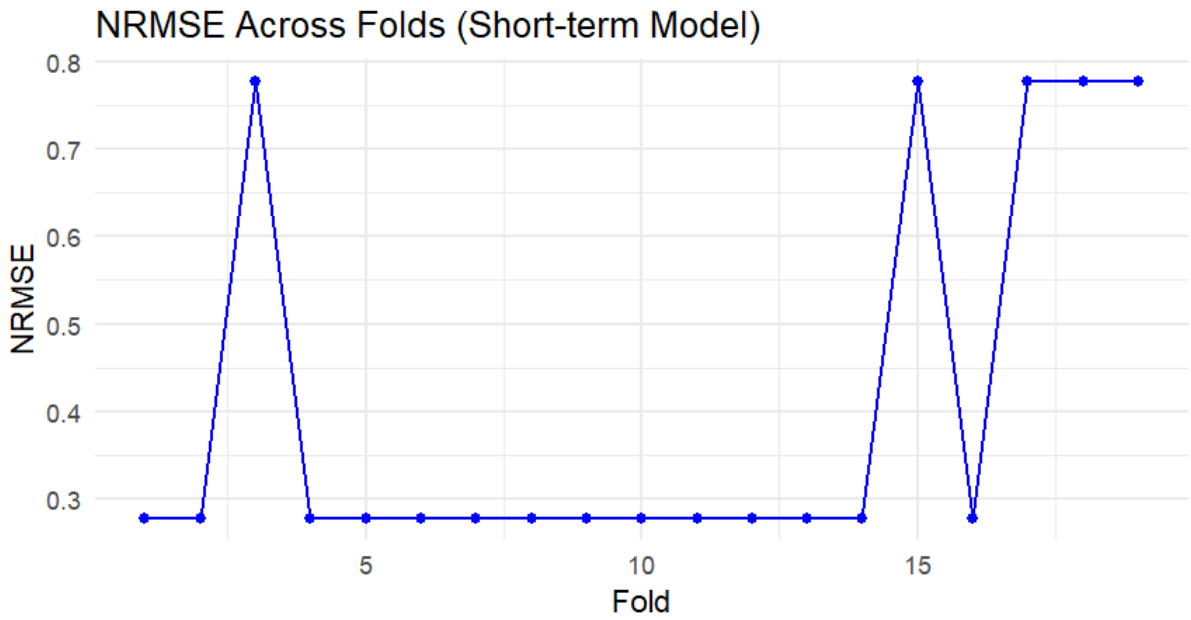

Figure 10 Distribution of normalized root mean squared error across folds for short term model

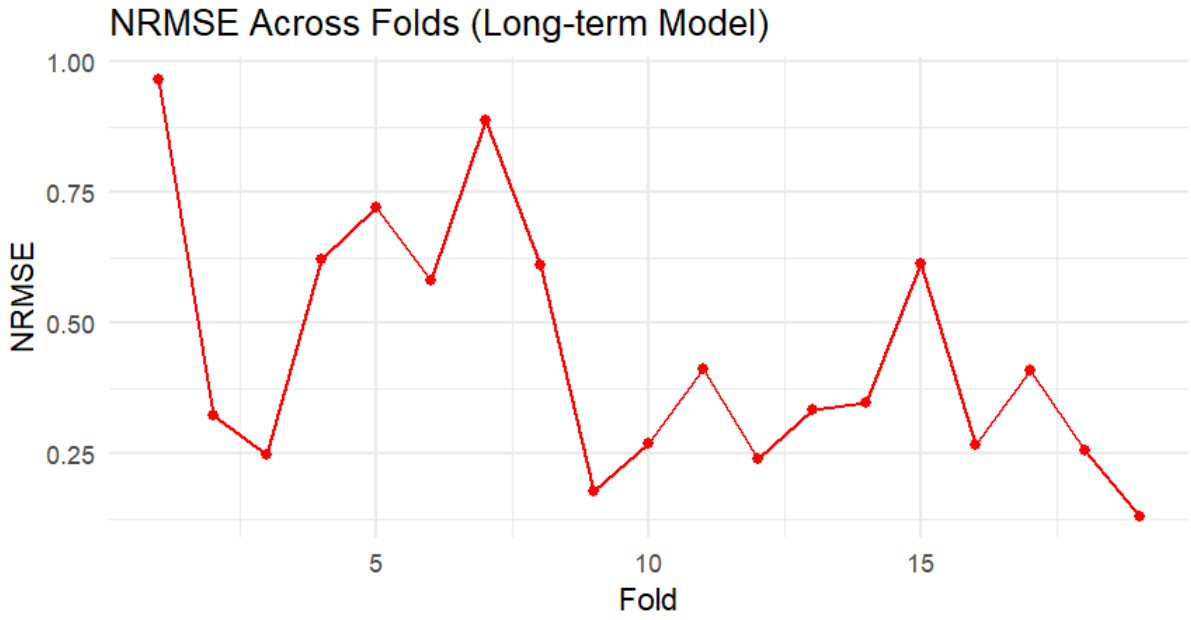

Figure 11 Distribution of normalized root mean squared error across folds for long term model

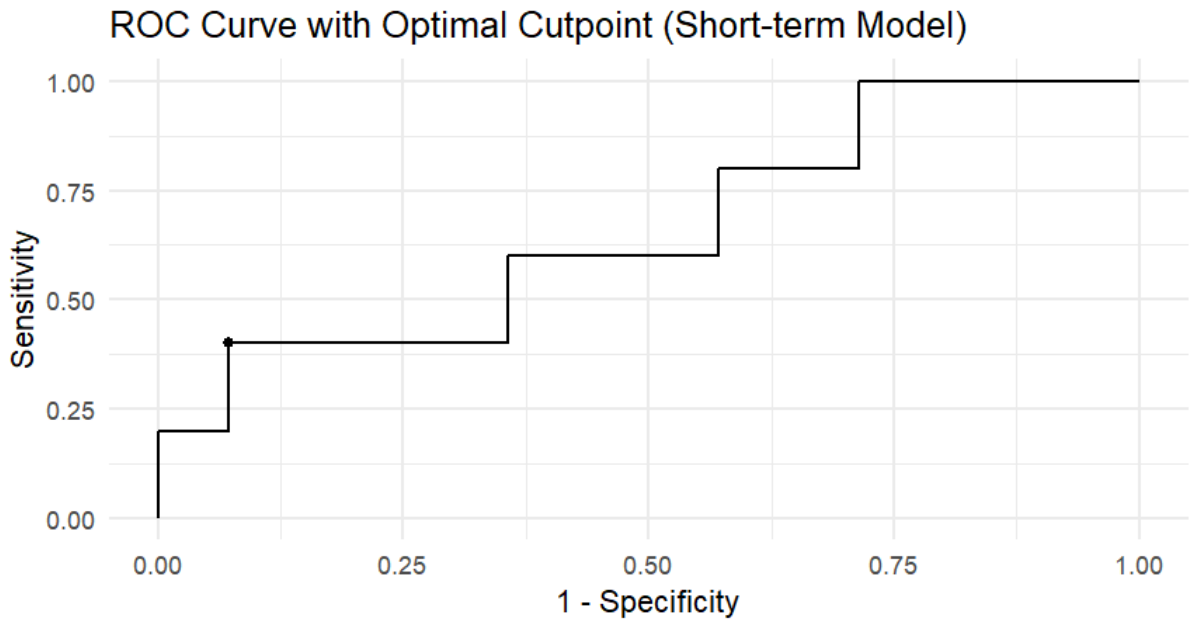

Figure 12 ROC curve with optimal cutpoint for short term model

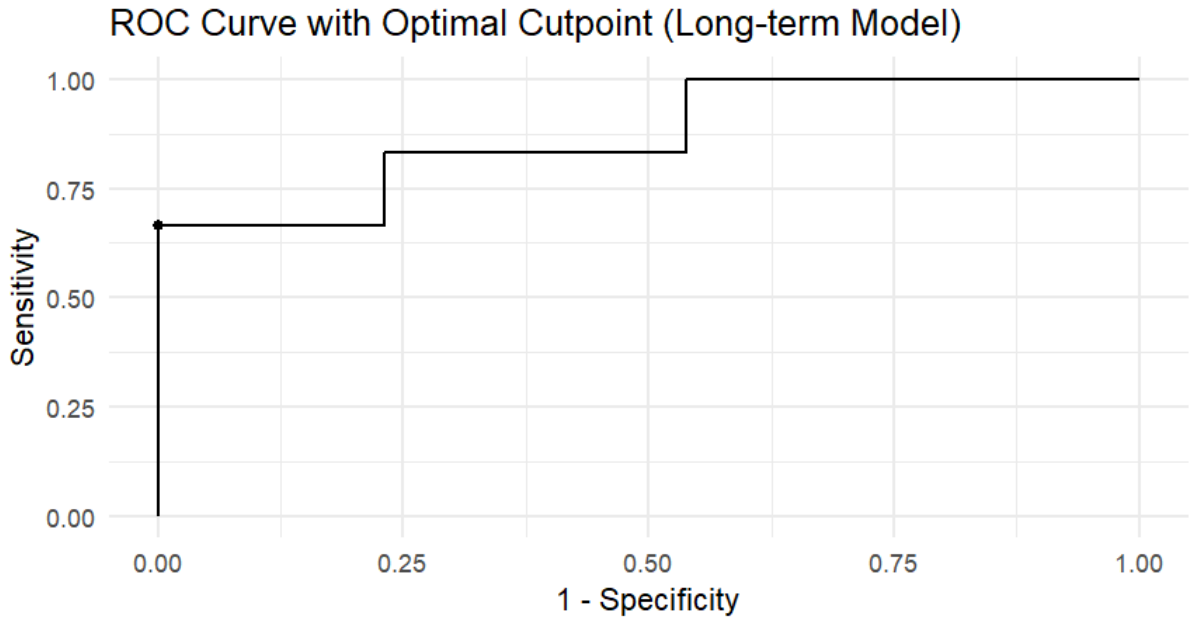

Figure 13 ROC curve with optimal cutpoint for long term model

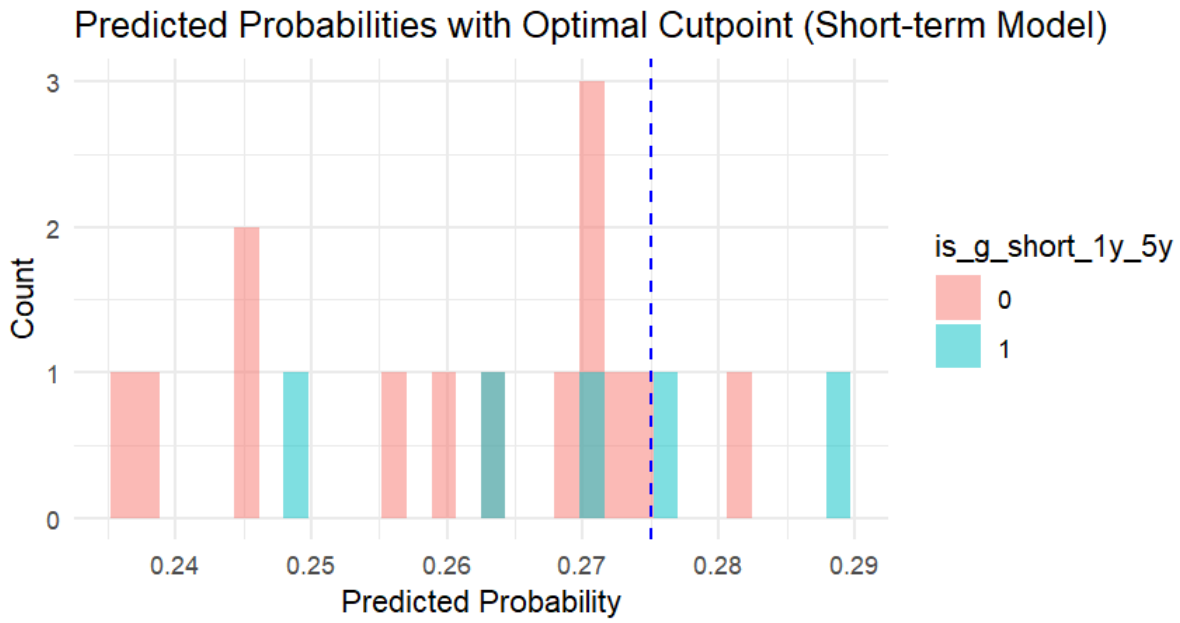

Figure 14 Distribution of predicted probabilities with optimal cutpoint (short term)

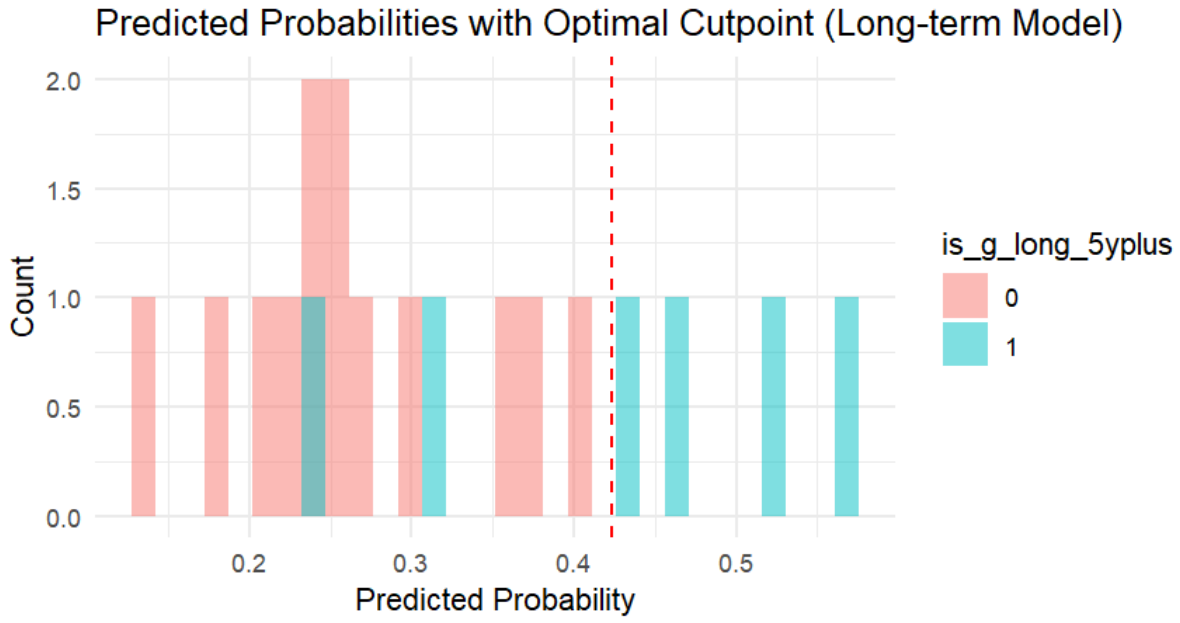

Figure 15 Distribution of predicted probabilities with optimal cutpoint (long term)

6 DISCUSSION

This study used metadata and brain glucose uptake results from 18F-fluorodeoxyglucose positron emission tomography (FDG-PET) scans and machine learning was employed to predict onset of neurological diseases (including dementia) in the short term (1–5 years) and long term (>5 years) in an aging population. The relationship between age and brain glucose uptake (BGU) was also assessed and group classifications based on BGU patterns were explored. The findings of this study provide important understanding into the capability of brain FDG-PET for neurological disease prediction, the important role of age in brain metabolism, and the challenges of differentiating diagnostic groups using BGU.

The first aim was to refine the AIVO database to allow the use of FDG-PET data to classify subjects into three groups: Control (331 participants), Memory Disorder (306 participants), and Other (728 participants). Significant demographic differences were found across groups, with the Memory Disorder group being older (mean age 63.19 years, Table 2), having a higher proportion of females (58.17%, Table 6), and lower weight and height compared to the Control group (Tables 8 and 9). These differences are in accordance with the known risk factors for dementia such as older age and female gender which are both associated with higher Alzheimer's occurrence (Mosconi et al., 2008), (GDB 2019 Dementia Forecasting Collaborators, 2022). However, no significant differences in BGU were found between the Control and Memory Disorder groups in either anatomical (Table 11) or functional (Table 12) brain regions after False Discovery Rate (FDR) correction. This lack of significant BGU differences was unexpected because regions like the temporal lobe, limbic lobe, and default mode network are typically affected in memory disorders like Alzheimer's (Mosconi et al., 2005), (Drzezga et al 2003), (Herholz et al., 2002). Several factors may explain this. The sample size for BGU analysis was bound to 147 participants with complete data which likely reduced the statistical power to detect differences. The Memory Disorder group included heterogeneous conditions (e.g., Mild Cognitive Impairment, Alzheimer's, other dementias), which may have introduced variability in BGU patterns by masking group level differences. The significant age and gender differences (Tables 7 and 10) likely influenced BGU, as older age and female gender are associated with lower brain metabolism (Mosconi et al., 2005). Without controlling these confounders in the group comparisons, neurological disease specific metabolic changes may have been masked. In my opinion, the absence of significant BGU differences does not degrade the value of FDG-PET for group classification but highlights the challenges of detecting metabolic changes. The trends towards a group difference between Control and Memory

Disorder groups in the anterior cerebellum and visual network suggest that differences may exist in regions less associated with neurological diseases and that machine learning approaches, more data-driven, can detect differences not tested with a pure hypothesis-driven method. Future studies should control for age and gender as covariates in BGU analyses and focus on more homogeneous diagnostic groups (e.g., Alzheimer's only) to enhance detection of dementia-specific patterns. Additionally combining FDG-PET with other imaging modalities e.g. amyloid PET, MRI could improve classification accuracy (Sperling et al., 2011).

The second aim was to examine the relationship between age and BGU across different brain regions. The results (Tables 13 and 14) showed significant negative correlations in all regions with Spearman's ρ values ranging from -0.166 (pons) to -0.451 (frontoparietal network). The top 10 regions (Table 14) included functional networks e.g., frontoparietal, default mode, ventral attention and anatomical regions e.g., fronto-temporal space, limbic lobe with ρ values from -0.396 to -0.451 and FDR-adjusted p-values confirming the significance (4.50×10^{-7} to 2.76×10^{-6}). These results confirms that as age increases then the brain metabolism decreases across the brain with the most prominent effects in the regions crucial for cognitive functions like executive control, memory and attention. The frontoparietal region ($\rho = -0.451$) and default mode region ($\rho = -0.439$) showed the strongest correlations suggesting that these cognitive networks are especially vulnerable to a metabolic decline related to increasing age. This is in accordance with the other research finding that aging effects cognitive regions and increases risk of having dementia (Corrada et al., 2010). Regions relevant to dementia such as the temporal lobe ($\rho = -0.361$) and posterior cingulate ($\rho = -0.352$) showed significant but weaker correlations showing that age affects these regions but to a lesser extent than cognitive networks. The weaker correlation in the pons ($\rho = -0.166$, $p = 0.0479$) suggests that brainstem regions involved in basic functions like autonomic regulation are less affected by increasing age. These findings are significant because they highlight the effect of aging on brain metabolism which is important for dementia studies. The stronger correlations in the default mode region and the limbic region are important because these regions are early sites of pathology in Alzheimer's (Mosconi et al., 2009). The Memory Disorder group showing older age seems to be contributed to the lower BGU in these regions which explains why the group differences were not significant in Tables 11 and 12.

The third aim of the study was to test that how well a machine learning model can predict neurological diseases (including dementia) in the short term (1–5 years) and long term (>5 years) using FDG-PET imaging metadata. The long term model worked well (AUC = 0.8718, sensitivity = 0.67, specificity = 1.00, accuracy = 0.89, balanced accuracy = 0.83, Table 15)

showing that it can predict dementia after 5 years as compared to the short term model that had only moderate performance (AUC = 0.6571, sensitivity = 0.40, specificity = 0.93, accuracy = 0.79, balanced accuracy = 0.66). The ROC curves also clearly show that the long-term model performed better. Its curve is closer to the top-left corner compared to the short-term model's curve which is closer to the diagonal line showing near chance performance. The short term model had high specificity (93%) and the long-term model had perfect specificity (100%) means that both models are good at identifying people who do not develop in the future neurological disorders. The short-term model's cautious predictions (optimal cutpoint = 0.275) caused many false negatives. The predicted probability plots show that the long-term model separated neurological diseases from non-neurological cases better. Its higher cutpoint (0.4236) gave a better balance between sensitivity and specificity. One important choice in this study was to include all 25 BGU regions in the LASSO models instead of using only the top 10 regions that had the strongest correlation with age (Table 14). Neurological diseases could be affected in many brain areas not just those mostly affected by aging. The feature importance graph showed that the bilateral inferior temporal gyri and the intercept were the top features in the long-term model. Interestingly the inferior temporal gyri only had a moderate correlation (ρ = -0.309, Table 13) with age. This region is known to be an early marker of Alzheimer's disease, showing hypometabolism even before clinical symptoms appear (Mosconi et al., 2009). By including all 25 regions, LASSO was able to choose features that are important to neurological diseases, which might have been missed if only the top 10 regions were used. LASSO automatically selects important features by shrinking the less useful ones to zero which makes it a good choice for handling many features and avoiding bias (Bucci et al.,2014). The finding that the inferior temporal gyri were important is meaningful as it shows the biological relevance of this region for neurologic disease onset prediction. The fact that it was selected over regions with stronger age correlations shows that neurologic diseases-related metabolic changes are different from normal aging and hints towards using a full set of features in machine learning models. None the less the findings should be tested in larger datasets to confirm that they are reliable.

With 1,447 subjects (data from AIVO), it was a good basis for the study. Gathering information of all the variables was very difficult and time consuming as well, for the reason that the data does not have a common repository where we can find it all summed up, we needed to research many data sources. Data collection consumed the greatest amount of time in this study; however, we were unable to uncover all of the missing data. Permission to use the data was also very hard to obtain since sensitive information needed to be handled. The pseudo anonymized

data could be used for the analysis which also took almost 3 months. The use of strong statistical methods, such as non-parametric Kruskal-Wallis, Dunn's, Fisher's Exact tests, and FDR correction, contributed to ensuring the robustness of results even in the presence of not normally distributed data. We employed LASSO regression with nested Leave-One-Out Cross-Validation (LOOCV) in the studies' analyses to mitigate overfitting and permitting equitable feature selection of the most informative inputs to predict neurological diseases onset (including dementia) in short- and long-term. These high correlations recorded between age and BGU (Table 14) illustrate the considerable reduction in brain metabolism with advanced age. This could justify the use of FDG PET to investigate the effects of ageing on brain metabolism and risk for neurodegenerative diseases (Mosconi et al., 2008). There are many limitations of this study too that should be taken into consideration. We attribute this, (among others), to the small sample size available for predictive modeling (among the subjects with BGU information only 41 participants had future diagnoses available), that may have affected the sensitivity of the short-term model, sensitivity (40%), (also potentially driving the perfect specificity (100%) for the long-term model, i.e. overfitting). The low sample size for BGU analysis (N = 147 cases with complete data) affected the possibility to observe group differences between controls and patients with memory disorders (Tables 11 and 12) because the missingness of this variable decreased the power of analyses. Demographic characteristics in order of older age and greater proportion of females in the Memory Disorder group were not adjusted in the BGU analyses and probably influenced the findings.

7 CONCLUSION

The study had many limitations, but it was able to highlight that if we use FDG-PET data with machine learning methods like LASSO regression then it can produce good predictions at least for long-term disease risk. No strong differences between groups could be found based on the BGU but all the correlations between age and BGU were negative and not even a single was positive which explains that how aging affects brain metabolism, especially in cognitive areas.

All 25 BGU regions were included in the models (in a data-driven fashion), so that the patterns that are related to neurological diseases onset would show up, and they did, even with a small sample. This means that machine learning models are important tools for research in neuroimaging even with small sample size. This research can be improved to get better results in the future, for instance including a larger sample or a validation cohort and by considering different confounders for the group comparisons.

The next steps for the study can be to address the data gaps and adding more types of information to the sample, like MRI or amyloid PET, to increase accuracy of prediction. Psychological test scores or even lifestyle factors can be included from the AIVO database for future study. When these improvements will be made, this will improve the models' predictive power and will help to yield better results.

REFERENCES

Anchisi, D., Borroni, B., Franceschi, M., Kerrouche, N., Kalbe, E., Beuthien-Beumann, B., Cappa, S., Lenz, O., Ludecke, S., Marcone, A., Mielke, R., Ortelli, P., Padovani, A., Pelati, O., Pupi, A., Scarpini, E., Weisenbach, S., Herholz, K., Salmon, E., … Perani, D. (2005). Heterogeneity of brain glucose metabolism in mild cognitive impairment and clinical progression to Alzheimer disease. *Archives of Neurology*, *62*(11), 1728–1733. https://doi.org/10.1001/archneur.62.11.1728

Benussi, A., Karikari, T. K., Ashton, N., Gazzina, S., Premi, E., Benussi, L., Ghidoni, R., Rodriguez, J. L., Emeršič, A., Simrén, J., Binetti, G., Fostinelli, S., Giunta, M., Gasparotti, R., Zetterberg, H., Blennow, K., & Borroni, B. (2020). Diagnostic and prognostic value of serum NfL and p-Tau181 in frontotemporal lobar degeneration. *Journal of Neurology, Neurosurgery, and Psychiatry*, *91*(9), 960–967. https://doi.org/10.1136/jnnp-2020-323487

Blondell, S. J., Hammersley-Mather, R., & Veerman, J. L. (2014). Does physical activity prevent cognitive decline and dementia?: A systematic review and meta-analysis of longitudinal studies. *BMC Public Health*, *14*, 510. https://doi.org/10.1186/1471-2458-14-510

Boellaard, R., Delgado-Bolton, R., Oyen, W. J. G., Giammarile, F., Tatsch, K., Eschner, W., Verzijlbergen, F. J., Barrington, S. F., Pike, L. C., Weber, W. A., Stroobants, S., Delbeke, D., Donohoe, K. J., Holbrook, S., Graham, M. M., Testanera, G., Hoekstra, O. S., Zijlstra, J., Visser, E., … European Association of Nuclear Medicine (EANM). (2015). FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *European Journal of Nuclear Medicine and Molecular Imaging*, *42*(2), 328–354. https://doi.org/10.1007/s00259-014-2961-x

Braak, H., & Del Tredici, K. (2015). The preclinical phase of the pathological process underlying sporadic Alzheimer's disease. *Brain: A Journal of Neurology*, *138*(Pt 10), 2814–2833. https://doi.org/10.1093/brain/awv236

Caminiti, S. P., Ballarini, T., Sala, A., Cerami, C., Presotto, L., Santangelo, R., Fallanca, F., Vanoli, E. G., Gianolli, L., Iannaccone, S., Magnani, G., Perani, D., & BIOMARKAPD Project. (2018). FDG-PET and CSF biomarker accuracy in prediction

of conversion to different dementias in a large multicentre MCI cohort. *NeuroImage. Clinical*, *18*, 167–177. https://doi.org/10.1016/j.nicl.2018.01.019

Chincarini, A., Bosco, P., Calvini, P., Gemme, G., Esposito, M., Olivieri, C., Rei, L., Squarcia, S., Rodriguez, G., Bellotti, R., Cerello, P., De Mitri, I., Retico, A., & Nobili, F. (2011). Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer's disease. *NeuroImage*, *58*(2), 469–480. https://doi.org/10.1016/j.neuroimage.2011.05.083

Corrada, M. M., Brookmeyer, R., Paganini-Hill, A., Berlau, D., & Kawas, C. H. (2010). Dementia incidence continues to increase with age in the oldest old: The 90+ study. *Annals of Neurology*, *67*(1), 114–121. https://doi.org/10.1002/ana.21915

Craig-Schapiro, R., Perrin, R. J., Roe, C. M., Xiong, C., Carter, D., Cairns, N. J., Mintun, M. A., Peskind, E. R., Li, G., Galasko, D. R., Clark, C. M., Quinn, J. F., D'Angelo, G., Malone, J. P., Townsend, R. R., Morris, J. C., Fagan, A. M., & Holtzman, D. M. (2010). YKL-40: A novel prognostic fluid biomarker for preclinical Alzheimer's disease. *Biological Psychiatry*, *68*(10), 903–912. https://doi.org/10.1016/j.biopsych.2010.08.025

Davatzikos, C., Bhatt, P., Shaw, L. M., Batmanghelich, K. N., & Trojanowski, J. Q. (2011). Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiology of Aging*, *32*(12), 2322.e19-27. https://doi.org/10.1016/j.neurobiolaging.2010.05.023

de Leon, M. J., Convit, A., Wolf, O. T., Tarshish, C. Y., DeSanti, S., Rusinek, H., Tsui, W., Kandil, E., Scherer, A. J., Roche, A., Imossi, A., Thorn, E., Bobinski, M., Caraos, C., Lesbre, P., Schlyer, D., Poirier, J., Reisberg, B., & Fowler, J. (2001). Prediction of cognitive decline in normal elderly subjects with 2-[(18)F]fluoro-2-deoxy-D-glucose/poitron-emission tomography (FDG/PET). *Proceedings of the National Academy of Sciences of the United States of America*, *98*(19), 10966–10971. https://doi.org/10.1073/pnas.191044198

Dickerson, B. C., Bakkour, A., Salat, D. H., Feczko, E., Pacheco, J., Greve, D. N., Grodstein, F., Wright, C. I., Blacker, D., Rosas, H. D., Sperling, R. A., Atri, A., Growdon, J. H., Hyman, B. T., Morris, J. C., Fischl, B., & Buckner, R. L. (2009). The cortical signature of Alzheimer's disease: Regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. *Cerebral Cortex (New York, N.Y.: 1991)*, *19*(3), 497–510. https://doi.org/10.1093/cercor/bhn113

Dimitriadis, S. I., Liparas, D., & Tsolaki, M. N. (2018). Random forest feature selection, fusion and ensemble strategy: Combining multiple morphological MRI measures to discriminate among healhy elderly, MCI, cMCI and alzheimer's disease patients: From the alzheimer's disease neuroimaging initiative (ADNI) database. *Journal of Neuroscience Methods*, *302*, 14–23. https://doi.org/10.1016/j.jneumeth.2017.12.010

Drzezga, A., Lautenschlager, N., Siebner, H., Riemenschneider, M., Willoch, F., Minoshima, S., Schwaiger, M., & Kurz, A. (2003). Cerebral metabolic changes accompanying conversion of mild cognitive impairment into Alzheimer's disease: A PET follow-up study. *European Journal of Nuclear Medicine and Molecular Imaging*, *30*(8), 1104–1113. https://doi.org/10.1007/s00259-003-1194-1

GBD 2019 Dementia Forecasting Collaborators (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, e105-e125. - References—Scientific Research Publishing. (n.d.). Retrieved July 29, 2025, from https://www.scirp.org/reference/referencespapers?referenceid=3676701

Global status report on the public health response to dementia. https://www.who.int/publications/i/item/9789240033245

Guerreiro, R., Wojtas, A., Bras, J., Carrasquillo, M., Rogaeva, E., Majounie, E., Cruchaga, C., Sassi, C., Kauwe, J. S. K., Younkin, S., Hazrati, L., Collinge, J., Pocock, J., Lashley, T., Williams, J., Lambert, J.-C., Amouyel, P., Goate, A., Rademakers, R., … Alzheimer Genetic Analysis Group. (2013). TREM2 variants in Alzheimer's disease. *The New England Journal of Medicine*, *368*(2), 117–127. https://doi.org/10.1056/NEJMoa1211851

Hansson, O., Seibyl, J., Stomrud, E., Zetterberg, H., Trojanowski, J. Q., Bittner, T., Lifke, V., Corradini, V., Eichenlaub, U., Batrla, R., Buck, K., Zink, K., Rabe, C., Blennow, K., Shaw, L. M., Swedish BioFINDER study group, & Alzheimer's Disease Neuroimaging Initiative. (2018). CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and 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), 1470–1481. https://doi.org/10.1016/j.jalz.2018.01.010

Hardy, J. A., & Higgins, G. A. (1992). Alzheimer's disease: The amyloid cascade hypothesis. *Science (New York, N.Y.)*, *256*(5054), 184–185. https://doi.org/10.1126/science.1566067

Hason, L., & Krishnan, S. (2022). Spontaneous speech feature analysis for alzheimer's disease screening using a random forest classifier. *Frontiers in Digital Health*, *4*, 901419. https://doi.org/10.3389/fdgth.2022.901419

Heneka, M. T., Carson, M. J., El Khoury, J., Landreth, G. E., Brosseron, F., Feinstein, D. L., Jacobs, A. H., Wyss-Coray, T., Vitorica, J., Ransohoff, R. M., Herrup, K., Frautschy, S. A., Finsen, B., Brown, G. C., Verkhratsky, A., Yamanaka, K., Koistinaho, J., Latz, E., Halle, A., … Kummer, M. P. (2015). Neuroinflammation in Alzheimer's disease. *The Lancet. Neurology*, *14*(4), 388–405. https://doi.org/10.1016/S1474-4422(15)70016-5

Herholz, K., Salmon, E., Perani, D., Baron, J. C., Holthoff, V., Frölich, L., Schönknecht, P., Ito, K., Mielke, R., Kalbe, E., Zündorf, G., Delbeuck, X., Pelati, O., Anchisi, D., Fazio, F., Kerrouche, N., Desgranges, B., Eustache, F., Beuthien-Baumann, B., … Heiss, W. D. (2002). Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *NeuroImage*, *17*(1), 302–316. https://doi.org/10.1006/nimg.2002.1208

Iadecola, C. (2013). The pathobiology of vascular dementia. *Neuron*, *80*(4), 844–866. https://doi.org/10.1016/j.neuron.2013.10.008

Islam, J., & Zhang, Y. (2018). Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. *Brain Informatics*, *5*(2), 2. https://doi.org/10.1186/s40708-018-0080-3

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., … Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *14*(4), 535–562. https://doi.org/10.1016/j.jalz.2018.02.018

Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., & Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet. Neurology*, *9*(1), 119–128. https://doi.org/10.1016/S1474-4422(09)70299-6

Klunk, W. E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D. P., Bergström, M., Savitcheva, I., Huang, G., Estrada, S., Ausén, B., Debnath, M. L., Barletta, J., Price, J. C., Sandell, J., Lopresti, B. J., Wall, A., Koivisto, P., Antoni, G., … Långström, B.

(2004). Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Annals of Neurology*, *55*(3), 306–319. https://doi.org/10.1002/ana.20009

Kuiper, J. S., Zuidersma, M., Oude Voshaar, R. C., Zuidema, S. U., van den Heuvel, E. R., Stolk, R. P., & Smidt, N. (2015). Social relationships and risk of dementia: A systematic review and meta-analysis of longitudinal cohort studies. *Ageing Research Reviews*, *22*, 39–57. https://doi.org/10.1016/j.arr.2015.04.006

Kunkle, B. W., Grenier-Boley, B., Sims, R., Bis, J. C., Damotte, V., Naj, A. C., Boland, A., Vronskaya, M., van der Lee, S. J., Amlie-Wolf, A., Bellenguez, C., Frizatti, A., Chouraki, V., Martin, E. R., Sleegers, K., Badarinarayan, N., Jakobsdottir, J., Hamilton-Nelson, K. L., Moreno-Grau, S., … Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium (GERAD/PERADES),. (2019). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nature Genetics*, *51*(3), 414–430. https://doi.org/10.1038/s41588-019-0358-2

Lanoiselée, H.-M., Nicolas, G., Wallon, D., Rovelet-Lecrux, A., Lacour, M., Rousseau, S., Richard, A.-C., Pasquier, F., Rollin-Sillaire, A., Martinaud, O., Quillard-Muraine, M., de la Sayette, V., Boutoleau-Bretonniere, C., Etcharry-Bouyx, F., Chauviré, V., Sarazin, M., le Ber, I., Epelbaum, S., Jonveaux, T., … Campion, D. (2017). APP, PSEN1, and PSEN2 mutations in early-onset Alzheimer disease: A genetic screening study of familial and sporadic cases. *PLoS Medicine*, *14*(3), e1002270. https://doi.org/10.1371/journal.pmed.1002270

Lebedev, A. V., Westman, E., Van Westen, G. J. P., Kramberger, M. G., Lundervold, A., Aarsland, D., Soininen, H., Kłoszewska, I., Mecocci, P., Tsolaki, M., Vellas, B., Lovestone, S., Simmons, A., & Alzheimer's Disease Neuroimaging Initiative and the AddNeuroMed consortium. (2014). Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage. Clinical*, *6*, 115–125. https://doi.org/10.1016/j.nicl.2014.08.023

Liu, M., Zhang, D., & Shen, D. (2012). Ensemble sparse classification of Alzheimer's disease. *NeuroImage*, *60*(2), 1106–1116. https://doi.org/10.1016/j.neuroimage.2012.01.055

Magnin, B., Mesrob, L., Kinkingnéhun, S., Pélégrini-Issac, M., Colliot, O., Sarazin, M., Dubois, B., Lehéricy, S., & Benali, H. (2009). Support vector machine-based

classification of Alzheimer's disease from whole-brain anatomical MRI. *Neuroradiology*, *51*(2), 73–83. https://doi.org/10.1007/s00234-008-0463-x

Minoshima, S., Foster, N. L., Sima, A. A. F., Frey, K. A., Albin, R. L., & Kuhl, D. E. (2001). Alzheimer's disease versus dementia with Lewy bodies: Cerebral metabolic distinction with autopsy confirmation. *Annals of Neurology*, *50*(3), 358–365. https://doi.org/10.1002/ana.1133

Minoshima, S., Frey, K. A., Koeppe, R. A., Foster, N. L., & Kuhl, D. E. (1995). A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine*, *36*(7), 1238–1248.

Mosconi, L., Mistur, R., Switalski, R., Tsui, W. H., Glodzik, L., Li, Y., Pirraglia, E., De Santi, S., Reisberg, B., Wisniewski, T., & de Leon, M. J. (2009b). FDG-PET changes in brain glucose metabolism from normal cognition to pathologically verified Alzheimer's disease. *European Journal of Nuclear Medicine and Molecular Imaging*, *36*(5), 811–822. https://doi.org/10.1007/s00259-008-1039-z

Mosconi, L., Tsui, W.-H., De Santi, S., Li, J., Rusinek, H., Convit, A., Li, Y., Boppana, M., & de Leon, M. J. (2005). Reduced hippocampal metabolism in MCI and AD: Automated FDG-PET image analysis. *Neurology*, *64*(11), 1860–1867. https://doi.org/10.1212/01.WNL.0000163856.13524.08

Olsson, B., Lautner, R., Andreasson, U., Öhrfelt, A., Portelius, E., Bjerke, M., Hölttä, M., Rosén, C., Olsson, C., Strobel, G., Wu, E., Dakin, K., Petzold, M., Blennow, K., & Zetterberg, H. (2016). CSF and blood biomarkers for the diagnosis of Alzheimer's disease: A systematic review and meta-analysis. *The Lancet. Neurology*, *15*(7), 673–684. https://doi.org/10.1016/S1474-4422(16)00070-3

Ossenkoppele, R., Reimand, J., Smith, R., Leuzy, A., Strandberg, O., Palmqvist, S., Stomrud, E., Zetterberg, H., Alzheimer's Disease Neuroimaging Initiative, Scheltens, P., Dage, J. L., Bouwman, F., Blennow, K., Mattsson-Carlgren, N., Janelidze, S., & Hansson, O. (2021). Tau PET correlates with different Alzheimer's disease-related features compared to CSF and plasma p-tau biomarkers. *EMBO Molecular Medicine*, *13*(8), e14398. https://doi.org/10.15252/emmm.202114398

Palmqvist, S., Janelidze, S., Quiroz, Y. T., Zetterberg, H., Lopera, F., Stomrud, E., Su, Y., Chen, Y., Serrano, G. E., Leuzy, A., Mattsson-Carlgren, N., Strandberg, O., Smith, R., Villegas, A., Sepulveda-Falla, D., Chai, X., Proctor, N. K., Beach, T. G., Blennow, K., … Hansson, O. (2020). Discriminative Accuracy of Plasma Phospho-tau217 for

Alzheimer Disease vs Other Neurodegenerative Disorders. *JAMA*, *324*(8), 772–781. https://doi.org/10.1001/jama.2020.12134

A diagnostic approach in Alzheimer`s disease using three-dimensional stereotactic surface projections of Fluorine-18-FDG PET. https://www.researchgate.net/publication/15408156_A_diagnostic_approach_in_Alzheimers_disease_using_three-dimensional_stereotactic_surface_projections_of_Fluorine-18-FDG_PET

FDG-PET changes in brain glucose metabolism from normal cognition to pathologically verified Alzheimer's disease. *ResearchGate*. https://doi.org/10.1007/s00259-008-1039-z

Prince, M. J., Wu, F., Guo, Y., Gutierrez Robledo, L. M., O'Donnell, M., Sullivan, R., & Yusuf, S. (2015). The burden of disease in older people and implications for health policy and practice. *Lancet (London, England)*, *385*(9967), 549–562. https://doi.org/10.1016/S0140-6736(14)61347-7

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: The DSM-5 approach. *Nature Reviews. Neurology*, *10*(11), 634–642. https://doi.org/10.1038/nrneurol.2014.181

Scarmeas, N., Stern, Y., Tang, M.-X., Mayeux, R., & Luchsinger, J. A. (2006). Mediterranean diet and risk for Alzheimer's disease. *Annals of Neurology*, *59*(6), 912–921. https://doi.org/10.1002/ana.20854

Shivamurthy, V. K. N., Tahari, A. K., Marcus, C., & Subramaniam, R. M. (2015). Brain FDG PET and the diagnosis of dementia. *AJR. American Journal of Roentgenology*, *204*(1), W76-85. https://doi.org/10.2214/AJR.13.12363

Silverman, D. H., Small, G. W., Chang, C. Y., Lu, C. S., Kung De Aburto, M. A., Chen, W., Czernin, J., Rapoport, S. I., Pietrini, P., Alexander, G. E., Schapiro, M. B., Jagust, W. J., Hoffman, J. M., Welsh-Bohmer, K. A., Alavi, A., Clark, C. M., Salmon, E., de Leon, M. J., Mielke, R., … Phelps, M. E. (2001). Positron emission tomography in evaluation of dementia: Regional brain metabolism and long-term outcome. *JAMA*, *286*(17), 2120–2127. https://doi.org/10.1001/jama.286.17.2120

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., Iwatsubo, T., Jack, C. R., Kaye, J., Montine, T. J., Park, D. C., Reiman, E. M., Rowe, C. C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M. C., Thies, B., Morrison-Bogorad, M., … Phelps, C. H. (2011). Toward defining the preclinical stages of Alzheimer's disease:

Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *7*(3), 280–292. https://doi.org/10.1016/j.jalz.2011.03.003

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, *47*(10), 2015–2028. https://doi.org/10.1016/j.neuropsychologia.2009.03.004

Suk, H.-I., Lee, S.-W., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2014). Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage*, *101*, 569–582. https://doi.org/10.1016/j.neuroimage.2014.06.077

Sweeney, M. D., Zhao, Z., Montagne, A., Nelson, A. R., & Zlokovic, B. V. (2019). Blood-Brain Barrier: From Physiology to Disease and Back. *Physiological Reviews*, *99*(1), 21–78. https://doi.org/10.1152/physrev.00050.2017

Tarkin, J. M., Joshi, F. R., & Rudd, J. H. F. (2014). PET imaging of inflammation in atherosclerosis. *Nature Reviews. Cardiology*, *11*(8), 443–457. https://doi.org/10.1038/nrcardio.2014.80

Villemagne, V. L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K. A., Salvado, O., Szoeke, C., Macaulay, S. L., Martins, R., Maruff, P., Ames, D., Rowe, C. C., Masters, C. L., & Australian Imaging Biomarkers and Lifestyle (AIBL) Research Group. (2013). Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: A prospective cohort study. *The Lancet. Neurology*, *12*(4), 357–367. https://doi.org/10.1016/S1474-4422(13)70044-9

World health statistics 2023. (2023). World Health Organization. https://www.who.int/publications/b/69040

Young, J., Modat, M., Cardoso, M. J., Mendelson, A., Cash, D., Ourselin, S., & Alzheimer's Disease Neuroimaging Initiative. (2013). Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. *NeuroImage. Clinical*, *2*, 735–745. https://doi.org/10.1016/j.nicl.2013.05.004

Bucci M, Almkvist O, Bluma M, Ashton NJ, Savitcheva I, Chiotis K, Di Molfetta G, Blennow K, Zetterberg H, Nordberg A. Profiling plasma biomarkers, particularly pTau217 and pTau217/Aβ42, and their relation to cognition in memory clinic patients. Journal of Neurochemistry. 2025 (accepted, ahead of print).

Bucci M*, Bluma M*, Savitcheva I, Ashton NJ, Chiotis K, Matton A, Kivipelto M, Di Molfetta G, Blennow K, Zetterberg H, Nordberg A. Profiling of plasma biomarkers in

the context of memory assessment in a tertiary memory clinic. Transl Psychiatry. 2023 Jul 25;13(1):268. *Equal contribution