

UNIVERSITY OF PAVIA – IUSS SCHOOL FOR ADVANCED STUDIES PAVIA

Department of Brain and Behavioral Sciences (DBBS)

MSc in Psychology, Neuroscience and Human Sciences



**UNIVERSITÀ
DI PAVIA**



IUSS

LATENT TRAJECTORY ANALYSIS OF IMMEDIATE, 30-MINUTE AND ONE-WEEK DELAYED RECALL AS A SENSITIVE COGNITIVE MARKER OF EARLY ALZHEIMER'S DISEASE PATHOLOGY IN COGNITIVELY UNIMPAIRED INDIVIDUALS

Supervisors:

Prof. Catricalà Eleonora

Prof. Sánchez-Benavides Gonzalo

Candidate

Ana Perez Mena

MAT. 549637

Academic year 2025/2026

TABLE OF CONTENTS

ABSTRACT	III
LIST OF ABBREVIATIONS	V
INTRODUCTION.....	VII
1. Alzheimer’s Disease	1
1.1 Definition and Epidemiology	1
1.2 Neuropathological Alterations	2
1.2.1 Amyloid- β Plaques	3
1.2.2 Tau Neurofibrillary Tangles	3
1.2.3 Synaptic Loss, Neurodegeneration, and Glial Responses	4
2. The AD Continuum: Staging Frameworks.....	6
2.1 From Clinical to Biological Definition: A/T/N and the Preclinical Stage	6
2.2 Risk Factors.....	8
2.2.1 Age.....	9
2.2.2 Genetic Risk: <i>APOE</i> - ϵ 4 Genotype and Family History	9
2.2.3 Sex	10
2.2.4 Education and Cognitive Reserve.....	10
2.2.5 Modifiable Risk Factors	11
2.3 Fluid Biomarker Measures	12
2.3.1 CSF A β 42/A β 40 Ratio.....	12
2.3.2 CSF p-tau181/A β 42 Ratio	13
2.3.3 Plasma p-tau217	14
2.4 Neuroimaging Markers	14
2.4.1 Structural Magnetic Resonance Imaging in Preclinical AD	14
2.4.2 The Hippocampus and Its Subfields: Early Vulnerability of CA1	15
2.4.3 AD Cortical Signatures.....	16
2.4.4 Amyloid- β PET and Tau-PET Imaging.....	16
2.5 Cognitive and Neuropsychological Markers of Preclinical AD.....	17
2.5.1 Subjective Cognitive Decline Assessment: MyCog Questionnaire	17
2.5.2 Memory Assessment: The Free and Cued Selective Reminding Test.....	17
3. Long-Term Memory Assessment for Preclinical AD Detection: Current Studies Review ..	19
3.1 Definition of Accelerated Long-Term Forgetting	19
3.1.1 Origins of the Construct	20
3.1.2 Rationale Behind the Use of One-Week Delayed Recall	21
3.2 Search Strategy.....	22
3.3 Current Evidence.....	25
3.3.1 ALF in Symptomatic AD and Mild Cognitive Impairment	25
3.3.2 ALF in Individuals at Genetic Risk	26
3.3.3 ALF in Cognitively Unimpaired Individuals.....	28
3.3.4 Methodological Heterogeneity and the Limits of Ratio-Based Scoring.....	29
3.4 Gap of Knowledge	31
4. The Present Study: Aims and Hypotheses.....	33
4.1 From the Research Gap to the Present Design.....	33

4.2 Aims	33
4.3 Hypotheses	34
4.4 Expected Contribution.....	35
5. Methods.....	37
5.1 Study Population: The ALFA+ Cohort	37
5.2 Sample Selection and Description.....	38
5.3 Neuropsychological Assessment.....	42
5.3.1 The FCSRT Protocol	43
5.3.2 The One-Week Delayed Recall by Telephone	45
5.4 Fluid Biomarkers and A β /AT Status Definition	46
5.5 Neuroimaging Acquisition and Processing	47
5.5.1 Hippocampal Subfield Segmentation	48
5.5.2 Cortical Thickness and AD Signatures.....	49
5.6 Operationalization of Long-Term Forgetting.....	49
5.6.1 Growth Mixture Modeling of FCSRT Trajectories.....	50
5.6.2 Model Selection and the Minimum Class-Size Rule.....	51
5.6.3 The Continuous ALF Ratio	52
5.7 Statistical Analysis	53
5.7.1 Cross-Sectional and Longitudinal Models	53
5.7.2 Correction for Multiple Comparisons.....	55
5.7.3 Sensitivity Analyses	55
6. Results	56
6.1 Latent Trajectory Classes.....	56
6.1.1 Model Selection and Fit.....	57
6.1.2 Description of the Three Classes	60
6.2 Associations With Fluid Biomarkers and Subjective Cognitive Complaints	62
6.3 Cross-Sectional Imaging Associations.....	64
6.4 Longitudinal Hippocampal Atrophy	64
6.5 Convergence With the Continuous ALF Ratio	67
6.6 Sensitivity Analyses	68
7. Discussion.....	70
7.1 Main Findings	70
7.2 Regional Specificity of Longitudinal Structural Change	72
7.3 Trajectory Classes and Continuous Ratio	73
7.4 Clinical and Research Implications.....	75
7.5 Limitations	75
8. Conclusion.....	79
REFERENCES.....	82
APPENDICES	110

ABSTRACT

Background

Accelerated long-term forgetting - disproportionate memory loss over days despite normal short-delay retention - has been suggested as an early cognitive manifestation of preclinical Alzheimer's disease (AD). Existing retention-ratio approaches overlook the full acquisition-to-consolidation pattern. We aimed to characterize distinct learning-retention profiles and explore their associations with early AD pathology in cognitively unimpaired adults.

Methods

A total of 322 cognitively unimpaired older adults (mean age 67, 60% female, 57% *APOE*- $\epsilon 4$ carriers, 95% with family history of AD) completed the Free and Cued Selective Reminding Test face-to-face (16-word list: three immediate recall trials, a 30-minute delayed recall). An additional, unannounced one-week delayed recall was administered by phone. We used growth mixture modeling across these five timepoints to identify distinct trajectory classes and explored their associations with AD biomarkers (CSF $A\beta 42/A\beta 40$, p-tau181/ $A\beta 42$; plasma p-tau217) using general linear models, and retrospective longitudinal MRI (bilateral total medial temporal lobe and hippocampal subfield volumes derived from a dedicated high-resolution inversion-recovery acquisition processed using the ASHS pipeline, three acquisitions across 6.6 years) employing linear mixed-effects models, adjusted for age, sex, and education, and intracranial volume for volumetric outcomes. Associations were corrected for multiple comparisons within outcome families by the Benjamini-Hochberg procedure (p_{FDR}).

Results

Three distinct trajectory classes were identified: High ($n = 201$; 62.4%), Intermediate ($n = 81$; 25.2%), and Low performers ($n = 40$; 12.4%). Intermediate performers showed preserved learning and 30-minute recall but disproportionate forgetting at one week compared to High performers. Plasma p-tau217 levels were significantly associated with classes ($p_{FDR} = .021$), with Low performers showing higher concentrations than Intermediate ($\beta = 0.49$; $p = .012$), and High performers ($\beta = 0.50$; $p = .006$). CSF p-tau181/ $A\beta 42$ followed the same pattern ($p_{FDR} = .021$), with Low performers differing from High performers ($\beta = 0.58$; $p = .006$) and Intermediate performers not ($\beta = 0.22$; $p = .161$), while CSF $A\beta 42/A\beta 40$ was not associated

with class. Longitudinal models revealed significant class $\times$ time interactions for bilateral total medial temporal lobe volume ($p_{\text{FDR}} = .034$); Low performers showed significantly greater atrophy than High performers ($\beta = -0.056$; $p = .015$), corresponding to $2.3\times$ greater annual atrophy. In particular, CA1 subfield revealed the strongest effect ($p_{\text{FDR}} = .019$); both Low ($\beta = -0.048$; $p = .004$) and Intermediate performers ($\beta = -0.029$; $p = .036$) showed significantly greater annual decline than High performers ($3.0\times$ and $2.2\times$ faster annual atrophy, respectively).

Conclusions

Learning and forgetting trajectories including delayed recalls beyond standard intervals relate to early AD pathology, in particular plasma p-tau217 elevation, and accelerated hippocampal atrophy. Further studies exploring the ability of combined metrics of plasma p-tau217 and long-term recall trials to predict short-term cognitive decline are warranted.

Keywords: accelerated long-term forgetting, preclinical Alzheimer’s disease, memory consolidation, growth mixture modeling, free and cued selective reminding test, plasma p-tau217, hippocampal volume

LIST OF ABBREVIATIONS

A β	amyloid- β
A β 40	amyloid- β peptide of 40 amino acids
A β 42	amyloid- β peptide of 42 amino acids
AD	Alzheimer's disease
ALF	accelerated long-term forgetting
ALFA	ALzheimer and FAMilies study
AMPA	α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
<i>APOE</i>	apolipoprotein E (gene)
<i>APOE</i> - ϵ 4	apolipoprotein E ϵ 4 allele
APP	amyloid precursor protein
ASHS	Automatic Segmentation of Hippocampal Subfields
A/T	amyloid and tau biomarker status
A/T/N	amyloid, tau and neurodegeneration classification system
BIC	Bayesian information criterion
CA1-CA4	cornu ammonis, fields 1 to 4
CAIDE	Cardiovascular Risk Factors, Aging and Dementia score
CI	confidence interval
COLBOS	Colombia-Boston biomarker study
CSF	cerebrospinal fluid
DNA	deoxyribonucleic acid
FCSRT	Free and Cued Selective Reminding Test
FCSRT-IR	Free and Cued Selective Reminding Test, Immediate Recall version
FDG	fluorodeoxyglucose
FDR	false discovery rate
GFAP	glial fibrillary acidic protein
IQR	interquartile range
LDL	low-density lipoprotein
MCI	mild cognitive impairment
MRI	magnetic resonance imaging
MTL	medial temporal lobe
NfL	neurofilament light chain
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association
NMDA	N-methyl-D-aspartate
PET	positron emission tomography
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
<i>PSEN1</i>	presenilin 1 (gene)

<i>PSEN2</i>	presenilin 2 (gene)
p-tau	phosphorylated tau
p-tau181	tau phosphorylated at threonine 181
p-tau217	tau phosphorylated at threonine 217
SCD	subjective cognitive decline
SCD-Q	Subjective Cognitive Decline Questionnaire
SD	standard deviation
sTREM2	soluble triggering receptor expressed on myeloid cells 2
TREM2	triggering receptor expressed on myeloid cells 2
UGOT	University of Gothenburg
WMS-IV	Wechsler Memory Scale, Fourth Edition
XNAT	Extensible Neuroimaging Archive Toolkit
YKL-40	chitinase-3-like protein 1

INTRODUCTION

Alzheimer's disease (AD) is increasingly understood as a biological continuum that begins years, and potentially decades, before clinically evident cognitive impairment emerges (Jack et al., 2024; Sperling et al., 2011). Amyloid- β and tau accumulation, followed by synaptic dysfunction and neurodegeneration, may therefore precede measurable symptoms by a substantial period. This shift to a biological, biomarker-based definition of AD (Jack et al., 2024) has made the identification of subtle cognitive change in the preclinical phase a major priority for research.

Cognitive tests are traditionally designed to identify clinically meaningful impairment, but their ability to detect small differences among cognitively unimpaired individuals is often limited. This limitation is especially relevant for episodic memory, which is one of the cognitive domains most closely linked to the early pathophysiology of AD, where performance near the upper limit of a test can reduce the observable variability and may conceal the earliest disease-related changes.

In the present study, episodic memory was assessed using the Free and Cued Selective Reminding Test (FCSRT), a test designed to improve its assessment by controlling encoding and combining free and cued retrieval (Grober & Buschke, 1987). Controlled encoding ensures that the material has been learned, while semantic cues separate difficulties of retrieval from difficulties of availability or consolidation. Its standard protocol, however, probes delayed recall no later than 30 minutes after encoding, an interval at which cognitively unimpaired individuals approach the ceiling of the scale (Lindbergh et al., 2021; Saloner et al., 2018).

One way around this is to assess long-term forgetting. Accelerated long-term forgetting (ALF) refers to a pattern in which learning and short-delay recall are relatively preserved but information is lost at an unusually rapid rate over subsequent days or weeks (Elliott et al., 2014). The defining feature of ALF is therefore not necessarily an initial memory deficit, but a dissociation between the amount of information acquired, the amount retained over a conventional short delay, and the amount that remains available after a longer interval.

A one-week delay is particularly relevant to this question for both theoretical and psychometric reasons. Memory consolidation is thought to involve processes that continue over days and weeks (McClelland et al., 1995; Squire & Alvarez, 1995), whereas a 30-

minute delay captures only its earlier stages, and one-week testing reduces ceiling effects without the floor effects and attrition of longer intervals (Molinare et al., 2025; Pizzoni et al., 2026). One-week recall has been related to medial temporal structures, cortical measures, genetic risk and AD biomarkers (Bartels et al., 2025; Saloner et al., 2018; Weston et al., 2018; Zimmermann & Butler, 2018), but the literature is heterogeneous in material, learning criteria, retention intervals, administration and scoring (Elliott et al., 2014; Geurts et al., 2015).

Most studies have quantified long-term forgetting either as a continuous retention ratio or by applying a predefined cutoff to classify individuals as impaired or unimpaired. Neither characterizes the structure of an individual's learning-to-retention profile. In particular, a single ratio cannot readily distinguish generally weaker memory performance from a specific pattern of preserved learning and short-delay recall followed by disproportionate loss at one week. Similarly, a cutoff may identify the presence of lower retention without revealing whether individuals differ in the overall level of performance, in the shape of the trajectory, or in both. No consensus exists on how forgetting should be quantified when initial learning differs (Della Sala et al., 2024; Geurts et al., 2015).

These limitations provide the rationale for the present thesis. Its central objective was to identify distinct latent trajectories of learning and forgetting across the five FCSRT timepoints and to test their association with biological markers of early AD, focusing not on the level of recall at each timepoint but on the relationship between acquisition, short-delay retention and one-week recall.

The study sample comprised 322 cognitively unimpaired adults who completed three immediate learning trials and a 30-minute delayed recall in person, followed by an unannounced one-week recall by telephone, a mode that yields results comparable to in-person testing (Allen et al., 2019). Growth mixture modeling was used to identify groups of individuals with similar learning-retention trajectories. The resulting trajectory classes were subsequently examined in relation to fluid biomarkers of AD pathology, including cerebrospinal fluid (CSF) A β 42/A β 40, CSF p-tau181/A β 42, and plasma p-tau217, as well as biomarker-defined amyloid and tau status. Subjective cognitive complaints were additionally evaluated to explore whether objectively derived patterns of learning and forgetting were accompanied by perceived cognitive changes. Retrospective neuroimaging analyses examined structural brain measures implicated in long-term memory and early AD-related neurodegeneration, including total medial temporal lobe (MTL) volume and individual hippocampal subfield volumes, cortical thickness in regions relevant to AD and

long-term forgetting, and longitudinal rates of structural change estimated across three magnetic resonance imaging (MRI) visits. Particular emphasis was placed on the hippocampus, and especially the CA1 subfield, given their biological relevance to early AD-related neurodegeneration and their established involvement in long-term memory processes (Bartels et al., 2025).

Three trajectory classes were identified. Two of them differed mainly in the overall level of recall, whereas the third was defined by the relationship between its timepoints rather than by its level: learning and 30-minute recall were close to those of the highest-performing class, but recall after one week fell below that of the lowest, a dissociation that a 30-minute protocol cannot detect. Class membership was significantly associated with plasma p-tau217, with the CSF p-tau181/A β 42 ratio and with the rate of MTL and CA1 atrophy over the years preceding the memory assessment, whereas volumes measured at a single timepoint did not differ across classes.

The thesis is organized as follows. Chapter 1 will introduce Alzheimer's disease, focusing on its principal neuropathological alterations. Chapter 2 will present the biological staging frameworks together with the fluid, neuroimaging, and cognitive markers relevant to the preclinical phase. Chapter 3 will review long-term memory assessment and accelerated long-term forgetting across the AD continuum, with particular attention to the methodological limitations of current approaches. Chapter 4 will define the aims and hypotheses of the present study. Chapter 5 will describe the cohort, memory assessment protocol, biomarker measures, neuroimaging variables, and statistical procedures. Chapter 6 will present the results, which will then be interpreted and discussed in Chapter 7 in relation to the study hypotheses. Chapter 8 will conclude the thesis by summarizing its main contributions and identifying directions for future research.

1. Alzheimer's Disease

In 1906, the German psychiatrist and neuropathologist Alois Alzheimer presented the case of his patient Auguste Deter, who had developed progressive memory loss, disorientation, and language disturbance from the age of 51. Post-mortem examination of her cerebral cortex later revealed the plaques and neurofibrillary tangles that became the neuropathological hallmarks of the disease named after him (Alzheimer, 1907). More than a century later, the conceptualization of Alzheimer's disease has undergone a profound transformation. Once understood primarily as a clinicopathological entity whose definitive diagnosis required post-mortem confirmation, AD is now understood as a biological continuum that can be detected in vivo through biomarkers, even before overt cognitive impairment emerges. (Jack et al., 2018, 2024; Kobayashi et al., 2026; Van Der Molen et al., 2025). This shift reflects growing evidence that AD pathology begins to accumulate long before clinical symptoms emerge, during a prolonged preclinical phase that may extend for one to two decades or longer (Dubois et al., 2024; Kergoat et al., 2025; Koychev et al., 2020; Therriault et al., 2024).

In contemporary research and clinical trial design, biomarker-based frameworks have made the early, asymptomatic stages of AD a central focus, because they allow investigators to detect disease-related changes before objective cognitive impairment emerges and to enrich prevention and disease-modifying trials (Jack et al., 2018, 2024; Therriault et al., 2024; Thibault et al., 2026). This emphasis is further supported by evidence that preclinical AD is biologically heterogeneous, with variability in early tau deposition, neurodegeneration, and risk of progression across cognitively unimpaired individuals (Barcellos et al., 2025; Xie et al., 2025; Young et al., 2022; Zhou et al., 2025).

This chapter first outlines the epidemiology and contemporary definition of AD and then examines the major pathophysiological alterations that occur at the cellular, proteinaceous, and synaptic levels.

1.1 Definition and Epidemiology

Alzheimer's disease is a progressive, irreversible neurodegenerative disorder that is biologically defined by the presence of amyloid- β plaques and tau neurofibrillary tangles, with synaptic dysfunction and neuronal loss contributing to clinical decline (Jack et al., 2018; Knopman et al., 2021; Kobayashi et al., 2026; Therriault et al., 2024). In research

frameworks, AD is conceptualized as a biological continuum rather than a syndrome defined solely by symptoms, and biomarker abnormalities can be detected before cognitive impairment becomes clinically evident (Dubois et al., 2024; Jack et al., 2018, 2024; Van Der Molen et al., 2025). Dementia, by contrast, is the clinical stage in which cognitive impairment becomes severe enough to compromise independence and daily functioning, rather than the definition of AD itself (Dubois et al., 2016; Jack et al., 2024). The preclinical phase of AD often begins 10-20 years before symptom onset, and some models place the earliest neuropathological changes as far as 25 years before clinical manifestation (Kergoat et al., 2025; Koychev et al., 2020; Therriault et al., 2024).

Dementia is a major global public health challenge, and AD is its leading cause worldwide (Livingston et al., 2024; Rawat et al., 2022; Therriault et al., 2022). Recent global estimates indicate that 57.4 million people were living with dementia in 2019, and this number is projected to rise to 152.8 million by 2050. The projected increase is geographically uneven, with the largest relative growth expected in north Africa, the Middle East, and eastern sub-Saharan Africa, largely driven by population aging and population growth (Nichols et al., 2022). In parallel, the annual global societal cost of dementia was estimated at US\$1.313 trillion in 2019, with roughly half attributable to informal care (Wimo et al., 2023). These expenses include direct medical costs, formal social care by professional providers, and the immense contribution of informal caregiving. Family members and informal caregivers provide billions of hours of unpaid care, which imposes a severe, often unquantified psychological, physical, and financial toll on families. Despite the severity of these figures, up to three-quarters of individuals living with dementia globally have never received a formal diagnosis, primarily due to the clinical overlap with other neurodegenerative conditions, a lack of standardized cognitive screening in primary care, and the persistent misconception that cognitive decline is a normal part of biological aging (Gauthier, 2021). Underdiagnosis is not only a question of service provision. It also reflects the difficulty of separating early AD from other causes of cognitive decline, and of separating either from normal aging, using designed instruments (McKhann et al., 2011).

1.2 Neuropathological Alterations

The definitive neuropathological diagnosis of the disease relies on the post-mortem or in vivo biomarker-based verification of both extracellular senile plaques and intracellular neurofibrillary tangles, as established by consensus guidelines (Hyman et al., 2012). These

proteinopathies are not merely independent hallmarks; they are highly synergistic and initiate a complex pathological cascade that drives neuroinflammation, progressive synaptic disorganization, white matter degradation, and the loss of vulnerable neuronal populations, culminating in macroscopic cerebral atrophy (Busche & Hyman, 2020; Chen & Yu, 2023; Meftah & Gan, 2023; Pereira et al., 2021; Tzioras et al., 2023).

1.2.1 Amyloid- β Plaques

Amyloid plaques are extracellular deposits of A β , a peptide released by the sequential proteolysis of the amyloid precursor protein (APP) (S. Li & Selkoe, 2020; Perez-Nievas & Serrano-Pozo, 2018; Rosenberry et al., 2022; Serrano-Pozo et al., 2011). Cleavage by α -secretase falls within the A β sequence and precludes its formation; cleavage by β -secretase, followed by the γ -secretase complex, releases it (Jankowsky et al., 2004). Because γ -secretase cuts imprecisely within the membrane, peptides of differing length are produced. A β 40 predominates and remains largely soluble, whereas the longer A β 42, a minority species under physiological conditions, is considerably more hydrophobic and aggregates readily (Gutierrez & Limon, 2022; Kim & Bezprozvanny, 2021).

Aggregation proceeds through soluble oligomers and protofibrils to insoluble fibrils, which form the cores of mature plaques. The distinction between these species matters for the relationship between pathology and cognition. Total plaque burden correlates only weakly with the severity of cognitive impairment, whereas the soluble oligomeric forms are those associated with synaptic dysfunction (Im & Choi, 2024; Jankowsky et al., 2004; Kim & Bezprozvanny, 2021; Rosenberry et al., 2022).

Deposition follows a reproducible topographic sequence. Thal et al. (2002) described five phases, beginning in the neocortex, extending to allocortical structures including the entorhinal cortex and hippocampal formation, then to subcortical nuclei, brainstem and finally cerebellum. This pattern is broadly distinct from tau spread, which begins more medially and then extends outward, so the two lesion trajectories together provide the anatomical basis for later biomarker staging models (Serrano-Pozo et al., 2011; Tolar et al., 2024).

1.2.2 Tau Neurofibrillary Tangles

Tau is a microtubule-associated protein that is predominantly localized to axons in mature neurons, where it stabilizes microtubules and supports axonal transport (Barbier et al., 2019; Serrano-Pozo et al., 2011). Its interaction with microtubules is regulated by

phosphorylation, and in Alzheimer's disease tau becomes abnormally hyperphosphorylated, loses its normal microtubule-binding and assembly-promoting functions, and contributes to microtubule disruption (Alonso et al., 1994; Barbier et al., 2019; J. Wang et al., 2007; J.-Z. Wang et al., 2012). In this altered state, tau is the major protein component of paired helical filaments that form neurofibrillary tangles, while related tau-positive lesions also appear as neuropil threads within neuronal processes (Serrano-Pozo et al., 2011; J. Wang et al., 2007; J.-Z. Wang et al., 2012).

Unlike amyloid burden, tau pathology follows a relatively stereotyped anatomical progression and is more closely related to cognitive impairment, particularly memory dysfunction (Arriagada et al., 1992; Guillozet et al., 2003; Nelson et al., 2012). Braak and Braak (1991) described six stages of neurofibrillary change: stages I-II involve the transentorhinal region, stages III-IV extend into limbic regions including the hippocampal formation, and stages V-VI involve widespread neocortical association areas. Early involvement of the entorhinal cortex, CA1/subiculum, and amygdala has been emphasized in clinicopathological studies, although limited early tau pathology can also be found in cognitively unimpaired older individuals and is therefore not diagnostic in itself (Arriagada et al., 1992; Braak & Braak, 1991; Nelson et al., 2012).

The correspondence between Braak stage and cognition is close, although not absolute. In the Nun Study, Braak stage was significantly related to cognitive state and delayed memory performance near death (Riley et al., 2002). Other clinicopathological studies similarly found that tangle burden, and in some analyses CA1 neuronal loss, predicts cognitive status more strongly than amyloid load (Bennett et al., 2004; Giannakopoulos et al., 2003; Gómez-Isla et al., 1997). Taken together, these findings support the view that early medial temporal tau pathology provides an anatomical basis for the emergence of memory impairment before more widespread cognitive decline.

1.2.3 Synaptic Loss, Neurodegeneration, and Glial Responses

A compelling body of neurobiological evidence supports the characterization of AD as an early-stage synaptopathy, wherein progressive synaptic failure and loss constitute the neuroanatomical correlates most strongly associated with cognitive decline, preceding the physical loss of neuronal cell bodies (Serrano-Pozo et al., 2011). Soluble amyloid- β oligomers bind with high affinity to synaptic membranes, disrupting glutamatergic transmission by inducing the internalization of AMPA and NMDA receptors, impairing the induction of long-term potentiation, and shifting the synaptic balance toward long-term

depression (Selkoe, 2002). This synaptic dysfunction is exacerbated by neuroinflammatory responses mediated by resident microglia and reactive astrocytes, which cluster around amyloid- β plaques and release pro-inflammatory cytokines, chemokines, and reactive oxygen species (Serrano-Pozo et al., 2011).

In the early, preclinical stages of the disease, microglial responses serve a protective role, facilitating the clearance of amyloid- β and forming a physical barrier that isolates plaques from the surrounding neuropil (Condello et al., 2015). This microglial state is regulated by the triggering receptor expressed on myeloid cells 2 (TREM2), a surface sensor for lipids and apolipoproteins. Under pathological conditions, TREM2 is cleaved from the microglial membrane, releasing its soluble form (sTREM2) into the cerebrospinal fluid (Morenas-Rodríguez et al., 2022; Y. Wang et al., 2015; Yeh et al., 2016). Elevated levels of sTREM2 in early stages are associated with a younger neuroimaging-derived brain age, suggesting a beneficial microglial phenotype that promotes tissue resilience (Cumplido-Mayoral et al., 2025). In the same cohort, higher baseline CSF sTREM2 also predicts a more favorable rate of global cognitive change over three years of follow-up, independently of AD pathology (Warmenhoven et al., 2024). However, as the disease progresses, chronic exposure to fibrillar aggregates drives microglia into a dysfunctional, dystrophic phenotype characterized by a loss of phagocytic capacity and the aberrant pruning of healthy synapses through complement-mediated pathways (Streit et al., 2020).

Simultaneously, reactive astrogliosis occurs, which can be monitored in vivo using biofluid biomarkers such as glial fibrillary acidic protein (GFAP), chitinase-3-like protein 1 (YKL-40), and S100 calcium-binding protein B. These markers elevate in response to cortical amyloid- β accumulation and correlate with longitudinal cognitive decline (Craig-Schapiro, 2012).

2. The AD Continuum: Staging Frameworks

The contemporary neurobiological paradigm conceptualizes AD not as a sudden onset clinical dementia, but as a slow, progressive biological continuum spanning several decades (Jack et al., 2018; Sperling et al., 2011). In vivo biomarker studies have demonstrated that the core neuropathological hallmarks of the disease develop in a stereotypical chronological sequence long before clinical symptoms become manifest. Objective staging frameworks are essential to characterize this long preclinical phase, allowing for the classification of individuals based on biological measures rather than heterogeneous clinical phenotypes. This chapter reviews the transition from syndromic to biological diagnostic criteria, examines non-modifiable and modifiable risk factors, and evaluates the fluid, neuroimaging, and cognitive markers that define the preclinical continuum of the disease.

2.1 From Clinical to Biological Definition: A/T/N and the Preclinical Stage

For over three decades, the clinical diagnosis of the disease was guided by the 1984 NINCDS-ADRDA consensus criteria, which defined the condition strictly as a clinical-pathological syndrome requiring the presence of progressive dementia (McKhann et al., 2011). Under this historical framework, an in vivo diagnosis of probable disease was purely syndromic and relied on the clinical exclusion of other potential causes of cognitive decline, while a definitive diagnosis could only be established post-mortem via autopsy. This clinical-pathological conflation was highly problematic for research, as up to thirty percent of patients clinically diagnosed with dementia lacked the characteristic amyloid and tau pathology at autopsy, while many cognitively unimpaired older individuals exhibited extensive plaque and tangle pathology without demonstrating any objective clinical deficits.

To resolve these limitations, Jack et al. (2016) proposed the descriptive, unbiased A/T/N biomarker classification system. This framework groups established biomarkers into three distinct categories based on the specific neuropathological process they measure, independent of the patient's clinical presentation. The A category measures amyloid- β deposition using amyloid positron emission tomography (PET) imaging or reduced concentrations of the CSF A β 42/A β 40 ratio. The T category measures the accumulation of hyperphosphorylated tau neurofibrillary tangles using tau PET or elevated concentrations of

phosphorylated tau in the CSF or plasma. The N category measures downstream structural or metabolic neurodegeneration or neuronal injury using structural MRI, FDG-PET, or elevated biofluid concentrations of neurofilament light chain and total tau. Under this descriptive system, individuals are assigned a biomarker profile based on a binary (+/-) rating for each category (such as A+T-N-), allowing for the biological characterization of participants across different clinical cohorts.

This framework was updated by the Alzheimer's Association Workgroup, which published revised diagnostic and staging criteria that formalize the disease strictly as a biological process starting with biomarker evidence of neuropathological change (Jack et al., 2024). These revised criteria classify biomarkers into Core 1, Core 2, and Non-Core categories. Core 1 biomarkers represent highly accurate biofluid assays and imaging modalities that are interchangeable for the biological diagnosis, including amyloid PET, CSF A β 42/A β 40, CSF p-tau181/A β 42, and plasma p-tau217. A biological diagnosis is established by demonstrating abnormality on any Core 1 biomarker. Core 2 biomarkers measure AD-specific tau pathology but may have lower diagnostic accuracy in early stages or reflect later pathological events, including tau PET and CSF microtubule-binding region tau. Non-Core biomarkers include markers of neurodegeneration, such as structural MRI, FDG-PET, or NfL, and biomarkers of secondary pathways, such as astrocytic reactivity (CSF/plasma GFAP) or vascular brain injury measured via white matter hyperintensities.

The 2024 revised criteria also establish a biological severity scheme based on the sequence of biomarker changes, moving from Stage A (initial amyloid abnormality only) to Stage D (advanced, extensive neocortical tau PET uptake and severe downstream neurodegeneration). This biological staging is integrated with a seven-stage numeric clinical staging scheme (Stages 0 to 6) to form a dual-axis classification system.

Clinical Stage 0 represents asymptomatic individuals carrying deterministic genetic mutations with normal biomarkers. Clinical Stage 1 represents asymptomatic individuals who perform within expected limits on objective cognitive tests but show biomarker evidence of the disease. Clinical Stage 2 (Transitional Decline) represents individuals with normal performance on objective cognitive tests who report subjective cognitive decline (SCD) or exhibit subtle, transitional cognitive changes. Clinical Stage 3 represents individuals meeting clinical criteria for mild cognitive impairment (MCI), characterized by objective cognitive impairment but preserved independence in daily functioning. Clinical Stages 4, 5, and 6 represent mild, moderate, and severe dementia, respectively (Jack et al., 2024).

This biological framework builds on early preclinical criteria proposed by Sperling et al. (2011), which defined Preclinical Stage 1 as asymptomatic amyloidosis, Preclinical Stage 2 as amyloidosis plus neurodegeneration, and Preclinical Stage 3 as amyloidosis, neurodegeneration, and subtle cognitive decline. Jack et al. (2013) proposed a hypothetical, dynamic biomarker model where biomarkers become abnormal in a sequential, non-linear manner, with amyloid- β biomarkers reaching abnormality first, followed by tau hyperphosphorylation, neurodegeneration, and finally clinical symptoms. This continuous model of pathogenesis has been validated in longitudinal studies and underpins contemporary clinical trials targeting preclinical and early symptomatic stages (Therriault et al., 2022). Figure 1 sets out the resulting continuum, with the clinical stages of the revised criteria grouped into the preclinical, mild cognitive impairment and dementia phases.

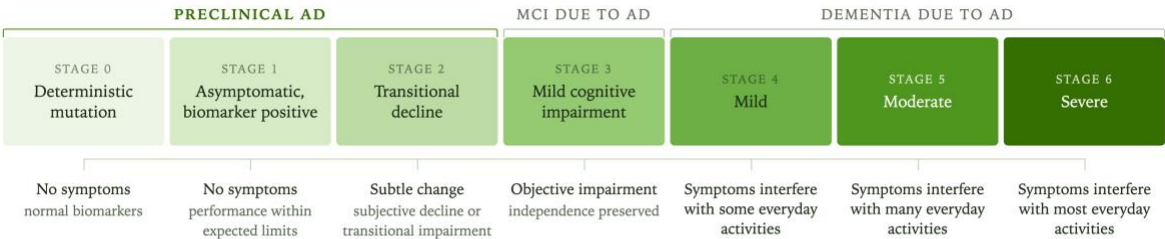


Figure 1. The Alzheimer’s disease continuum. Clinical stages 0 to 6 of the 2024 revised criteria, grouped into the preclinical, mild cognitive impairment and dementia phases. Stages follow Jack et al. (2024); the preclinical grouping follows Sperling et al. (2011).

2.2 Risk Factors

Sporadic late-onset Alzheimer’s disease arises from the interaction of non-modifiable factors, principally age and genotype, with modifiable factors acting across the lifespan. Heritability of the sporadic form has been estimated at between 58 and 79% in twin studies (Gatz et al., 2006). The factors described below are those relevant to the present analyses: age, sex and years of education are covariates in every model reported in Chapter 6, and the apolipoprotein E (*APOE*) $\epsilon 4$ genotype enters as a sensitivity analysis.

2.2.1 Age

Age is the strongest non-modifiable risk factor for sporadic Alzheimer's disease, and dementia prevalence increases sharply after age 65 (Livingston et al., 2024). Aging and Alzheimer's disease are nonetheless distinct processes. In cognitively unimpaired older adults, limited tau deposition in the MTL, particularly in the transentorhinal and entorhinal cortex, is common and can occur independently of amyloid- β , while age-related reductions in synaptic density and cortical thickness are also detectable before overt clinical impairment (LaPoint et al., 2017; Maass et al., 2018; Mecca et al., 2022; Schöll et al., 2016; Scott et al., 2020; Wuestefeld et al., 2023). Any marker intended to detect preclinical disease must therefore distinguish pathological change from the age-related neurobiological changes on which it is superimposed. The cohort examined here was recruited within a restricted age range, as described in Section 5.1, which limits the contribution of age without eliminating it; accordingly, age at assessment was included as a covariate in all models reported in Chapter 6.

2.2.2 Genetic Risk: *APOE*- ϵ 4 Genotype and Family History

The *APOE* gene on chromosome 19 is the principal genetic risk factor for sporadic late-onset Alzheimer's disease (Liu et al., 2013). Three common alleles, ϵ 2, ϵ 3 and ϵ 4, differ at two residues, and the substitutions alter the lipid-binding properties of the resulting isoforms (Huang & Mahley, 2014). The ϵ 4 allele raises risk in a dose-dependent manner: one copy is associated with a two- to threefold increase in lifetime risk and two copies with a substantially larger increase, together with an earlier age at onset (Reiman et al., 2020). The ϵ 2 allele is protective and ϵ 3 is treated as the reference.

The mechanism most directly implicated is isoform-dependent clearance of soluble A β from the brain parenchyma, which is least efficient for the ApoE4 isoform. Cognitively unimpaired ϵ 4 carriers accordingly show greater cortical amyloid deposition from midlife onward, and ϵ 4 carriage is also associated with white matter microstructural change (Peter et al., 2025). Genome-wide association studies have identified many further loci of individually small effect, implicating microglial lipid metabolism, immune response and endocytosis (Bellenguez et al., 2022).

In the cohort studied here, the genetic architecture has been characterized directly: polygenic risk and *APOE* genotype relate to the biomarker profile of participants who remain cognitively unimpaired (Vilor-Tejedor et al., 2024). Parental history of dementia was one of

the criteria by which the cohort was assembled (Molinuevo et al., 2016), so $\epsilon 4$ carriage is more frequent here than in the general population, and genotype is treated in Chapter 6 as a sensitivity analysis rather than as a covariate in the primary models.

2.2.3 Sex

Women account for a majority of people living with AD (Andrew & Tierney, 2018). Greater female longevity explains part of this excess, but not all of it (Fisher et al., 2018), and the interaction between sex and the measurement of memory has a specific bearing on the present study (Emrani & Sundermann, 2025). In a meta-analysis of the literature on cognitive sex differences, Hirnstein et al. (2023) found a consistent female advantage in verbal fluency and in verbal-episodic memory, present across the lifespan and robust across task formats. The advantage is small to moderate in size but highly reliable, and it has a direct diagnostic consequence, because the tests used to detect early AD are predominantly verbal.

Sundermann et al. (2016) interpreted this advantage as a form of sex-specific cognitive reserve: at a given level of underlying pathology, women perform better on verbal memory tests than men, and therefore reach the threshold for objective impairment later. The corollary, examined directly by Sundermann et al. (2016) is that a diagnostic system calibrated on verbal memory will underestimate the extent of AD pathology in women. Consistent with a later crossing of the diagnostic threshold, sex differences have also been reported in the rate at which story recall declines during the preclinical stage (Cowman et al., 2025).

The implication for a study on long-term forgetting is twofold. Sex must be adjusted for in every analysis of a verbal memory outcome, as it is throughout Chapter 6. And a measure that is sensitive within the normal range, rather than at the threshold of impairment, is precisely the kind of measure least affected by the masking effect that the female verbal advantage produces (Novozhilova et al., 2025).

2.2.4 Education and Cognitive Reserve

Cognitive reserve denotes the observation that the clinical expression of a given burden of neuropathology varies between individuals, and that the variation is predicted by educational attainment, occupational complexity and lifelong cognitive engagement (Livingston et al., 2024). The construct makes a specific prediction: at any given level of clinical severity, individuals with higher reserve will be found to carry more pathology, and

once symptoms emerge, they will decline more rapidly, because the pathological burden at onset is greater (Van Loenhoud et al., 2019).

Two consequences follow for the present work. The first is methodological. Educational attainment predicts performance on verbal memory measures independently of the underlying pathology (M. G. Jansen et al., 2021), and reference data for the test employed here indicate an association between educational level and performance on every index of the test (Brugulat-Serrat et al., 2021). Years of education is accordingly included as a covariate in all models reported in Chapter 6, although it should be regarded as a proxy for reserve rather than as a direct measure of it. The second consequence concerns the sensitivity of the assessment. On the reserve account, individuals with greater reserve are expected to sustain performance within normal limits at a given level of pathology, so that the point at which a conventional instrument detects a departure from those limits is displaced towards a more advanced stage of the disease process (Stern, 2012). In a cohort defined by the absence of objective impairment, this consideration applies to the sample as a whole, and it bears on the choice of memory measure discussed in Section 2.5.

2.2.5 Modifiable Risk Factors

The Lancet Commission on dementia prevention estimated that fourteen modifiable risk factors together account for approximately 45% of dementia risk worldwide, distributed across the lifespan: low educational attainment in early life; hearing loss, high LDL cholesterol, traumatic brain injury, hypertension, excessive alcohol intake and obesity in midlife; and smoking, depression, social isolation, physical inactivity, untreated vision loss, air pollution and diabetes in later life (Livingston et al., 2024). This figure is not the sum of independent contributions. The exposures co-occur within individuals, and analyses that weight population attributable fractions for their overlap yield a combined estimate smaller than the sum of the fractions computed separately (Stephan et al., 2024). An umbrella review of the same literature identifies a comparable set of factors, while noting that the evidence available for the transitional stages preceding dementia is considerably thinner than that available for dementia itself (Jones et al., 2025). Whether modification of these exposures alters the cognitive trajectory has also been examined experimentally: a two-year multidomain intervention combining diet, exercise, cognitive training and vascular risk monitoring produced a benefit on cognition in older adults at elevated risk (Ngandu et al., 2015), and Son et al. (2025) subsequently examined which combinations of factors yield the largest predicted benefit when targeted together.

The pathways through which these exposures act are heterogeneous, and two of them bear on the interpretation of cognitive measures. The first is cerebrovascular and metabolic. Blood pressure lowering has been associated with a reduced incidence of dementia and cognitive impairment in a meta-analysis of randomized trials (Hughes et al., 2020) and diabetes, depressive symptoms, frailty and multimorbidity have been associated with incident mild cognitive impairment in adults who were cognitively normal at baseline (He et al., 2024). The second is sensory. Adult-onset hearing loss (Yu et al., 2024) and untreated visual impairment (Kuźma et al., 2021) are each associated with incident cognitive impairment, and both additionally constrain the sensory access on which cognitive testing depends, so that their contribution to measured performance cannot be attributed to memory alone.

The third concerns sleep. Chronic insomnia in cognitively unimpaired older adults has been associated with subsequent cognitive outcomes, greater amyloid burden, and white matter alterations (Carvalho et al., 2025). Buongiorno et al. (2025) further linked abnormal nocturnal blood pressure patterns to impaired perivascular clearance, providing a potential connection between sleep-related vascular regulation and pathways involved in A β removal. Sleep is also closely involved in the consolidation and stabilization of newly acquired memories, suggesting that disturbances in sleep may influence cognition through mechanisms that extend beyond general AD risk (Zavec et al., 2023).

2.3 Fluid Biomarker Measures

Fluid biomarkers quantify the balance between production and clearance of the core Alzheimer's proteins, and they become abnormal earlier than any structural measure (Jia et al., 2024). Three were used in the present study, all measured at the second study visit: the CSF A β 42/A β 40 ratio, the CSF p-tau181/A β 42 ratio, and plasma p-tau217. All three are Core 1 biomarkers under the current criteria, meaning that abnormality on any one of them establishes the biological diagnosis (Jack et al., 2024).

2.3.1 CSF A β 42/A β 40 Ratio

As A β 42 aggregates into plaques within the parenchyma, less of it is cleared into the cerebrospinal fluid, and its concentration there falls. Concentration alone is an unreliable index, however, because total amyloid production varies widely between individuals. A β 40 is abundant and its concentration is largely unaffected by the disease, so expressing A β 42

relative to A β 40 normalizes for that variation and reclassifies individuals who would otherwise be misassigned on a single-analyte threshold. The ratio is consequently preferred to A β 42 alone (Jack et al., 2016, 2024).

That preference is supported by direct comparison against amyloid imaging. In a sample of 200 participants with and without cognitive impairment, Lewczuk et al. (2016) found the ratio to agree with the amyloid PET classification in 89.4% of cases, against 74.9% for A β 42 alone, and the discordant cases to be asymmetrically distributed: abnormality in CSF with a normal scan was more frequent than the reverse, as would be expected if the fluid measure changes before fibrillar deposition becomes detectable by imaging. What such abnormality signifies at the individual level is more circumscribed. Amyloid pathology is common in the absence of dementia: pooling 55 studies and 2,914 participants with normal cognition, W. J. Jansen et al. (2015) estimated its prevalence to rise from 10.4% at age 50 to 43.8% at age 90, and to be substantially higher among *APOE*- ϵ 4 carriers. A positive result therefore identifies membership of a population at elevated risk rather than an individual in whom decline has begun, which is the distinction the current criteria preserve by separating the biological from the clinical axis of staging (Jack et al., 2024).

2.3.2 CSF p-tau181/A β 42 Ratio

Phosphorylated tau in cerebrospinal fluid rises in response to amyloid pathology and before neurofibrillary tangles are detectable by imaging, so it indexes an active phosphorylation process rather than the accumulated tangle burden. Combining it with A β 42 in a single ratio incorporates both proteinopathies into one measure, which improves its correspondence with the neuropathological definition of the disease and its stability across assay platforms; the ratio is classified as a Core 1 biomarker on that basis (Jack et al., 2024)

The ratio has been characterized directly in unimpaired samples. Cacciaglia et al. (2025) examined it against a set of downstream markers in 234 cognitively unimpaired middle-aged adults and found it associated with cortical amyloid deposition and with regional gray matter volume, but not with glucose metabolism. The structural association was positive rather than negative, and it was weaker, in the regions that atrophy earliest, among participants positive for both amyloid and tau, so that the relation between the ratio and brain structure is not constant across the continuum. The choice of phosphorylation site nonetheless carries a limitation. Threonine 181 is not the site with the greatest specificity for Alzheimer's pathology: CSF p-tau217 discriminates the clinical stages over a wider dynamic

range and corresponds more closely to tau imaging (Janelidze et al., 2020). That is the reason the blood-based marker described in the next section is measured at threonine 217.

2.3.3 Plasma p-tau217

Blood-based markers have altered the practical reach of biomarker assessment, since they require neither lumbar puncture nor a scanner (Karikari et al., 2022). Among them, phosphorylated tau at threonine 217 has the strongest diagnostic performance, approaching that of cerebrospinal fluid and imaging measures (Ashton et al., 2024).

Milà-Alomà et al. (2022) characterized plasma p-tau217 in the cohort studied here, showing that it behaves as a state marker of amyloid- β pathology in the preclinical stage: it rises early, tracks cortical amyloid burden, and separates amyloid-positive from amyloid-negative cognitively unimpaired individuals. Its availability in a simple blood sample makes it the biomarker with the widest coverage in the present sample.

2.4 Neuroimaging Markers

Imaging measures index the structural and molecular consequences of AD with an anatomical specificity that fluid markers cannot provide (Khalili et al., 2026). The present study uses structural MRI in two forms, hippocampal subfield volumetry and cortical thickness, and does not use PET; the latter is nonetheless described below.

Each of the two measures is supported by a specific body of evidence. Volumetric studies of the hippocampal subfields converge on CA1 and the subiculum as the sectors reduced earliest and most consistently (De Flores et al., 2015), a pattern that holds when subfields are compared across successive stages of the continuum (Zhao et al., 2019) and when unimpaired participants are compared with those meeting criteria for MCI (Qu et al., 2023). For the cortical measures, entorhinal and parahippocampal thickness both decrease across the continuum, but only entorhinal thickness is reduced at the earliest stage examined (Hari et al., 2023); each is treated as a separate outcome in Chapter 6.

2.4.1 Structural Magnetic Resonance Imaging in Preclinical AD

Structural magnetic resonance imaging is the standard measure of the N category (Jack et al., 2016). Its limitation in the preclinical stage is that macroscopic atrophy is a late consequence: by the time lobar volume loss is detectable, substantial pathology has accumulated (Gordon et al., 2018; McDade et al., 2018). Two strategies address this. The

first is anatomical specificity, measuring the structures affected earliest rather than the brain as a whole. Killiany et al. (2002) showed that entorhinal cortex and hippocampal measurements predict progression in individuals who are not yet impaired, and Ahn et al. (2011) mapped the relation between regional cortical morphology and specific neuropsychological deficits across MCI and AD.

The second strategy is to summarize many regional measurements into a single index sensitive to deviation from the normal aging trajectory (Verdi et al., 2024). Machine learning applied to structural images yields an estimate of biological brain age whose difference from chronological age is associated with AD biomarkers and with neurodegeneration, and behaves differently in women and men (Cumplido-Mayoral et al., 2023). Beyond gray matter, white matter microstructural abnormality is detectable early and contributes to cognitive difference independently of gray matter measures (Peter et al., 2025).

Neither approach fully resolves a key challenge in the preclinical stage: cross-sectional brain volume reflects the cumulative effects of multiple lifelong influences, only one of which is AD (Xie et al., 2020). The analyses reported in Chapter 6 therefore treat the rate of change across repeated scans, rather than volume at a single timepoint, as the quantity of interest.

2.4.2 The Hippocampus and Its Subfields: Early Vulnerability of CA1

Hippocampal volume loss is the classical imaging marker of AD, and automated measurement of change across serial scans is well established (Leung et al., 2010). The hippocampus is not homogeneous, however. It comprises histologically distinct subfields, the cornu ammonis sectors CA1 to CA4, the dentate gyrus and the subiculum, which differ in cytoarchitecture, connectivity and vulnerability. Treating the structure as a single volume averages across that heterogeneity and dilutes the signal from the sectors affected first.

The relevant asymmetry is that tau pathology reaches CA1 and the subiculum at Braak stages III and IV, before CA3 and the dentate gyrus (Braak & Braak, 1991). Subfield segmentation methods applied to dedicated high-resolution acquisitions make this distinction accessible in vivo (Yushkevich et al., 2015), and the method used in the present study is described in Section 5.5.1.

Several findings connect this anatomy to delayed recall specifically. Bartels et al. (2025) reported that long-term retrieval performance was associated with CA1 volume in older adults and in individuals at risk of dementia. Ma et al. (2023) mapped the cortical and hippocampal network supporting long-term delayed recall and found performance

constrained by its structural and functional connectivity. Li et al. (2020) reported associations between regional gray matter morphology and objective memory performance in participants with SCD. The convergence of these results on CA1, and on retention over long rather than short delays, is the basis for the regional prediction tested in Chapter 6.

2.4.3 AD Cortical Signatures

Cortical thinning in AD concentrates in a distributed set of regions comprising medial and inferior temporal cortex, the temporal pole, the angular gyrus, the superior parietal lobule, the superior frontal gyrus and the supramarginal gyrus. Dickerson et al. (2009) showed that the average thickness of these regions, taken as a single measure, discriminates patients from controls, relates to symptom severity, and is reduced in cognitively unimpaired amyloid-positive individuals. The regions are defined on a standard cortical parcellation (Desikan et al., 2006) and extracted using automated surface reconstruction (Fischl, 2012).

A signature measure trades anatomical resolution for statistical power (Rusinek et al., 2026), which is the correct trade in a cognitively unimpaired sample where regional effects are small (Fletcher et al., 2021). Its reduction has also been related to subjective complaints in the absence of objective impairment, alongside gray matter volume and amyloid burden (Valech et al., 2019). In the present study the AD enters as an outcome, and the corresponding summary for the regions that thin in normal aging serves as its covariate, so that the estimate reflects thinning specific to the disease-related regions.

2.4.4 Amyloid- β PET and Tau-PET Imaging

Positron emission tomography visualizes the two core pathologies directly. Amyloid imaging became possible with the introduction of the Pittsburgh Compound-B tracer, which binds fibrillar A β and revealed cortical retention in patients and in a proportion of cognitively unimpaired older adults (Klunk et al., 2004). Longitudinal studies established that neocortical amyloid accumulates over roughly two decades and approaches a plateau early, so that amyloid burden distinguishes those on the continuum from those who are not, but discriminates poorly between stages once positivity is reached (Villemagne et al., 2013).

Tau tracers, by contrast, follow the anatomical progression described by Braak, and regional tau signal corresponds more closely than amyloid to cortical thinning, hypometabolism and cognitive performance (Therriault et al., 2023). This asymmetry underlies the position of tau-PET as a Core 2 biomarker in the current criteria (Jack et al., 2024).

2.5 Cognitive and Neuropsychological Markers of Preclinical AD

Cognition represents the ultimate clinical outcome against which AD is defined (Albert, 2011), with episodic memory typically being the earliest cognitive domain affected (Mortamais et al., 2017), in line with the anatomical progression described in Chapter 1. In the preclinical stage, however, the principal challenge is one of measurement rather than conceptual relevance (Jutten et al., 2021): instruments developed to identify clinically significant impairment are, by design, relatively insensitive to subtle variation within the cognitively normal range (Kueper et al., 2018). The present study therefore draws on two complementary sources of information: a self-report questionnaire, which registers change as the participant experiences it and so does not depend on measured performance at all (Chapman et al., 2021), and a performance-based memory test, whose sensitivity within this range depends on how it is administered (Rami et al., 2023).

2.5.1 Subjective Cognitive Decline Assessment: MyCog Questionnaire

Subjective cognitive decline denotes a persistent self-experienced worsening of cognitive capacity relative to a previous level, in an individual who performs within normal limits on objective testing. The conceptual framework was refined by Jessen et al. (2020), who set out the features that raise the probability of underlying AD, including onset after the age of 60, recent onset, memory as the domain of complaint, and concern about the decline.

The instrument used here is the self-report component, MyCog, of the Subjective Cognitive Decline Questionnaire, which pairs a participant questionnaire with an informant version and was validated in a Spanish-speaking sample (Rami et al., 2014). Scores on the questionnaire have been related to gray matter volume and to amyloid burden in cognitively unimpaired participants, which supports their interpretation as an early clinical signal rather than as a reflection of mood alone (Valech et al., 2019). The relation between subjective complaint and objectively measured long-term forgetting is examined in Chapter 6.

2.5.2 Memory Assessment: The Free and Cued Selective Reminding Test

The Free and Cued Selective Reminding Test was designed to separate the components of episodic memory that a conventional list-learning task confounds. Its distinguishing feature is controlled encoding: each item is presented with a semantic category cue and identified by the participant before learning proceeds, so that the material is known to have been encoded and the cue available at retrieval is the one used at encoding

(Buschke, 1984). Retrieval is then tested first without the cue and, for items not recalled, with it, yielding free recall, cued recall and total recall scores.

The design supports a specific inference. If a failure of free recall is corrected by the cue, the difficulty lies in retrieval, the pattern typical of normal aging and of disorders affecting frontal systems. If the cue does not restore performance, the information is not available to be retrieved, a pattern consistent with impaired storage or consolidation and commonly associated with AD-type memory impairment (Grober & Buschke, 1987). Cerami et al. (2017) evaluated delayed cued recall against a structured framework for biomarker development and concluded that it meets the criteria for a gateway measure in the early detection of AD.

The Spanish-language version has been extensively normed. Peña-Casanova et al. (2009) provided norms within the NEURONORMA project, Grau-Guinea et al. (2021) developed and established the equivalence of a parallel version B for repeated administration, and Bonete-López et al. (2020) validated a recognition task following the standard administration.

One constraint of the standard protocol is decisive for what follows. Retention is assessed after a delay of thirty minutes, an interval chosen to fit within a single session rather than for any property of memory. In a cognitively unimpaired sample, performance at that delay approaches the ceiling of the scale, so that the variance in which early decline would be visible is compressed (Lindbergh et al., 2021). Extending the interval to one week is the modification examined in the following chapter.

3. Long-Term Memory Assessment for Preclinical AD

Detection: Current Studies Review

This chapter brings together the existing literature on long-term delayed recall as a potential indicator of early Alzheimer's disease, the theoretical framework upon which the current study is grounded. It begins by defining accelerated long-term forgetting and elucidating the origins of the construct. Subsequently, it outlines the methodology employed to identify pertinent studies, synthesizes the evidence presented across populations at varying risk levels, and sets out the unresolved questions that persist about this topic.

3.1 Definition of Accelerated Long-Term Forgetting

Standard neuropsychological protocols typically assess retention after an interval of ten to thirty minutes. This choice is conventional rather than principled: it reflects the practical constraints of a single testing session rather than any property of the memory systems being evaluated. ALF denotes a pattern of memory dysfunction that this convention fails to capture. Individuals showing ALF encode new information normally and retain it over minutes but lose it at an abnormally rapid rate over the following days and weeks. The defining feature is therefore a dissociation rather than a deficit: learning and short-delay recall fall within normal limits, and the impairment becomes visible only when the retention interval exceeds the standard protocol. Interestingly, two interpretations of this dissociation coexist in the literature: one viewpoint treats it as qualitatively distinct from short-interval forgetting, while the other considers it a milder manifestation of the same underlying process. Distinguishing between these accounts therefore requires evidence capable of establishing whether the two forms of forgetting are genuinely dissociable. Mayes et al. (2019) proposed that this issue could be resolved by demonstrating a double dissociation between accelerated forgetting and hippocampal amnesia. Specifically, individuals with accelerated forgetting and those with hippocampal amnesia should show a cross-over pattern between initial and delayed recall relative to matched controls, rather than the single dissociation reported in most studies. To date, only two studies have directly tested this prediction, and both observed the expected cross-over pattern, a finding that is not readily accounted for by the quantitative explanation.

3.1.1 Origins of the Construct

ALF was first characterized in the context of epilepsy rather than Alzheimer's disease. Kapur et al. (1997) reported a patient with temporal lobe epilepsy who retained new information normally over hours and days but lost it at an abnormal rate thereafter, and interpreted the pattern as evidence for multiple-stage consolidation processes; they termed the phenomenon long-term amnesia. Blake et al. (2000) subsequently demonstrated the same dissociation at group level in temporal lobe epilepsy, arguing that acquisition and initial consolidation are preserved while an extended consolidation process, operating over days and weeks, is disrupted. The label "accelerated long-term forgetting" became established through this work and through subsequent studies of transient epileptic amnesia, in which the phenomenon co-occurs with remote memory loss in patients whose routine neuropsychological profile is unremarkable (Butler et al., 2010; Butler & Zeman, 2008). A subsequent methodological review of the epilepsy literature established both the reliability of the effect and the heterogeneity of the paradigms used to elicit it (Elliott et al., 2014).

Two features of this literature made the paradigm suitable for research on Alzheimer's disease. The first is the logic of the manipulation: extending the retention interval exposes a consolidation failure that a thirty-minute delay cannot detect, so a longer delay functions as a stress test of a process that standard protocols overlook. The second is that the effect appears in individuals whose conventional testing is normal, which perfectly aligns with preclinical stages of AD described in Chapter 2.

The transfer of the paradigm to Alzheimer's disease progressed from symptomatic to presymptomatic populations. Early studies compared patients with dementia or amnesic mild cognitive impairment with healthy controls (Degenszajn et al., 2001; Manes et al., 2008; Walsh et al., 2014). Weston et al. (2018) then demonstrated the effect in asymptomatic carriers of autosomal dominant mutations who were, on average, seven years from expected symptom onset. It is crucial to emphasize that what is borrowed from the literature on epilepsy is the logic of the paradigm, not a claim of shared pathophysiology: the mechanisms responsible for accelerated forgetting in temporal lobe epilepsy and in Alzheimer's disease do not necessarily share the same underlying causes, and no study has yet tested this hypothesis.

3.1.2 Rationale Behind the Use of One-Week Delayed Recall

Three factors influence the decision to use a one-week interval. The first is theoretical. Systems-level consolidation is believed to occur over days and weeks as the hippocampal dependence of a new memory trace gradually shifts to a distributed neocortical representation (McClelland et al., 1995; Squire & Alvarez, 1995). A thirty-minute delay only captures the initial phase of this process, while a one-week delay examines the timeframe during which slow consolidation is thought to take place. Reviewing evidence from both human and animal studies, Ohno (2026) argues that accelerated forgetting may arise from a disruption of systems-level consolidation rather than from impaired hippocampal encoding. More specifically, the dysfunction may involve the cortical regions to which memory traces are gradually transferred, particularly medial prefrontal areas such as the anterior cingulate cortex. In transgenic models, synaptic plasticity in these regions is impaired, together with local amyloid deposition, before comparable hippocampal alterations emerge. From this perspective, preserved initial encoding followed by disproportionate forgetting over longer delays would be an expected cognitive profile during the presymptomatic stage.

The second consideration is psychometric, and for a cognitively unimpaired cohort it is the stronger argument. At thirty minutes, healthy participants perform near ceiling, and ceiling effects compress variance precisely where early decline would appear; several studies reviewed below explicitly report this limitation (Degenszajn et al., 2001; Lindbergh et al., 2021). Saloner et al. (2018) found that recall at one week, but not at thirty minutes, was associated with MTL volumes in neurologically normal older adults, and Manoli et al. (2018) reported that extending assessment to one-week increased sensitivity to age-related differences. However, the trade-off operates in the opposite direction as well: intervals of a month or more introduce floor effects and attrition. For instance, Molinare et al. (2025) found associations with tau burden and cortical thickness at one week that were absent at six months in the same participants. Meta-analytic evidence converges on the same interval. Pooling sixteen studies across the Alzheimer's disease continuum, Pizzoni et al. (2026) found no significant group difference at twenty-four hours ($g = 0.55$, 95% CI [-0.03, 1.12]), a robust and homogeneous effect at one week ($g = 0.63$, 95% CI [0.40, 0.87], $I^2 = 0\%$) that was of comparable magnitude in at-risk groups considered separately ($g = 0.61$, 95% CI [0.37, 0.86]), and effects that persisted at one month or longer but weakened in those same at-risk groups. One week is therefore both the most practical interval and the one at which the effect is currently best documented.

The third consideration is feasibility. A one-week delay requires a second contact, and it is this practical constraint, rather than any theoretical objection, that has limited the implementation of the paradigm in large observational cohorts. A brief telephone re-administration of material previously learned in person can substantially reduce this additional burden. However, this change in testing context may raise concerns because it removes some environmental and nonverbal cues available during encoding. However, Allen et al. (2019) found no meaningful differences in one-week recall or retention between telephone and in-person testing, supporting the validity of remote assessment. This issue is nevertheless relevant, as van der Werf et al. (2016) identified differences in administration mode between groups as a potential source of bias. In the present study, all participants are re-tested using the same procedure, preventing administration mode from systematically confounding group comparisons. Furthermore, validation work on one-week batteries specifically designed for cognitively unimpaired samples has demonstrated that this approach is psychometrically viable (Rami et al., 2023). The protocol adopted in the present study is described in Section 5.3.2.

3.2 Search Strategy

A structured literature search was carried out in order to identify studies that had examined long-term delayed recall as a marker of early Alzheimer's disease pathology. The search and the selection process followed the reporting recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement (Page et al., 2021).

PubMed, Scopus and Web of Science were consulted, in order to obtain a comprehensive and comparative coverage of the literature. The following combination of key terms was used and adapted to the query syntax of each database: “accelerated long-term forgetting” OR “accelerated forgetting” OR “forgetting” OR “long-term forgetting” OR “delayed recall” OR “long delay” OR “list learning” OR “story recall” OR “retrieval performance” OR “memory consolidation” AND “Alzheimer” OR “Alzheimer disease” OR “mild cognitive impairment” OR “MCI” OR “memory complaints” OR (“presymptomatic” AND “cognitive decline”).

The same filters were applied in each database: publication date between 2001 and 2026, and exclusion of reviews, meta-analyses and animal studies. The last search was run on August 9, 2026 and was complemented by a manual screening of the reference lists of

the retrieved articles and of previously published reviews. Titles and abstracts of all records were screened by a single reviewer, and those considered relevant were assessed in full text.

Studies were eligible if they examined cognitively unimpaired individuals, or populations at varying levels of risk for Alzheimer's disease and related conditions. These included MCI and amnesic MCI, SCD and subjective memory decline, carriers and non-carriers of genetic mutations (*APOE-ε4*, *APP*, *PSEN1/2*), individuals with positive or negative amyloid status, and members of families affected by autosomal dominant AD. This deliberately broad definition of the target population was adopted because accelerated long-term forgetting may emerge differently depending on baseline cognitive status and underlying risk; the aim was to establish whether it behaves as a consistent marker of presymptomatic decline across heterogeneous groups.

Eligible paradigms had to include four elements: a study phase; a baseline assessment of learning of the study material, operationalized as immediate recall; at least one assessment of retention of the same material after an extended delay of no less than 24 hours, measured either by free recall or by recognition; and an adequate comparison group.

Two of these criteria are broader than those usually applied in this literature, and warrant comment. Recognition was admitted alongside free recall because several otherwise eligible paradigms report recognition as their primary or complementary retention outcome (Molinare et al., 2025; Ruggeri et al., 2024; Yang et al., 2021); restricting the criterion to free recall would have excluded the only large study of remote, digital multi-day testing in cognitively unimpaired older adults. The comparison group criterion was likewise defined broadly, as either a healthy control group matched for age and socio-educational level or a within-cohort contrast defined by genotype, amyloid status or severity of subjective decline. A strict matched case-control requirement would have excluded single-cohort correlational designs (Lindbergh et al., 2021; Lu et al., 2025; Saloner et al., 2018; Wearn et al., 2020) whose structure is closest to that of the present study.

The selection process is summarized in Figure 2. Twenty-one studies met the inclusion criteria; their design, sample, memory paradigm, retention intervals and main outcomes are reported in Table A1 (Appendix A). Reviews and meta-analyses are excluded from that count by the criteria above, but three prior syntheses cover overlapping ground and are referred to where relevant: Rodini et al. (2022) systematically reviewed nineteen studies of accelerated forgetting in neurodegenerative disease, García-Martínez et al. (2023) reviewed the Alzheimer's literature narratively, and Pizzoni et al. (2026) provide the only

meta-analysis, covering twenty-eight studies and 1,399 participants across the Alzheimer’s continuum.

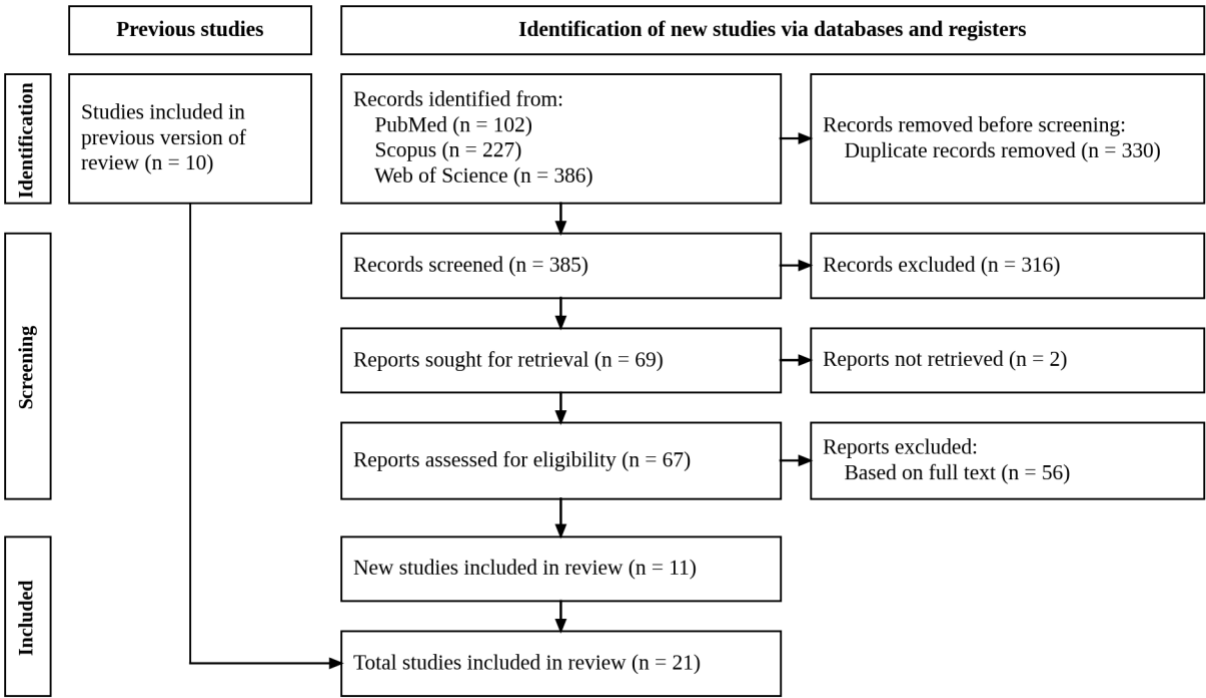


Figure 2. PRISMA 2020 flow diagram for the identification, screening and inclusion of studies.

One caveat applies to the evidence base as a whole. Several of the included studies draw on the same cohorts, and their samples are therefore partially overlapping: the autosomal dominant Alzheimer’s disease cohort of the Dementia Research Centre, University College London (Magill et al., 2026; O’Connor et al., 2020; Parker et al., 2025; Weston et al., 2018), the Harvard Aging Brain Study and affiliated Mass General Brigham cohorts (Molinare et al., 2025), and the COLBOS *PSEN1* E280A kindred (You et al., 2026). These studies report distinct outcomes, ranging from cross-sectional group differences to structural correlates and longitudinal prediction, and each contributes independent information. The number of independent participants is nonetheless lower than the sum of the reported sample sizes, and the four University College London reports in particular should be read as successive analyses of a single presymptomatic cohort rather than as independent replications.

3.3 Current Evidence

The studies identified are heterogeneous in population, material, retention interval and scoring, and a study-by-study analysis would obscure rather than elucidate the pattern. Consequently, this section is structured around the established literature rather than individual reports. Sections 3.3.1 to 3.3.3 follow the disease continuum, encompassing symptomatic Alzheimer's disease and mild cognitive impairment, populations defined by genetic risk, and cognitively unimpaired individuals characterized by biomarkers. Section 3.3.4 then examines the methodological variations across these studies and assesses their potential constraints on the conclusions drawn, with particular emphasis on the practice of scoring forgetting as a ratio.

3.3.1 ALF in Symptomatic AD and Mild Cognitive Impairment

Evidence from symptomatic populations represents the earliest line of research in this field. While it establishes the existence of the phenomenon, it only indirectly speaks to its potential for early detection.

Degenszajn et al. (2001) conducted a study comparing patients with Alzheimer's disease and healthy controls on the FCSRT and found large group differences at both thirty minutes and twenty-four hours, with perfect discrimination. The study is instructive for a reason other than the one initially intended: there was no difference in performance between thirty minutes and twenty-four hours, suggesting that the observed variations were not due to accelerated forgetting but rather an amnesic deficit. Controls recalled all sixteen items at every interval, and the twenty-four-hour delay was probably too short to expose a consolidation failure.

Manes et al. (2008) extended the interval to six weeks in individuals with memory complaints and amnesic MCI. Both groups performed normally at immediate and thirty-minute testing but were significantly impaired at six weeks, with group $\times$ time interactions for verbal ($F_{4,46} = 4.39, p = .005$) and visual material ($F_{4,46} = 2.77, p = .036$). This design isolates the effect of interest, but its primary limitation is the absence of intermediate timepoints.

Walsh et al. (2014) presented the most convincing demonstration in amnesic MCI. They employed a story recall task with a 90% learning criterion and observed a significant group $\times$ time interaction across a week ($F_{1,31,27} = 11.62, p = .001$), and critically, a steeper forgetting slope in the amnesic group ($F_{1,26} = 10.32, p = .003$) rather than merely a lower

overall score. Enforcing a learning criterion is what licenses the inference: acquisition was equated, so divergence over the retention interval can be attributed to consolidation.

Three recent studies push the finding below the threshold of objective impairment. Rodini et al. (2024) found a lower twenty-four-hour to thirty-minute saving index in severe SCD than in controls (84.2% vs. 94.3%, $p = .023$), discriminating severe from mild SCD with an area under the curve of 0.77. Ruggeri et al. (2024) found poorer one-week recall and recognition in participants with subjective memory decline despite intact short-delay performance ($F_{1,94} = 12.06$, $p = .001$). Bartels et al. (2025) reported that a long-term retrieval rate differentiated controls from SCD and MCI participants ($p = .007$) and that FCSRT-IR long-term retrieval was associated with hippocampal CA1 volume ($r = 0.42$, $p = .017$).

Stamate et al. (2020) warrant separate mention because they found no evidence of accelerated forgetting from immediate recall to one month. Their learning criterion was 70%, lower than in any other study reviewed, and repeated testing improved one-month recall substantially in both patients and controls. The result constrains the generality of the claim and raises the methodological question taken up in Section 3.3.4: without a stringent learning criterion, differences observed at a long delay may be confounded with differences in acquisition.

3.3.2 ALF in Individuals at Genetic Risk

Populations defined by genetic risk provide the strongest evidence that ALF precedes symptom onset, for the straightforward reason that carrier status identifies individuals who will develop the disease before any clinical change is detectable.

Weston et al. (2018) tested asymptomatic carriers of *PSEN1* and *APP* mutations who were on average seven years from expected symptom onset and who performed normally on learning, on thirty-minute recall, and on a full conventional battery. At one week, carriers had lost substantially more material than non-carriers across all three modalities: word list ($\Delta = -30.94$ percentage points, 95% CI [-45.16, -16.73], $p = .0002$), story ($\Delta = -20.10$, 95% CI [-33.28, -6.91], $p = .0048$) and figure ($\Delta = -15.41$, 95% CI [-26.88, -3.93], $p = .012$). The severity of forgetting correlated with proximity to estimated onset. This remains the anchor finding of the field.

O'Connor et al. (2020) situated the effect within a data-driven staging model of presymptomatic decline, showing intact thirty-minute recall alongside markedly poorer one-week retention (carriers 66.8%, IQR 59.8-81.9; non-carriers 93.1%, IQR 85.1-95.2; $p = .0004$) a decade before expected onset. Yang et al. (2021) replicated the dissociation in a

Chinese autosomal dominant cohort, with normal thirty-minute recall and impaired one-week retention for lists (53.3% vs. 66.7%, $p < .001$), figures (38.9% vs. 58.3%, $p = .002$) and recognition (70.0% vs. 80.0%, $p = .006$). Notably, forgetting was correlated with estimated years to onset ($r = -.69$, $p = .002$).

Two studies move beyond group comparison. Parker et al. (2025) examined the neuroanatomical correlates of the effect and found that verbal ALF was associated with prefrontal cortical thinning across four contiguous regions bilaterally, most associations surviving correction ($p_{FDR} = 0.00769$). Interestingly, there were no associations found with hippocampal, parahippocampal, or entorhinal measures, nor was there any correlation between thirty-minute recall and gray matter structure in any region. Magill et al. (2026) followed twenty asymptomatic carriers for a median of seven years. Nine became symptomatic, and baseline ALF distinguished progressors from non-progressors for list ($p = .03$) and story ($p < .001$) but not figure ($p = .21$), discriminating with areas under the curve of 0.73 and 0.82 respectively. Neither learning, nor thirty-minute recall, nor any of eight conventional cognitive tests separated the two groups. This is the only study in the present review with a clinical endpoint, and it is the strongest available evidence that the measure carries prognostic value, rather than just group-level information.

You et al. (2026) proposed a candidate mechanism in the Colombian *PSEN1* E280A kindred. Non-demented carriers showed a significant time $\times$ carrier status interaction ($\beta = -3$, 95% CI [-4.8, -1.2], $p = .001$), driven by the seven-day interval ($\beta = -1.9$, 95% CI [-3.2, -0.56], $p = .0008$), with learning ($p = .16$) and twenty-minute recall ($p = .19$) unimpaired; the effect persisted after excluding carriers with MCI. Reduced parietal sleep spindle power was associated both with forgetting and with precuneus tau burden ($\beta = -0.36$, 95% CI [-0.6, -0.1], $p = .008$), suggesting a route by which early tau accumulation might disrupt sleep-dependent consolidation.

Evidence from probabilistic rather than deterministic genetic risk speaks more directly to the sporadic form of the disease, in which the great majority of preclinical cases arise. Zimmermann and Butler (2018) found no genotype differences in encoding or thirty-minute recall in cognitively normal adults, but greater one-week forgetting in *APOE*- $\epsilon 4$ carriers, increasing linearly with allele dose ($p = .021$). Tort-Merino et al. (2021) found greater forgetting at three months in $\epsilon 4$ carriers ($F_{1,19} = 4.99$, $p = .037$), correlating with the CSF A β 42/p-tau ratio ($r = -.614$, $p = .003$).

3.3.3 ALF in Cognitively Unimpaired Individuals

This is the section most directly relevant to the present study, and the one in which the evidence is thinnest.

Saloner et al. (2018) tested 140 neurologically normal older adults and found that story recall at one week was associated with MTL volumes (partial $r = 0.24$, $p = .006$), specifically entorhinal ($r = 0.27$, $p = .001$) and hippocampal ($r = 0.21$, $p = .02$). No comparable association was found at thirty minutes, where performance was near ceiling. The effect sizes were small to medium, and the study was cross-sectional.

Wearn et al. (2020) followed 46 cognitively unimpaired older adults for one year and found that one-month word list recall predicted subsequent cognitive decline (Δ ACE-III, $R^2 = .123$, $p = .018$), classifying decliners with an area under the curve of .752 that rose to .918 when combined with hippocampal subfield volumes. No biomarker data were available, so the decline observed cannot be attributed to Alzheimer's pathology specifically.

Lindbergh et al. (2021) provide the strongest longitudinal evidence available in this population: 207 cognitively normal adults followed annually for up to eight years. Poorer one-week story recall predicted steeper decline in episodic memory ($b = 0.01$, $p = .003$), executive functioning ($b = 0.006$, $p = .011$) and MTL volume ($b = 0.004$, $p = .037$). The authors report ceiling effects at thirty minutes, which is itself part of the argument for the longer interval.

Tort-Merino et al. (2021) were the first to combine subjective complaints with fluid biomarkers. In 52 cognitively unimpaired participants stratified by SCD-Q score, the high-complaint group showed greater forgetting at three months ($F_{1,44} = 4.72$, $p = .035$); within that group, amyloid-positive participants performed worse than amyloid-negative participants at one-week free recall ($F_{1,24} = 4.49$, $p = .044$) and cued recall ($F_{1,24} = 7.10$, $p = .008$), and SCD-Q scores predicted higher forgetting rates ($\beta = -0.212$, $p < .05$). The biomarker subsample was 25 participants, which is the main limitation.

Lu et al. (2025) studied 429 cognitively normal members of a birth cohort, 306 with complete biomarker data. Amyloid-positive participants retained less of a visual figure at one week than amyloid-negative participants (90% vs. 95%, $p < .01$), and whole-brain atrophy interacted with amyloid status in predicting retention. The verbal measure, however, showed no association with amyloid status, and thirty-minute recall neither differentiated amyloid groups nor predicted atrophy. The study therefore supports the value of the extended interval while demonstrating that verbal and visual paradigms do not converge.

Molinare et al. (2025) provide the closest methodological precedent to the present work: 104 cognitively unimpaired older adults, of whom 16.3% were amyloid-positive, learned face-name pairs remotely on seven consecutive days and were tested for recognition at one week and again at six months. Better initial learning was associated with less forgetting ($\beta = 0.52$, 95% CI [0.34, 0.71], $p < .001$). One-week forgetting showed no association with global amyloid burden ($\beta = -0.07$, 95% CI [-0.23, 0.09], $p = .418$) but was associated with medial temporal tau ($\beta = -0.18$, 95% CI [-0.29, -0.07], $p = .001$), neocortical tau ($\beta = -0.16$, 95% CI [-0.31, -0.01], $p = .032$) and cortical thickness in AD-signature regions ($\beta = 0.33$, 95% CI [0.13, 0.52], $p = .001$). No association with any biomarker survived at six months. The tau findings require an important qualification supplied by the authors themselves: excluding the single participant with the highest tau value rendered both associations non-significant ($\beta = -0.06$, 95% CI [-0.20, 0.09], $p = .436$). The study is best read as establishing feasibility and generating hypotheses rather than as establishing a tau-forgetting relationship.

In summary, these studies suggest that extended-delay recall identifies variance that short-delay recall does not, and that this variance correlates with medial temporal structure, cortical thinning, and, more tentatively, tau. However, it remains unclear whether this correlation holds in a large, biomarker-characterized, longitudinally followed cohort of cognitively unimpaired individuals, utilizing a protocol that could be realistically deployed at scale.

3.3.4 Methodological Heterogeneity and the Limits of Ratio-Based Scoring

Five methodological choices vary across the studies reviewed, and each limits comparability. Eleven of the twenty-one studies did not predefine a learning criterion, while the remainder range from 70% (Stamate et al., 2020) to 90% (Lindbergh et al., 2021; Saloner et al., 2018; Walsh et al., 2014); without equated acquisition, differences at the long delay are confounded with differences in learning. Retention intervals span twenty-four hours to six months, and results are not commensurable across that range, as Molinare et al. (2025) demonstrated directly by finding at one week what was absent at six months in the same sample. Verbal and visual paradigms do not converge: Parker et al. (2025) found structural correlates for verbal ALF only, and Lu et al. (2025) found an amyloid association for the visual but not the verbal measure.

The fourth choice concerns scoring. Most studies express forgetting as the proportion of material retained at the long delay relative to short-delay recall. This ratio has three weaknesses. It is unstable when the denominator is small, so participants with poorer short-delay recall generate inflated or uninterpretable values; it conflates forgetting with acquisition wherever a learning criterion is not enforced. A third limitation concerns a more fundamental issue. Expressing retention as a proportion of short-delay recall presupposes that the rate of forgetting depends on how much was learned initially, so that dividing by the short-delay score removes a genuine source of variance. That premise is contested. Reviewing the evidence, Della Sala et al. (2024) concluded that when forgetting is measured in absolute terms, and ceiling and floor effects are avoided, the forgetting curves of high and low learners run in parallel, so that the rate of loss is largely independent of the initial level. On that reading the ratio adjusts for something that does not require adjustment while introducing a scaling artifact, and the artifact is largest exactly where the measure is most often applied, since proportional change is, in those authors' account, particularly problematic when the groups compared differ in baseline performance. Geurts et al. (2015) reach the same impasse from the opposite direction, noting that several methods exist for adjusting long-delay performance for unequal initial learning and that no consensus has emerged on whether or how the adjustment should be made. These limitations motivate the two complementary operationalizations adopted in the present study, described in Section 5.6.

A fifth choice is rarely made explicit: whether the same material is probed at every retention interval. Repeated retrieval is itself a learning event, and Della Sala et al. (2024) argue that a group difference in how much repeated testing strengthens the trace would be sufficient to produce an apparent difference in forgetting rate; they raise the possibility for Weston et al. (2018) specifically. Probing a different subset of the material at each interval removes the confound. Among the studies reviewed here, Stamate et al. (2020) is the one that did so, and also the one that found no accelerated forgetting, although its learning criterion differed from that of every other study as well, so the two explanations cannot be separated.

Finally, the anatomical findings do not concur. Bartels et al. (2025), Saloner et al. (2018), Wearn et al. (2020) and Lindbergh et al. (2021) report medial temporal associations, whereas Parker et al. (2025) found prefrontal associations and none in the MTL. Whether slow consolidation depends on medial temporal or on neocortical integrity remains open.

3.4 Gap of Knowledge

The evidence reviewed above supports three main conclusions. ALF is detectable in individuals whose learning and short-delay recall are normal, across populations ranging from symptomatic Alzheimer's disease to asymptomatic mutation carriers and cognitively unimpaired older adults. It relates to AD structural and molecular markers, including medial temporal volume, cortical thinning in AD-signature regions and regional tau burden. In the one cohort followed to a clinical endpoint, it predicted which asymptomatic carriers would develop symptoms over the following seven years, where eight conventional cognitive tests did not (Magill et al., 2026).

Several limitations necessitate caution in interpreting these conclusions. The sample sizes are relatively small: excluding the two birth-cohort studies, the median sample size across the reviewed studies is well below one hundred, and the presymptomatic evidence rests substantially on a single cohort analyzed four times. The designs are predominantly cross-sectional, which limits the ability to establish a direct temporal relationship between forgetting and pathology. The evidence is weighted towards autosomal dominant disease, in which the timing of pathology is known but the generalizability to sporadic Alzheimer's disease is assumed rather than demonstrated. Moreover, there is no consensus on how forgetting should be quantified, with the ratio scoring used by most studies carrying the weaknesses described in the previous section.

Two further gaps are specific to the preclinical stage. Few studies combine extended-delay recall with a full biomarker characterization in cognitively unimpaired participants; those that do are either small (Tort-Merino, Valech, et al., 2021) or have low proportions of amyloid-positive participants (Molinare et al., 2025). Additionally, almost none of the studies relate forgetting to longitudinal imaging outcomes. The association between forgetting and brain structure is typically established at a single timepoint, making it unclear whether individuals who forget more rapidly also exhibit more rapid atrophy. Commenting on Weston et al. (2018), Reiman (2018) set out what would be needed next: confirmation that the effect generalizes from autosomal dominant to late-onset preclinical disease, clarification of how forgetting relates to biomarker measures of amyloid, tau and neurodegeneration, and protocols that minimize the rehearsal and learning effects introduced by repeated assessment. The first two remain substantially open eight years later.

It is crucial to clearly delineate what would and would not be novel. Remote administration is not inherently new: Ruggeri et al. (2024) describe ecological remote

administration, Molinare et al. (2025) provide a fully remote and digital approach, and Magill et al. (2026) administered the seven-day test remotely. Nor is the one-week interval new, having been used in the majority of the studies reviewed. No previous study has combined all of the following: a one-week delayed recall of material learned during a standard in-person visit, administered via a brief telephone call; a cognitively unimpaired cohort characterized by both fluid and imaging biomarkers; a sample of several hundred participants; retrospective longitudinal follow-up enabling the estimation of atrophy rates; and a data-driven modeling of individual learning-retention trajectories rather than a single ratio threshold applied to the entire sample.

Every study reviewed treats forgetting as a continuous quantity to be correlated, or applies a cut-off to define impaired and unimpaired groups. Neither approach can identify a subgroup whose profile is characterized by normal learning, normal thirty-minute recall and disproportionate loss at one week, because such a profile is defined by the shape of an individual trajectory and not by the value of any single score. Whether such a subgroup exists in a cognitively unimpaired population, and whether it differs from the remainder of the cohort on biomarkers of Alzheimer's disease, is the question the present study addresses. Specific aims and hypotheses will be set out throughout the next chapter.

4. The Present Study: Aims and Hypotheses

4.1 From the Research Gap to the Present Design

The previous chapter concluded with a question that the existing literature cannot answer: whether a subgroup characterized by normal acquisition and normal short-delay recall but disproportionate loss at one week exists in a cognitively unimpaired population, and whether it differs from the remainder of the cohort on markers of AD pathology. The present chapter sets out how that question was operationalized¹.

The FCSRT protocol administered at the third visit of the ALFA+ study yields five points on a single learning-retention curve: three immediate learning trials, a thirty-minute delayed recall, and an unannounced one-week delayed recall administered by telephone. A retention score, regardless of its scaling, is a single number and locates a participant along a single dimension. The five timepoints locate the same participant along three: the amount of material acquired, the amount that survived thirty minutes, and the amount that survived a week. It is the relation among these three quantities, rather than the value of any one of them, that defines the profile of interest. Latent trajectory modeling estimates this relationship at the individual level and groups participants by the resulting shape, enabling the identification of a subgroup that is unremarkable at every single timepoint yet can be distinguished from the rest of the sample. This is the property that the present analyses exploit, and the reason the study models the protocol as a whole rather than its endpoints.

4.2 Aims

The overall aim was to identify distinct latent trajectories of learning and forgetting across immediate, thirty-minute and one-week recall in cognitively unimpaired adults from the ALFA+ cohort, and to determine whether these trajectories were associated with biological markers of early AD. Given that the sample was cognitively unimpaired by design, the focus is the preclinical stage of AD. Four primary objectives were pursued.

¹ Part of the work reported in this thesis was presented as a poster at the Alzheimer's Association International Conference (AAIC) 2026, on the earlier data release of the same cohort ($N = 313$): <https://epostersonline.com/aaic2026/node/39946>

1. **Trajectory identification.** To identify latent classes of memory performance across the five FCSRT timepoints using growth mixture modeling. The number of classes was not specified a priori; it was determined by model fit and a prespecified minimum class-size rule.

2. **Trajectory characterization.** To describe the distribution of one-week recall in relation to thirty-minute recall, and to establish whether the classes differ in acquisition, in thirty-minute retention and in one-week forgetting, and in particular whether they differ in the shape of the trajectory and not only in its level.

3. **Fluid biomarker associations.** To examine whether class membership is associated with CSF A β 42/A β 40, CSF p-tau181/A β 42, plasma p-tau217 and biomarker-defined amyloid status and tau positivity among amyloid-positive participants, using general linear and logistic models adjusted for age, sex and years of education.

4. **Neuroimaging associations.** To test whether the classes differ in brain structure at the time of assessment and in the rate of structural change estimated retrospectively across three MRI visits, using linear models and linear mixed-effects models with a class-by-time interaction, respectively. The same regions were included in both analyses: the hippocampus and its subfields, entorhinal and parahippocampal cortex, the AD cortical signature, four prefrontal regions in which associations with forgetting have also been reported, and the lateral occipital cortex as a control region where no such associations are anticipated.

Prior to conducting the analyses, the primary objectives and the corresponding models were specified before the analyses were run. A secondary aim was introduced subsequently. It concerned convergence, specifically determining whether the conventional continuous ALF ratio identifies the same associations as the latent classes. It is reported as secondary, and the distinction is maintained in the presentation of the results.

4.3 Hypotheses

Six primary hypotheses were tested.

(H1) One-week recall would exhibit greater dispersion and a reduced ceiling effect compared with thirty-minute recall, thereby capturing variance that the standard protocol does not (Manoli et al., 2018; Saloner et al., 2018).

(H2) Trajectory modeling would distinguish at least one class whose acquisition and thirty-minute recall lay within the range of the more proficient participants, but whose one-week recall was disproportionately low; the hypothesis concerns the existence of such a

dissociated profile, not the number of classes, and would not be supported by classes ordered identically at all five timepoints and therefore differing solely in level (O'Connor et al., 2020; Weston et al., 2018).

(H3) Trajectories characterized by disproportionate loss at one week, whether or not they were also lower performing at acquisition and at thirty minutes, would show higher plasma p-tau217 concentrations and a less favorable CSF profile, namely lower A β 42/A β 40 and higher p-tau181/A β 42, than the best-retaining trajectory (Lu et al., 2025; Tort-Merino, Valech, et al., 2021; Weston et al., 2018).

(H4) The same trajectories would show smaller MTL volume and thinner entorhinal and parahippocampal cortex, and no difference in the lateral occipital control region (Lindbergh et al., 2021; Saloner et al., 2018).

(H5) Within the hippocampus, the association would be more pronounced in CA1 compared with the remaining subfields (Bartels et al., 2025).

(H6) The same trajectories would show a faster rate of MTL volume loss and of entorhinal and parahippocampal cortex thinning across the three MRI visits, with no comparable acceleration in the lateral occipital control region (Lindbergh et al., 2021; Wearn et al., 2020).

Three further questions were examined without a directional prediction of comparable confidence and are reported as exploratory throughout. The first concerns the prefrontal regions included in the imaging outcome set. Parker et al. (2025) reported associations between ALF and prefrontal thinning, with no corresponding effect in the hippocampus or adjacent cortex. These regions were therefore retained so that a null hippocampal finding would adjudicate between the medial temporal and the neocortical account of slow consolidation, rather than leaving the question open. The second concerns subjective cognitive complaints and is exploratory because the studies linking complaints to forgetting selected participants on that basis, whereas here complaints are an outcome in an unselected cohort (Rodini et al., 2024; Ruggeri et al., 2024). Convergence between the classes and the ratio was expected, but no prediction was made as to which would prove more sensitive.

4.4 Expected Contribution

The study aims to examine associations rather than diagnostic utility. It does not assess the prediction of clinical progression, and a lower-performing trajectory class is not

equivalent to a diagnosis of Alzheimer's disease. Its expected contribution is methodological as well as substantive. The study investigates whether a five-minute telephone call appended to an existing protocol can index long-term consolidation in several hundred cognitively unimpaired adults characterized by both fluid and imaging biomarkers. Additionally, it explores whether modeling the shape of the learning-retention curve can recover information that a single retention score cannot. The following chapter provides details about the participants, the memory assessment, the biomarker measures, the MRI acquisition and processing, and the statistical procedures employed to test these aims and hypotheses.

5. Methods

5.1 Study Population: The ALFA+ Cohort

The present study was conducted within ALFA+, the nested longitudinal cohort of the ALFA (ALzheimer and FAMilies) study, a research platform established in 2013 at the Barcelonaβeta Brain Research Center (Barcelona, Spain) to characterize the pathophysiology of the preclinical stage of Alzheimer's disease (Minguillón et al., 2020; Molinuevo et al., 2016). The ALFA parent cohort comprises 2,743 cognitively unimpaired individuals aged 45 to 74 years at recruitment, largely adult offspring of patients with AD, and is enriched for family history of the disease and for *APOE*-ε4 carriership (Molinuevo et al., 2016; Vilor-Tejedor et al., 2024). Its baseline visit included a full cognitive evaluation, clinical history, lifestyle questionnaires, blood sampling for DNA extraction, and magnetic resonance imaging (Molinuevo et al., 2016).

ALFA+ was designed as the deep-phenotyping arm of this platform. A total of 450 participants were invited on the basis of an individual Alzheimer's disease risk profile derived from an algorithm combining parental history of AD, *APOE* genotype, verbal episodic memory score, and the CAIDE (Cardiovascular Risk Factors, Aging and Dementia) score (Milà-Alomà et al., 2022). Inclusion further required previous participation in ALFA, an age between 45 and 65 years at the time of inclusion in ALFA, and a long-term commitment to the study, including agreement to undergo MRI, PET, and lumbar puncture (Milà-Alomà et al., 2022). Each ALFA+ visit was organized into consecutive sessions completed within approximately three to six months: a first session (S1) comprising the clinical interview, blood sampling, cognitive testing and MRI, followed by a lumbar puncture session, an extended neuropsychological and imaging session, and amyloid and FDG PET sessions; follow-up visits took place approximately every three years (Minguillón et al., 2020). All participants provided written informed consent, and both the ALFA and the ALFA+ study protocols were approved by the Independent Ethics Committee "Parc de Salut Mar", Barcelona².

² ALFA: ClinicalTrials.gov NCT01835717. ALFA+: protocol ALFA-FPM-0311, ClinicalTrials.gov NCT02485730. Both studies were conducted in accordance with the Declaration of Helsinki (World Medical Association, 2013) and with Spanish Law 14/(2007), of 3 July, on Biomedical Research.

Two features of this design bear on the interpretation of what follows. The cohort is enriched for familial risk and *APOE*- ϵ 4 carriership and is therefore not representative of the general population. Selection also drew in part on verbal episodic memory performance, which restricts variance in the cognitive domain examined here.

The analyses reported here draw on all three ALFA+ visits, but not on the same measures at each. The FCSRT protocol incorporating the one-week telephone recall was introduced at the third visit (V3) and is therefore available at that visit only. Fluid biomarkers were obtained at the second visit (V2). Structural MRI was acquired at all three visits, with mean intervals of 3.21 years between V1 and V2 and 3.35 years between V2 and V3, allowing rates of atrophy to be estimated retrospectively over a mean interval of 6.6 years ending at V3. Each of these measures is described in the sections that follow.

5.2 Sample Selection and Description

The measures described above were not obtained in every member of the cohort, and the present work therefore rests on a subset of ALFA+. Of the 450 participants enrolled in ALFA+, 341 had completed the third study visit at the time of the present analysis. The one-week delayed recall by telephone formed part of the third-visit protocol for 336 of them; the remaining five were not scheduled for the call. Of those 336, 322 completed it, and these 322 participants constitute the analytic sample of this thesis. The design of the study and the sequence of the analyses are summarized in Figure 3.

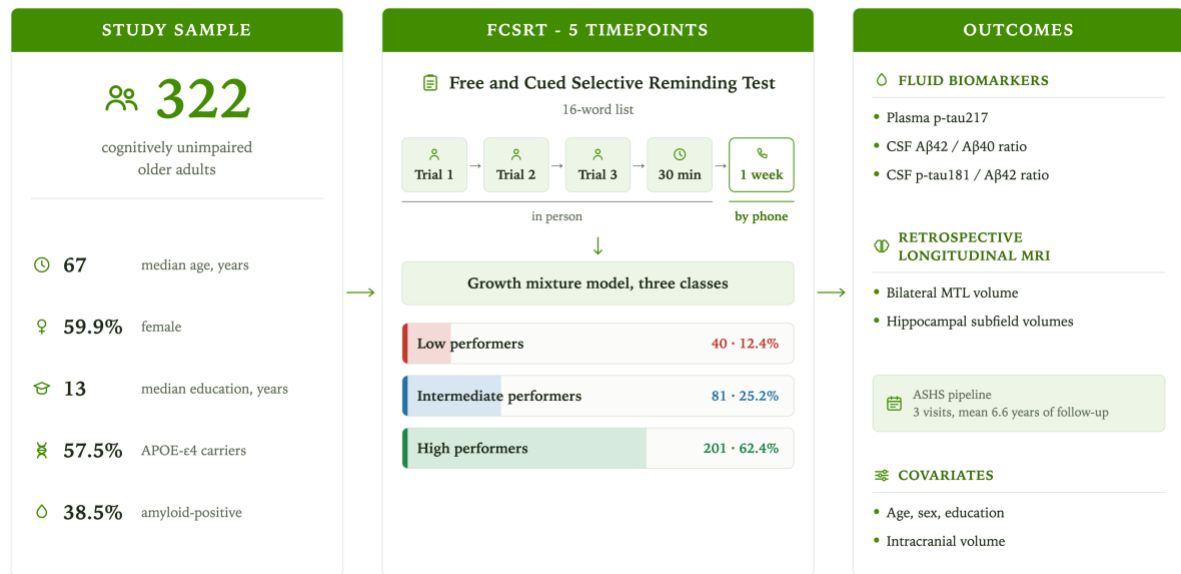


Figure 3. Study design and analysis pipeline.

Inclusion rested on a single requirement: that the participant had completed the one-week delayed recall by telephone. No further criterion was applied. Because the telephone call was the last component of the FCSRT protocol to be delivered, everyone who completed it had already completed the three immediate learning trials and the thirty-minute delayed recall during the preceding face-to-face session. All five FCSRT timepoints are therefore observed in every participant, and no one contributes a partial trajectory to the models described in Section 5.6.

Fourteen of the 336 participants for whom the call was scheduled did not undergo it. In five, the FCSRT had been administered less than a year earlier at a tau-PET or intermediate study visit, and repeating it would have confounded long-term retention with recent exposure to the same word list. In a further five, the seven-to-ten-day window could not be met because it fell within a holiday period. One participant could not be reached after repeated attempts, one call was omitted through a scheduling oversight, and one participant declined. The fourteenth was excluded on clinical grounds, after amnesic impairment was observed during the face-to-face session itself.

Three restrictions that are common in this literature were deliberately not imposed. No learning criterion was applied, so participants were not required to reach a threshold of acquisition before their retention intervals were considered. No FCSRT score was removed as an outlier. And completeness of biomarker or imaging data was not required: each participant contributes to those models for which the relevant measurement exists. The first

of these choices departs from a substantial part of the work reviewed in Section 3.3.4, in which participants who fail to acquire the list are removed before forgetting is estimated. Retaining them preserves the lower end of the learning distribution, which is precisely where the poorest-performing trajectory is defined.

Coverage of the biological measures is consequently incomplete and unequal across outcomes, and the effective sample size differs from model to model. Plasma p-tau217 was available in 309 participants (96.0%), CSF A β 42/A β 40 in 286 (88.8%) and CSF p-tau181/A β 42 in 240 (74.5%); biomarker-defined amyloid status was available in the same 286 (88.8%) and A/T status in 234 (72.7%), of whom the 95 amyloid-positive participants carry the tau contrast. Subjective cognitive complaints were recorded in 321 participants (99.7%). Hippocampal subfield volumes at the third visit were available in 314 (97.5%) and the cortical AD signature in 312 (96.9%). For the longitudinal imaging analyses, 306 participants (95.0%) contributed all three MRI visits, ten (3.1%) contributed two and six (1.9%) contributed one. The characteristics of the sample, overall and by trajectory class, are given in Table 1.

One feature of the sample bears on the interpretation of everything that follows. Participants enter only if the telephone recall was administered, and the reasons for non-administration were almost entirely logistical. One was not: the single exclusion made on clinical grounds removed the participant in whom memory impairment was most evident, so that the sample is depleted, if only marginally, of precisely the profile the study sets out to detect. Together with the restriction of variance in verbal episodic memory that follows from the ALFA+ selection algorithm described in Section 5.1, this makes the present sample a conservative one: an association between long-term forgetting and preclinical AD pathology observed here is more likely to be attenuated than inflated.

Table 1. Cohort characteristics stratified by latent trajectory class of one-week delayed FCSRT total recall. Continuous variables are given as median [interquartile range] unless stated otherwise.

	Available n	Overall	Low performers(<i>n</i> = 40, 12.4%)	Intermediate performers(<i>n</i> = 81, 25.2%)	High performers(<i>n</i> = 201, 62.4%)
Age, median [IQR]		67.1 [64.0-71.0]	70.7 [67.6-72.9]	69.2 [65.5-71.7]	66.1 [63.1-69.7]
Female sex, <i>n</i> (%)		193 (59.9%)	23 (57.5%)	41 (50.6%)	129 (64.2%)
Education, years, median [IQR]		13.0 [11.0-17.0]	12.0 [11.0-17.0]	14.0 [11.0-17.0]	15.0 [11.0-17.0]
<i>APOE</i> -ε4 carriers, <i>n</i> (%)		185 (57.5%)	20 (50.0%)	43 (53.1%)	122 (60.7%)
Aβ-positive ¹ , <i>n</i> (%)	286	110 (38.5%)	11 (33.3%)	31 (42.5%)	68 (37.8%)
A-T-, <i>n</i> (%)	234	139 (59.4%)	16 (61.5%)	34 (53.1%)	89 (61.8%)
A+T-, <i>n</i> (%)	234	67 (28.6%)	4 (15.4%)	22 (34.4%)	41 (28.5%)
A+T+, <i>n</i> (%)	234	28 (12.0%)	6 (23.1%)	8 (12.5%)	14 (9.7%)
MyCog questionnaire score, median [IQR]	321	4.0 [1.0-8.0]	5.0 [2.0-8.2]	4.0 [2.0-9.0]	3.0 [1.0-7.2]
MyCog questionnaire score (memory subscale), median [IQR]	321	1.0 [0.0-3.0]	1.5 [0.0-4.0]	1.0 [0.0-4.0]	1.0 [0.0-3.0]
Plasma p-tau217 ² (pg/mL), median [IQR]	309	0.210 [0.167-0.292]	0.249 [0.189-0.465]	0.213 [0.173-0.276]	0.206 [0.162-0.287]
CSF Aβ42/Aβ40 ² ratio, median [IQR]	286	0.079 [0.060-0.092]	0.079 [0.046-0.086]	0.077 [0.057-0.090]	0.080 [0.061-0.094]
CSF p-tau181/Aβ42 ² ratio, median [IQR]	240	0.0116 [0.0092-0.0172]	0.0123 [0.0103-0.0322]	0.0129 [0.0099-0.0209]	0.0105 [0.0090-0.0159]
Total MTL volume at third visit, mm ³ , mean (<i>SD</i>)	314	12,583 (1,255)	12,132 (1,470)	12,649 (1,217)	12,638 (1,218)
CA1 volume at third visit, mm ³ , mean (<i>SD</i>)	314	2,702 (326)	2,603 (390)	2,711 (288)	2,716 (326)
Immediate Total Recall (sum of 3 trials, 0-48), median [IQR]		46.0 [43.0-47.0]	38.0 [36.0-40.0]	46.0 [43.0-47.0]	46.0 [45.0-47.0]
30-minute delayed Total Recall (0-16), median [IQR]		16.0 [15.0-16.0]	13.0 [12.0-14.2]	15.0 [14.0-16.0]	16.0 [15.0-16.0]
One-week delayed Total Recall (0-16), median [IQR]		11.0 [9.0-13.0]	9.0 [7.0-9.2]	7.0 [6.0-9.0]	12.0 [11.0-13.0]
ALF ratio (total one-week recall / total 30-min recall * 100, %), median [IQR]		68.8 [57.4-81.2]	62.1 [53.7-75.0]	50.0 [40.0-56.2]	80.0 [68.8-87.5]

Note. FCSRT = Free and Cued Selective Reminding Test; ALF = accelerated long-term forgetting; IQR = interquartile range; MTL = medial temporal lobe. Available *n* is reported only for measures that were not obtained in the full sample; where the column is blank, the measure is available in all 322 participants. Percentages for biomarker variables are computed over participants with the relevant measurement.

¹ Aβ-positive status defined as a CSF Aβ42/Aβ40 ratio < 0.071.

² Biomarker assays: plasma p-tau217, CSF Aβ42 and CSF Aβ40 measured by Elecsys platform (NeuroToolKit kit, Roche; UGOT laboratory); CSF p-tau181 measured by Elecsys platform (standard kit; UGOT laboratory).

5.3 Neuropsychological Assessment

Neuropsychological assessment in ALFA and ALFA+ was administered at the Barcelonaβeta Brain Research Center by trained neuropsychologists, and followed a two-stage structure: a screening stage intended to establish cognitive normality, and an experimental battery intended to measure cognition with enough resolution to track its change over time (Molinuevo et al., 2016). Because testing took place in a bilingual region, participants were assessed in their preferred language, Catalan or Spanish.

The screening stage served an exclusionary purpose. It was designed to confirm cognitive normality and to identify functional, psychiatric, neurological or medical conditions liable to affect cognition (Molinuevo et al., 2016), and comprised the Mini-Mental State Examination (Folstein et al., 1975), the Memory Impairment Screen (Buschke et al., 1999), the Time-Orientation subtest of the Barcelona Test II (Peña-Casanova, 1991) and the Clinical Dementia Rating scale (Morris, 1997); mood symptoms were screened with the Goldberg Anxiety and Depression Scale (Goldberg et al., 1987). These instruments define the sample rather than describe it. They establish that participants are cognitively unimpaired, and none of them enters the present analyses as an outcome.

The experimental battery served the opposite purpose. Screening instruments are built to detect impairment, and in a sample selected to be unimpaired they behave accordingly: several show marked ceiling effects and leave little usable variance. ALFA therefore added a battery chosen for sensitivity to subtle change over longitudinal follow-up, covering premorbid intelligence, attention and working memory, verbal episodic memory, executive function, language and visual processing (Molinuevo et al., 2016). The instruments are listed in Table 2. The reasoning behind that choice is the same one that motivates the present study at the level of a single test: where a measure is compressed against its ceiling, the variance that carries early decline is not absent but unmeasured, and recovering it requires changing the measurement rather than the sample.

Table 2. Cognitive domains and instruments of the ALFA neuropsychological battery.

Domain	Instruments
Premorbid intelligence	Test de Acentuación de Palabras (TAP) (Del Ser et al., 1997)
Screening	Mini-Mental State Examination
Attention and working memory	WAIS-IV Digit Span (Wechsler, 2008); WMS-IV Symbol Span; Trail Making Test A (Reitan & Wolfson, 2009)
Verbal episodic memory	Free and Cued Selective Reminding Test; Memory Binding Test; WMS-IV Logical Memory
Executive function	Trail Making Test B (Reitan & Wolfson, 2009); Five Digits Test (Sedó, 2007); WAIS-IV Coding; WAIS-IV Matrix Reasoning; NIH Toolbox Flanker (Eriksen & Eriksen, 1974)
Language	Semantic (animal) fluency
Visual processing	WAIS-IV Visual Puzzles; Judgment of Line Orientation (RBANS) (Randolph et al., 1998)

Note. WAIS-IV = Wechsler Adult Intelligence Scale, fourth edition; WMS-IV = Wechsler Memory Scale, fourth edition; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status. Of the instruments listed, only the FCSRT enters the analyses reported in this thesis.

Three of these instruments assess verbal episodic memory: the FCSRT, the Memory Binding Test (Buschke, 2014) and the Logical Memory subtest of the WMS-IV (Wechsler, 1945). The Memory Binding Test was included precisely because the FCSRT was judged insufficiently sensitive to the earliest episodic memory changes of Alzheimer's disease (Molinuevo et al., 2016). The FCSRT is therefore neither the only episodic memory measure available in the cohort nor the one the protocol regarded as most sensitive, and what distinguishes the present work is not the instrument but the interval over which it is scored.

Within ALFA+ this protocol was extended and repeated at each visit alongside the fluid and imaging measures described in the sections that follow (Minguillón et al., 2020). Two of its components bear directly on the present work. The FCSRT, administered at the third visit, is the measure of interest and is described in detail below. Subjective cognitive complaints were recorded at the same visit with the MyCog questionnaire described in Section 2.5.1, from which the total score and the memory subscale were taken; these enter the analyses as outcomes rather than as predictors, for the reason set out in Section 4.3.

5.3.1 The FCSRT Protocol

The FCSRT assesses verbal episodic memory under conditions of controlled learning and cued retrieval, a design intended to ensure effective encoding and to separate a genuine

storage deficit from a failure of spontaneous retrieval (Buschke, 1984; Grober & Buschke, 1987). ALFA+ uses the Spanish validated version A, administered according to the protocol of the NEURONORMA project (Peña-Casanova et al., 2009). Only a Spanish version of the test exists, so that although the rest of the assessment could be conducted in Catalan or Spanish, the FCSRT was always administered in Spanish.

The list comprises sixteen words, printed four to a card, each belonging to a distinct semantic category that later serves as its retrieval cue. Encoding is controlled rather than assumed: the participant reads each card aloud and then identifies every word in response to its category cue, and identification of all sixteen items is a precondition for the test to proceed. This procedure is carried out once, before the first recall trial.

Three immediate recall trials follow. Each is preceded by twenty seconds of serial subtraction by threes, from a different starting number on each occasion, so that recall is not contaminated by rehearsal of the last items presented. Free recall is allowed up to ninety seconds and is discontinued after fifteen seconds without a response; the category cue is then supplied for every item not spontaneously produced, with five seconds allowed for each. On the first two trials only, any word still not produced after cueing is given to the participant before the next trial begins.

A delayed trial is administered between twenty-five and thirty-five minutes after the end of the third immediate trial, an interval occupied by the remaining tests of the battery, so that the delay is filled with interfering activity rather than left free, and is referred to throughout as the thirty-minute recall. Its structure reproduces the immediate trials without the reminding step: free recall first, then the category cue for each item not produced. Free and cued recall are scored separately, and their sum gives the total recall score for the trial, which ranges from 0 to 16; summed across the three immediate trials, immediate total recall ranges from 0 to 48. The three immediate total recall scores and the thirty-minute total recall score constitute the first four of the five timepoints modeled in this study. The choice of total recall is supported by evidence from the same cohort, in which it is the efficiency of semantic cueing, rather than free recall, that declines over three years in participants with abnormal amyloid and p-tau (Sánchez-Benavides et al., 2023). The full administration sequence is summarized in Figure 4.

Reference data for the FCSRT derived from ALFA+ participants with normal CSF biomarker levels show that education is associated with performance on every index of the test, that women outperform men, and that age has only a marginal effect, confined to

delayed free recall (Brugulat-Serrat et al., 2021). The covariate set used throughout the present analyses follows from this.



Figure 4. FCSRT administration protocol.

5.3.2 The One-Week Delayed Recall by Telephone

The one-week delayed recall is the element that distinguishes this protocol from the standard administration. Between seven and ten days after the face-to-face session, participants received a brief, unannounced telephone call and were asked to recall the same sixteen-item list learned at the visit. The call was not announced in advance: participants were not told at the face-to-face session that they would be contacted, so nothing in the protocol invited them to rehearse the list during the interval. The structure of the trial reproduces that of the thirty-minute delayed trial exactly: a period of free recall, followed by the category cue for each item not spontaneously produced, yielding free, cued and total scores on the same 0 to 16 scale. No material was re-presented and nothing was re-learned; the call measures only what survived the interval. The whole procedure takes approximately five minutes.

A forty-item recognition task was administered over the telephone immediately after the cued recall, and is available for 303 of the 322 participants. Recognition after a long delay distinguishes a failure of storage from a failure of retrieval and is therefore of direct theoretical relevance to ALF. It is not analyzed in the present work, which is confined to the recall trajectory, and is noted here because it remains available for future analysis.

The interval was closely adhered to. Across the 322 participants the median elapsed time between the face-to-face session and the telephone call was 7 days (interquartile range 7 to 7; mean 7.2, *SD* 0.57), 272 participants (84.5%) were called at exactly seven days, and

320 (99.4%) fell within the protocol window of seven to ten days; the two exceptions were called at six and at eleven days. Variation in elapsed time is therefore bounded to a range within which differences in retention of the magnitude reported in Chapter 6 cannot plausibly be produced by the delay itself.

5.4 Fluid Biomarkers and A β /AT Status Definition

Fluid biomarkers were obtained at the second ALFA+ visit, on average 3.3 years before the cognitive assessment analyzed here. This ordering is a property of the cohort rather than a design choice, and it has a consequence worth stating plainly: the biomarker measurements precede the memory measurement, so an association between them cannot be read as a concurrent correlation, and the direction of the temporal relation is the reverse of the one usually assumed when cognition is treated as an outcome of pathology.

Cerebrospinal fluid was obtained by lumbar puncture, and all fluid biomarkers were assayed on the Elecsys platform at the University of Gothenburg laboratory. CSF A β 42, CSF A β 40 and plasma p-tau217 were measured with the Roche NeuroToolKit kit, and CSF p-tau181 with the standard Elecsys kit. Two ratios were derived: CSF A β 42/A β 40, which indexes amyloid pathology and is the first CSF marker to change along the Alzheimer's continuum, and CSF p-tau181/A β 42, which pairs tau phosphorylation with amyloid burden and therefore combines a standard-kit measurement with a NeuroToolKit one. The second ratio is more sensitive than either marker alone in preclinical samples (Milà-Alomà et al., 2020). Plasma p-tau217 was retained as the blood-based marker because in this cohort it is among the plasma markers most strongly associated with amyloid PET retention in early-accumulating regions, and reaches abnormal levels before overt plaque pathology is established (Milà-Alomà et al., 2022).

Amyloid and tau positivity were defined with the thresholds derived in this cohort by two-component Gaussian mixture modeling, each cut-off being the mean plus two standard deviations of the non-pathological distribution: a participant is amyloid-positive when the CSF A β 42/A β 40 ratio falls below 0.071 and tau-positive when CSF p-tau181 exceeds 24 pg/mL (Milà-Alomà et al., 2020). A/T status combines the two into three groups, A-T-, A+T- and A+T+. The fourth logical combination, A-T+, is not assigned to a group: it is uncommon in a cognitively unimpaired sample and does not lie on the amyloid-initiated pathway the staging model describes, so participants in that cell contribute to no A/T comparison.

Amyloid PET was available for a subset of participants as a binary visual read of flutemetamol rather than as a continuous measure, and is used only as second-modality confirmation, not as a primary amyloid variable. The two modalities are not interchangeable at this stage of the continuum: the CSF definition is the more inclusive of the two in this sample, and almost every discordant participant is positive on CSF with a negative visual read rather than the reverse.

All three continuous biomarkers are strongly right-skewed and were transformed to base-ten logarithms before analysis. The logarithms were not standardized at this stage; standardization is applied inside each model, so that regression coefficients are expressed in standard deviations of the outcome and are comparable across biomarkers measured in different units.

5.5 Neuroimaging Acquisition and Processing

T1-weighted structural MRI was acquired at all three ALFA+ visits on a single Philips Ingenia CX 3T scanner at the Barcelonaβeta Brain Research Center. The sequence is a high-resolution three-dimensional T1-weighted acquisition with a repetition time of 9.9 ms, an echo time of 4.6 ms, a flip angle of 8 degrees and an isotropic voxel of 0.75 mm (Cumplido-Mayoral et al., 2023). Because the scanner and the protocol were unchanged across visits, between-visit differences in volume and thickness are not confounded by a change of platform, which is the principal threat to a retrospective longitudinal design of this kind. The mean interval was 3.21 years between the first and second visits and 3.35 years between the second and third, giving approximately 6.6 years of retrospective follow-up ending at the visit at which memory was assessed.

Acquisition and quality control follow the centralized framework of the ALFA neuroimaging platform, in which every session is archived and processed through an XNAT-based infrastructure, automatic validators are applied sequence by sequence, and the output of each processing pipeline is reviewed visually before release (Huguet et al., 2021). Acquisitions that did not pass this review were withheld, so the pattern of missing imaging data is largely documented rather than arbitrary. Five participants have no segmentation at the first visit, and those same five together with one more at the second, because the acquisition revealed an incidental finding that made it unsuitable for research use; one further participant has no second-visit segmentation because of a movement artefact; and three had not yet been scanned for the third visit when the data were extracted. Hippocampal

segmentations were therefore available for 316 participants at the first visit and 314 at each of the second and third; the mixed-effects models use every available acquisition.

Each imaging session was processed through two independent segmentation pipelines, and their output supports three families of outcomes. Hippocampal subfields were segmented with the Automatic Segmentation of Hippocampal Subfields (ASHS) pipeline, which runs on the dedicated inversion-recovery acquisition; cortical thickness was estimated with FreeSurfer on the T1 volume, using the Desikan-Killiany atlas; and from that same parcellation two cortical signatures, one specific to Alzheimer's disease and one indexing normal aging, were computed as weighted averages of thickness. Both pipelines are launched and archived through the same pipeline engine as part of the platform described above, so the derivatives of a given session are released together or not at all. The regions of interest were fixed in advance and are identical in the cross-sectional and the longitudinal analyses, so that the two differ only in whether they model the level or the rate of change of the same quantity.

Models of volumetric outcomes adjust for intracranial volume, estimated by FreeSurfer at the first visit and held constant across visits. Using the baseline estimate rather than a visit-specific one is deliberate: intracranial volume is fixed in adulthood, so visit-to-visit variation in its estimate reflects measurement error, and allowing it to vary would introduce that error into the very term intended to remove head-size variation.

5.5.1 Hippocampal Subfield Segmentation

Hippocampal subfields were segmented automatically with the ASHS pipeline (Yushkevich et al., 2015) which labels the hippocampal subfields and the adjacent medial temporal cortices in each hemisphere. The segmentation is driven by a dedicated high-resolution inversion-recovery acquisition rather than by the T1 volume used for cortical reconstruction. Four labels are analyzed as volumes in their own right: CA1, CA2, CA3, and the dentate gyrus taken together with CA4.

Two conventions govern how the segmentations were reduced to the variables entering the models, and both were verified against the source exports. Bilateral volumes are the sum of the left and right labels. The total MTL volume is the sum of the nine labeled structures: the five hippocampal subfields (CA1, CA2, CA3, the dentate gyrus with CA4, and the subiculum) together with the entorhinal cortex, the two perirhinal subdivisions (Brodmann areas 35 and 36) and the parahippocampal cortex, excluding the miscellaneous, sulcus and intracranial-volume entries, which are not anatomical structures of interest.

5.5.2 Cortical Thickness and AD Signatures

Cortical reconstruction and parcellation were performed with FreeSurfer v7 (Fischl, 2012), using the Desikan-Killiany atlas (Desikan et al., 2006). Average thickness was extracted for seven regions selected a priori: entorhinal and parahippocampal cortex, as the cortical targets of early Alzheimer's pathology; four prefrontal regions - caudal middle frontal, rostral middle frontal, superior frontal and pars triangularis - included because Parker et al. (2025) reported associations between accelerated forgetting and prefrontal thinning in the absence of medial temporal effects; and the lateral occipital cortex as a control region in which no association is expected. Bilateral thickness is the mean of the two hemispheres, not their sum, since thickness is not an additive quantity. Two participants could not complete the third-visit MRI session, because of claustrophobia or discomfort in the scanner, and were recalled shortly afterwards to finish it. In both cases the T1-weighted volume had already been acquired before the session was interrupted, so the interrupted session also yields cortical derivatives; these were discarded and those of the complete session retained.

Two composite measures were also taken, both computed as weighted averages of cortical thickness over a fixed set of regions. The AD cortical signature follows the definition of Dickerson et al. (2009) and covers the nine regions in which thinning most reliably separates Alzheimer's disease from normal aging: medial temporal cortex, inferior temporal gyrus, temporal pole, angular gyrus, superior frontal gyrus, superior parietal lobule, supramarginal gyrus, precuneus and inferior frontal sulcus. The aging signature is the corresponding summary for the regions in which thinning accompanies normal aging. It enters only as a covariate in the models of the disease signature, in place of intracranial volume, so that the estimate reflects thinning specific to the disease-related regions over and above the thinning expected at that age.

5.6 Operationalization of Long-Term Forgetting

Forgetting was operationalized in two complementary ways, and the distinction between them is the methodological core of the study. The first treats the five FCSRT scores as five observations of a single underlying curve and asks whether the sample divides into subgroups that differ in the form of that curve, that is in how much is acquired over the three learning trials, how much of it is still available after thirty minutes, and how much of it survives the week. The second reduces retention to a single number, the ratio of one-week to thirty-minute total recall, expressed as a percentage. The two are reported side by side

throughout Chapter 6; neither is treated as a validation of the other, since they answer different questions, and Section 3.3.4 gives the reasons for not relying on the ratio alone.

5.6.1 Growth Mixture Modeling of FCSRT Trajectories

Latent trajectories were estimated with growth mixture models fitted by maximum likelihood in the *lcmm* package for R. The data were arranged in long form, with five observations per participant: total recall at immediate trials one, two and three, at thirty minutes, and at one week, indexed by an ordinal timepoint variable running from one to five. This indexing treats the five occasions as equally spaced positions in the protocol rather than as points on a physical time axis. The intervals between them are in fact very unequal, from the minutes separating the three learning trials to the seven days preceding the final one, and a model fitted on elapsed time would be governed almost entirely by that last interval. The curve therefore describes the shape of the learning and retention profile across the five occasions of the protocol, and its coefficients are not rates of loss per unit of time.

A growth mixture model assumes that the population comprises a finite number of latent classes, each with its own mean trajectory, and that individuals vary around their class mean according to random effects (Muthén & Shedden, 1999). This distinguishes it from latent class growth analysis, in which within-class variance is constrained to zero and every member of a class is assumed to follow the class mean exactly (Nagin, 1999). The distinction matters here: constraining within-class variance forces heterogeneity to be absorbed by adding classes, and would inflate the number of classes required to describe the same data (Jung & Wickrama, 2008). Both families were nevertheless fitted, so that the choice between them rested on fit rather than on assumption.

Fourteen specifications were estimated in total, crossing the functional form of the trajectory with the random-effects structure and with the number of classes: latent class growth models with one to four classes; growth mixture models with a random intercept and one to four classes; growth mixture models with a random intercept and slope and one to three classes; and quadratic models with quadratic random effects and one to three classes. All multi-class models allowed class-specific proportional variance. Each specification was fitted by a grid search of fifty random starting-value sets, with a fixed seed, each run for a maximum of ten iterations from the starting values before the best was carried forward to convergence.

This procedure applies to the primary outcome, the trajectory of total recall culminating in the one-week trial, whose selection, fit and class descriptions are reported in

Section 6.1. Five secondary trajectory outcomes were modeled by the same procedure: the trajectories of total recall restricted to the three immediate trials and to the immediate trials plus the thirty-minute recall, and the three corresponding free-recall trajectories. Being ancillary to the question the study asks, these are reported in Appendix B.

5.6.2 Model Selection and the Minimum Class-Size Rule

Model selection required a rule beyond the information criteria, for a reason that is specific to mixture models and that materially affected the result. The likelihood of a growth mixture model increases as a class contracts onto a handful of participants with near-zero residual variance (Hipp & Bauer, 2006). Such solutions attain the highest log-likelihood and the lowest Bayesian information criterion (BIC) while being substantively empty: a class of three participants describes three people, not a subgroup of the population (Bauer & Curran, 2003). The grid search implemented in *lcm* returns only the single best-likelihood solution across its random starts and discards the remainder, so it can return a degenerate solution without any indication that it has done so.

The following rule was therefore specified: *among converged solutions, discard any whose smallest class contains fewer than 5% of the sample, then select the solution with the highest log-likelihood among those remaining*. The threshold is conventional rather than derived. Jung and Wickrama (2008) treat a class holding less than 1% of the sample as grounds for rejecting a solution, and the 5% adopted here is deliberately stricter. It is applied uniformly to the primary and to all secondary outcomes. Its effect here was decisive, and is documented in Section 6.1.1: the unconstrained best-likelihood solution places seven participants in its smallest class, and a solution better still places three.

Applying the rule to the primary outcome required more than a single estimation. The initial grid search converged on a solution the rule excludes, so the same specification was re-estimated from a denser grid of two hundred random starting-value sets and, separately, from a chosen set of starting values. The dense grid returned the excluded solution again; the chosen start reached both a higher log-likelihood and an admissible partition, and it is that solution which is reported.

To establish that selection did not depend on favorable starting values, the retained specification was refitted from eighty independent random starts, retaining every converged solution rather than only the best (Hipp & Bauer, 2006). This audit serves two purposes. It quantifies how often each distinct optimum is reached, and it reveals whether the retained solution is an artifact of the optimizer's path.

One further decision must be stated explicitly rather than presented as an outcome of the data. The number of classes was fixed at three for the primary outcome and applied uniformly to all secondary outcomes. For the primary outcome this coincides with the solution preferred by the information criteria among admissible models. For three of the four secondary outcomes that yielded an admissible solution, however, the Bayesian information criterion would have preferred two classes, and the third class is retained as an a priori structural choice made for comparability across outcomes. The secondary results in Appendix B should be read accordingly.

Class assignment used modal posterior probability. Classes were then ordered by their mean score and labeled Low, Intermediate and High performers, and the resulting variable enters all subsequent models as an ordered factor. The consequence for interpretation is that the primary coefficient for class is the linear polynomial contrast, which tests whether the outcome varies monotonically across the ordered classes, rather than a comparison against an arbitrary reference category. Pairwise contrasts between classes are reported separately where the ordered contrast is significant. Entropy is reported as a measure of classification separation, together with the mean posterior probability of the assigned class; neither is a fit statistic, and neither was used for selection.

Because assignment is modal, class membership enters every subsequent model as though it were observed, and the uncertainty quantified by the posterior probabilities is not carried forward. The consequence is documented: naive assignment attenuates the association between class membership and an external variable rather than inflating it, and leaves the standard errors conditional on the classification (Bolck et al., 2004; Vermunt, 2010). Corrections that propagate this uncertainty into the second stage were not applied here, so the associations reported in Chapter 6 are associations with the assigned class rather than with latent class membership itself.

5.6.3 The Continuous ALF Ratio

The second operationalization is the most conventional one in previous literature. The ALF ratio expresses one-week total recall as a percentage of thirty-minute total recall, so that a value of 100 denotes complete retention over the interval and lower values denote proportionally greater loss. The scale is not bounded above: four participants recalled more at one week than at thirty minutes and therefore exceed 100. It is analyzed as a continuous predictor, with no threshold applied, and it enters exactly the same models as the trajectory

classes, adjusted for the same covariates and subjected to the same correction for multiple comparisons.

The three weaknesses of ratio scoring identified in Section 3.3.4 apply here and are not remedied by treating the ratio as continuous. The ratio is unstable when its denominator is small, so participants with poorer thirty-minute recall generate less reliable values; in the present sample the denominator is high and its variance is compressed, which limits but does not eliminate the problem. The ratio also conflates forgetting with acquisition, since two participants with identical ratios may differ substantially in how much they learned. And the ratio presupposes that retention depends on the amount initially learned, which is the premise Della Sala et al. (2024) contest; if that premise fails, dividing by thirty-minute recall introduces a scaling artifact rather than removing a confound. The trajectory model does not share these properties, because acquisition is represented explicitly as part of the curve rather than divided out of it. The two approaches are reported in parallel for precisely this reason: agreement between them is informative, and so is disagreement.

5.7 Statistical Analysis

All analyses were carried out in R v4.5.2 (R Core Team, 2025). Mixed-effects models were fitted with lme4 v1.1.38 (Bates et al., 2015) and tested with lmerTest v3.1.3 (Kuznetsova et al., 2017); marginal means, slopes and contrasts were obtained with emmeans v2.0.0 (Lenth & Piaskowski, 2025); trajectory models were fitted with lcmm v2.2.2 (Proust-Lima et al., 2017).

Every model adjusts for age at the third visit, sex and years of education. Models of volumetric outcomes additionally adjust for intracranial volume at baseline, and models of the AD cortical signature adjust for the aging signature in place of intracranial volume.

All models are fitted to complete cases and no values were imputed. The number of participants contributing to each model therefore follows the coverage reported in Section 5.2, and is given with each result.

5.7.1 Cross-Sectional and Longitudinal Models

Continuous biomarker and cross-sectional imaging outcomes were analyzed with linear models in which the predictor is the trajectory class or the ALF ratio, and binary outcomes with logistic regression. These are amyloid status and, within the amyloid-positive subgroup, tau status. Confining the tau contrast to amyloid-positive participants keeps it

distinct from the amyloid contrast; dichotomizing the three-level A/T variable instead would compare A-T- against all other participants and so repeat the amyloid test rather than test tau. Continuous outcomes were standardized within each model, so that coefficients are expressed in standard deviations of the outcome and are comparable across measures recorded in different units.

Rates of structural change were estimated with linear mixed-effects models fitted to the three MRI visits in long form, with a random intercept per participant and time expressed in years from the first visit. The term of interest is the interaction between predictor and time, which tests whether the classes differ in the rate at which the structure changes rather than in its level; the main effect of class in these models corresponds to the difference at the first visit, on average 6.6 years (*SD* 0.6) before the memory assessment, and is not the quantity of interest. Restricted maximum likelihood was used throughout. Degrees of freedom for the model coefficients, which supply the *p*-values entering the correction for multiple comparisons, were approximated by the Satterthwaite method, while the marginal slopes and their pairwise contrasts were obtained with the Kenward-Roger approximation (Kenward & Roger, 1997). The bobyqa optimizer (Powell, 2009) was used for all fits.

Two random-effects structures were fitted to every longitudinal outcome: a random intercept alone, and a random intercept together with a random slope for time. The results reported in Chapter 6 use the random intercept. That structure was the one specified when the models were run, and with three observations per participant an individual slope rests on two intervals, so the simpler structure was retained as primary and the richer one examined as a sensitivity analysis. The comparison between the two is reported in Section 6.6. The distinction matters, because the slope variance is supported by the data and admitting it widens the standard error of the class-by-time interaction. Annual rates of change are reported in native units, obtained from the same models before standardization, so that the magnitude of the effect can be judged in cubic millimeters per year and not only in standard deviations.

Effect sizes for the longitudinal outcomes are reported as the standardized difference between class means at the final visit, computed from the residual standard deviation of the fitted model. Two quantities must not be confused. The rate of atrophy is estimated across three occasions and is comparatively precise; the difference accumulated by any one visit is a single contrast and is not. A significant difference in slope alongside an effect size whose confidence interval includes zero is therefore not a contradiction but a description of two different quantities, and both are reported.

5.7.2 Correction for Multiple Comparisons

Outcomes were assigned in advance to three families: fluid biomarkers and subjective cognitive complaints, comprising seven outcomes; cross-sectional imaging, comprising thirteen regions; and longitudinal imaging, comprising the same thirteen regions tested as a predictor-by-time interaction. The Benjamini-Hochberg false discovery rate (FDR) procedure was applied within each family, separately for the trajectory classes and for the ALF ratio (Benjamini & Hochberg, 1995). Corrected p -values are reported throughout Chapter 6 and denoted p_{FDR} ; for tests that enter the correction, an uncorrected p -value is quoted only where the corresponding corrected value is given alongside it.

Exactly one p -value per outcome enters the correction, and it is always the one from the base model. The sensitivity variants described in the next section do not contribute additional tests, and the pairwise contrasts between classes are reported without correction and interpreted only as a decomposition of an ordered contrast that has already survived it. Family membership was fixed before the models were run; no outcome was moved between families, and none was added or removed after inspecting the results.

5.7.3 Sensitivity Analyses

Three sets of sensitivity analyses were run. Each base model was refitted with *APOE*- $\epsilon 4$ genotype added as a covariate, with an interaction between the predictor and sex, and with both. The purpose of the genotype variant is to establish whether an association with biomarkers or structure is carried by the strongest single genetic risk factor rather than by the memory trajectory; the purpose of the sex variant is to test whether an effect present in the whole sample is driven by one sex. Neither variant contributes a p -value to the correction for multiple comparisons, and both are reported as sensitivity analyses rather than as results in their own right.

A further check concerns the stability of the trajectory solution itself. The multi-start audit described in Section 5.6.2 bears on whether the classes are a property of the data or of the estimation procedure, and its results are reported with the model selection in Section 6.1.1.

A third check concerns the random-effects structure of the longitudinal models. Both structures described in Section 5.7.1 were fitted to every longitudinal outcome and are compared in Section 6.6. The comparison is reported in full because the two specifications yield different standard errors for the predictor-by-time interaction.

6. Results

The results follow the order in which the analyses were specified. Section 6.1 begins with the property of the one-week measure that makes the trajectory models informative at all, and then reports those models, from the selection of the retained solution to the description of the three classes it defines. Sections 6.2 to 6.4 test the classes against the three families of outcomes in turn: fluid biomarkers and subjective cognitive complaints, brain structure at the time of assessment, and the rate of structural change estimated retrospectively across the three MRI visits. Section 6.5 repeats the same models with the continuous ALF ratio in place of the classes, so that the two operationalizations of forgetting can be compared directly, and Section 6.6 reports the sensitivity analyses. Throughout, *p*-values are given both uncorrected and after the Benjamini-Hochberg correction applied within each family, and an association is described as surviving only where the corrected value falls below .05. The hypotheses set out in Section 4.3 are addressed at the point where the relevant evidence appears.

6.1 Latent Trajectory Classes

Before the trajectory models are reported, one property of the data conditions everything that follows. At the conventional retention intervals, the FCSRT is insufficiently demanding for this cohort: most participants obtain the maximum score, so the test discriminates poorly among them, and a measure on which participants barely differ cannot covary with anything else. Extending the interval to one week removes that limitation. At the third immediate trial the standard deviation of total recall is 1.04 and 222 of 322 participants (68.9%) score the maximum of 16; at thirty minutes the standard deviation is 1.46 and 173 (53.7%) still score the maximum; at one week the standard deviation is 2.98 and only 11 participants (3.4%) do so. Dispersion therefore roughly doubles when the retention interval is extended, and the distribution does not simply collapse against the lower bound instead: 10 participants (3.1%) score four or below, and the scores run the length of the scale rather than banking at one end of it. H1, which predicted greater dispersion and a reduced ceiling effect at one week than at thirty minutes, is supported.

This is not a minor psychometric observation. More than half of a cognitively unimpaired cohort is indistinguishable on the standard thirty-minute measure, which is to say that for those participants the standard protocol returns no information about retention.

The variance that the trajectory models exploit is variance that only the extended interval makes visible.

6.1.1 Model Selection and Fit

Fourteen specifications were fitted to the five FCSRT timepoints. Their fit statistics are given in Table 3. The quadratic family separates sharply from the linear one: the three quadratic specifications occupy the three best positions by Bayesian information criterion, and the gap between the worst quadratic model and the best linear model exceeds 1,100 BIC units. Within the quadratic family, three classes are preferred over two by 17.1 BIC units and over one by more than 100. Three of the fourteen specifications failed to satisfy the convergence criteria and are excluded from consideration regardless of their fit.

Table 3. Fit of the fourteen candidate specifications, ordered by Bayesian information criterion. The retained solution is given in the final row.

Specification	Classes	Parameters	Log-likelihood	BIC	Entropy	Smallest class
Quadratic GMM, 3 classes	3	20	-2,953.25	6,021.99	0.717	2.2%
Quadratic GMM, 2 classes	2	15	-2,976.24	6,039.09	0.931	8.4%
Quadratic GMM, 1 class	1	10	-3,033.35	6,124.45	-	100%
LCGA, 3 classes	3	9	-3,618.51	7,288.98	0.862	0.6%
GMM random intercept, 2 classes	2	8	-3,623.75	7,293.70	0.611	13.7%
LCGA, 2 classes	2	6	-3,633.82	7,302.29	0.853	10.2%
GMM random intercept, 3 classes †	3	12	-3,617.16	7,303.61	0.714	4.0%
GMM intercept + slope, 2 classes	2	10	-3,623.50	7,304.75	0.645	13.4%
LCGA, 4 classes	4	12	-3,618.51	7,306.31	0.474	0%
GMM intercept + slope, 3 classes †	3	14	-3,616.03	7,312.90	0.761	3.1%
GMM random intercept, 4 classes †	4	16	-3,617.21	7,326.81	0.714	0%
GMM intercept + slope, 1 class	1	6	-3,656.06	7,346.78	-	100%
GMM random intercept, 1 class	1	4	-3,664.11	7,351.31	-	100%
LCGA, 1 class	1	3	-3,688.34	7,394.00	-	100%
Retained solution						
Quadratic GMM, 3 classes	3	20	-2,951.02	6,017.53	0.732	12.4%

Note. GMM = growth mixture model; LCGA = latent class growth analysis. † did not satisfy the convergence criteria. Entropy is not defined for single-class models. The retained solution has a higher log-likelihood and a lower BIC than the same specification as returned by the grid search, and a smallest class six times larger; see the text and Table 4.

The row that requires explanation is the difference between the first and the last. Both are the three-class quadratic growth mixture model; they differ only in which local optimum the estimation reached. The grid search over fifty random starting values returned a solution whose smallest class contains seven participants, 2.2% of the sample. That solution is inadmissible under the minimum class-size rule of Section 5.6.2. Re-estimating the same specification located a second optimum with a higher log-likelihood (-2,951.02 against -2,953.25), a lower BIC (6,017.53 against 6,021.99) and a smallest class of 40 participants, 12.4% of the sample. The retained solution is thus better on every criterion simultaneously.

The multi-start audit quantifies the problem the rule is designed to solve (Table 4). Across eighty independent random starts, five distinct optima were reached, all satisfying the convergence criteria. The most frequently reached solution, arrived at by 60 of 80 starts, places seven participants in its smallest class. The highest-likelihood solution overall, better than the retained one by 9.4 log-likelihood points and reached by 10 starts, places three participants - 0.9% of the sample - in its smallest class, and does so by merging the two best-retaining groups into a single class of 280. Its entropy of 0.952 illustrates why entropy must not be used for selection: a solution that assigns almost everyone to one class separates that class from a handful of outliers almost perfectly.

Table 4. Multi-start audit of the three-class quadratic growth mixture model: the distinct optima reached across eighty independent random starts. All five satisfied the convergence criteria.

Log-likelihood	Starts reaching it	Smallest class, n (%)	Entropy	Admissible
-2,941.63	10	3 (0.9%)	0.952	no
-2,951.02	3	40 (12.4%)	0.732	yes - retained
-2,952.28	1	28 (8.7%)	0.668	yes
-2,953.22	6	4 (1.2%)	0.725	no
-2,953.25	60	7 (2.2%)	0.717	no

Note. Admissibility is determined by the minimum class-size rule of Section 5.6.2: a solution is discarded if its smallest class contains fewer than 5% of the sample. Among admissible solutions the one with the highest log-likelihood is retained.

Two consequences follow. First, the retained solution is reached by 3 of 80 starts, so it does not depend on informed starting values even though it was originally located by starting the optimizer from a previous solution. Second, the solution that a single grid search would return is degenerate, and it is returned by three quarters of random starts. Had model selection rested on the information criteria alone, applied to the output of a single grid search, the analysis would have reported a spurious class of seven participants. This is a general property of mixture models rather than a peculiarity of these data.

The retained model converged with all convergence criteria satisfied, all standard errors finite and non-zero, and no negative variance components. Entropy is 0.732, and the mean posterior probability of the assigned class is 0.901, 0.881 and 0.878 for the three classes respectively, with an overall mean maximum posterior probability of 0.883.

Classification is therefore good without being so sharp as to suggest that the classes are artifacts of a degenerate solution.

6.1.2 Description of the Three Classes

The three classes and their mean trajectories are shown in Figure 5 and tabulated in Table 5. They are labeled by their standing on the one-week trial: Low performers ($n = 40$, 12.4%), Intermediate performers ($n = 81$, 25.2%) and High performers ($n = 201$, 62.4%).

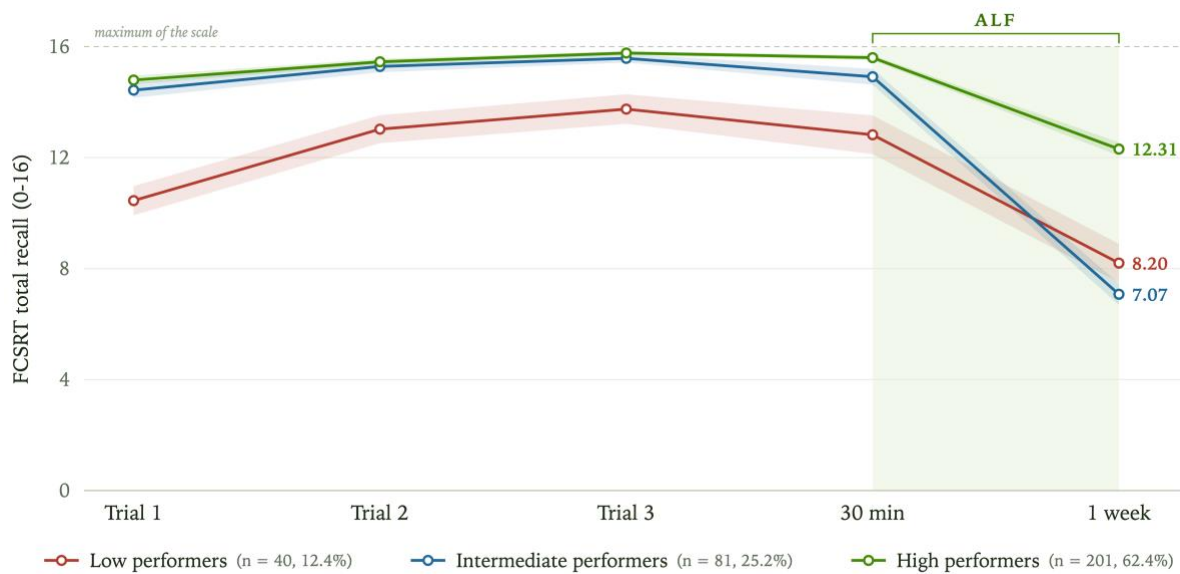


Figure 5. Latent trajectory classes of FCSRT total recall across the five timepoints.

The Low performers score below both other classes at every timepoint up to thirty minutes. They acquire less material (10.45 at the first trial against 14.80 for the High class) and retain less at thirty minutes (12.82 against 15.60). At one week they recall 8.20 against 12.31 for the High class. This is a profile of uniformly weaker verbal episodic memory, and it is the profile that a conventional protocol would already detect.

The Intermediate class is the substantive finding. Its members learn normally: 14.43, 15.28 and 15.58 across the three immediate trials, values that differ from the High performers by roughly a third of a point and lie far above the Low performers. They retain normally at thirty minutes, scoring 14.91 against 15.60 for the High class. At one week they recall 7.07 of 16 - the lowest mean of the three classes, below even the Low performers, and 5.24 points below the High class. Expressed as a proportion of what they still held at thirty minutes, they

retain 47.5% (*SD* 11.6), against 64.0% (*SD* 16.8) for the Low performers and 79.0% (*SD* 11.3) for the High performers.

This is a dissociation and not a matter of degree. A class that is ordered second of three at every timepoint through thirty minutes, and last of three at one week, cannot be described as a milder version of the Low class; the ordering itself changes. H2, which predicted at least one class whose acquisition and thirty-minute recall lay within the range of the more proficient participants, but whose one-week recall was disproportionately low, is therefore supported; three classes ordered identically at all five timepoints would have contradicted it.

Cross-tabulating the one-week classes against the trajectory classes derived from the same protocol truncated at thirty minutes, 43 participants classified as High performers on the thirty-minute trajectory are Intermediate performers at one week. That is 13.4% of the cohort whose learning and short-delay recall are unremarkable and whose one-week retention is not - a group invisible to a standard protocol by construction. The correlation between thirty-minute and one-week total recall is 0.44, so the two intervals share less than a fifth of their variance.

Table 5. FCSRT total recall by trajectory class at each timepoint, mean (*SD*), and ALF ratio.

Class	n (%)	Trial 1	Trial 2	Trial 3	30 min	1 week	ALF ratio, %
Low performers	40 (12.4)	10.45 (1.69)	13.03 (1.62)	13.75 (1.72)	12.82 (2.25)	8.20 (2.23)	64.0 (16.8)
Intermediate performers	81 (25.2)	14.43 (1.25)	15.28 (0.95)	15.58 (0.70)	14.91 (1.32)	7.07 (1.77)	47.5 (11.6)
High performers	201 (62.4)	14.80 (1.11)	15.45 (0.79)	15.77 (0.51)	15.60 (0.65)	12.31 (1.72)	79.0 (11.3)
Whole sample	322 (100)	14.2 (1.9)	15.1 (1.2)	15.5 (1.0)	15.1 (1.5)	10.5 (3.0)	69.2 (18.2)

Note. All scores are total recall, on a 0-16 scale. ALF ratio is one-week total recall as a percentage of thirty-minute total recall, mean (*SD*).

The classes also differ on demographic and biological characteristics. These are given by class in Table 1, and the models that test these differences with adjustment for covariates follow in Sections 6.2 to 6.4. The Low performers are the oldest (median 70.7 years against 66.1 for the High class) and the least educated (median 12.0 against 15.0 years), which is the expected ordering and one reason why all subsequent models adjust for age and

education. The Intermediate class lies between the two on both, closer to the High class on education than on age.

6.2 Associations With Fluid Biomarkers and Subjective Cognitive Complaints

Seven outcomes constitute the first correction family, defined in Section 5.7.2. Five derive from fluid biomarkers: plasma p-tau217 and the CSF A β 42/A β 40 and p-tau181/A β 42 ratios, each analyzed as a continuous concentration, and two dichotomous statuses derived from cerebrospinal fluid, amyloid positivity and, within the amyloid-positive subgroup, tau positivity. The remaining two are the MyCog total and memory scores, which quantify self-reported cognitive complaints and are not markers of pathology. Results are given in Table 6. Coverage varies across these outcomes, from 309 participants for plasma p-tau217 to 95 for the tau contrast, but the variation is unrelated to the predictor: missingness on the three fluid measures is independent of trajectory class ($p = .92$, $.40$ and $.36$). Two outcomes survive correction for multiple comparisons: plasma p-tau217, for which the ordered contrast across trajectory classes gives $p = .0055$ and $p_{\text{FDR}} = .021$, and CSF p-tau181/A β 42, which gives $p = .0059$ and $p_{\text{FDR}} = .021$.

Table 6. Associations between trajectory class and the family of fluid biomarkers and subjective cognitive complaints. p -values are for the linear ordered contrast across classes, from models adjusted for age, sex and years of education.

Outcome	n	p	p_{FDR}
Plasma p-tau217	309	.0055	.021*
CSF A β 42/A β 40	286	.181	.211
CSF p-tau181/A β 42	240	.0059	.021*
Amyloid status	286	.468	.468
Tau positivity (A+T+ vs A+T-)	95	.033	.065
MyCog total	321	.064	.089
MyCog memory	321	.037	.065

Note. Continuous biomarkers were log-transformed and standardized within model. Amyloid status and tau positivity were analyzed by logistic regression, the latter within the amyloid-positive subgroup. Correction is by the Benjamini-Hochberg procedure within the family of seven. * survives correction at the 5% level.

The pairwise decomposition of the p-tau217 effect is shown in Figure 6 and does not follow the ordering of the classes on one-week recall. As set out in Section 5.7.2, pairwise contrasts are reported without correction, since they decompose an ordered contrast that has already survived it. The Low performers have higher concentrations than the Intermediate class ($\beta = 0.49$, 95% CI [0.11, 0.87], $p = .012$) and than the High class ($\beta = 0.50$, 95% CI [0.15, 0.85], $p = .006$), both expressed in standard deviations of log-transformed p-tau217. The Intermediate and High classes are indistinguishable ($\beta = 0.008$, 95% CI [-0.26, 0.27], $p = .95$).

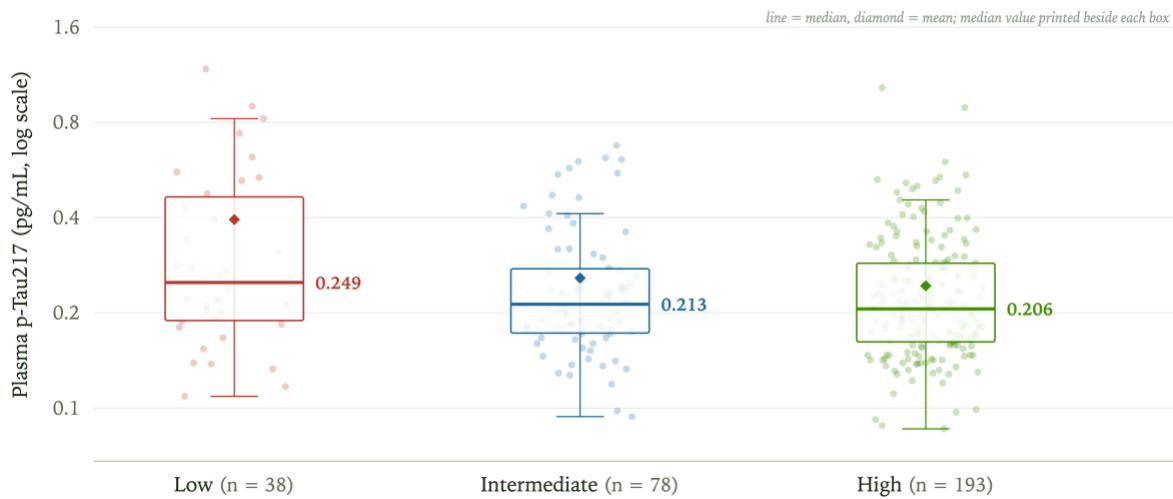


Figure 6. Plasma p-tau217 concentration by trajectory class ($n = 309$). Boxes show median and interquartile range, diamonds the mean, points individual participants.

Hypothesis H3 predicted that trajectories showing disproportionate loss at one week, whether or not also lower at acquisition, would show higher p-tau217 and a less favorable CSF profile than the best-retaining trajectory. It holds for the Low performers but not for the Intermediate class, so H3 is only partially supported, and what fails is the comparison the study was designed to make. CSF p-tau181/A β 42 repeats this shape rather than adding to it: the Low performers differ from the High performers ($\beta = 0.58$ standard deviations, $p = .006$) while the Intermediate class does not ($\beta = 0.22$, $p = .161$). The CSF A β 42/A β 40 ratio shows no association with trajectory class.

6.3 Cross-Sectional Imaging Associations

Thirteen outcomes constitute the cross-sectional imaging family: total MTL volume and the CA1, CA2, CA3 and dentate gyrus subfields, entorhinal and parahippocampal thickness, the AD cortical signature, four prefrontal regions, and the lateral occipital control region. None is associated with trajectory class or with the ALF ratio after correction.

Before correction, the strongest associations for the classes are total MTL volume ($p = .074$), pars triangularis ($p = .080$) and the dentate gyrus ($p = .114$), none of which approaches significance within a family of thirteen (all $p_{\text{FDR}} \geq .49$). The control region behaves as intended, with $p = .992$ and $p_{\text{FDR}} = .992$. Hypotheses H4 and H5, which predicted smaller MTL volume and thinner entorhinal and parahippocampal cortex in the poorer-retaining trajectories, with a more pronounced effect in CA1, are not supported at the level of structural volume measured at a single timepoint.

The descriptive values in Section 5.2 explain why. Bilateral total MTL volume is 12,132 mm³ in the Low class against 12,638 mm³ in the High class, a difference of about 4%, while the standard deviation within each class exceeds 1,200 mm³. The between-class difference is small relative to the between-person variation in head size and MTL volume, and a cross-sectional comparison adjusted for intracranial volume has little power to resolve it. The result of Section 6.4 suggests that this is a limitation of the cross-sectional contrast rather than an absence of structural difference.

6.4 Longitudinal Hippocampal Atrophy

The same thirteen regions were tested as a class-by-time interaction in mixed-effects models fitted to the three MRI visits, spanning 6.6 years of retrospective follow-up on average ending at the visit at which memory was assessed. Two outcomes survive correction: bilateral CA1 ($p = .0014$, $p_{\text{FDR}} = .019$) and total MTL volume ($p = .0053$, $p_{\text{FDR}} = .034$). These are not independent of one another, since total MTL volume is the sum of the segmented structures defined in Section 5.5.1 and therefore includes CA1. Results are given in Table 7.

Table 7. Associations between trajectory class and the rate of structural change across three MRI visits. *p*-values are for the class-by-time interaction, from models adjusted for age, sex, years of education and, for volumetric outcomes, baseline intracranial volume.

Region	<i>p</i>	<i>p</i> _{FDR}
Bilateral MTL, total	.0053	.034*
Bilateral CA1	.0014	.019*
CA2	.621	.682
CA3	.472	.682
Dentate gyrus / CA4	.020	.087
Entorhinal thickness	.273	.592
Parahippocampal thickness	.764	.764
AD cortical signature	.445	.682
Caudal middle frontal thickness	.174	.564
Rostral middle frontal thickness	.246	.592
Superior frontal thickness	.329	.610
Pars triangularis thickness	.630	.682
Lateral occipital thickness (control)	.620	.682

Note. Correction is by the Benjamini-Hochberg procedure within the family of thirteen. The lateral occipital cortex was specified a priori as a control region in which no association was expected. MTL = medial temporal lobe. * survives correction at the 5% level.

The magnitude of the CA1 effect is shown in Figure 7. Annual change in bilateral CA1 volume is -23.9 mm³ per year in the Low class, -17.5 in the Intermediate class and -7.9 in the High class, corresponding to 0.86%, 0.62% and 0.29% of baseline volume per year. The Low performers therefore lose CA1 volume 3.01 times faster than the High performers, and the Intermediate performers 2.20 times faster.

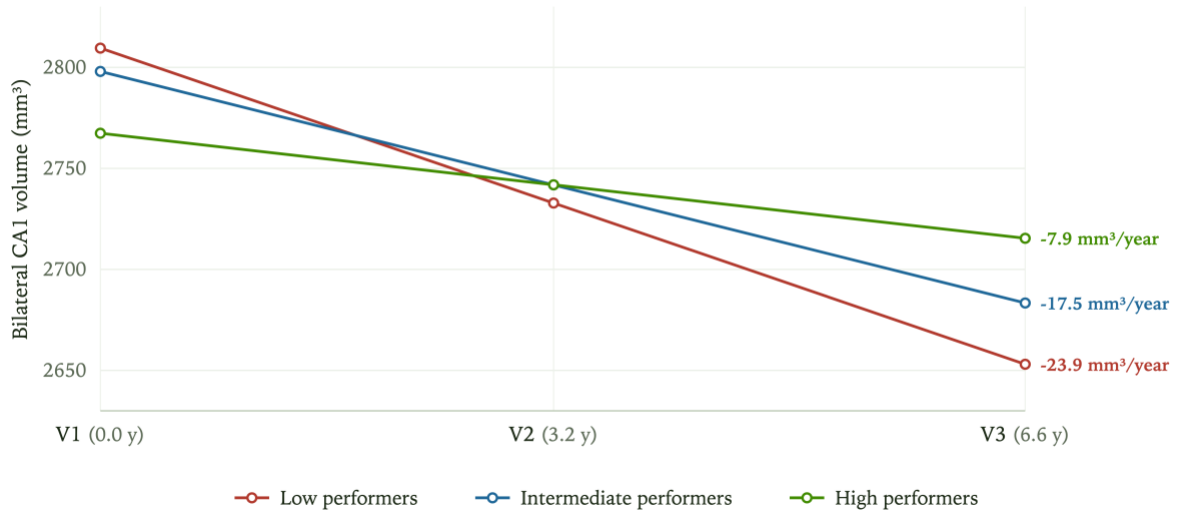


Figure 7. Model-implied trajectories of bilateral CA1 volume by trajectory class across the three MRI visits, with covariates held at their sample means and annual rates of change annotated.

The pairwise slope contrasts identify which comparison carries the effect. The Low and High classes differ ($\beta = -0.048$, 95% CI $[-0.083, -0.013]$, $p = .004$, expressed in standard deviations of CA1 volume per year), and so do the Intermediate and High classes ($\beta = -0.029$, 95% CI $[-0.056, -0.002]$, $p = .036$). The Low and Intermediate classes do not differ from each other ($p = .489$). The second of these is the contrast the study was designed to detect: a group whose learning and thirty-minute recall are within the normal range loses CA1 volume more than twice as fast as the best-retaining group. This contrast, and the two associations that survive correction, depend on the random-effects structure of the models; Section 6.6 reports what becomes of them when a random slope for time is admitted.

Total MTL volume shows the same ordering with a weaker gradient: -136.1 , -83.1 and -59.7 mm³ per year for the Low, Intermediate and High classes, corresponding to fold ratios against the High class of 2.28 and 1.39. Only the Low versus High contrast reaches significance ($\beta = -0.056$, 95% CI $[-0.103, -0.009]$, $p = .015$). The attenuation follows from the nesting: total volume dilutes the CA1 effect in the subfields that show none, so the composite is less strongly associated than the region that drives it. The dentate gyrus shows a similar pattern - fold ratios of 2.20 and 1.87 - but does not survive correction ($p = .020$, $p_{FDR} = .087$), and it should not be reported as a positive finding. The lateral occipital control region shows no difference in rate of change ($p = .620$, $p_{FDR} = .682$), with fold ratios of 1.09 and 0.92, consistent with regional specificity rather than with generalized atrophy.

Hypothesis H6 predicted faster MTL volume loss and faster entorhinal and parahippocampal thinning in the poorer-retaining trajectories, with no acceleration in the control region. It is supported for MTL volume and for CA1, and not supported for entorhinal or parahippocampal thickness, neither of which approaches significance. H5, which predicted a more pronounced effect in CA1 than in the remaining subfields, is partially supported in the longitudinal domain, where it was not supported cross-sectionally: the effect is strongest in CA1, while the dentate gyrus shows the same pattern without surviving correction. The difference between the two subfields was not tested, so regional specificity rests on the contrast with the control regions, in which no acceleration is observed, rather than on a contrast between subfields.

Expressed as the standardized difference between class means at the final visit, the effect sizes are moderate and their confidence intervals include zero: for CA1, $d = -0.47$ [-1.24, 0.29] for Low versus High and $d = -0.24$ [-0.82, 0.34] for Intermediate versus High; for total MTL volume, $d = -0.40$ [-0.94, 0.13] and $d = -0.17$ [-0.58, 0.24]. This is not in conflict with the significant slope contrasts. A rate estimated across three occasions is estimated precisely; the difference accumulated by one particular visit is a single contrast and is not. The two are different quantities, and both statements are true of these data.

6.5 Convergence With the Continuous ALF Ratio

The same models were run with the continuous ALF ratio in place of the trajectory classes. Three outcomes survive correction, and they are not the same three (Table 8). Among the subjective complaint scores and the fluid biomarkers, the MyCog total and memory scores survive (both $p_{FDR} = .016$); CSF p-tau181/A β 42 falls just short ($p = .023$, $p_{FDR} = .054$), and plasma p-tau217, which survives for the classes, does not ($p = .298$, $p_{FDR} = .348$). Among the longitudinal outcomes, bilateral CA1 survives ($\beta = 0.015$ per standard deviation of the ratio per year, 95% CI [0.005, 0.025], $p = .0026$, $p_{FDR} = .033$). As for the classes, no measure of structure at the third visit is associated with it.

Table 8. Outcomes surviving correction for multiple comparisons under each operationalization of forgetting. Corrected *p*-values are shown; a dash indicates that the outcome did not survive correction within its family.

Outcome	Trajectory classes	ALF ratio
Plasma p-tau217	.021	-
CSF p-tau181/A β 42	.021	-
MyCog total	-	.016
MyCog memory	-	.016
Bilateral CA1, rate of change	.019	.033
Bilateral medial temporal lobe, rate of change	.034	-
Cross-sectional MRI	-	-

The convergence is therefore partial, and the pattern of disagreement is itself informative. The two operationalizations agree on the one finding that is common to both - the rate of CA1 atrophy - and diverge everywhere else, in a direction that follows from what each measure represents. The ratio divides one-week recall by thirty-minute recall and thereby removes the level of performance, retaining only the proportion lost; it is associated with subjective complaints and with the CSF marker that combines tau with amyloid. The classes retain the level as well as the shape, and the two fluid biomarkers they identify, plasma p-tau217 and CSF p-tau181/A β 42, both prove on decomposition to be carried by the Low performers rather than by the Intermediate class.

No prediction had been made as to which measure would prove more sensitive, and these results do not establish that either is. What they show is that the two are not interchangeable: the ratio alone would have missed both fluid biomarker associations and the MTL volume result, and the classes alone would have missed the association with subjective memory complaints.

6.6 Sensitivity Analyses

Adding *APOE*- ϵ 4 genotype as a covariate does not alter the pattern of results, and neither does adding an interaction between the predictor and sex. The associations that survive correction in the base models remain significant in both variants, and no association that fails to survive in the base model becomes significant in either. Neither variant contributed a *p*-value to the correction for multiple comparisons, and neither is reported as a result in its own right.

The random-effects structure of the longitudinal models was examined directly. Refitting every longitudinal outcome with a random slope for time alongside the random intercept improves fit (likelihood-ratio tests on two degrees of freedom: $\chi^2 = 72.6$ for CA1, 105.5 for total MTL volume, 56.2 for the dentate gyrus; all $p < .001$; no fit singular) and widens the standard error of the class-by-time interaction by 28% to 35%. The estimates themselves are essentially unchanged: the Intermediate versus High contrast for CA1 is -9.5 mm³ per year under the random intercept and -10.4 under the random slope. Under the wider standard errors, the two longitudinal associations reported in Section 6.4 fall short of the corrected threshold (CA1 $p_{\text{FDR}} = .073$, total MTL volume $p_{\text{FDR}} = .073$). The longitudinal findings are therefore consistent in direction and magnitude across both specifications, but under the wider standard errors they no longer clear the correction across thirteen regions.

7. Discussion

The previous chapter presented the results of the statistical models; the present chapter considers their interpretation and implications. Section 7.1 integrates the evidence across the six hypotheses and identifies the three principal findings on which the discussion is based. Section 7.2 examines the specificity of the imaging finding, addressing both its restriction to a single hippocampal subfield and its emergence in longitudinal change rather than cross-sectional volume. Section 7.3 considers what the divergence between the two operationalizations of forgetting reveals about the constructs captured by each measure. Section 7.4 defines the scope of interpretation supported by the findings and the claims that remain beyond the evidence provided by the present study. Finally, Section 7.5 addresses the principal limitations of the study, including one that is particularly relevant to the interpretation of the longitudinal findings.

7.1 Main Findings

Of the six primary hypotheses specified in Section 4.3, the findings supported H1 and H2, provided partial support for H3, H5 and H6, and did not support H4. Extending the retention interval from 30 minutes to one week approximately doubled the dispersion of total recall and reduced the ceiling effect from 53.7% to 3.4%. The retained trajectory model identified an Intermediate class whose acquisition and 30-minute recall were close to those of the High-performing class, but whose one-week recall was the lowest of the three classes. The associations with plasma p-tau₂₁₇ and the CSF p-tau₁₈₁/A β ₄₂ ratio were driven by the Low-performing class rather than by the Intermediate class. No cross-sectional measure of brain structure differed significantly between classes after correction for multiple comparisons. The predicted neuroanatomical pattern was therefore observed only longitudinally, most clearly in CA1, with weaker support for total MTL volume and no corresponding association with entorhinal or parahippocampal cortical thickness. The lateral occipital control region showed no comparable association.

Three findings are central to the study. First, the conventional 30-minute retention interval showed limited discriminatory value in this cohort, whereas the one-week interval revealed substantially greater variability in performance. More than half of the participants achieved the maximum score at 30 minutes, compared with only 3.4% at one week. The

extended interval therefore exposed individual differences that were largely compressed at the conventional delay.

Second, the additional variability observed at one week was structured rather than random. The trajectory model identified an Intermediate class comprising 81 participants, or approximately one quarter of the cohort, whose acquisition and 30-minute recall were relatively preserved but whose one-week recall was the lowest of the three classes. In a trajectory model truncated at 30 minutes, 43 of these participants were classified as High performers, whereas the inclusion of the one-week assessment placed them in the Intermediate class. The one-week assessment therefore identified a subgroup that would not have been distinguished by the conventional protocol alone.

Third, the trajectory pattern showed a regional neuroanatomical correlate. In the primary random-intercept models, which covered an average of 6.6 years across the three MRI visits, the Low-performing class lost bilateral CA1 volume 3.01 times faster than the High-performing class, while the Intermediate class lost CA1 volume 2.20 times faster. A similar but weaker gradient was observed for total MTL volume. These findings suggest that the one-week trajectory may carry information about longitudinal hippocampal change, particularly in CA1.

The interpretation of the Intermediate class requires caution. Neither of the fluid biomarker associations that survived correction was specifically carried by this class: plasma p-tau217 and CSF p-tau181/A β 42 differed between the Low-performing and High-performing classes, whereas the Intermediate and High-performing classes were not distinguishable on either measure. Amyloid positivity also did not differ significantly across classes, occurring in 33.3% of the Low-performing class, 42.5% of the Intermediate class and 37.8% of the High-performing class. The fluid biomarker findings therefore point to a pattern of poorer overall memory performance in the Low-performing class rather than to a consistent biomarker profile corresponding to the dissociated Intermediate trajectory.

The longitudinal findings were somewhat sensitive to the random-effects structure. Adding a random slope for time improved model fit and left the estimates similar in direction and magnitude, although it increased the standard errors of the class-by-time interactions. The longitudinal pattern therefore remained similar across model specifications, while the strength of the statistical evidence varied with the complexity of the model. These findings should be interpreted cautiously and considered in need of confirmation in future studies.

Taken together, the results support a narrower interpretation of the one-week assessment. One-week delayed recall appears to provide complementary information that is

not fully captured by standard short-delay assessment, particularly with regard to longitudinal hippocampal change. The findings do not establish uniformly greater sensitivity across biological, structural and subjective outcomes. The value of the extended interval may therefore lie in complementing, rather than replacing, conventional measures of learning and short-delay recall.

7.2 Regional Specificity of Longitudinal Structural Change

The imaging analyses pointed to a regionally specific longitudinal pattern. The clearest associations involved bilateral CA1 and total MTL volume, where trajectory classes differed in the rate of volume change after correction for multiple comparisons. This pattern was not observed in cross-sectional measures of brain structure at the third visit, nor across most of the other predefined regions. The dentate gyrus showed a similar direction of association, but the result did not survive correction and is therefore not interpreted as a positive finding. No comparable association was observed for CA2, CA3, entorhinal or parahippocampal cortical thickness, the Alzheimer's disease cortical signature, or the four prefrontal regions. The lateral occipital cortex, specified a priori as a control region, showed no evidence of a difference in the rate of change.

This pattern should nevertheless be interpreted in light of two important considerations. First, the associations were restricted to two of the thirteen regions examined. Because total MTL volume includes CA1 among its constituent structures, the two findings should be interpreted as related measures of MTL change rather than as entirely separate pieces of evidence. Second, the associations emerged in the rate of structural change rather than in cross-sectional measures of brain structure at the third visit. The absence of a cross-sectional association should therefore not be interpreted as evidence that the relationship between long-term forgetting and hippocampal structure is absent.

CA1 represents a biologically plausible target for this association. Neuropathological changes are thought to begin in the transentorhinal region and subsequently extend into the hippocampal formation, with CA1 and the subiculum among the earliest affected targets (Braak & Braak, 1991). Consistent with this framework, clinicopathological studies have reported that CA1 neuronal loss may be more closely related to cognitive impairment than amyloid burden (Bennett et al., 2004; Giannakopoulos et al., 2003; Gómez-Isla et al., 1997). Bartels et al. (2025) also reported an association between long-term retrieval performance

and CA1 volume in older adults and individuals at risk of dementia, providing the rationale for examining CA1 in the present study.

The present findings extend this literature in one respect and qualify it in another. They extend previous cross-sectional evidence by suggesting that the relationship between long-term forgetting and CA1 may also be reflected in the rate of volume loss over time. This longitudinal dimension could not be evaluated in the study by Bartels et al. (2025). At the same time, the absence of a cross-sectional association with CA1 in the present sample indicates that this relationship may not be detectable through brain-structure measures obtained at a single time point.

The secondary trajectory analyses reported in Appendix B further support a distinction between concurrent structural measures and longitudinal change. The full trajectory including the one-week recall was the only trajectory predictor associated with corrected longitudinal changes in bilateral CA1 and total MTL volume. This pattern suggests that the different retention windows may capture partly distinct aspects of memory performance.

7.3 Trajectory Classes and Continuous Ratio

The two operationalizations of long-term forgetting converged on one outcome but diverged across the remaining outcome domains. In the primary analyses, both the trajectory classes and the continuous ALF ratio were associated with the rate of bilateral CA1 atrophy. Neither operationalization, however, was associated with cross-sectional measures of brain structure at the third visit. Beyond this common longitudinal finding, the trajectory classes were associated with plasma p-tau217 and the CSF p-tau181/A β 42 ratio, but not with subjective cognitive complaints. The continuous ALF ratio showed the opposite pattern for subjective cognition, being associated with both the MyCog total score and its memory subscale, but not with plasma p-tau217. Its association with the CSF p-tau181/A β 42 ratio was weaker and did not survive correction.

The divergence between the two operationalizations is consistent with the information retained by each measure. The ALF ratio expresses one-week recall as a proportion of 30-minute recall and therefore focuses on relative loss while largely removing the overall level of performance. The trajectory model preserves both the level and the shape of performance across acquisition, short-delay retention and one-week retention. This

distinction helps explain why the fluid biomarker associations were carried primarily by the Low-performing class, which also showed poorer acquisition, whereas the ratio captured associations with subjective cognitive complaints that were not detected by the class-based analysis.

These findings also bear on the methodological debate surrounding ratio-based measures of forgetting. Della Sala et al. (2024) argued that expressing retention as a proportion of short-delay recall may adjust for initial performance, whereas Geurts et al. (2015) suggested that ratio scores can introduce a scaling artefact. The present study cannot resolve this theoretical issue because it was not designed to test whether the relationship between long-term forgetting and biological outcomes depends on initial learning. It does nevertheless show that, within this sample, the choice of operationalization affected which associations were detected. The results support their use as complementary approaches, provided that their different interpretations are made explicit.

The trajectory classes also capture information that cannot be reduced to the continuous ratio. The Intermediate class had the lowest mean one-week recall, with a mean score of 7.07 out of 16, below the Low-performing class mean of 8.20. It also had the lowest retention proportion, retaining 47.5% of the material recalled at 30 minutes, compared with 64.0% in the Low-performing class and 79.0% in the High-performing class. At the same time, the Intermediate class showed relatively preserved acquisition and 30-minute recall, whereas the Low-performing class performed worse at every earlier time point. The two classes therefore differed not only in the amount of material retained, but also in the shape of their learning-retention trajectories.

Taken together, these findings suggest that trajectory classes and the continuous ALF ratio provide related but non-interchangeable descriptions of long-term forgetting. The ratio offers a concise measure of proportional retention, whereas trajectory modeling distinguishes overall performance level from the pattern of change across the protocol. The convergence of both measures on CA1 atrophy supports the relevance of long-term forgetting to longitudinal hippocampal change, while their divergence across biomarkers and subjective cognition indicates that each operationalization contributes information that the other does not capture.

7.4 Clinical and Research Implications

Beyond its analytical implications, the study also provides preliminary evidence about the practical feasibility of extending memory assessment over a one-week interval. Of the 336 participants scheduled to complete the telephone recall, 322 did so, and the reasons for non-completion were predominantly logistical. The call lasted approximately five minutes, required no additional visit and used material that had already been learned during the face-to-face session. Previous work has found no meaningful difference between telephone and in-person administration at one week (Allen et al., 2019), and the present protocol used the same procedure for all participants. One-week delayed recall therefore appears feasible as an addition to an existing face-to-face cognitive assessment.

The one-week assessment may also provide information that is largely absent from the standard retention interval. At 30 minutes, more than half of the participants achieved the maximum total-recall score, whereas only a small proportion did so after one week. The extended interval consequently revealed substantially greater variability in a cohort that was cognitively unimpaired at conventional assessment. This supports the use of longer retention intervals when the aim is to characterize individual differences in long-term retention rather than to assess only short-delay memory performance.

These findings have potential value for research on preclinical AD, particularly in studies that seek to identify subtle differences in memory consolidation before conventional cognitive impairment becomes evident. The results suggest that a brief delayed assessment may complement standard neuropsychological measures and may be especially informative when combined with longitudinal neuroimaging.

Overall, the findings support the feasibility and research utility of adding a one-week telephone recall to an established cognitive protocol. They do not, on their own, justify using the measure for diagnosis or screening. Further studies in cohorts selected independently of episodic memory performance, with prospective longitudinal follow-up and clinical endpoints, will be needed to determine whether the trajectory-defined patterns are reproducible and clinically meaningful.

7.5 Limitations

Some limitations should be considered when interpreting the present findings. The study was conducted in a cognitively unimpaired research cohort with a relatively restricted

range of age and cognitive performance. The sample is therefore not representative of the general population, and the findings may not generalize directly to individuals with objective cognitive impairment or to populations with different educational, linguistic, and sociodemographic characteristics. In addition, the analytic sample included only participants who completed the one-week telephone assessment. Although most missing assessments were attributable to logistical or procedural reasons, one participant was excluded because clinically relevant memory impairment had already been observed during the face-to-face assessment. The resulting sample may consequently have been slightly depleted of the most pronounced memory profiles.

A more substantial limitation concerns the observational and partially retrospective design. The memory assessment was conducted at the third study visit, whereas the fluid biomarkers were obtained at the preceding visit, approximately 3.3 years earlier. The biomarker values therefore did not represent measurements collected concurrently with the memory assessment. Although this temporal ordering is compatible with the possibility that earlier biological abnormalities are related to later memory performance, it does not establish that the observed biomarker differences caused the forgetting trajectories. Similarly, the MRI analyses estimated structural change retrospectively across three visits. These models identify differences in the rate of hippocampal change between trajectory classes, but they do not demonstrate that one-week forgetting predicts future atrophy or subsequent clinical decline.

The availability of biological and neuroimaging measures was incomplete and unequal across outcomes. Plasma p-tau217 was available for 309 participants, CSF A β 42/A β 40 for 286, CSF p-tau181/A β 42 for 240, and A/T classification for 234 participants. The comparison of tau positivity within the amyloid-positive subgroup included only 95 individuals. Although missingness in the principal fluid measures did not differ significantly across trajectory classes, the outcome-specific sample sizes reduce statistical precision, particularly for subgroup analyses and dichotomous biomarker comparisons. All models were based on complete cases and no values were imputed. Consequently, the possibility that unmeasured participant characteristics influenced the availability of particular assessments cannot be completely excluded.

The biological characterization was also necessarily incomplete. The study combined fluid biomarkers with structural MRI, but did not include tau-PET in the primary analyses or direct measures of several processes that may influence long-term memory, including sleep-dependent consolidation, synaptic dysfunction, neuroinflammation, vascular injury,

and cerebrovascular disease. Amyloid PET was available only for a subset of participants and was used as a binary secondary confirmation rather than as a continuous primary measure. The associations with plasma p-tau217 and the CSF p-tau181/A β 42 ratio should therefore not be interpreted as evidence that tau-related pathology is the only mechanism contributing to long-term forgetting. Conversely, the absence of a significant association with amyloid positivity does not demonstrate that amyloid is unrelated to the observed memory profiles.

The assessment procedure introduces additional measurement considerations. The one-week recall was administered by telephone, which increased feasibility but provided less control over environmental conditions than an in-person assessment. Participants may have differed in the presence of distractions, interruptions, or other contextual factors during the call. The delay was tightly controlled for most participants, although two assessments occurred outside the prespecified seven-to-ten-day window. More importantly, the same material was tested at several time points. Repeated retrieval may strengthen memory traces, and individuals may differ in the extent to which they benefit from prior testing. The protocol therefore cannot completely separate the effects of repeated retrieval from the underlying rate of forgetting. In addition, the present analyses focused on total recall and did not examine the recognition measure administered after the telephone assessment. It is consequently not possible to determine whether poor one-week performance primarily reflects impaired storage, inefficient retrieval, or both.

The trajectory methodology also requires careful interpretation. The five assessment points were indexed as equally spaced occasions, despite the fact that the intervals between them varied considerably, from minutes between immediate trials to approximately one week before the final assessment. The resulting trajectories therefore describe the shape of performance across the testing protocol rather than rates of forgetting per unit of time. In addition, latent classes are statistical approximations and should not be regarded as naturally occurring diagnostic categories. Their number and composition depend on the functional form, random-effects structure, starting values, and selection criteria used in the mixture model.

The minimum class-size threshold of 5% was substantively motivated but ultimately conventional. Although three classes were preferred for the primary outcome, some secondary trajectory analyses were better described by two classes according to the Bayesian information criterion. Three classes were nevertheless retained in the secondary analyses to preserve comparability across outcomes. This decision facilitates interpretation across

models, but it also means that the secondary classes should not be considered entirely independent empirical discoveries.

A further limitation is that class assignment was based on the modal posterior probability, and uncertainty surrounding individual classification was not propagated into the subsequent biomarker and neuroimaging models. The reported associations therefore concern assigned classes rather than fully corrected estimates of association with latent class membership. Although modal assignment generally attenuates associations with external variables, it may also obscure distinctions between closely related groups. Future analyses should consider three-step procedures or other approaches that incorporate classification uncertainty directly into the outcome models.

The longitudinal findings should be interpreted with some caution. The primary analyses used a random-intercept structure, whereas models including a random slope for time provided a better fit but produced wider standard errors. Although the estimated associations involving CA1 and total MTL volume were similar in direction and magnitude across the two specifications, they did not remain statistically significant after correction for multiple comparisons in the more complex models. Further studies using prospective data and appropriately specified longitudinal models will be needed to clarify the robustness and significance of this association.

Finally, the study was not designed to establish diagnostic accuracy, clinical utility, or prognostic value. No prospective clinical conversion endpoint was available, and the analyses cannot determine whether the Intermediate trajectory predicts subsequent cognitive decline or incident Alzheimer's disease. The identified trajectories should consequently not be interpreted as diagnostic categories, and the one-week assessment should not be used clinically without independent replication. The principal limitation is that the present design cannot establish whether the observed associations are reproducible, specific to AD pathology, or predictive of clinically meaningful outcomes. Future research should address these questions through prospective longitudinal studies incorporating contemporaneous multimodal biomarkers, recognition measures, independent samples, and clinically relevant endpoints.

8. Conclusion

This thesis examined whether latent trajectory analysis of immediate, 30-minute, and one-week delayed recall can characterize distinct patterns of learning and forgetting and identify cognitive profiles associated with early Alzheimer's disease pathology in cognitively unimpaired adults. The study was based on the premise that long-term memory performance should not be reduced to a single retention score. Instead, the relationship between acquisition, short-delay retention, and delayed recall may provide a more informative description of individual differences in memory consolidation.

The results support the value of modeling the complete learning-to-retention trajectory. The conventional 30-minute assessment showed limited discriminatory capacity in this cohort because more than half of the participants achieved the maximum total-recall score. By contrast, the one-week assessment substantially reduced the ceiling effect and revealed greater variability in performance. This finding indicates that delayed recall beyond the standard interval can expose differences in long-term retention that are not apparent during the earlier stages of the memory protocol.

Growth mixture modeling identified three distinct trajectory classes: Low, Intermediate, and High performers, demonstrating meaningful heterogeneity in both overall memory performance and the pattern of change across the assessment protocol. The Low-performing class showed lower performance from the learning trials through the one-week assessment, whereas the High-performing class consistently showed the strongest performance. The Intermediate class displayed relatively preserved acquisition and 30-minute recall but greater loss over the longer delay, illustrating how the one-week assessment can reveal aspects of memory performance that are not captured by conventional short-delay testing. Taken together, these findings indicate that modeling the full learning-to-retention trajectory provides a more differentiated characterization of memory function than relying on a single delayed-recall score or retention ratio.

The biological associations provide preliminary support for the relevance of these trajectories to early AD-related changes. Plasma p-tau₂₁₇ and the CSF p-tau₁₈₁/A β ₄₂ ratio were associated with trajectory class after correction for multiple comparisons. These associations were primarily driven by the Low-performing class, which showed higher biomarker levels than the High-performing class. In contrast, the Intermediate class did not differ significantly from the High-performing class on these measures, and amyloid

positivity was not significantly associated with trajectory class. The findings therefore do not support a simple interpretation in which the dissociated Intermediate profile represents a specific biomarker-defined stage of AD. Rather, they suggest that different aspects of the memory trajectory may relate differently to underlying biological processes.

Among the biological measures examined, plasma p-tau217 is particularly relevant from a translational perspective. Its association with memory trajectory links the cognitive findings to a blood-based marker of AD-related pathology. Because plasma p-tau217 can be measured from a blood sample without lumbar puncture or specialized imaging equipment, it has practical advantages over CSF-based assessment and may be more suitable for large-scale research and screening-oriented protocols, pending further validation. Its broad availability in the present study also supports its potential usefulness when combined with sensitive cognitive measures.

The neuroimaging analyses further suggest that the information captured by long-term memory trajectories may be related to longitudinal hippocampal change. The strongest association involved bilateral CA1 volume, for which both the Low-performing and Intermediate classes showed faster decline than the High-performing class. A weaker pattern was also observed for total MTL volume. In contrast, no corrected cross-sectional association was observed at the time of memory assessment, and no comparable association emerged for the other predefined cortical or hippocampal regions. These results are consistent with the possibility that long-term memory trajectories are more closely related to the rate of structural change.

The comparison between latent trajectories and the conventional ALF ratio showed that the two approaches provide complementary information. Both converged on the rate of CA1 atrophy, but they differed in their associations with fluid biomarkers and subjective cognitive complaints. The ALF ratio summarizes proportional loss relative to 30-minute recall, whereas latent trajectory analysis retains information about both overall performance and the shape of the learning-retention curve. This distinction is important because a participant with poor one-week recall may have either a generally weak memory profile or a specific difficulty maintaining information over the longer interval. Trajectory analysis can preserve this difference, whereas a single ratio may not fully represent it.

The findings also have practical implications for cognitive assessment. The one-week telephone recall was brief, required no additional in-person visit, and was completed by the large majority of scheduled participants. This supports its feasibility as an addition to an established neuropsychological protocol. Its most appropriate role, however, is likely to be

complementary. The extended assessment should not replace measures of learning and short-delay recall, but may provide additional information about long-term retention.

Overall, this thesis provides evidence that latent trajectory analysis of immediate, 30-minute, and one-week delayed recall can reveal distinct patterns of memory performance in cognitively unimpaired adults. The results support the potential of this approach as a candidate sensitive cognitive marker of early AD-related change, particularly when long-term memory trajectories are considered alongside plasma p-tau217 and longitudinal hippocampal measures. Nevertheless, the present findings establish associations rather than diagnostic or prognostic validity. The identified trajectory classes should not be interpreted as diagnostic categories, and the study does not determine whether the Intermediate profile predicts subsequent cognitive decline or clinical conversion.

Future research should replicate these findings in independent cohorts, use contemporaneous multimodal biomarkers, examine recognition performance alongside recall, and include prospective cognitive and clinical follow-up. Particular attention should be given to whether combined measures of plasma p-tau217 and long-term recall improve the prediction of short-term cognitive decline. This combination may be especially valuable because it integrates an accessible biological measure with a cognitive assessment designed to capture memory changes that remain undetected at shorter retention intervals.

In conclusion, the present study supports a multidimensional approach to the assessment of memory consolidation in preclinical AD. Modeling the complete trajectory from acquisition to one-week retention provides a more differentiated account of individual memory performance than a single short-delay or proportional retention score. The association with plasma p-tau217, together with the observed relationship between long-term memory trajectories and longitudinal CA1 change, provides a rationale for further investigating combined cognitive and blood-based markers. Their diagnostic and prognostic value remains to be established, but the present findings indicate that long-term delayed recall may contribute meaningfully to the detection and characterization of early Alzheimer's disease-related changes.

REFERENCES

- Ahn, H.-J., Seo, S. W., Chin, J., Suh, M. K., Lee, B. H., Kim, S. T., Im, K., Lee, J.-M., Lee, J. H., Heilman, K. M., & Na, D. L. (2011). The cortical neuroanatomy of neuropsychological deficits in mild cognitive impairment and Alzheimer's disease: A surface-based morphometric analysis. *Neuropsychologia*, 49(14), 3931–3945. <https://doi.org/10.1016/j.neuropsychologia.2011.10.010>
- Albert, M. S. (2011). Changes in cognition. *Neurobiology of Aging*, 32, S58–S63. <https://doi.org/10.1016/j.neurobiolaging.2011.09.010>
- Allen, R., Kemp, S., Morson, S., Wells, C., Grindheim, K., & Baddeley, A. (2019). Does telephone testing of long-term memory retention and forgetting influence performance in young and older adults? An examination using the Crimes Test. *The Neuropsychologist*, 8.
- Alonso, A. C., Zaidi, T., Grundke-Iqbal, I., & Iqbal, K. (1994). Role of abnormally phosphorylated tau in the breakdown of microtubules in Alzheimer disease. *Proceedings of the National Academy of Sciences*, 91(12), 5562–5566. <https://doi.org/10.1073/pnas.91.12.5562>
- Alzheimer, A. (1907). *Über eine eigenartige Erkrankung der Hirnrinde*. 64(1), 46–148.
- Andrew, M. K., & Tierney, M. C. (2018). The puzzle of sex, gender and Alzheimer's disease: Why are women more often affected than men? *Women's Health*, 14, 1745506518817995. <https://doi.org/10.1177/1745506518817995>
- Arriagada, P. V., Growdon, J. H., Hedley-Whyte, E. T., & Hyman, B. T. (1992). Neurofibrillary tangles but not senile plaques parallel duration and severity of Alzheimer's disease. *Neurology*, 42(3), 631–631. <https://doi.org/10.1212/WNL.42.3.631>
- Ashton, N. J., Brum, W. S., Di Molfetta, G., Benedet, A. L., Arslan, B., Jonaitis, E., Langhough, R. E., Cody, K., Wilson, R., Carlsson, C. M., Vanmechelen, E., Montoliu-Gaya, L., Lantero-Rodriguez, J., Rahmouni, N., Tissot, C., Stevenson, J., Servaes, S., Therriault, J., Pascoal, T., ... Zetterberg, H. (2024). Diagnostic Accuracy of a Plasma Phosphorylated Tau 217 Immunoassay for Alzheimer Disease Pathology. *JAMA Neurology*, 81(3), 255. <https://doi.org/10.1001/jamaneurol.2023.5319>
- Barbier, P., Zejneli, O., Martinho, M., Lasorsa, A., Belle, V., Smet-Nocca, C., Tsvetkov, P. O., Devred, F., & Landrieu, I. (2019). Role of Tau as a Microtubule-Associated

- Protein: Structural and Functional Aspects. *Frontiers in Aging Neuroscience*, 11, 204. <https://doi.org/10.3389/fnagi.2019.00204>
- Barcellos, L. F. B., Ferrari-Souza, J. P., De Jesus Vanni, I. J., Valer, M. A. A., Bieger, A., Leffa, D. T., Lussier, F. Z., Brum, W. S., Aguzzoli, C., Corin, A., Bastiani, M. A. D., Carello-Collar, G., Borelli, W. V., Rahmouni, N., Therriault, J., Trudel, L., Macedo, A. C., Ferreira, P. C. L., Povala, G., ... Zimmer, E. R. (2025). Imaging-based biological staging associates with brain atrophy in preclinical Alzheimer's disease. *Alzheimer's & Dementia*, 21(S2), e107663. https://doi.org/10.1002/alz70856_107663
- Bartels, C., Tzu-Yueh Chen, J., Belz, M., Yakupov, R., Düzel, E., Glanz, W., Lüsebrink, F., Dechent, P., Kleineidam, L., Stark, M., Spottke, A., Coenjaerts, M., Fließbach, K., Schneider, A., Rostamzadeh, A., Jessen, F., H. Schott, B., Wiltfang, J., Frommann, I., ... Goya-Maldonado, R. (2025). Long-term retrieval performance is associated with CA1 hippocampal volume in older adults and individuals at risk for dementia. *Alzheimer's Research & Therapy*, 17(1), 195. <https://doi.org/10.1186/s13195-025-01833-4>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods*, 8(3), 338. <https://doi.org/doi.org/10.1037/1082-989X.8.3.338>
- Bellenguez, C., Küçükali, F., Jansen, I. E., Kleineidam, L., Moreno-Grau, S., Amin, N., Naj, A. C., Campos-Martin, R., Grenier-Boley, B., & Andrade, V. (2022). New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nature Genetics*, 54(4), 412–436. <https://doi.org/doi.org/10.1038/s41588-022-01024-z>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bennett, D. A., Schneider, J. A., Wilson, R. S., Bienias, J. L., & Arnold, S. E. (2004). Neurofibrillary Tangles Mediate the Association of Amyloid Load With Clinical Alzheimer Disease and Level of Cognitive Function. *Archives of Neurology*, 61(3), 378. <https://doi.org/10.1001/archneur.61.3.378>

- Blake, R. V., Wroe, S. J., Breen, E. K., & McCarthy, R. A. (2000). Accelerated forgetting in patients with epilepsy: Evidence for an impairment in memory consolidation. *Brain*, 123(3), 472–483. <https://doi.org/10.1093/brain/123.3.472>
- Bolck, A., Croon, M., & Hagenaaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, 12(1), 3–27. <https://doi.org/10.1093/pan/mp001>
- Bonete-López, B., Oltra-Cucarella, J., Marín, M., Antón, C., Balao, N., López, E., & Macià, E. S. (2020). *Validation and Norms for a Recognition Task for the Spanish Version of the Free and Cued Selective Reminding Test*. <https://doi.org/doi.org/10.1093/arclin/aaa117>
- Braak, H., & Braak, E. (1991). Demonstration of Amyloid Deposits and Neurofibrillary Changes in Whole Brain Sections. *Brain Pathology*, 1(3), 213–216. <https://doi.org/10.1111/j.1750-3639.1991.tb00661.x>
- Brugulat-Serrat, A., Cañas-Martínez, A., Canals-Gispert, L., Marne, P., Gramunt, N., Milà-Alomà, M., Suárez-Calvet, M., Arenaza-Urquijo, E. M., Grau-Rivera, O., González-de-Echávarri, J. M., Minguillon, C., Fauria, K., Kollmorgen, G., Suridjan, I., Zetterberg, H., Blennow, K., Gispert, J. D., Molinuevo, J. L., Sánchez-Benavides, G., & ALFA study. (2021). Enhancing the Sensitivity of Memory Tests: Reference Data for the Free and Cued Selective Reminding Test and the Logical Memory Task from Cognitively Healthy Subjects with Normal Alzheimer’s Disease Cerebrospinal Fluid Biomarker Levels. *Journal of Alzheimer’s Disease: JAD*, 84(1), 119–128. <https://doi.org/10.3233/JAD-210640>
- Buongiorno, M., Sánchez-Benavides, G., Caruana, G., Elias-Mas, A., Artero, C., Cullell, N., Delgado, P., Giraldo, D. M., Marzal-Espí, C., Grau-Rivera, O., de la Sierra, A., Delgado-Sanchez, A., Ray, N. J., & Krupinski, J. (2025). Abnormal sleep blood pressure patterns are associated with the diffusion tensor imaging along the perivascular space index in cognitively impaired individuals. *Frontiers in Aging Neuroscience*, 17. <https://doi.org/10.3389/fnagi.2025.1578270>
- Busche, M. A., & Hyman, B. T. (2020). Synergy between amyloid- β and tau in Alzheimer’s disease. *Nature Neuroscience*, 23(10), 1183–1193. <https://doi.org/10.1038/s41593-020-0687-6>
- Buschke, H. (1984). Cued recall in Amnesia. *Journal of Clinical Neuropsychology*, 6(4), 433–440. <https://doi.org/10.1080/01688638408401233>
- Buschke, H. (2014). Rationale of the memory binding test. *Dementia and Memory*, 55–71.

- Buschke, H., Kuslansky, G., Katz, M., Stewart, W. F., Sliwinski, M. J., Eckholdt, H. M., & Lipton, R. B. (1999). Screening for dementia with the Memory Impairment Screen. *Neurology*, 52(2), 231–231. <https://doi.org/10.1212/WNL.52.2.231>
- Butler, C. R., Muhlert, N., & Zeman, A. (2010). Accelerated long-term forgetting. In *Forgetting* (pp. 225–252). Psychology Press. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780203851647-15/accelerated-long-term-forgetting-christopher-butler-nils-muhlert-adam-zeman>
- Butler, C. R., & Zeman, A. Z. (2008). Recent insights into the impairment of memory in epilepsy: Transient epileptic amnesia, accelerated long-term forgetting and remote memory impairment. *Brain*, 131(9), 2243–2263. <https://doi.org/10.1093/brain/awn127>
- Cacciaglia, R., Shekari, M., Salvadó, G., Milà-Alomà, M., Falcon, C., Sánchez-Benavides, G., Minguillón, C., Fauria, K., Grau-Rivera, O., & Molinuevo, J. L. (2025). The CSF p-tau/ β -amyloid 42 ratio correlates with brain structure and fibrillary β -amyloid deposition in cognitively unimpaired individuals at the earliest stages of pre-clinical Alzheimer's disease. *Brain Communications*, 7(1), fcae451. <https://doi.org/doi.org/10.1093/braincomms/fcae451>
- Carvalho, D. Z., Kolla, B. P., McCarter, S. J., St. Louis, E. K., Machulda, M. M., Przybelski, S. A., Fought, A. J., Lowe, V. J., Somers, V. K., Boeve, B. F., Petersen, R. C., Jack, C. R., Graff-Radford, J., Varga, A. W., & Vemuri, P. (2025). Associations of Chronic Insomnia, Longitudinal Cognitive Outcomes, Amyloid-PET, and White Matter Changes in Cognitively Normal Older Adults. *Neurology*, 105(7), e214155. <https://doi.org/10.1212/WNL.00000000000214155>
- Cerami, C., Dubois, B., Boccardi, M., Monsch, A. U., Demonet, J. F., & Cappa, S. F. (2017). Clinical validity of delayed recall tests as a gateway biomarker for Alzheimer's disease in the context of a structured 5-phase development framework. *Neurobiology of Aging*, 52, 153–166. <https://doi.org/10.1016/j.neurobiolaging.2016.03.034>
- Chapman, S., Sunderaraman, P., Joyce, J. L., Azar, M., Colvin, L. E., Barker, M. S., McKeague, I., Kreisl, W. C., & Cosentino, S. (2021). Optimizing Subjective Cognitive Decline to Detect Early Cognitive Dysfunction. *Journal of Alzheimer's Disease*, 80(3), 1185–1196. <https://doi.org/10.3233/JAD-201322>
- Chen, Y., & Yu, Y. (2023). Tau and neuroinflammation in Alzheimer's disease: Interplay mechanisms and clinical translation. *Journal of Neuroinflammation*, 20(1), 165. <https://doi.org/10.1186/s12974-023-02853-3>

- Condello, C., Yuan, P., Schain, A., & Grutzendler, J. (2015). Microglia constitute a barrier that prevents neurotoxic protofibrillar A β 42 hotspots around plaques. *Nature Communications*, 6(1), 6176. <https://doi.org/10.1038/ncomms7176>
- Cowman, D., Langhough, R., Olson, H., Basche, K., Sanson-Miles, L., Bruno, D., Hermann, B., Christian, B. T., Betthausen, T. J., Johnson, S. C., & Mueller, K. D. (2025). Sex differences in story recall decline in preclinical Alzheimer's disease. *Brain Communications*, 7(3), fcaf169. <https://doi.org/10.1093/braincomms/fcaf169>
- Craig-Schapiro, R. (2012). *Identification of Novel Fluid Biomarkers for Alzheimer's Disease*. <https://doi.org/10.7936/K7PR7T12>
- Cumplido-Mayoral, I., García-Prat, M., Operto, G., Falcon, C., Shekari, M., Cacciaglia, R., Milà-Alomà, M., Lorenzini, L., Ingala, S., Meije Wink, A., Mutsaerts, H. J., Minguillón, C., Fauria, K., Molinuevo, J. L., Haller, S., Chetelat, G., Waldman, A., Schwarz, A. J., Barkhof, F., ... OASIS study. (2023). Biological brain age prediction using machine learning on structural neuroimaging data: Multi-cohort validation against biomarkers of Alzheimer's disease and neurodegeneration stratified by sex. *eLife*, 12, e81067. <https://doi.org/10.7554/eLife.81067>
- Cumplido-Mayoral, I., Sánchez-Benavides, G., Vilor-Tejedor, N., López-Martos, D., Brugulat-Serrat, A., Milà-Alomà, M., Falcon, C., Cacciaglia, R., Minguillón, C., Fauria, K., Kollmorgen, G., Quijano-Rubio, C., Molinuevo, J. L., Grau-Rivera, O., Suárez-Calvet, M., Vilaplana, V., & Gispert, J. D. (2025). Neuroimaging-derived biological brain age and its associations with glial reactivity and synaptic dysfunction cerebrospinal fluid biomarkers. *Molecular Psychiatry*, 30(8), 3718–3728. <https://doi.org/10.1038/s41380-025-02961-x>
- De Flores, R., La Joie, R., & Chételat, G. (2015). Structural imaging of hippocampal subfields in healthy aging and Alzheimer's disease. *Neuroscience*, 309, 29–50. <https://doi.org/10.1016/j.neuroscience.2015.08.033>
- Degenszajn, J., Caramelli, P., Caixeta, L., & Nitrini, R. (2001). Encoding process in delayed recall impairment and rate of forgetting in Alzheimer's disease. *Arquivos de Neuro-Psiquiatria*, 59, 171–174. <https://doi.org/10.1590/S0004-282X2001000200003>
- Del Ser, T., González-Montalvo, J.-I., Martínez-Espinosa, S., Delgado-Villapalos, C., & Bermejo, F. (1997). Estimation of Premorbid Intelligence in Spanish People with the Word Accentuation Test and Its Application to the Diagnosis of Dementia. *Brain and Cognition*, 33(3), 343–356. <https://doi.org/10.1006/brcg.1997.0877>

- Della Sala, S., Baddeley, A., Peng, N., & Logie, R. (2024). Assessing long-term forgetting: A pragmatic approach. *Cortex*, 170, 80–89. <https://doi.org/10.1016/j.cortex.2023.11.009>
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3), 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021>
- 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*, 19(3), 497–510. <https://doi.org/10.1093/cercor/bhn113>
- Dubois, B., Hampel, H., Feldman, H. H., Scheltens, P., Aisen, P., Andrieu, S., Bakardjian, H., Benali, H., Bertram, L., Blennow, K., Broich, K., Cavedo, E., Crutch, S., Dartigues, J., Duyckaerts, C., Epelbaum, S., Frisoni, G. B., Gauthier, S., Genthon, R., ... Proceedings of the Meeting of the International Working Group (IWG) and the American Alzheimer’s Association on “The Preclinical State of AD”; July 23, 2015; Washington DC, USA. (2016). Preclinical Alzheimer’s disease: Definition, natural history, and diagnostic criteria. *Alzheimer’s & Dementia*, 12(3), 292–323. <https://doi.org/10.1016/j.jalz.2016.02.002>
- Dubois, B., Villain, N., Schneider, L., Fox, N., Campbell, N., Galasko, D., Kivipelto, M., Jessen, F., Hanseeuw, B., Boada, M., Barkhof, F., Nordberg, A., Froelich, L., Waldemar, G., Frederiksen, K. S., Padovani, A., Planche, V., Rowe, C., Bejanin, A., ... Frisoni, G. B. (2024). Alzheimer Disease as a Clinical-Biological Construct—An International Working Group Recommendation. *JAMA Neurology*, 81(12), 1304. <https://doi.org/10.1001/jamaneurol.2024.3770>
- Elliott, G., Isaac, C. L., & Muhlert, N. (2014). Measuring forgetting: A critical review of accelerated long-term forgetting studies. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 54(100), 16–32. <https://doi.org/10.1016/j.cortex.2014.02.001>

- Emrani, S., & Sundermann, E. E. (2025). Sex/gender differences in the clinical trajectory of Alzheimer's disease: Insights into diagnosis and cognitive reserve. *Frontiers in Neuroendocrinology*, 77, 101184. <https://doi.org/10.1016/j.yfrne.2025.101184>
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143–149. <https://doi.org/10.3758/BF03203267>
- Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62(2), 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>
- Fisher, D. W., Bennett, D. A., & Dong, H. (2018). Sexual dimorphism in predisposition to Alzheimer's disease. *Neurobiology of Aging*, 70, 308–324. <https://doi.org/10.1016/j.neurobiolaging.2018.04.004>
- Fletcher, E., Gavett, B., Crane, P., Soldan, A., Hohman, T., Farias, S., Widaman, K., Groot, C., Renteria, M. A., Zahodne, L., DeCarli, C., Mungas, D., & for the Alzheimer's Disease Neuroimaging Initiative. (2021). A robust brain signature region approach for episodic memory performance in older adults. *Brain*, 144(4), 1089–1102. <https://doi.org/10.1093/brain/awab007>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". *Journal of Psychiatric Research*, 12(3), 189–198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
- García-Martínez, M., Sánchez-Juan, P., & Butler, C. R. (2023). A review of accelerated long-term forgetting in Alzheimer's disease: Current situation and prospects. *Neuropsychology*, 37(6), 673–682. <https://doi.org/10.1037/neu0000827>
- Gatz, M., Reynolds, C. A., Fratiglioni, L., Johansson, B., Mortimer, J. A., Berg, S., Fiske, A., & Pedersen, N. L. (2006). Role of genes and environments for explaining Alzheimer disease. *Archives of General Psychiatry*, 63(2), 168–174. <https://doi.org/10.1001/archpsyc.63.2.168>
- Gauthier, S. (2021). Functional assessment. *Alzheimer's Disease International, Editor. World Alzheimer Report*. <https://www.alzint.org/u/World-Alzheimer-Report-2021-Chapter-04.pdf>
- Geurts, S., van der Werf, S. P., & Kessels, R. P. C. (2015). Accelerated forgetting? An evaluation on the use of long-term forgetting rates in patients with memory problems. *Frontiers in Psychology*, 6, 752. <https://doi.org/10.3389/fpsyg.2015.00752>
- Giannakopoulos, P., Herrmann, F. R., Bussi re, T., Bouras, C., K vari, E., Perl, D. P., Morrison, J. H., Gold, G., & Hof, P. R. (2003). Tangle and neuron numbers, but not

- amyloid load, predict cognitive status in Alzheimer's disease. *Neurology*, 60(9), 1495–1500. <https://doi.org/10.1212/01.WNL.0000063311.58879.01>
- Goldberg, D. P., Bridges, K., Duncan-Jones, P., & Grayson, D. (1987). Dimensions of neuroses seen in primary-care settings. *Psychological Medicine*, 17(2), 461–470. <https://doi.org/10.1017/S0033291700025022>
- Gómez-Isla, T., Hollister, R., West, H., Mui, S., Growdon, J. H., Petersen, R. C., Parisi, J. E., & Hyman, B. T. (1997). Neuronal loss correlates with but exceeds neurofibrillary tangles in Alzheimer's disease. *Annals of Neurology*, 41(1), 17–24. <https://doi.org/10.1002/ana.410410106>
- Gordon, B. A., Blazey, T. M., Su, Y., Hari-Raj, A., Dincer, A., Flores, S., Christensen, J., McDade, E., Wang, G., Xiong, C., Cairns, N. J., Hassenstab, J., Marcus, D. S., Fagan, A. M., Jack, C. R., Hornbeck, R. C., Paumier, K. L., Ances, B. M., Berman, S. B., ... Benzinger, T. L. S. (2018). Spatial patterns of neuroimaging biomarker change in individuals from families with autosomal dominant Alzheimer's disease: A longitudinal study. *The Lancet Neurology*, 17(3), 241–250. [https://doi.org/10.1016/S1474-4422\(18\)30028-0](https://doi.org/10.1016/S1474-4422(18)30028-0)
- Grau-Guinea, L., Pérez Enríquez, C., García-Escobar, G., Arrondo Elizarán, C., Pereira Cutiño, B., Florido Santiago, M., Piqué Candini, J., Planas, A., Paez, M., Peña Casanova, J., & Sánchez-Benavides, G. (2021). Development, equivalence study, and normative data of version B of the Spanish-language Free and Cued Selective Reminding Test. *Neurologia*, 36(5), 353–360. <https://doi.org/10.1016/j.nrleng.2018.02.001>
- Grober, E., & Buschke, H. (1987). Genuine memory deficits in dementia. *Developmental Neuropsychology*, 3(1), 13–36. <https://doi.org/10.1080/87565648709540361>
- Guillozet, A. L., Weintraub, S., Mash, D. C., & Mesulam, M. M. (2003). Neurofibrillary tangles, amyloid, and memory in aging and mild cognitive impairment. *Archives of Neurology*, 60(5), 729–736. <https://doi.org/10.1001/archneur.60.5.729>
- Gutierrez, B. A., & Limon, A. (2022). Synaptic Disruption by Soluble Oligomers in Patients with Alzheimer's and Parkinson's Disease. *Biomedicines*, 10(7), 1743. <https://doi.org/10.3390/biomedicines10071743>
- Hari, E., Kurt, E., Ulasoglu-Yildiz, C., Bayram, A., Bilgic, B., Demiralp, T., & Gurvit, H. (2023). Morphometric analysis of medial temporal lobe subregions in Alzheimer's disease using high-resolution MRI. *Brain Structure and Function*, 228(8), 1885–1899. <https://doi.org/10.1007/s00429-023-02683-2>

- He, C. Y., Zhou, Z., Kan, M. M., Chan, D. H., Wong, A. C., Mok, K. H., Lam, F. M., Chan, S. C., Cheung, C. K., & Yeung, M. K. (2024). Modifiable risk factors for mild cognitive impairment among cognitively normal community-dwelling older adults: A systematic review and meta-analysis. *Ageing Research Reviews*, 99, 102350. <https://doi.org/10.1016/j.arr.2024.102350>
- Hipp, J. R., & Bauer, D. J. (2006). 'Local Solutions in the Estimation of Growth Mixture Models': Correction to Hipp and Bauer (2006). <https://doi.org/10.1037/1082-989X.11.3.305>
- Hirnstien, M., Stuebs, J., Moè, A., & Hausmann, M. (2023). Sex/Gender Differences in Verbal Fluency and Verbal-Episodic Memory: A Meta-Analysis. *Perspectives on Psychological Science*, 18(1), 67–90. <https://doi.org/10.1177/17456916221082116>
- Huang, Y., & Mahley, R. W. (2014). Apolipoprotein E: Structure and function in lipid metabolism, neurobiology, and Alzheimer's diseases. *Neurobiology of Disease*, 72, 3–12. <https://doi.org/10.1016/j.nbd.2014.08.025>
- Hughes, D., Judge, C., Murphy, R., Loughlin, E., Costello, M., Whiteley, W., Bosch, J., O'Donnell, M. J., & Canavan, M. (2020). Association of blood pressure lowering with incident dementia or cognitive impairment: A systematic review and meta-analysis. *Jama*, 323(19), 1934–1944. <https://doi.org/10.1001/jama.2020.4249>
- Huguet, J., Falcon, C., Fusté, D., Girona, S., Vicente, D., Molinuevo, J. L., Gisbert, J. D., Operto, G., & for the ALFA Study. (2021). Management and Quality Control of Large Neuroimaging Datasets: Developments From the Barcelonaβeta Brain Research Center. *Frontiers in Neuroscience*, 15, 633438. <https://doi.org/10.3389/fnins.2021.633438>
- Hyman, B. T., Phelps, C. H., Beach, T. G., Bigio, E. H., Cairns, N. J., Carrillo, M. C., Dickson, D. W., Duyckaerts, C., Frosch, M. P., Masliah, E., Mirra, S. S., Nelson, P. T., Schneider, J. A., Thal, D. R., Thies, B., Trojanowski, J. Q., Vinters, H. V., & Montine, T. J. (2012). National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's & Dementia*, 8(1), 1–13. <https://doi.org/10.1016/j.jalz.2011.10.007>
- Im, D., & Choi, T. S. (2024). Distinctive contribution of two additional residues in protein aggregation of Aβ42 and Aβ40 isoforms. *BMB Reports*, 57(6), 263–272. <https://doi.org/10.5483/BMBRep.2024-0044>
- Jack, C. R., Andrews, J. S., Beach, T. G., Buracchio, T., Dunn, B., Graf, A., Hansson, O., Ho, C., Jagust, W., McDade, E., Molinuevo, J. L., Okonkwo, O. C., Pani, L., Rafii, M. S.,

- Scheltens, P., Siemers, E., Snyder, H. M., Sperling, R., Teunissen, C. E., & Carrillo, M. C. (2024). Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimer's & Dementia*, 20(8), 5143–5169. <https://doi.org/10.1002/alz.13859>
- 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., ... Silverberg, N. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14(4), 535–562. <https://doi.org/10.1016/j.jalz.2018.02.018>
- Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Feldman, H. H., Frisoni, G. B., Hampel, H., Jagust, W. J., Johnson, K. A., Knopman, D. S., Petersen, R. C., Scheltens, P., Sperling, R. A., & Dubois, B. (2016). A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology*, 87(5), 539–547. <https://doi.org/10.1212/WNL.0000000000002923>
- Jack, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., Shaw, L. M., Vemuri, P., Wiste, H. J., & Weigand, S. D. (2013). Tracking pathophysiological processes in Alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*, 12(2), 207–216. [https://doi.org/10.1016/S1474-4422\(12\)70291-0](https://doi.org/10.1016/S1474-4422(12)70291-0)
- Janelidze, S., Stomrud, E., Smith, R., Palmqvist, S., Mattsson, N., Airey, D. C., Proctor, N. K., Chai, X., Shcherbinin, S., & Sims, J. R. (2020). Cerebrospinal fluid p-tau217 performs better than p-tau181 as a biomarker of Alzheimer's disease. *Nature Communications*, 11(1), 1683. <https://doi.org/10.1038/s41467-020-15436-0>
- Jankowsky, J. L., Fadale, D. J., Anderson, J., Xu, G. M., Gonzales, V., Jenkins, N. A., Copeland, N. G., Lee, M. K., Younkin, L. H., Wagner, S. L., Younkin, S. G., & Borchelt, D. R. (2004). Mutant presenilins specifically elevate the levels of the 42 residue β -amyloid peptide in vivo: Evidence for augmentation of a 42-specific γ secretase. *Human Molecular Genetics*, 13(2), 159–170. <https://doi.org/10.1093/hmg/ddh019>
- Jansen, M. G., Geerligs, L., Claassen, J. A. H. R., Overdorp, E. J., Brazil, I. A., Kessels, R. P. C., & Oosterman, J. M. (2021). Positive Effects of Education on Cognitive Functioning Depend on Clinical Status and Neuropathological Severity. *Frontiers in Human Neuroscience*, 15, 723728. <https://doi.org/10.3389/fnhum.2021.723728>

- Jansen, W. J., Ossenkoppele, R., Knol, D. L., Tijms, B. M., Scheltens, P., Verhey, F. R., Visser, P. J., Group, A. B. S., Aalten, P., & Aarsland, D. (2015). Prevalence of cerebral amyloid pathology in persons without dementia: A meta-analysis. *Jama*, 313(19), 1924–1938. <https://doi.org/10.1001/jama.2015.4668>
- Jessen, F., Amariglio, R. E., Buckley, R. F., van der Flier, W. M., Han, Y., Molinuevo, J. L., Rabin, L., Rentz, D. M., Rodriguez-Gomez, O., & Saykin, A. J. (2020). The characterisation of subjective cognitive decline. *The Lancet Neurology*, 19(3), 271–278. [https://doi.org/10.1016/S1474-4422\(19\)30368-0](https://doi.org/10.1016/S1474-4422(19)30368-0)
- Jia, J., Ning, Y., Chen, M., Wang, S., Yang, H., Li, F., Ding, J., Li, Y., Zhao, B., Lyu, J., Yang, S., Yan, X., Wang, Y., Qin, W., Wang, Q., Li, Y., Zhang, J., Liang, F., Liao, Z., & Wang, S. (2024). Biomarker Changes during 20 Years Preceding Alzheimer's Disease. *New England Journal of Medicine*, 390(8), 712–722. <https://doi.org/10.1056/NEJMoa2310168>
- Jones, A., Ali, M. U., Mayhew, A., Aryal, K., Correia, R. H., Dash, D., Manis, D. R., Rehman, A., O'Connell, M. E., & Taler, V. (2025). Environmental risk factors for all-cause dementia, Alzheimer's disease dementia, vascular dementia, and mild cognitive impairment: An umbrella review and meta-analysis. *Environmental Research*, 270, 121007. <https://doi.org/doi.org/10.1016/j.envres.2025.121007>
- Jung, T., & Wickrama, K. A. S. (2008). An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. *Social and Personality Psychology Compass*, 2(1), 302–317. <https://doi.org/10.1111/j.1751-9004.2007.00054.x>
- Jutten, R. J., Sikkes, S. A. M., Amariglio, R. E., Buckley, R. F., Properzi, M. J., Marshall, G. A., Rentz, D. M., Johnson, K. A., Teunissen, C. E., Van Berckel, B. N. M., Van Der Flier, W. M., Scheltens, P., Sperling, R. A., Papp, K. V., for the Alzheimer Disease Neuroimaging Initiative, & National Alzheimer's Coordinating Center, the Harvard Aging Brain Study, and the Alzheimer Dementia Cohort. (2021). Identifying Sensitive Measures of Cognitive Decline at Different Clinical Stages of Alzheimer's Disease. *Journal of the International Neuropsychological Society*, 27(5), 426–438. <https://doi.org/10.1017/S1355617720000934>
- Kapur, N., Millar, J., Colbourn, C., Abbott, P., Kennedy, P., & Docherty, T. (1997). Very Long-Term Amnesia in Association with Temporal Lobe Epilepsy: Evidence for Multiple-Stage Consolidation Processes. *Brain and Cognition*, 35(1), 58–70. <https://doi.org/10.1006/brcg.1997.0927>

- Karikari, T. K., Ashton, N. J., Brinkmalm, G., Brum, W. S., Benedet, A. L., Montoliu-Gaya, L., Lantero-Rodriguez, J., Pascoal, T. A., Suarez-Calvet, M., & Rosa-Neto, P. (2022). Blood phospho-tau in Alzheimer disease: Analysis, interpretation, and clinical utility. *Nature Reviews Neurology*, 18(7), 400–418. <https://doi.org/doi.org/10.1038/s41582-022-00665-2>
- Kenward, M. G., & Roger, J. H. (1997). Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. *Biometrics*, 53(3), 983. <https://doi.org/10.2307/2533558>
- Kergoat, M., Villalpando, J., Leclerc, B., Le, M. T., Hudon, C., & Bolduc, A. (2025). A comparaisn of clinical diagnostic classification criteria used in longitudinal cohort studies of the Alzheimer’s disease continuum: A systematic review. *Alzheimer’s & Dementia*, 21(S6), e101123. https://doi.org/10.1002/alz70860_101123
- Khalili, E., Sodaei, F., Noroozian, M., Sheldrick-Michel, T. M., & Vafaei, M. S. (2026). Structural MRI in frontotemporal dementia and Alzheimer’s disease: Stage-dependent atrophy patterns. *Journal of Neural Transmission*. <https://doi.org/10.1007/s00702-026-03158-4>
- Killiany, R. J., Hyman, B. T., Gomez-Isla, T., Moss, M. B., Kikinis, R., Jolesz, F., Tanzi, R., Jones, K., & Albert, M. S. (2002). MRI measures of entorhinal cortex vs hippocampus in preclinical AD. *Neurology*, 58(8), 1188–1196. (world). <https://doi.org/10.1212/WNL.58.8.1188>
- Kim, M., & Bezprozvanny, I. (2021). Conformational Models of APP Processing by Gamma Secretase Based on Analysis of Pathogenic Mutations. *International Journal of Molecular Sciences*, 22(24), 13600. <https://doi.org/10.3390/ijms222413600>
- 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>
- Knopman, D. S., Amieva, H., Petersen, R. C., Chételat, G., Holtzman, D. M., Hyman, B. T., Nixon, R. A., & Jones, D. T. (2021). Alzheimer disease. *Nature Reviews Disease Primers*, 7(1), 33. <https://doi.org/10.1038/s41572-021-00269-y>
- Kobayashi, G., Hirata, K., Ono, M., Kasuga, K., & Takado, Y. (2026). The 2024 NIA-AA biological definition of Alzheimer’s disease: Linking biomarkers to clinical practice. *Frontiers in Dementia*, 5, 1736297. <https://doi.org/10.3389/frdem.2026.1736297>

- Koychev, I., Hofer, M., & Friedman, N. (2020). Correlation of Alzheimer Disease Neuropathologic Staging with Amyloid and Tau Scintigraphic Imaging Biomarkers. *Journal of Nuclear Medicine*, 61(10), 1413–1418. <https://doi.org/10.2967/jnumed.119.230458>
- Kueper, J. K., Speechley, M., & Montero-Odasso, M. (2018). The Alzheimer's Disease Assessment Scale–Cognitive Subscale (ADAS-Cog): Modifications and Responsiveness in Pre-Dementia Populations. A Narrative Review. *Journal of Alzheimer's Disease*, 63(2), 423–444. <https://doi.org/10.3233/JAD-170991>
- Kuźma, E., Littlejohns, T. J., Khawaja, A. P., Llewellyn, D. J., Ukoumunne, O. C., & Thiem, U. (2021). Visual Impairment, Eye Diseases, and Dementia Risk: A Systematic Review and Meta-Analysis. *Journal of Alzheimer's Disease*, 83(3), 1073–1087. <https://doi.org/10.3233/JAD-210250>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13). <https://doi.org/10.18637/jss.v082.i13>
- LaPoint, M. R., Chhatwal, J. P., Sepulcre, J., Johnson, K. A., Sperling, R. A., & Schultz, A. P. (2017). The association between tau PET and retrospective cortical thinning in clinically normal elderly. *NeuroImage*, 157, 612–622. <https://doi.org/10.1016/j.neuroimage.2017.05.049>
- Lenth, R. V., & Piaskowski, J. (2025). emmeans: Estimated marginal means, aka least-squares means. *R Package Version*, 1(2).
- Leung, K. K., Barnes, J., Ridgway, G. R., Bartlett, J. W., Clarkson, M. J., Macdonald, K., Schuff, N., Fox, N. C., & Ourselin, S. (2010). Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. *NeuroImage*, 51(4), 1345–1359. <https://doi.org/10.1016/j.neuroimage.2010.03.018>
- Lewczuk, P., Matzen, A., Blennow, K., Parnetti, L., Molinuevo, J. L., Eusebi, P., Kornhuber, J., Morris, J. C., & Fagan, A. M. (2016). Cerebrospinal Fluid A $\beta_{42/40}$ Corresponds Better than A β_{42} to Amyloid PET in Alzheimer's Disease. *Journal of Alzheimer's Disease*, 55(2), 813–822. <https://doi.org/10.3233/JAD-160722>
- Ley 14/2007, de 3 de julio, de Investigación biomédica, Boletín Oficial del Estado, núm. 159, 4 de julio de 2007, pp. 28826–28848 (España). (2007).
- Li, Q., Ren, A., Cui, L., Sun, X., Zhong, Y., Wang, B., Miao, Y., & Guo, Q. (2020). Associations Between Cerebral Morphology and Objective Memory Performance in

- Patients With Subjective Cognitive Decline*. <https://doi.org/doi.org/10.21203/rs.3.rs-55458/v1>
- Li, S., & Selkoe, D. J. (2020). A mechanistic hypothesis for the impairment of synaptic plasticity by soluble A β oligomers from Alzheimer's brain. *Journal of Neurochemistry*, 154(6), 583–597. <https://doi.org/10.1111/jnc.15007>
- Lindbergh, C. A., Walker, N., La Joie, R., Weiner-Light, S., Staffaroni, A. M., Casaletto, K. B., Elahi, F., Walters, S. M., You, M., Cotter, D., Asken, B., Apple, A. C., Tsoy, E., Neuhaus, J., Fonseca, C., Wolf, A., Cobigo, Y., Rosen, H., Kramer, J. H., & Hillblom Aging Network. (2021). Worth the Wait: Delayed Recall after 1 Week Predicts Cognitive and Medial Temporal Lobe Trajectories in Older Adults. *Journal of the International Neuropsychological Society: JINS*, 27(4), 382–388. <https://doi.org/10.1017/S1355617720001009>
- Liu, C.-C., Kanekiyo, T., Xu, H., & Bu, G. (2013). Apolipoprotein E and Alzheimer disease: Risk, mechanisms and therapy. *Nature Reviews Neurology*, 9(2), 106–118. <https://doi.org/10.1038/nrneurol.2012.263>
- Livingston, G., Huntley, J., Liu, K. Y., Costafreda, S. G., Selbæk, G., Alladi, S., Ames, D., Banerjee, S., Burns, A., & Brayne, C. (2024). Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *The Lancet*, 404(10452), 572–628.
- Lu, K., Baker, J., Nicholas, J. M., Street, R. E., Keuss, S. E., Coath, W., James, S.-N., Keshavan, A., Weston, P. S., & Murray-Smith, H. (2025). Associations between accelerated forgetting, amyloid deposition and brain atrophy in older adults. *Brain*, 148(4), 1302–1315. <https://doi.org/10.1093/brain/awae316>
- Ma, J., Zheng, M.-X., Wu, J.-J., Xing, X.-X., Xiang, Y.-T., Wei, D., Xue, X., Zhang, H., Hua, X.-Y., Guo, Q.-H., & Xu, J.-G. (2023). Mapping the long-term delayed recall-based cortex-hippocampus network constrained by the structural and functional connectome: A case-control multimodal MRI study. *Alzheimer's Research & Therapy*, 15(1), 61. <https://doi.org/10.1186/s13195-023-01197-7>
- Maass, A., Lockhart, S. N., Harrison, T. M., Bell, R. K., Mellinger, T., Swinnerton, K., Baker, S. L., Rabinovici, G. D., & Jagust, W. J. (2018). Entorhinal Tau Pathology, Episodic Memory Decline, and Neurodegeneration in Aging. *The Journal of Neuroscience*, 38(3), 530–543. <https://doi.org/10.1523/JNEUROSCI.2028-17.2017>
- Magill, N., Tank, R., Ferguson, D., Alston, D., O'Connor, A., Rice, H., Lu, K., Crutch, S., Fox, N. C., & Weston, P. S. (2026). Accelerated long-term forgetting as a predictor

- of clinical onset in presymptomatic autosomal dominant Alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 97(6), 493–496. <https://doi.org/10.1136/jnnp-2025-336750>
- Manes, F., Serrano, C., Calcagno, M. L., Cardozo, J., & Hodges, J. (2008). Accelerated forgetting in subjects with memory complaints. A new form of Mild Cognitive Impairment? *Journal of Neurology*, 255(7), 1067–1070. <https://doi.org/10.1007/s00415-008-0850-6>
- Manoli, R., Chartaux-Danjou, L., Mecheri, H., Delecroix, H., Noulhiane, M., & Moroni, C. (2018). Extension to one week of verbal memory consolidation assessment. *Neuropsychology*, 32(3), 366–373. <https://doi.org/10.1037/neu0000403>
- Mayes, A. R., Hunkin, N. M., Isaac, C., & Muhlert, N. (2019). Are there distinct forms of accelerated forgetting and, if so, why? *Cortex, Accelerated Long-Term Forgetting (ALF): A Novel and Important Form of Memory Impairment*, 110, 115–126. <https://doi.org/10.1016/j.cortex.2018.04.005>
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), 419–457. <https://doi.org/10.1037/0033-295X.102.3.419>
- McDade, E., Wang, G., Gordon, B. A., Hassenstab, J., Benzinger, T. L. S., Buckles, V., Fagan, A. M., Holtzman, D. M., Cairns, N. J., Goate, A. M., Marcus, D. S., Morris, J. C., Paumier, K., Xiong, C., Allegri, R., Berman, S. B., Klunk, W., Noble, J., Ringman, J., ... Koeppe, R. (2018). Longitudinal cognitive and biomarker changes in dominantly inherited Alzheimer disease. *Neurology*, 91(14). <https://doi.org/10.1212/WNL.00000000000006277>
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 263–269. <https://doi.org/10.1016/j.jalz.2011.03.005>
- Mecca, A. P., Chen, M.-K., O'Dell, R. S., Naganawa, M., Toyonaga, T., Godek, T. A., Harris, J. E., Bartlett, H. H., Zhao, W., Banks, E. R., Ni, G. S., Rogers, K., Gallezot, J. D.,

- Ropchan, J., Emery, P. R., Nabulsi, N. B., Vander Wyk, B. C., Arnsten, A. F. T., Huang, Y., ... Van Dyck, C. H. (2022). Association of entorhinal cortical tau deposition and hippocampal synaptic density in older individuals with normal cognition and early Alzheimer's disease. *Neurobiology of Aging*, *111*, 44–53. <https://doi.org/10.1016/j.neurobiolaging.2021.11.004>
- Meftah, S., & Gan, J. (2023). Alzheimer's disease as a synaptopathy: Evidence for dysfunction of synapses during disease progression. *Frontiers in Synaptic Neuroscience*, *15*, 1129036. <https://doi.org/10.3389/fnsyn.2023.1129036>
- Milà-Alomà, M., Ashton, N. J., Shekari, M., Salvadó, G., Ortiz-Romero, P., Montoliu-Gaya, L., Benedet, A. L., Karikari, T. K., Lantero-Rodriguez, J., Vanmechelen, E., Day, T. A., González-Escalante, A., Sánchez-Benavides, G., Minguillon, C., Fauria, K., Molinuevo, J. L., Dage, J. L., Zetterberg, H., Gispert, J. D., ... Blennow, K. (2022). Plasma p-tau231 and p-tau217 as state markers of amyloid- β pathology in preclinical Alzheimer's disease. *Nature Medicine*. <https://doi.org/10.1038/s41591-022-01925-w>
- Milà-Alomà, M., Salvadó, G., Gispert, J. D., Vilor-Tejedor, N., Grau-Rivera, O., Sala-Vila, A., Sánchez-Benavides, G., Arenaza-Urquijo, E. M., Crous-Bou, M., González-de-Echávarri, J. M., Minguillon, C., Fauria, K., Simon, M., Kollmorgen, G., Zetterberg, H., Blennow, K., Suárez-Calvet, M., Molinuevo, J. L., & Study, for the A. (2020). Amyloid beta, tau, synaptic, neurodegeneration, and glial biomarkers in the preclinical stage of the Alzheimer's continuum. *Alzheimer's & Dementia*, *16*(10), 1358–1371. <https://doi.org/10.1002/alz.12131>
- Minguillón, C., Fauria, K., Soteras, A., Menchon, T., Meléndez, X., Freixa, J., Sánchez-Benavides, G., Gispert, J. D., Molinuevo, J. L., & ALFA Study. (2020). ALFA+: A cohort study to understand and model the preclinical stage of Alzheimer's disease: Epidemiology / Innovative methods in epidemiology (i.e., assessment methods, design, recruitment strategies, statistical methods, etc.). *Alzheimer's & Dementia*, *16*(S10), e045935. <https://doi.org/10.1002/alz.045935>
- Molinare, C. P., Soberanes, D., Dubbelman, M., Hsieh, S., Johnson, K. A., Rentz, D. M., Sperling, R. A., Marshall, G. A., Amariglio, R. E., Papp, K. V., & Jutten, R. J. (2025). Using remote, digital, multi-day testing to characterize long-term forgetting in cognitively unimpaired older adults. *Alzheimer's & Dementia*, *21*(3), e70047. <https://doi.org/10.1002/alz.70047>
- Molinuevo, J. L., Gramunt, N., Gispert, J. D., Fauria, K., Esteller, M., Minguillon, C., Sánchez-Benavides, G., Huesa, G., Morán, S., Dal-Ré, R., & Camí, J. (2016). The

- ALFA project: A research platform to identify early pathophysiological features of Alzheimer's disease. *Alzheimer's & Dementia (New York, N. Y.)*, 2(2), 82–92. <https://doi.org/10.1016/j.trci.2016.02.003>
- Morenas-Rodríguez, E., Li, Y., Nuscher, B., Franzmeier, N., Xiong, C., Suárez-Calvet, M., Fagan, A. M., Schultz, S., Gordon, B. A., Benzinger, T. L. S., Hassenstab, J., McDade, E., Feederle, R., Karch, C. M., Schlepckow, K., Morris, J. C., Kleinberger, G., Nellgard, B., Vöglein, J., ... Xu, X. (2022). Soluble TREM2 in CSF and its association with other biomarkers and cognition in autosomal-dominant Alzheimer's disease: A longitudinal observational study. *The Lancet Neurology*, 21(4), 329–341. [https://doi.org/10.1016/S1474-4422\(22\)00027-8](https://doi.org/10.1016/S1474-4422(22)00027-8)
- Morris, J. C. (1997). Clinical Dementia Rating: A Reliable and Valid Diagnostic and Staging Measure for Dementia of the Alzheimer Type. *International Psychogeriatrics*, 9, 173–176. <https://doi.org/10.1017/S1041610297004870>
- Mortamais, M., Ash, J. A., Harrison, J., Kaye, J., Kramer, J., Randolph, C., Pose, C., Albala, B., Ropacki, M., Ritchie, C. W., & Ritchie, K. (2017). Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. *Alzheimer's & Dementia*, 13(4), 468–492. <https://doi.org/10.1016/j.jalz.2016.06.2365>
- Muthén, B., & Shedden, K. (1999). Finite Mixture Modeling with Mixture Outcomes Using the EM Algorithm. *Biometrics*, 55(2), 463–469. <https://doi.org/10.1111/j.0006-341X.1999.00463.x>
- Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods*, 4(2), 139. <https://doi.org/10.1037/1082-989X.4.2.139>
- Nelson, P. T., Alafuzoff, I., Bigio, E. H., Bouras, C., Braak, H., Cairns, N. J., Castellani, R. J., Crain, B. J., Davies, P., Tredici, K. D., Duyckaerts, C., Frosch, M. P., Haroutunian, V., Hof, P. R., Hulette, C. M., Hyman, B. T., Iwatsubo, T., Jellinger, K. A., Jicha, G. A., ... Beach, T. G. (2012). Correlation of Alzheimer Disease Neuropathologic Changes With Cognitive Status: A Review of the Literature. *Journal of Neuropathology & Experimental Neurology*, 71(5), 362–381. <https://doi.org/10.1097/NEN.0b013e31825018f7>
- Ngandu, T., Lehtisalo, J., Solomon, A., Levälahti, E., Ahtiluoto, S., Antikainen, R., Bäckman, L., Hänninen, T., Jula, A., & Laatikainen, T. (2015). A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): A

- randomised controlled trial. *The Lancet*, 385(9984), 2255–2263. [https://doi.org/10.1016/S0140-6736\(15\)60461-5](https://doi.org/10.1016/S0140-6736(15)60461-5)
- Nichols, E., Steinmetz, J. D., Vollset, S. E., Fukutaki, K., Chalek, J., Abd-Allah, F., Abdoli, A., Abualhasan, A., Abu-Gharbieh, E., Akram, T. T., Al Hamad, H., Alahdab, F., Alanezi, F. M., Alipour, V., Almustanyir, S., Amu, H., Ansari, I., Arabloo, J., Ashraf, T., ... Vos, T. (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the Global Burden of Disease Study 2019. *The Lancet Public Health*, 7(2), e105–e125. [https://doi.org/10.1016/S2468-2667\(21\)00249-8](https://doi.org/10.1016/S2468-2667(21)00249-8)
- Novozhilova, S., Fonov, V., Shafie, N., Klein, D., Villeneuve, S., Collins, D. L., & the ADNI and PREVENT-AD research groups. (2025). *Sex specific trajectories of verbal memory decline suggest masking of Alzheimer's Disease symptoms in females*. *Neurology*. <https://doi.org/10.1101/2025.11.07.25339777>
- O'Connor, A., Weston, P. S., Pavisic, I. M., Ryan, N. S., Collins, J. D., Lu, K., Crutch, S. J., Alexander, D. C., Fox, N. C., & Oxtoby, N. P. (2020). *Quantitative detection and staging of presymptomatic cognitive decline in familial Alzheimer's disease: A retrospective cohort analysis*. <https://doi.org/10.1186/s13195-020-00695-2>
- Ohno, M. (2026). Accelerated long-term forgetting and relevant biomarkers for early detection in preclinical Alzheimer's disease. *Neuroscience*, 611, 124–133. <https://doi.org/10.1016/j.neuroscience.2026.06.020>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). *The PRISMA 2020 statement: An updated guideline for reporting systematic reviews*. <https://doi.org/10.1136/bmj.n71>
- Parker, C. S., Young, C., Magill, N., Lu, K., Crutch, S. J., Fox, N. C., & Weston, P. S. J. (2025). Accelerated forgetting in presymptomatic Alzheimer's: Mediation by prefrontal cortical degeneration. *Brain Communications*, 7(6). <https://doi.org/10.1093/braincomms/fcaf478>
- Peña-Casanova, J. (1991). Programa integrado de exploración neuropsicológica - test barcelona: Bases teóricas, objetivos y contenidos. *Revista de Logopedia, Foniatría y Audiología*, 11(2), 66–79. [https://doi.org/10.1016/S0214-4603\(91\)75505-8](https://doi.org/10.1016/S0214-4603(91)75505-8)
- Peña-Casanova, J., Gramunt-Fombuena, N., Quiñones-Úbeda, S., Sánchez-Benavides, G., Aguilar, M., Badenes, D., Molinuevo, J. L., Robles, A., Barquero, M. S., & Payno,

- M. (2009). Spanish multicenter normative studies (NEURONORMA Project): Norms for the Rey–Osterrieth complex figure (copy and memory), and free and cued selective reminding test. *Archives of Clinical Neuropsychology*, 24(4), 371–393. <https://doi.org/doi.org/10.1093/arclin/acp041>
- Pereira, J. B., Janelidze, S., Ossenkoppele, R., Kvartsberg, H., Brinkmalm, A., Mattsson-Carlgen, N., Stomrud, E., Smith, R., Zetterberg, H., Blennow, K., & Hansson, O. (2021). Untangling the association of amyloid- β and tau with synaptic and axonal loss in Alzheimer's disease. *Brain*, 144(1), 310–324. <https://doi.org/10.1093/brain/awaa395>
- Perez-Nievas, B. G., & Serrano-Pozo, A. (2018). Deciphering the Astrocyte Reaction in Alzheimer's Disease. *Frontiers in Aging Neuroscience*, 10, 114. <https://doi.org/10.3389/fnagi.2018.00114>
- Peter, C., Sathe, A., Shashikumar, N., Pechman, K. R., Workmeister, A. W., Jackson, T. B., Huo, Y., Mukherjee, S., Mez, J., Dumitrescu, L. C., Gifford, K. A., Bolton, C. J., Gaynor, L. S., Risacher, S. L., Beason-Held, L. L., An, Y., Arfanakis, K., Erus, G., Davatzikos, C., ... Alzheimer's Disease Neuroimaging Initiative (ADNI), the B. of C. D. A. N. A. (BIOCARD) S. T., and the Alzheimer's Disease Sequencing Project (ADSP). (2025). White Matter Abnormalities and Cognition in Aging and Alzheimer Disease. *JAMA Neurology*, 82(8), 825–836. <https://doi.org/10.1001/jamaneurol.2025.1601>
- Pizzoni, L., Piero, A. D., Blasutto, B., Casagrande, M., Piccardi, L., & Guariglia, C. (2026). Accelerated long-term forgetting in Alzheimer's disease continuum: A systematic review and meta-analysis. *Journal of Alzheimer's Disease*, 111(4), 1459–1482. <https://doi.org/10.1177/13872877261443537>
- Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06*, University of Cambridge, Cambridge, 26(26–46), 1.
- Proust-Lima, C., Philipps, V., & Liqueur, B. (2017). Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: The R Package **lcmm**. *Journal of Statistical Software*, 78(2). <https://doi.org/10.18637/jss.v078.i02>
- Qu, H., Ge, H., Wang, L., Wang, W., & Hu, C. (2023). Volume changes of hippocampal and amygdala subfields in patients with mild cognitive impairment and Alzheimer's disease. *Acta Neurologica Belgica*, 123(4), 1381–1393. <https://doi.org/10.1007/s13760-023-02235-9>

- R Core Team. (2025). *R: A language and environment for statistical computing*.
- Rami, L., León, M., Valech, N., Coll-Adrós, N., Bosch, B., Olives, J., Salinero, A., Pérez-Millan, A., Molinuevo, J. L., Sánchez-Valle, R., & Tort-Merino, A. (2023). Design and validation of the 1-week memory battery for assessing episodic memory and accelerated long-term forgetting in cognitively unimpaired subjects. *Neuropsychology*, 37(6), 683–697. <https://doi.org/10.1037/neu0000879>
- Rami, L., Mollica, M. A., Garcia-Sanchez, C., Saldana, J., Sanchez, B., Sala, I., Valls-Pedret, C., Castellví, M., Olives, J., & Molinuevo, J. L. (2014). *The Subjective Cognitive Decline Questionnaire (SCD-Q): A Validation Study*. <https://doi.org/10.3233/JAD-132027>
- Randolph, C., Tierney, M. C., Mohr, E., & Chase, T. N. (1998). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Preliminary Clinical Validity. *Journal of Clinical and Experimental Neuropsychology*, 20(3), 310–319. <https://doi.org/10.1076/jcen.20.3.310.823>
- Rawat, P., Sehar, U., Bisht, J., Selman, A., Culberson, J., & Reddy, P. H. (2022). Phosphorylated Tau in Alzheimer’s Disease and Other Tauopathies. *International Journal of Molecular Sciences*, 23(21), 12841. <https://doi.org/10.3390/ijms232112841>
- Reiman, E. M. (2018). Long-term forgetting in preclinical Alzheimer’s disease. *The Lancet Neurology*, 17(2), 104–105. [https://doi.org/10.1016/S1474-4422\(17\)30458-1](https://doi.org/10.1016/S1474-4422(17)30458-1)
- Reiman, E. M., Arboleda-Velasquez, J. F., Quiroz, Y. T., Huentelman, M. J., Beach, T. G., Caselli, R. J., Chen, Y., Su, Y., Myers, A. J., & Hardy, J. (2020). Exceptionally low likelihood of Alzheimer’s dementia in APOE2 homozygotes from a 5,000-person neuropathological study. *Nature Communications*, 11(1), 667. <https://doi.org/10.1038/s41467-019-14279-8>
- Reitan, R. M., & Wolfson, D. (2009). The Halstead–Reitan neuropsychological test battery for adults—Theoretical, methodological, and validation bases. *Neuropsychological Assessment of Neuropsychiatric and Neuromedical Disorders*, 1, 3–24.
- Riley, K. P., Snowden, D. A., & Markesbery, W. R. (2002). Alzheimer’s neurofibrillary pathology and the spectrum of cognitive function: Findings from the Nun Study. *Annals of Neurology*, 51(5), 567–577. <https://doi.org/10.1002/ana.10161>
- Rodini, M., Bonarota, S., Serra, L., Caltagirone, C., & Carlesimo, G. A. (2024). Could Accelerated Long-Term Forgetting Be a Feature of the Higher Rate of Memory Complaints Associated with Subjective Cognitive Decline? An Exploratory Study.

- Journal of Alzheimer's Disease*, 100(4), 1165–1182. <https://doi.org/10.3233/JAD-240218>
- Rodini, M., De Simone, M. S., Caltagirone, C., & Carlesimo, G. A. (2022). Accelerated long-term forgetting in neurodegenerative disorders: A systematic review of the literature. *Neuroscience & Biobehavioral Reviews*, 141, 104815. <https://doi.org/10.1016/j.neubiorev.2022.104815>
- Rosenberry, T. L., Zhou, H.-X., Stagg, S. M., & Paravastu, A. K. (2022). Oligomer Formation by Amyloid- β 42 in a Membrane-Mimicking Environment in Alzheimer's Disease. *Molecules*, 27(24), 8804. <https://doi.org/10.3390/molecules27248804>
- Ruggeri, M., Ricci, M., Gerace, C., & Blundo, C. (2024). Accelerated long-term forgetting: From subjective memory decline to a defined clinical entity. *Aging, Neuropsychology, and Cognition*, 31(6), 1058–1069. <https://doi.org/10.1080/13825585.2024.2317924>
- Rusinek, H., Bokacheva, L., Chen, H., Masurkar, A., Osorio, R., Betensky, R., Vedvyas, A., Chodosh, J., Shao, Y., Shepherd, T., Marsh, K., & Wisniewski, T. (2026). Test-retest reliability of FreeSurfer measures of neurodegeneration. *NeuroImage*, 333, 121920. <https://doi.org/10.1016/j.neuroimage.2026.121920>
- Saloner, R., Casaletto, K. B., Marx, G., Dutt, S., Vanden Bussche, A. B., You, M., Fox, E., Stiver, J., & Kramer, J. H. (2018). Worth the Wait: Performance on a 1-Week Delayed Recall Task is Associated with Medial Temporal Lobe Structures in Neurologically Normal Older Adults. *The Clinical Neuropsychologist*, 32(3), 456–467. <https://doi.org/10.1080/13854046.2017.1370134>
- Sánchez-Benavides, G., López-Martos, D., Suarez-Calvet, M., Milà-Alomà, M., Gispert, J. D., Minguillon, C., Zetterberg, H., Blennow, K., & Grau-Rivera, O. (2023). Longitudinal decrements in the sensitivity of cueing in the Free and Cued Selective Reminding Test predict A+T+ status in cognitively unimpaired individuals. *Alzheimer's & Dementia*, 19(S18), e079555. <https://doi.org/10.1002/alz.079555>
- Schöll, M., Lockhart, S. N., Schonhaut, D. R., O'Neil, J. P., Janabi, M., Ossenkoppele, R., Baker, S. L., Vogel, J. W., Faria, J., Schwimmer, H. D., Rabinovici, G. D., & Jagust, W. J. (2016). PET Imaging of Tau Deposition in the Aging Human Brain. *Neuron*, 89(5), 971–982. <https://doi.org/10.1016/j.neuron.2016.01.028>
- Scott, M. R., Hampton, O. L., Buckley, R. F., Chhatwal, J. P., Hanseeuw, B. J., Jacobs, H. Il., Properzi, M. J., Sanchez, J. S., Johnson, K. A., Sperling, R. A., & Schultz, A. P. (2020). Inferior temporal tau is associated with accelerated prospective cortical

- thinning in clinically normal older adults. *NeuroImage*, 220, 116991. <https://doi.org/10.1016/j.neuroimage.2020.116991>
- Sedó, M. A. (2007). *The Five-Digits Test*. TEA Ediciones.
- Selkoe, D. J. (2002). Alzheimer's Disease Is a Synaptic Failure. *Science*, 298(5594), 789–791. <https://doi.org/10.1126/science.1074069>
- Serrano-Pozo, A., Frosch, M. P., Masliah, E., & Hyman, B. T. (2011). Neuropathological alterations in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, 1(1), a006189. <https://doi.org/10.1101/cshperspect.a006189>
- Son, S., Speechley, M., Zou, G., & Montero-Odasso, M. (2025). Identifying the optimal combinations of modifiable dementia risk factors to target in multidomain intervention – Three-year longitudinal findings from the Canadian longitudinal study on aging. *The Journal of Prevention of Alzheimer's Disease*, 12(9), 100321. <https://doi.org/10.1016/j.tjpad.2025.100321>
- 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*, 7(3), 280–292. <https://doi.org/10.1016/j.jalz.2011.03.003>
- Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: A neurobiological perspective. *Current Opinion in Neurobiology*, 5(2), 169–177. [https://doi.org/10.1016/0959-4388\(95\)80023-9](https://doi.org/10.1016/0959-4388(95)80023-9)
- Stamate, A., Logie, R. H., Baddeley, A. D., & Della Sala, S. (2020). Forgetting in Alzheimer's disease: Is it fast? Is it affected by repeated retrieval? *Neuropsychologia*, 138, 107351. <https://doi.org/10.1016/j.neuropsychologia.2020.107351>
- Stephan, B. C., Cochrane, L., Kafadar, A. H., Brain, J., Burton, E., Myers, B., Brayne, C., Naheed, A., Anstey, K. J., & Ashor, A. W. (2024). Population attributable fractions of modifiable risk factors for dementia: A systematic review and meta-analysis. *The Lancet Healthy Longevity*, 5(6), e406–e421. [https://doi.org/10.1016/S2666-7568\(24\)00061-8](https://doi.org/10.1016/S2666-7568(24)00061-8)
- Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology*, 11(11), 1006–1012. [https://doi.org/10.1016/S1474-4422\(12\)70191-6](https://doi.org/10.1016/S1474-4422(12)70191-6)

- Streit, W. J., Khoshbouei, H., & Bechmann, I. (2020). Dystrophic microglia in late-onset Alzheimer's disease. *Glia*, 68(4), 845–854. <https://doi.org/10.1002/glia.23782>
- Sundermann, E. E., Maki, P. M., Rubin, L. H., Lipton, R. B., Landau, S., Biegon, A., & For the Alzheimer's Disease Neuroimaging Initiative. (2016). Female advantage in verbal memory: Evidence of sex-specific cognitive reserve. *Neurology*, 87(18), 1916–1924. <https://doi.org/10.1212/WNL.00000000000003288>
- Thal, D. R., Rüb, U., Orantes, M., & Braak, H. (2002). Phases of A β -deposition in the human brain and its relevance for the development of AD. *Neurology*, 58(12), 1791–1800. <https://doi.org/10.1212/WNL.58.12.1791>
- Therriault, J., Schindler, S. E., Salvadó, G., Pascoal, T. A., Benedet, A. L., Ashton, N. J., Karikari, T. K., Apostolova, L., Murray, M. E., Verberk, I., Vogel, J. W., La Joie, R., Gauthier, S., Teunissen, C., Rabinovici, G. D., Zetterberg, H., Bateman, R. J., Scheltens, P., Blennow, K., ... Rosa-Neto, P. (2024). Biomarker-based staging of Alzheimer disease: Rationale and clinical applications. *Nature Reviews Neurology*, 20(4), 232–244. <https://doi.org/10.1038/s41582-024-00942-2>
- Therriault, J., Vermeiren, M., Servaes, S., Tissot, C., Ashton, N. J., Benedet, A. L., Karikari, T. K., Lantero-Rodriguez, J., Brum, W. S., & Lussier, F. Z. (2023). Association of phosphorylated tau biomarkers with amyloid positron emission tomography vs tau positron emission tomography. *JAMA Neurology*, 80(2), 188–199. <https://doi.org/10.1001/jamaneurol.2022.4485>
- Therriault, J., Zimmer, E. R., Benedet, A. L., Pascoal, T. A., Gauthier, S., & Rosa-Neto, P. (2022). Staging of Alzheimer's disease: Past, present, and future perspectives. *Trends in Molecular Medicine*, 28(9), 726–741. <https://doi.org/10.1016/j.molmed.2022.05.008>
- Thibault, E. G., Del Carmen Montenegro, G., Becker, J., Price, J. C., Healy, B. C., Hanseeuw, B. J., Buckley, R. F., Jacobs, H. I. L., Properzi, M. J., Sperling, R. A., Johnson, K. A., & Farrell, M. E. (2026). Amyloid spatial extent with florbetapir-PET for early detection of preclinical Alzheimer's disease. *The Journal of Prevention of Alzheimer's Disease*, 13(5), 100529. <https://doi.org/10.1016/j.tjpad.2026.100529>
- Tolar, M., Hey, J. A., Power, A., & Abushakra, S. (2024). The Single Toxin Origin of Alzheimer's Disease and Other Neurodegenerative Disorders Enables Targeted Approach to Treatment and Prevention. *International Journal of Molecular Sciences*, 25(5), 2727. <https://doi.org/10.3390/ijms25052727>

- Tort-Merino, A., Laine, M., Valech, N., Olives, J., León, M., Ecay-Torres, M., Estanga, A., Martínez-Lage, P., Fortea, J., Sánchez-Valle, R., Rami, L., & Rodríguez-Fornells, A. (2021). Accelerated long-term forgetting over three months in asymptomatic APOE $\epsilon 4$ carriers. *Annals of Clinical and Translational Neurology*, 8(2), 477–484. <https://doi.org/10.1002/acn3.51245>
- Tort-Merino, A., Valech, N., Laine, M., Olives, J., León, M., Ecay-Torres, M., Estanga, A., Martínez-Lage, P., Fortea, J., Molinuevo, J. L., Sánchez-Valle, R., Rodríguez-Fornells, A., & Rami, L. (2021). Accelerated long-term forgetting in individuals with subjective cognitive decline and amyloid- β positivity. *International Journal of Geriatric Psychiatry*, 36(7), 1037–1049. <https://doi.org/10.1002/gps.5539>
- Tzioras, M., McGeachan, R. I., Durrant, C. S., & Spires-Jones, T. L. (2023). Synaptic degeneration in Alzheimer disease. *Nature Reviews Neurology*, 19(1), 19–38. <https://doi.org/10.1038/s41582-022-00749-z>
- Valech, N., Sanchez-Benavides, G., Tort-Merino, A., Coll-Padros, N., Olives, J., Leon, M., Falcon, C., Molinuevo, J. L., & Rami, L. (2019). *Associations Between the Subjective Cognitive Decline-Questionnaire's Scores, Gray Matter Volume, and Amyloid-Levels*. <https://doi.org/10.3233/JAD-190624>
- Van Der Molen, L. H., Boenink, M., Van Lente, H., & Richard, E. (2025). Changing definitions of disease: Transformations in the diagnostic criteria for Alzheimer's disease. *Alzheimer's & Dementia*, 21(4), e70133. <https://doi.org/10.1002/alz.70133>
- van der Werf, S. P., Geurts, S., & de Werd, M. M. E. (2016). Subjective Memory Ability and Long-Term Forgetting in Patients Referred for Neuropsychological Assessment. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00605>
- Van Loenhoud, A. C., Van Der Flier, W. M., Wink, A. M., Dicks, E., Groot, C., Twisk, J., Barkhof, F., Scheltens, P., Ossenkoppele, R., & for the Alzheimer's Disease Neuroimaging Initiative. (2019). Cognitive reserve and clinical progression in Alzheimer disease: A paradoxical relationship. *Neurology*, 93(4). <https://doi.org/10.1212/WNL.00000000000007821>
- Verdi, S., Rutherford, S., Fraza, C., Tosun, D., Altmann, A., Raket, L. L., Schott, J. M., Marquand, A. F., Cole, J. H., & for the Alzheimer's Disease Neuroimaging Initiative. (2024). Personalizing progressive changes to brain structure in Alzheimer's disease using normative modeling. *Alzheimer's & Dementia*, 20(10), 6998–7012. <https://doi.org/10.1002/alz.14174>

- Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, 18(4), 450–469. <https://doi.org/10.1093/pan/mpq025>
- Villemagne, V. L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K. A., Salvado, O., Szoek, C., Macaulay, S. L., Martins, R., & Maruff, P. (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](https://doi.org/10.1016/S1474-4422(13)70044-9)
- Vilor-Tejedor, N., Genius, P., Rodríguez-Fernández, B., Minguillón, C., Sadeghi, I., González-Escalante, A., Crous-Bou, M., Suárez-Calvet, M., Grau-Rivera, O., Bruguat-Serrat, A., Sánchez-Benavides, G., Esteller, M., Fauria, K., Molinuevo, J. L., Navarro, A., Gispert, J. D., for the Alzheimer's Disease Neuroimaging Initiative, & for the ALFA study. (2024). Genetic characterization of the ALFA study: Uncovering genetic profiles in the Alzheimer's *continuum*. *Alzheimer's & Dementia*, 20(3), 1703–1715. <https://doi.org/10.1002/alz.13537>
- Walsh, C. M., Wilkins, S., Bettcher, B. M., Butler, C. R., Miller, B. L., & Kramer, J. H. (2014). Memory consolidation in aging and MCI after 1 week. *Neuropsychology*, 28(2), 273–280. <https://doi.org/10.1037/neu0000013>
- Wang, J., Grundke-Iqbal, I., & Iqbal, K. (2007). Kinases and phosphatases and tau sites involved in Alzheimer neurofibrillary degeneration. *European Journal of Neuroscience*, 25(1), 59–68. <https://doi.org/10.1111/j.1460-9568.2006.05226.x>
- Wang, J.-Z., Xia, Y.-Y., Grundke-Iqbal, I., & Iqbal, K. (2012). Abnormal Hyperphosphorylation of Tau: Sites, Regulation, and Molecular Mechanism of Neurofibrillary Degeneration. *Journal of Alzheimer's Disease*, 33(s1), S123–S139. <https://doi.org/10.3233/JAD-2012-129031>
- Wang, Y., Cella, M., Mallinson, K., Ulrich, J. D., Young, K. L., Robinette, M. L., Gilfillan, S., Krishnan, G. M., Sudhakar, S., Zinselmeyer, B. H., Holtzman, D. M., Cirrito, J. R., & Colonna, M. (2015). TREM2 Lipid Sensing Sustains the Microglial Response in an Alzheimer's Disease Model. *Cell*, 160(6), 1061–1071. <https://doi.org/10.1016/j.cell.2015.01.049>
- Warmenhoven, N., Sánchez-Benavides, G., González-Escalante, A., Milà-Alomà, M., Shekari, M., López-Martos, D., Ortiz-Romero, P., Kollmorgen, G., Quijano-Rubio, C., Minguillón, C., Gispert, J. D., Vilor-Tejedor, N., Arenaza-Urquijo, E., Palpatzis, E., Ashton, N. J., Zetterberg, H., Blennow, K., Suárez-Calvet, M., Grau-Rivera, O., & ALFA Study. (2024). CSF glial biomarkers are associated with cognition in

- individuals at risk of Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 20(9), 5819–5832. <https://doi.org/10.1002/alz.13862>
- Wearn, A. R., Saunders-Jennings, E., Nurdal, V., Hadley, E., Knight, M. J., Newson, M., Kauppinen, R. A., & Coulthard, E. J. (2020). Accelerated long-term forgetting in healthy older adults predicts cognitive decline over 1 year. *Alzheimer's Research & Therapy*, 12(1), 119. <https://doi.org/10.1186/s13195-020-00693-4>
- Wechsler, D. (1945). A Standardized Memory Scale for Clinical Use. *The Journal of Psychology*, 19(1), 87–95. <https://doi.org/10.1080/00223980.1945.9917223>
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale: WAIS-IV Technical and Interpretive Manual*. Pearson.
- Weston, P. S., Nicholas, J. M., Henley, S. M., Liang, Y., Macpherson, K., Donnachie, E., Schott, J. M., Rossor, M. N., Crutch, S. J., & Butler, C. R. (2018). Accelerated long-term forgetting in presymptomatic autosomal dominant Alzheimer's disease: A cross-sectional study. *The Lancet Neurology*, 17(2), 123–132. [https://doi.org/10.1016/S1474-4422\(17\)30434-9](https://doi.org/10.1016/S1474-4422(17)30434-9)
- Wimo, A., Seeher, K., Cataldi, R., Cyhlarova, E., Dielemann, J. L., Frisell, O., Guerchet, M., Jönsson, L., Malaha, A. K., Nichols, E., Pedroza, P., Prince, M., Knapp, M., & Dua, T. (2023). The worldwide costs of dementia in 2019. *Alzheimer's & Dementia*, 19(7), 2865–2873. <https://doi.org/10.1002/alz.12901>
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, 310(20), 2191–2194. <https://doi.org/10.1001/jama.2013.281053>
- Wuestefeld, A., Pichet Binette, A., Berron, D., Spotorno, N., Van Westen, D., Stomrud, E., Mattsson-Carlgrén, N., Strandberg, O., Smith, R., Palmqvist, S., Glenn, T., Moes, S., Honer, M., Arfanakis, K., Barnes, L. L., Bennett, D. A., Schneider, J. A., Wisse, L. E. M., & Hansson, O. (2023). Age-related and amyloid-beta-independent tau deposition and its downstream effects. *Brain*, 146(8), 3192–3205. <https://doi.org/10.1093/brain/awad135>
- Xie, L., Das, S. R., Li, Y., Wisse, L. E. M., McGrew, E., Lyu, X., DiCalogero, M., Shah, U., Ilesanmi, A., Denning, A. E., Brown, C. A., Cohen, J., Sreepada, L., Dong, M., Sadeghpour, N., Khandelwal, P., Ittyerah, R., Ravikumar, S., Sadaghiani, S., ... for the Alzheimer's Disease Neuroimaging Initiative. (2025). A multi-cohort study of longitudinal and cross-sectional Alzheimer's disease biomarkers in cognitively

- unimpaired older adults. *Alzheimer's & Dementia*, 21(2), e14492. <https://doi.org/10.1002/alz.14492>
- Xie, L., Wisse, L. E. M., Das, S. R., Vergnet, N., Dong, M., Ittyerah, R., De Flores, R., Yushkevich, P. A., Wolk, D. A., & for the Alzheimer's Disease Neuroimaging Initiative. (2020). Longitudinal atrophy in early Braak regions in preclinical Alzheimer's disease. *Human Brain Mapping*, 41(16), 4704–4717. <https://doi.org/10.1002/hbm.25151>
- Yang, J., Kong, C., Jia, L., Li, T., Quan, M., Li, Y., Lyu, D., Li, F., Jin, H., Li, Y., Wang, Q., & Jia, J. (2021). Association of accelerated long-term forgetting and senescence-related blood-borne factors in asymptomatic individuals from families with autosomal dominant Alzheimer's disease. *Alzheimer's Research & Therapy*, 13(1), 107. <https://doi.org/10.1186/s13195-021-00845-0>
- Yeh, F. L., Wang, Y., Tom, I., Gonzalez, L. C., & Sheng, M. (2016). TREM2 Binds to Apolipoproteins, Including APOE and CLU/APOJ, and Thereby Facilitates Uptake of Amyloid-Beta by Microglia. *Neuron*, 91(2), 328–340. <https://doi.org/10.1016/j.neuron.2016.06.015>
- You, J. C., Bender, A. C., Berezuk, C., Noriega-Makarskyy, D. T., Baena, A., Munera, D., Pardilla-Delgado, E., Ramirez-Gomez, L., Aguillon, D., Giudicessi, A., Martinez, J. E., Cronin-Golomb, A., Lam, A. D., & Quiroz, Y. T. (2026). Long-term forgetting, sleep, and tau in autosomal-dominant Alzheimer's disease. *Alzheimer's & Dementia*, 22(2), e71235. <https://doi.org/10.1002/alz.71235>
- Young, C. B., Winer, J. R., Younes, K., Cody, K. A., Betthausen, T. J., Johnson, S. C., Schultz, A., Sperling, R. A., Greicius, M. D., Cobos, I., Poston, K. L., Mormino, E. C., Alzheimer's Disease Neuroimaging Initiative and the Harvard Aging Brain Study, Weiner, M. W., Aisen, P., Petersen, R., Jack, C. R., Jagust, W., Trojanowki, J. Q., ... Sperling, R. A. (2022). Divergent Cortical Tau Positron Emission Tomography Patterns Among Patients With Preclinical Alzheimer Disease. *JAMA Neurology*, 79(6), 592. <https://doi.org/10.1001/jamaneurol.2022.0676>
- Yu, R.-C., Proctor, D., Soni, J., Pikett, L., Livingston, G., Lewis, G., Schilder, A., Bamiou, D., Mandavia, R., & Omar, R. (2024). Adult-onset hearing loss and incident cognitive impairment and dementia—A systematic review and meta-analysis of cohort studies. *Ageing Research Reviews*, 98, 102346. <https://doi.org/10.1016/j.arr.2024.102346>
- Yushkevich, P. A., Pluta, J. B., Wang, H., Xie, L., Ding, S.-L., Gertje, E. C., Mancuso, L., Kliot, D., Das, S. R., & Wolk, D. A. (2015). Automated volumetry and regional

- thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment: Automatic Morphometry of MTL Subfields in MCI. *Human Brain Mapping*, 36(1), 258–287. <https://doi.org/10.1002/hbm.22627>
- Zavec, Z., Shah, V. D., Murillo, O. G., Vallat, R., Mander, B. A., Winer, J. R., Jagust, W. J., & Walker, M. P. (2023). NREM sleep as a novel protective cognitive reserve factor in the face of Alzheimer's disease pathology. *BMC Medicine*, 21(1), 156. <https://doi.org/10.1186/s12916-023-02811-z>
- Zhao, W., Wang, X., Yin, C., He, M., Li, S., & Han, Y. (2019). Trajectories of the Hippocampal Subfields Atrophy in the Alzheimer's Disease: A Structural Imaging Study. *Frontiers in Neuroinformatics*, 13, 13. <https://doi.org/10.3389/fninf.2019.00013>
- Zhou, Z., Wang, Q., Liu, L., Wang, Q., Zhang, X., Li, C., Liu, J., Wei, Y., Gao, J., Fu, L., & Wang, R. (2025). Investigating the A β and tau pathology in autosomal dominant Alzheimer's disease: Insights from hybrid PET/MRI and network mapping. *Alzheimer's Research & Therapy*, 17(1), 45. <https://doi.org/10.1186/s13195-025-01690-1>
- Zimmermann, J. F., & Butler, C. R. (2018). Accelerated long-term forgetting in asymptomatic APOE ϵ 4 carriers. *The Lancet Neurology*, 17(5), 394–395. [https://doi.org/10.1016/S1474-4422\(18\)30078-4](https://doi.org/10.1016/S1474-4422(18)30078-4)

APPENDICES

Appendix A Summary of the studies included in the review.

Table A1. Summary of selected studies providing evidence that long-term delayed recall may serve as a sensitive cognitive marker for early Alzheimer’s disease pathology detection.

Author(s), Year, Country	Study Design	Sample	Long-Term Delayed Recall Measure		Learning Criterion	Recall Intervals	Main Outcomes	Strengths	Limitations
			Verbal	Visual					
Bartels et al., 2025 Göttingen (Germany)	Cross-sectional	• $N = 59$ (17 HC, 24 SCD, 18 MCI) • Subsample of participants with MRI data: $N = 32$ (13 HC, 19 SCD)	CERAD/ADAS-Cog	FCSRT-IR (picture version)	Not predefined for the ADAS-Cog word list (three learning trials); the FCSRT-IR uses a controlled encoding procedure in which all 16 items must be retrieved in immediate cued recall	• Immediate • 30 min • Between 2 and 30 days	• FCSRT-IR LTR rate significantly differentiated HC vs. SCD and MCI HC = 0.218 ± 0.232; SCD = 0.087 ± 0.123; MCI = 0.045 ± 0.125 ($p = .007$) • FCSRT-IR LTR positively associated with hippocampal CA1 volume ($r \approx 0.41$, $p < .02$) and left CA3 ($r = 0.53$, $p = .002$)	• Multicenter sample • Harmonized MRI protocol	• Small MRI subsample • Floor effects: many participants (even HC) scored LTR = 0 after 1-30 days • Variable retrieval interval (2-30 days) may have influenced results • No biomarker confirmation ($A\beta$, tau)
Degenszajn et al., 2001 São Paulo (Brazil)	Cross-sectional	$N = 30$ (15 AD, 15 HC)	FCSRT		Not predefined	• Immediate • 30 min • 24 h	• AD patients showed significantly lower recall compared with HC 30-min delayed TR: AD = 8.3 ± 4.8; HC = 16.0 ± 0 ($p = .0001$) 24-h delayed TR: AD = 8.2 ± 4.7; HC = 16.0 ± 0 ($p = .0001$) • Significant decline from immediate (12.06) to 30 min (8.3) and 24 h (8.2) ($p < .01$), but no difference between 30 min and 24 h ($p > .05$)	Delayed recall assessment showed excellent diagnostic accuracy (100% sensitivity and 100% specificity)	• Small sample size • The short delay in LTR (24h) may have compromised the effectiveness of the measure • Ceiling effects in HC • No neuroimaging or biomarker confirmation

Author(s), Year, Country	Study Design	Sample	Long-Term Delayed Recall Measure		Learning Criterion	Recall Intervals	Main Outcomes	Strengths	Limitations
			Verbal	Visual					
Lindbergh et al., 2021 San Francisco (US)	Longitudinal (annual visits up to 8 years)	$N = 207$ CU	Story Recall Test		90% accuracy	• Immediate • 30 min • 1 week	Poorer one-week delayed recall predicted steeper decline in episodic memory, executive functioning, and medial temporal lobe (MTL) volumes. • Episodic memory: $b = 0.01$; $p = .003$ • Executive functioning: $b = 0.006$; $p = .011$ • MTL volume: $b = 0.004$; $p = .037$	• Large sample • Comprehensive neuroimaging and cognitive follow-up	Ceiling effects for 30-min recall
Lu et al., 2025 London (UK)	Longitudinal (baseline and ≈ 3 -year follow-up)	• $N = 429$ CU • Subsample of participants with complete biomarkers data: $N = 306$ ($A\beta^+ = 77$, $A\beta^- = 229$)	FNAME-12	AMIPB Complex Figure	Not predefined	• Immediate • 30 min • 1 week	$A\beta^+$ showed poorer one-week recall than $A\beta^-$. • AMIPB Complex Figure outline: $A\beta^+ = 90\%$ (95% CI 85, 94); $A\beta^- = 95\%$ (CI 92, 99); $p < .01$ • Whole-brain atrophy $\times$ amyloid interaction: -1.9 percentage-points one-week retention per ml/year atrophy (95% CI 0.5, 3.7); $p < .05$ • Hippocampal atrophy in $A\beta^+$ associated with lower one-week recall ($p < .05$) • FNAME-12 one-week recall showed no association with $A\beta$ status • 30-min recall did not differentiate amyloid groups or predict atrophy	• Large birth-cohort sample with multimodal biomarkers data • Item-level analysis (AMIPB Complex Figure outline vs. detail)	No tau-PET data available limits conclusions about tau vs. $A\beta$ effects

Author(s), Year, Country	Study Design	Sample	Long-Term Delayed Recall Measure		Learning Criterion	Recall Intervals	Main Outcomes	Strengths	Limitations
			Verbal	Visual					
Magill et al., 2026 London (UK)	Longitudinal (annual follow-up, median 7 years)	<i>N</i> = 20 asymptomatic ADAD mutation carriers (18 <i>PSEN1</i> , 2 <i>APP</i>) • All CDR 0 with normal performance on standard testing • Progressors <i>n</i> = 9; non-progressors <i>n</i> = 11	• Word list • Short story	Complex abstract figure	80% accuracy (list: 4-10 trials; story: 2-10 trials); no criterion for figure	• Learning • 30 min • 7 days (remote)	Baseline ALF predicted clinical onset; 9/20 developed symptoms (CDR global > 0). • List ALF: progressors = 30 (IQR 30-36.4); non-progressors = 58.3 (IQR 33-66.7); <i>p</i> = .03 • Story ALF: progressors = 58.8 (IQR 44-66); non-progressors = 81.2 (IQR 69.1-87.8); <i>p</i> < .001 • Figure ALF: 68.2 (IQR 55.6-81.3) vs 74.5 (IQR 64.5-85.9); <i>p</i> = .21 • Discrimination: story AUC = 0.82; list AUC = 0.73; figure AUC = 0.64 • Learning and 30-min recall did not differ between groups • No other cognitive test discriminated progressors	• Only study with a clinical outcome rather than group differences alone • Extended follow-up (median 7 years) • ALF specificity relative to conventional tests • 7-day recall administered remotely	• Small sample, reflecting the rarity of ADAD • ADAD only; replication in sporadic AD required • High resource demands (two sessions one week apart) • UCL cohort, overlapping Weston 2018, O'Connor 2020 and Parker 2025
Manes et al., 2008 Buenos Aires (Argentina)	Cross-sectional	<i>N</i> = 26 (9 CU, 10 memory complaint, 7 aMCI)	AMIPB, two stories recall	Rey-Osterrieth Complex Figure	Not predefined	• Immediate • 30 min • 6 weeks	Memory complaint and aMCI groups recalled normally at immediate and 30-min tests but showed significantly poorer recall at 6-week testing. • Logical Memory 6-week recall: group × time interaction $F_{4,46} = 4.39$; <i>p</i> = .005 • Rey figure 6-week recall: group × time interaction $F_{4,46} = 2.77$; <i>p</i> = .036	Evaluation of individuals with memory complaints	• Small sample size • 6-week interval only, no intermediate timepoints • No biomarker or imaging confirmation

Author(s), Year, Country	Study Design	Sample	Long-Term Delayed Recall Measure		Learning Criterion	Recall Intervals	Main Outcomes	Strengths	Limitations
			Verbal	Visual					
Molinare et al., 2025 Boston (US)	Cross-sectional (remote, digital, multi-day testing)	• $N = 104$ CU older adults (16.3% Aβ+) • CDR 0 subsample: $n = 93$ • Mass General Brigham cohorts (HABS, IADL, SCD)	Face-Name Matching Task (BRANCH); associative memory assessed by recognition		Not predefined; identical stimuli on seven consecutive days; learning curve quantified as AUC	• Daily learning across 7 days • 1 week • 6 months	• Better initial learning associated with less LTF: $\beta = 0.52$, 95% CI [0.34, 0.71]; $p < .001$ • No association with global amyloid: $\beta = -0.07$, 95% CI [-0.23, 0.09]; $p = .418$ • Medial temporal tau: $\beta = -0.18$, 95% CI [-0.29, -0.07]; $p = .001$ • Neocortical tau: $\beta = -0.16$, 95% CI [-0.31, -0.01]; $p = .032$ • Cortical thickness, AD meta-ROI: $\beta = 0.33$, 95% CI [0.13, 0.52]; $p = .001$ • No significant associations with any biomarker at 6 months	• Large CU sample with both amyloid and tau PET • Fully remote, digital administration • Comparison of two retention intervals (1 week and 6 months)	• Recognition-based outcome rather than free recall • Low proportion of Aβ+ participants (16.3%) • Tau associations did not survive exclusion of the single highest-tau participant (medial temporal $\beta = -0.06$, 95% CI [-0.20, 0.09]; $p = .436$) • HABS cohort, overlapping other BRANCH studies
O'Connor et al., 2020 London (UK)	Cross-sectional (event-based model of presymptomatic decline in FAD)	$N = 35$ (21 <i>APP</i> or <i>PSEN1</i> mutation carriers, 14 non-carriers)	AMIPB, stories and lists recall	Figure recall	Not predefined	• Immediate • 30 min • 1 week	Mutation carriers showed intact 30-min recall but significantly poorer one-week retention than non-carriers. Story one-week retention: carriers median 66.8% (IQR 59.8-81.9); non-carriers 93.1% (85.1-95.2); $p = .0004$	• Data-driven staging model of cognitive decline in presymptomatic FAD • Demonstrated ALF abnormalities a decade before symptom onset	No complete biomarker and imaging confirmation

Author(s), Year, Country	Study Design	Sample	Long-Term Delayed Recall Measure		Learning Criterion	Recall Intervals	Main Outcomes	Strengths	Limitations
			Verbal	Visual					
Parker et al., 2025 London (UK)	Cross-sectional (multimodal MRI: T1 and diffusion-weighted imaging)	$N = 30$ (18 presymptomatic ADAD mutation carriers, 12 non-carriers) - Carriers: EYO 7.29 ± 4.72; CDR 0	- Word list - Short story	Complex figure	80% accuracy (protocol as in Weston et al., 2018)	- Learning 30 min 7 days	ALF scores, median (IQR), carriers vs non-carriers: - List: 43.06 (30.00-61.54) vs 71.79 (58.55-81.82) - Story: 66.67 (58.82-83.02) vs 90.70 (78.80-96.71) - Figure: 72.41 (62.75-83.88) vs 87.55 (75.68-96.25) - In carriers, worse verbal ALF associated with lower prefrontal cortical thickness across four contiguous regions bilaterally; most survived correction (pFDR = 0.00769) - No association between ALF and medial temporal lobe structures - No association for visual (figure) ALF, nor between 30-min recall and gray matter in any ROI - Structural connectivity associations did not survive correction	- First study of the neuroanatomical correlates of ALF in presymptomatic ADAD - Multimodal design (cortical thickness and structural connectivity) - Explicit dissociation between ALF and short-interval recall	- Small sample; cross-sectional - Group differences previously reported in Weston et al., 2018 (same UCL cohort) - Generalizability from ADAD to sporadic AD not established
Rodini et al., 2024 Rome (Italy)	Cross-sectional	$N = 47$ (16 CU, 18 mild SCD, 13 severe SCD)	Short Story Test	Rey-Osterrieth Complex Figure	Not predefined	- Immediate 30 min 24 h 1 week	Severe SCD showed significantly poorer 24-h and one-week story recall compared with HC. - Short Story 24h-30min saving index: severe SCD = 84.2%; HC = 94.3%; $p = .023$ - ROC analysis: 24h-30min saving index discriminated severe SCD from mild SCD (AUC = 0.77) - EMQ correlated with 24h-30min saving index ($r = -0.46$; $p < .01$) - fMRI: ALF associated with reduced FC in DMN, FPN, and SAL	- SCD severity subgroups - Integration of cognitive, recognition, and fMRI measures	Small sample size with gender imbalance in severe SCD

Author(s), Year, Country	Study Design	Sample	Long-Term Delayed Recall Measure		Learning Criterion	Recall Intervals	Main Outcomes	Strengths	Limitations
			Verbal	Visual					
Ruggeri et al., 2024 Rome (Italy)	Cross-sectional	<i>N</i> = 96 (70 HC, 26 SMD)	Rey Auditory Verbal Learning Test	Rey-Osterrieth Complex Figure	Not predefined	• Immediate • 15 min • 1 week	SMD participants showed poorer one-week RAVLT recall and recognition, despite intact short-delay performance. • RAVLT one-week recall: HC = 6.7 ± 3.3; SMD = 4.0 ± 2.9; $F_{1,94} = 12.06$; $p = .001$ • RAVLT one-week recognition hits: HC = 12.4 ± 2.3; SMD = 11.1 ± 2.6; $F_{1,90} = 5.33$; $p = .023$	• Direct comparison of verbal and visual long-term recall • Ecological remote administration feasible for large-scale studies	Increased probability of Type I errors
Saloner et al., 2018 San Francisco (US)	Cross-sectional	<i>N</i> = 140 CU	• Story Recall Test • CVLT-II		90% accuracy	• Immediate • 30 min • 1 week	One-week story recall was associated with MTL volumes (partial $r = 0.24$; $p = .006$). • Story recall (18.4 ± 1.5 at 30 min; 12.8 ± 3.8 at one-week) • Specific associations with entorhinal ($r = 0.27$; $p = .001$) and hippocampal ($r = 0.21$; $p = .02$) volumes	Direct MRI correlations	Small to medium effect sizes
Stamate et al., 2020 Bucharest (Romania)	Cross-sectional (between-subjects: single test vs repeated testing)	• <i>N</i> = 75 (33 AD, 42HC) • Subsamples: AD repeated testing = 19, AD single test = 14; CU repeated testing = 21, CU single test = 21	Fables cued-recall test		70% accuracy	• Immediate • 24h • 1 week • 1 month	No evidence of ALF from immediate to one month. Repeated testing improved 1-month recall in both groups: • AD repeated = 5.1 ± 2.3; AD single = 2.0 ± 1.5; $p = .0002$ • HC repeated = 7.4 ± 2.5; HC single = 3.5 ± 1.6; $p = .0001$	Lower learning criteria than other studies	Some missing follow-ups
Tort-Merino et al., 2021a Barcelona (Spain)	Cross-sectional	<i>N</i> = 22 CU (11 <i>APOE</i> ε4 carriers, 11 <i>APOE</i> ε4 non-carriers)	AFE-T		Two learning sessions across 2 days; no predefined accuracy criterion	• Immediate • 1 week • 3 months • 6 months	<i>APOE</i> ε4 carriers showed greater forgetting at 3 months, particularly in cued recall. • Carriers = 7.6 ± 6.4, non-carriers = 13.3 ± 5.5; $F_{1,19} = 4.99$; $p = .037$ • 3-month cued forgetting correlated negatively with CSF Aβ42/ptau ratio ($r = -.614$; $p = .003$) and positively with SCD-Q complaints ($r = .636$; $p = .002$)	• Use of demanding associative memory task (AFE-T) sensitive to subtle deficits • Long follow-up (6 months)	• Small sample size • Longer learning session compared with other studies: two consecutive days, each with 7 learning runs (≈ 45 min per day)

Author(s), Year, Country	Study Design	Sample	Long-Term Delayed Recall Measure		Learning Criterion	Recall Intervals	Main Outcomes	Strengths	Limitations
			Verbal	Visual					
Tort-Merino et al., 2021b Barcelona (Spain)	Cross-sectional	$N = 52$ CU (Low SCD-Q = 21), High SCD-Q = 31) - Subsample of participants with complete CSF biomarkers data: $N = 25$ (Aβ+/Aβ-)	AFE-T		Two learning sessions across 2 days; no predefined accuracy criterion	- Immediate 1 week 3 months 6 months	High SCD-Q group showed greater forgetting at 3 months ($F_{1,44} = 4.72$; $p = .035$). Within high SCD-Q, Aβ+ performed worse than Aβ- on: - One-week free recall ($F_{1,24} = 4.49$; $p = .044$) - One-week cued recall ($F_{1,24} = 7.10$; $p = .008$) - One-week cued forgetting rate ($F_{1,24} = 5.13$; $p = .032$) 3-month cued recall ($F_{1,24} = 4.27$; $p = .049$) - SCD-Q scores predicted higher forgetting rates ($\beta = -0.212$; $p < .05$)	- Use of demanding associative memory task (AFE-T) sensitive to subtle deficits - Integration of subjective complaints and CSF biomarkers	- Small biomarker subsample - Dropouts at 3- and 6-month follow-up
Walsh et al., 2014 San Francisco (US)	Cross-sectional	$N = 30$ (15 HC, 15 aMCI)	Story Recall Task		90% accuracy	- Immediate 30 min 1 week	aMCI subjects showed ALF. - One-week recall: significant group $\times$ time interaction ($F_{1,31,27} = 11.62$; $p = .001$); aMCI recalled less overall ($F_{1,29} = 22.97$; $p < .001$) - Steeper forgetting slope in aMCI vs. HC across one-week ($F_{1,26} = 10.32$; $p = .003$)	Explicitly isolates consolidation deficits in aMCI	aMCI diagnostic criteria were slightly less strict than Petersen/International Consensus (possible heterogeneity)
Wearn et al., 2020 Bristol (UK)	Longitudinal (1 year follow-up)	$N = 46$ CU	- CVLT-II - RBMT-3 (story recall)	Rey-Osterrieth Complex Figure	75% accuracy	- Immediate 30 min 1 month	- One-month word list recall predicted one-year cognitive decline (ΔACE-III; $R^2 = .123$; $p = .018$) - ROC analysis: 1-month word list recall AUC = .752; $p = .007$; best classification when combined with hippocampal subfield volumes (AUC = .918; 93% sensitivity; 86% specificity)	Integration with MRI hippocampal subfield volumes	No biomarker confirmation

Author(s), Year, Country	Study Design	Sample	Long-Term Delayed Recall Measure		Learning Criterion	Recall Intervals	Main Outcomes	Strengths	Limitations
			Verbal	Visual					
Weston et al., 2018, London (UK)	Cross-sectional	<i>N</i> = 35 CU (21 ADAD mutation carriers, 14 non-carriers)	AMIPB (list and story recall)	AMIPB (figure recall)	80% accuracy	• Immediate • 30 min • 1 week	Carriers showed greater forgetting than non-carriers at one-week delayed recall. • Word list: $\Delta M = -30.94$ pp; 95% CI -45.16, -16.73; $p = .0002$ • Story: $\Delta M = -20.10$ pp; CI 33.28, -6.91; $p = .0048$ • Figure: $\Delta M = -15.41$ pp; CI 26.88, -3.93; $p = .012$ Severity of ALF correlated with proximity to estimated symptom onset ($p \leq .01$)	First demonstration of ALF in presymptomatic ADAD	EYO based on parental onset (proxy)
Yang et al., 2021 Beijing (China)	Cross-sectional	<i>N</i> = 39 (18 ADAD mutation carriers, 21 non-carriers)	WHO-UCLA-AVLT	Rey-Osterrieth Complex Figure	80% accuracy	• Immediate • 30 min • 1 week	Carriers showed normal 30-min recall but greater forgetting at one-week. • List recall: 53.3% vs. 66.7%; $p < .001$ • Figure recall: 38.9% vs. 58.3%; $p = .002$ • List recognition: 70.0% vs 80.0%; $p = .006$ • Forgetting correlated with EYO ($r = -.69$; $p = .002$) and with blood biomarkers ($\downarrow$THBS4, $\uparrow$GDF11, $\uparrow$CCL11; all $p \leq .007$).	Integration of ALF measures with senescence-related blood biomarkers	• EYO based on parental onset (proxy) • Blood assay specificity issues (e.g., GDF11 vs. myostatin)

Author(s), Year, Country	Study Design	Sample	Long-Term Delayed Recall Measure		Learning Criterion	Recall Intervals	Main Outcomes	Strengths	Limitations
			Verbal	Visual					
You et al., 2026 Boston (US) Medellín (Colombia)	Cross-sectional (COLBOS biomarker study)	$N = 52$ (28 non-demented <i>PSEN1</i> E280A carriers, 24 non-carriers) - Carriers FAST 1-3, therefore including some participants with MCI - PET subsamples: PiB $n = 31$; flortaucipir $n = 32$	NEUROPSI 12-Word List (three learning trials)		Not predefined (fixed at three learning trials)	- Learning 20 min 7 days	- Time $\times$ carrier status interaction: $\beta = -3$, 95% CI [-4.8, -1.2]; $p = .001$ - Learning: $\beta = -0.51$, 95% CI [-1.2, 0.2]; $p = .16$ 20-min recall: $\beta = -1$, 95% CI [-2.3, 0.28]; $p = .19$ 7-day recall: $\beta = -1.9$, 95% CI [-3.2, -0.56]; $p = .0008$ - Findings unchanged after excluding carriers with MCI - Carriers showed reduced parietal sleep spindle power ($\beta = -0.46$, 95% CI [-0.89, -0.03]; $p = .04$) - Spindle power inversely associated with precuneus tau ($\beta = -0.36$, 95% CI [-0.6, -0.1]; $p = .008$); no association with amyloid	- Integration of ALF, sleep microarchitecture and amyloid/tau PET - Proposes a candidate neurophysiological mechanism - ALF finding robust to exclusion of participants with MCI	- Verbal material only; no visual measure - Non-demented rather than strictly cognitively unimpaired sample - PET available in approximately 60% of the sample - High prevalence of diagnosed sleep apnea ($\approx 48\%$ in both groups)
Zimmermann & Butler, 2018 London (UK)	Cross-sectional	$N = 60$ CU (20 <i>APOE</i> $\epsilon 4/\epsilon 4$, 20 <i>APOE</i> $\epsilon 3/\epsilon 4$, 20 <i>APOE</i> $\epsilon 3/\epsilon 3$)	AMIPB (list and story recall)	AMIPB (figure recall)	80% accuracy	- Immediate 30 min 1 week	Encoding and 30-min recall did not differ across genotypes (all ns). At one week, $\epsilon 4$ status associated with greater forgetting ($p = .042$), increasing linearly with number of $\epsilon 4$ alleles ($p = .021$). Recognition at one week also impaired in $\epsilon 4$ carriers ($p = .011$; linear trend $p = .0043$). Relationship between learning rate and long-term forgetting: - Present in $\epsilon 3/\epsilon 3$ ($p = .0051$), - Reduced in $\epsilon 3/\epsilon 4$ ($p = .020$), - Absent in $\epsilon 4/\epsilon 4$ ($p = .17$).	Evidence of ALF in sporadic AD genetic risk	Relatively young sample (mean age ~ 51 y), well below typical symptom onset

Note. Several included studies draw on the same cohorts, and samples are therefore partially overlapping: the UCL autosomal dominant Alzheimer’s disease cohort (Weston, 2018; O’Connor, 2020; Parker, 2025; Magill, 2026), the Harvard Aging Brain Study and affiliated Mass General Brigham cohorts (Molinare, 2025), and the COLBOS *PSEN1* E280A kindred (You, 2026). These studies report distinct outcomes (cross-sectional group differences, structural

correlates and longitudinal prediction), but the number of independent participants is lower than the sum of the reported sample sizes. **Abbreviations:** ACE-III = Addenbrooke's Cognitive Examination, Third Edition; AD = Alzheimer's Disease; ADAD = Autosomal Dominant Alzheimer's Disease; AFE-T = Ancient Farming Equipment Test; ALF = Accelerated Long-term Forgetting; aMCI = amnesic Mild Cognitive Impairment; AMIPB = Adult Memory and Information Processing Battery; *APOE* = Apolipoprotein E; *APP* = Amyloid Precursor Protein; AUC = Area Under the Curve; AVLT = Auditory Verbal Learning Test; A β = Amyloid β ; BRANCH = Boston Remote Assessment for NeuroCognitive Health; CCL11 = CC Chemokine Ligand 11; CDR = Clinical Dementia Rating; CERAD/ADAS-Cog = Consortium to Establish a Registry for AD/AD Assessment Scale-Cognitive Subscale; COLBOS = Colombia-Boston biomarker study; CU = Cognitively Unimpaired; CVLT-II = California Verbal Learning Test, Second Edition; DMN = Default Mode Network; EMQ = Everyday Memory Questionnaire; EYO = Estimated Years to Onset; FAD = Familial Alzheimer's Disease; FAST = Functional Assessment Staging Tool; FC = Functional Connectivity; FCSRT = Free and Cued Selective Reminding Test; FCSRT-IR = Free and Cued Selective Reminding Task with Immediate Recall; FNAME-12 = Face Name Associative Memory Exam, 12 items; FPN = Frontoparietal Network; GDF11 = Growth Differentiation Factor 11; GDS = Geriatric Depression Scale; HABS = Harvard Aging Brain Study; HC = Healthy Controls; IADL = Instrumental Activities of Daily Living; IQR = Interquartile Range; LTF = Long-Term Forgetting; LTR = Long-Term Retrieval; MCI = Mild Cognitive Impairment; MTL = Medial Temporal Lobe; NEUROPSI = Neuropsi Atención y Memoria battery; ns = non-significant; PET = Positron Emission Tomography; pFDR = False Discovery Rate corrected *p*-value; PiB = Pittsburgh Compound B; *PSEN1* = Presenilin 1 gene; RAVLT = Rey Auditory Verbal Learning Test; RBMT-3 = Rivermead Behavioural Memory Test, Third Edition; ROCF = Rey-Osterrieth Complex Figure; ROI = Region of Interest; SAL = Salience Network; SCD = Subjective Cognitive Decline; SCD-Q = Subjective Cognitive Decline Questionnaire; SMD = Subjective Memory Decline; SRT = Story Recall Test; tau = tubulin associated unit; THBS4 = Thrombospondin-4; TR = Total Recall; UCL = University College London; WHO-UCLA-AVLT = World Health Organization-University of California Los Angeles Auditory Verbal Learning Test.

Appendix B Secondary trajectory outcomes and the specificity of the one-week retention window.

B.1 Secondary trajectory outcomes

The primary outcome models the trajectory of total recall across the three immediate trials, the thirty-minute recall and the one-week recall. Five secondary trajectory outcomes were modeled by the identical procedure described in Section 5.6: the trajectory of total recall truncated at the three immediate trials, the trajectory of total recall truncated at thirty minutes, and the three corresponding trajectories of free recall. Truncating the protocol in this way isolates what each retention window contributes, and comparing the resulting classes against the same biological outcomes tests whether the associations reported in Chapter 6 are specific to the one-week interval or are already present at shorter delays.

Four of the five yielded an admissible solution under the minimum class-size rule. Their fit and class structure are given in Table B1, and their class means at each timepoint in Table B2. Two points must be carried forward from Section 5.6.2. First, the number of classes was fixed at three a priori for comparability, and for three of these four outcomes the Bayesian information criterion would have preferred two classes; the third class is an analytic constraint and not a finding. Second, one outcome - the free-recall trajectory culminating at one week - produced no admissible solution at all, which is itself reported below as a result rather than as a failure.

Table B1. Retained solutions for the four secondary trajectory outcomes with an admissible fit. Class sizes are given as n (%).

Outcome	Timepoints	Specification	BIC	Entropy	Low	Intermediate	High	BIC-preferred classes
Total recall, learning	3	GMM, intercept + slope	2,845.0	0.732	31 (9.6)	97 (30.1)	194 (60.2)	2
Total recall, 30 minutes	4	Quadratic GMM	3,517.0	0.780	64 (19.9)	74 (23.0)	184 (57.1)	3
Free recall, learning	3	LCGA	4,230.1	0.761	51 (15.8)	218 (67.7)	53 (16.5)	2
Free recall, 30 minutes	4	Quadratic GMM	5,563.3	not recorded	39 (12.1)	180 (55.9)	103 (32.0)	2
Free recall, one week	5	no admissible solution	-	-	-	-	-	-

Note. GMM = growth mixture model; LCGA = latent class growth analysis. Classes are labeled Low, Intermediate and High performers, ordered by class mean; the label Intermediate is reserved for the two outcomes that include the one-week trial. The final column gives the number of classes the Bayesian information criterion would select if the three-class structure were not imposed. Entropy for the free-recall thirty-minute solution was not retained by the estimation log and would require refitting to recover.

Table B2. Mean total or free recall by class and timepoint for the four secondary trajectory outcomes.

Outcome	Class	Trial 1	Trial 2	Trial 3	30 min
Total recall, learning trials only (0-16 per trial)					
	Low performers	10.77	12.68	13.29	-
	Intermediate performers	12.95	14.68	15.43	-
	High performers	15.31	15.71	15.84	-
Total recall, learning + thirty minutes					
	Low performers	11.77	13.39	14.14	13.00
	Intermediate performers	13.27	15.04	15.61	15.30
	High performers	15.36	15.73	15.88	15.72
Free recall, learning trials only					
	Low performers	4.94	5.82	7.14	-
	Intermediate performers	7.37	9.64	10.84	-
	High performers	10.83	12.36	13.55	-
Free recall, learning + thirty minutes					
	Low performers	5.67	6.33	7.28	5.23
	Intermediate performers	6.40	9.48	10.85	10.68
	High performers	10.29	10.67	11.73	12.17

Note. None of these four solutions exhibits the crossing pattern that defines the Intermediate class of the primary outcome: in every case the three classes preserve their ordering at every timepoint, differing in level rather than in shape. This is the expected result if the dissociation reported in Section 6.1.2 is specific to the one-week interval.

B.2 The free-recall one-week trajectory

The trajectory of free recall culminating at one week produced no solution meeting the criteria of Section 5.6.2, and no class variable was therefore derived from it. Across sixty random starts for each three-class specification, the quadratic growth mixture model and the growth mixture model with a random

intercept failed to converge at every start; the model with a random intercept and slope converged at three starts of sixty, with a smallest class of 4.97%; and the latent class growth model converged at fifty-nine of sixty starts but with a smallest class of 2.80%. The only specification producing well-populated classes was a two-class latent class growth model, which fits approximately 1,200 BIC units worse than the best available model.

The explanation is a floor effect, and it is the mirror image of the ceiling effect reported in Section 6.1. Free recall excludes the cued component, and after a week many participants recall very few items without a cue. The distribution is compressed against zero, leaving too little variance in the shape of the curve for a mixture model to resolve. This is a substantive finding about the measure rather than a technical failure: it indicates that the cued component is necessary at the one-week interval, and that total recall, not free recall, is the appropriate score when the retention interval is extended to a week in a cognitively unimpaired sample. Studies planning to adopt this paradigm should note it.

B.3 Specificity of the one-week window

The two secondary outcomes based on total recall, namely the learning trajectory and the thirty-minute trajectory, were tested against the same three families of biological and subjective outcomes as the primary outcome. The same models, covariates and correction for multiple comparisons were used for each trajectory predictor. Table B3 compares the corrected p-values obtained for the three retention windows. The results indicate a differentiated pattern across retention windows, which qualifies the interpretation of the findings presented in Chapter 6. Associations varied according to the outcome domain examined, and the results did not indicate a uniform increase in sensitivity with longer retention intervals.

Table B3. Corrected p-values for the same outcomes under three trajectory predictors: classes derived from the three learning trials, from the learning trials plus thirty-minute recall, and from the full protocol including the one-week recall. Correction is by the Benjamini-Hochberg procedure within family, applied separately for each predictor.

Family and outcome	Learning trajectory	30-minute trajectory	One-week trajectory
Fluid biomarkers and subjective cognition			
Plasma p-tau217	.0006 *	.034 *	.021 *
CSF A β 42/A β 40	.042 *	.276	.211
CSF p-tau181/A β 42	.005 *	.121	.021 *
MyCog total	.267	.034 *	.089
MyCog memory	.124	.039 *	.065
Amyloid status	.547	.984	.468
Tau positivity (A+T+ vs A+T-)	.109	.276	.065
Cross-sectional MRI			
Bilateral MTL	.021 *	.905	.493
Bilateral CA1	.036 *	.905	.779
Dentate gyrus / CA4	.021 *	.905	.493
Rostral middle frontal thickness	.036 *	.710	.567
Entorhinal thickness	.917	.905	.992
Lateral occipital thickness (control)	.236	.905	.992
Longitudinal MRI (predictor $\times$ time)			
Bilateral CA1	.289	.141	.019 *
Bilateral MTL	.447	.141	.034 *
Dentate gyrus / CA4	.751	.381	.087
Rostral middle frontal thickness	.289	.313	.592
Lateral occipital thickness (control)	.751	.770	.682

Note. MTL = medial temporal lobe. * survives correction at the 5% level. Only the outcomes needed to show the pattern are listed. The three trajectory predictors were processed identically, using the same models, the same covariates and the same correction procedure.

Concurrent biological measures, including biomarker concentrations measured at the second visit and brain structure measured at the third visit, were more frequently associated with the learning trajectory than with the trajectories that included the thirty-minute or one-week assessments. In contrast, the rate of structural change across the three MRI visits was associated only with the full trajectory including the one-week recall. This pattern suggests that the different retention windows may capture partly distinct aspects of memory performance rather than representing progressively more sensitive versions of the same measure. Taken together, these findings suggest that one-week delayed recall provides complementary information that is not fully captured by standard short-delay assessment, particularly with regard to longitudinal hippocampal change.