Alzheimer's disease-associated biomarkers in delirium and cognitively normal older adults

Thesis by
Ane-Victoria Idland

Department of Geriatric Medicine
Institute of Clinical Medicine
Faculty of Medicine
University of Oslo

2020
Table of Contents

Acknowledgements .................................................................................................................... 5
Summary .................................................................................................................................... 7
Sammendrag ............................................................................................................................... 9
Abbreviations ........................................................................................................................... 11
Articles in the thesis ................................................................................................................. 13
1 Introduction ...................................................................................................................... 15
   1.1 Alzheimer’s disease ................................................................................................... 15
      1.1.1 Introduction ........................................................................................................ 15
      1.1.2 Neuropathology ................................................................................................. 16
      1.1.3 Pathophysiology ............................................................................................... 19
      1.1.4 Alzheimer’s disease-associated biomarkers ....................................................... 21
      1.1.5 Biomarkers in criteria for Alzheimer’s disease .................................................. 28
   1.2 Cognitively normal older adults ................................................................................ 30
      1.2.1 What is normal cognition? ................................................................................. 30
      1.2.2 Cognitive changes in cognitively normal adults ................................................ 31
      1.2.3 Brain atrophy in cognitively normal adults ........................................................ 32
      1.2.4 Amyloid and tau pathology in cognitively normal adults .................................. 33
      1.2.5 Other neuropathological changes in cognitively normal adults ......................... 36
      1.2.6 Resilience against neuropathology ..................................................................... 38
   1.3 Delirium ..................................................................................................................... 39
      1.3.1 Introduction ........................................................................................................ 39
      1.3.2 Delirium diagnosis ............................................................................................. 39
      1.3.3 Risk factors for delirium .................................................................................... 40
      1.3.4 Pathophysiology ............................................................................................... 41
      1.3.5 The interrelationship between delirium and dementia ....................................... 42
2 Aims of the thesis ............................................................................................................. 45
3 Materials and Methods ..................................................................................................... 46
   3.1 Cohorts and study designs ......................................................................................... 46
      3.1.1 Elective surgery cohort ....................................................................................... 46
      3.1.2 Hip fracture cohort ............................................................................................. 47
   3.2 Clinical assessments ................................................................................................... 49
Acknowledgements

The research presented in this thesis was carried out at the Department of Geriatric Medicine, Institute of Clinical Medicine, University of Oslo and Oslo University Hospital and also partly at the Department of Psychology, University of Oslo. The research was mainly funded by the Medical Student Research Program at the University of Oslo and the National Association for Public Health’s dementia research program, Norway.

Firstly, I would like to thank the participants in the studies and their next of kin/care takers for their participation in the projects. I do also want to thank the staff at the Departments of Urology, Gynecology, Anesthesiology, and Orthopedic Surgery at Oslo University Hospital, and the Departments of Anesthesiology and Orthopedic Surgery at Diakonhjemmet Hospital, for letting me recruit patients from your departments and for the collection of CSF samples.

My supervisors also deserve a big thank you. I would like to express my deepest gratitude to my principal supervisor, Leiv Otto Watne. Thank you for sharing all your knowledge about delirium, for letting me study your cohort of hip fracture patients, and for your ongoing support, enthusiasm, positivity and good sense of humor. To my co-supervisor Torgeir Bruun Wyller, thank you for giving me the opportunity to do research in the field of geriatrics when I was a medical student, and for your patience, guidance and for always being available. I do also want to thank you for the opportunity to work in the field of geriatrics after becoming a medical doctor. Special thanks should also be given to my co-supervisor Anders Martin Fjell, thank you for making more follow-ups and assessments of the elective surgery cohort possible, for introducing me to the field of brain aging, for inviting me to work in your young and lively research group, for challenging me to not always see brain and cognitive changes as pathological, and for your guidance and good advices. To my co-supervisor Anne Brækhus, thank you for your contributions in planning the project including elective surgery patients, for teaching me cognitive testing, for all discussions about cognitive test results, and for your commitment and support.

I would also like to thank all other co-authors for your efforts and valuable contributions to the manuscripts. I would especially express my gratitude to Roser Sala-Llonch, for being a great officemate, for processing MRI data, and doing statistical analyses, to Kaj Blennow and Henrik Zetterberg for doing interesting CSF analyses, to Randi Støen for enthusiastically...
collecting CSF samples and showing an interest in delirium research, and to Kristine Walhovd for allowing me to be a part of your research group and for stimulating academic discussions.

I do also want to thank my colleagues at “Loftet”, in Oslo Delirium Research Group, and in the research group for Lifespan Changes in Brain and Cognition for good companionship and support. Most especially, thanks to Knut Engedal for planning the project including elective surgery patients and teaching me cognitive testing, and to Nathalie Bodd Halaas and Shams Dakhil for administrating the latest follow-ups of the elective surgery cohort.

Finally, a special thanks to my family, in-laws, and friends. Thank you for always believing in me, for encouraging me and cheering me on. Most especially, thanks to my ever-supportive husband, Gaute, for introducing me to the Medical Student Research Program, for your help with data cleansing and technical issues, and for your patience and love.
Summary

**Background:** Higher age comes with a multitude of changes in the brain, but the role of different neuropathological processes in cognitively normal older adults is far from fully understood. A substantial number of cognitively normal older adults show biomarker evidence of brain pathologies often associated with Alzheimer’s disease (AD), but the implications are not fully established. Furthermore, cognitively normal older adults often show hippocampal atrophy and memory decline, which are hallmarks of AD dementia. Older age is also a risk factor for delirium, a condition characterized by an acute disturbance in attention, awareness, and cognition. Delirium has epidemiologically been strongly linked to dementia, however the neuropathological processes underlying this relationship are not understood.

**Aims:** The overall aim was to increase knowledge about the role of different neuropathological processes in delirium and in cognitively normal older adults using AD-associated cerebrospinal fluid (CSF) biomarkers reflecting core AD pathology (Aβ42 reflecting Aβ deposition, T-tau reflecting neurodegeneration, P-tau reflecting tauopathy) and novel biomarkers (NFL reflecting axonal damage, YKL-40 reflecting neuroinflammation, and FABP3 reflecting neuronal damage). More specifically, we wanted to assess relationships between biomarkers, test whether subgroups of individuals with similar biomarker profiles could be identified using clustering analyses, and examine the relationship between biomarkers and longitudinal hippocampal atrophy and relationships between biomarker-based subgroups and longitudinal hippocampal atrophy and memory change, respectively, in cognitively normal older adults. We also wanted to examine whether the core AD biomarkers are related to delirium.

**Methods:** CSF from cognitively normal older adults (n=99) was assessed for CSF Aβ42, T-tau, P-tau, NFL, YKL-40, and FABP3, and CSF from hip fracture patients (n=129) was assessed for CSF Aβ42, T-tau, and P-tau. Delirium was assessed pre- and postoperatively in hip fracture patients, and the diagnosis of dementia at admission was based upon clinical consensus. Hippocampal volume and memory in cognitively normal adults was assessed across multiple follow-up examinations over up to 6 years.

**Main results:** The novel biomarkers NFL, YKL-40, and FABP3 were linked to T-tau and P-tau, but not to Aβ42. Concentrations of NFL, YKL-40, and FABP3 differed between at least
two biomarker subgroups identified by clustering analyses with a relatively large effect size. High NFL levels predicted higher hippocampal atrophy rate in cognitively normal older adults, independently of core AD biomarkers, and also in subgroups unlikely to have preclinical AD. A clustering-based subgroup characterized by high concentrations of T-tau, P-tau, and FABP3 showed more memory decline than biomarkers groups with less abnormal biomarker levels, whereas biomarker groups based on only Aβ42 and P-tau showed no differences in memory trajectories. The core AD biomarkers (low CSF Aβ42, high CSF T-tau, low Aβ42/T-tau, low Aβ42/P-tau) were significantly associated with delirium in patients without dementia.

**Conclusion:** We have shown that CSF biomarkers previously associated with AD are also associated with delirium, hippocampal atrophy, and memory decline in individuals without dementia. Our findings suggest that neuronal damage, axonal damage, and neuroinflammation are accompanying tauopathy and neurodegeneration, but not Aβ deposition, in this group. Our results further suggest that it is meaningful to use the three novel biomarkers for characterization of brain states in cognitively normal older adults, and that addition of novel biomarkers improve prediction of memory decline compared to classification based on biomarkers of Aβ deposition and tauopathy only. Moreover, our findings suggest that a high degree of neurodegeneration and tauopathy is associated with more hippocampal atrophy and greater memory decline in cognitively normal adults. Lastly, our findings in the hip fracture cohort suggests that AD pathologies may underlie the interrelationship between delirium and dementia, and raises the question of whether delirium is an early symptom of AD.
Sammendrag

**Bakgrunn:** Det skjer en rekke ulike endringer i hjernen når vi blir eldre, men betydningen av de ulike nevropatologiske prosessene som foregår i hjernen hos kognitivt normale eldre er dårlig forstått. En stor andel av kognitivt normale eldre har unormale nivåer av biomarkører som ofte er assosiert med Alzheimers sykdom, men vi vet ikke nok om konsekvensene av disse unormale biomarkørene. Atrofi av hippocampus og gradvis dårligere hukommelse er karakteristisk for Alzheimers demens, men disse endringene er også vanlige hos eldre mennesker med normal kognitiv funksjon. Økende alder er også en risikofaktor for delirium, en tilstand som er karakterisert ved en akutt endring i oppmerksomhetsfunksjoner og kognitiv funksjon. Delirium har epidemiologisk en sterk sammenheng med demens, men de nevropatologiske prosessene som ligger bak denne sammenhengen er vet vi svært lite om.

**Mål:** Hovedmålet i denne doktorgradsavhandlingen var å øke kunnskapen om hvilke roller ulike nevropatologiske prosesser spiller ved delirium og hos kognitivt friske eldre individer ved bruk av Alzheimer-assosierede spinalvæskebiomarkører som reflekterer Alzheimer-patologi (Aβ42 reflekterer amyloidavleiring, T-tau reflekterer nevrodegenerasjon, P-tau reflekterer tauopati) og nyere biomarkører (NFL som reflekterer aksonskade, YKL-40 som reflekterer nevroinflammasjon og FABP3 som reflekterer nevronskade). Hos kognitivt normale eldre ønsket vi å undersøke sammenhenger mellom de ulike biomarkørene, om man kunne identifisere subgrupper av individer med like biomarkørprofiler ved hjelp av clusteringanalyser, samt undersøke sammenhenger mellom biomarkører og hippocampusatrofi over tid og mellom biomarkør-baserte subgrupper og hhv. hippocampusatrofi og hukommelsesendringer over tid. Vi ønsket også å undersøke om spinalvæskebiomarkører for Alzheimer-patologi er assosiert med delirium.

**Metode:** Spinalvæskebiomarkørene Aβ42, T-tau, P-tau, NFL, YKL-40 og FABP3 ble analysert hos kognitivt normale eldre (n=99), og spinalvæskebiomarkørene Aβ42, T-tau og P-tau ble analysert hos pasienter med hoftebrudd (n=129). Pasienter med hoftebrudd ble undersøkt for delirium pre- og postoperativt, og demensdiagnose på innleggsstidspunktet var basert på klinisk konsensus. Hippocampusvolum og hukommelsesfunksjon hos kognitivt normale eldre ble undersøkt på flere tidspunkter over en periode på inntil 6 år.
**Hovedresultater:** De nye biomarkørene NFL, YKL-40 og FABP3 var assosiert med T-tau og P-tau, men ikke med Aβ42. Relativt store forskjeller i konsentrasjoner av hver av biomarkørene NFL, YKL-40 og FABP3 skilte minst to av subgruppene som ble identifisert i clusteringanalyser. Høye NFL-nivåer var assosiert med høyere atrofirate av hippocampus hos kognitivt normale eldre. Sammenhengen mellom NFL og hippocampusatrofi var uavhengig av Aβ42 og P-tau, og den var også tilstede i subgrupper med lav sannsynlighet for preklinisk Alzheimers sykdom. En clustering-basert subgruppe karakterisert av høye nivåer av T-tau, P-tau og FABP3 hadde mer reduksjon av hukommelsesfunksjon over tid enn biomarkørgrupper med mindre unormale biomarkørnivåer. Det var imidlertid ingen forskjell i hukommelsesfunksjon over tid mellom biomarkørgrupper basert på kun Aβ42 og P-tau.

Spinalvæskebiomarkører for Alzheimers sykdom (lav Aβ42, høy T-tau, lav Aβ42/T-tau, lav Aβ42/P-tau) var signifikant assosiert med delirium blant pasienter uten demens.

**Konklusjon:** Vi har vist at spinalvæskebiomarkører som tidligere har vært assosiert med Alzheimers sykdom også er assosiert med delirium, hippocampusatrofi, og reduksjon av hukommelsesfunksjon hos personer uten demens. Funnene våre antyder at nevronskade, aksonskade og nevroinflammasjon ledsager tauopati og nevrodegenerasjon, men ikke amyloidavleiring, i denne gruppen. Funnene våre tyder også på at det er meningsfullt å bruke de tre nyere biomarkørene til å karakterisere hjernetilstander hos kognitivt normale eldre, samt at bruk av disse nye biomarkørene i tillegg til etablerte biomarkører for Alzheimer-patologi øker evnen til å predikere hukommelsesendring sammenlignet med klassifisering basert utelukkende på biomarkører for amyloidavleiring og tauopati. Videre antyder funnene våre at en høy grad av nevrodegenerasjon og tauopati er assosiert med mer hippocampusatrofi og reduksjon av hukommelsesfunksjon hos kognitivt normale eldre. Funnene våre i hoftebruddskohorten antyder at de nevropatologiske endringene som karakteriserer Alzheimers sykdom kan ligge bak sammenhengen mellom delirium og demens, og de reiser spørsmålet om delirium kan være et tidlig symptom på Alzheimers sykdom.
Abbreviations

A+  Amyloid positive measured by amyloid PET, CSF Aβ42, or Aβ42/Aβ40 ratio
Aβ  Amyloid-β
Aβ42  Amyloid-β with 42 amino acids
AD  Alzheimer's disease
APACHE II  Acute Physiology and Chronic Health Evaluation II
ApoE  Apolipoprotein E
APP  Amyloid Precursor Protein
CAM  The Confusion Assessment Method
CERAD  Consortium to Establish a Registry for Alzheimer’s Disease
CNS  Central Nervous System
CSF  Cerebrospinal Fluid
DLB  Lewy Body Dementia
FABP3  Fatty Acid-Binding Protein 3
FDG  Fluorodeoxyglucose
FTD  Frontotemporal Dementia
GAMMs  Generalized Additive Mixed Models
ICD-10  International Classification of Diseases, 10th edition
IQCODE-SF  Informant Questionnaire on Cognitive Decline in the Elderly short form
MCI  Mild Cognitive Impairment
MMSE  The Mini Mental Status Examination
MRI  Magnetic Resonance Imaging
MTL  Medial Temporal Lobe
N+  Neurodegeneration positive measured by structural MRI, FDG PET, or CSF T-tau
NFL  Neurofilament Light Protein
NFTs  Neurofibrillary Tangles
NIA-AA  National Institute on Aging-Alzheimer’s Association
PART  Primary Age-Related Tauopathy
PD  Parkinson’s disease
PET  Positron Emission Tomography
POCD  Postoperative Cognitive Dysfunction
P-tau  Phosphorylated Tau Protein
SD  Standard deviation
SNAP  Suspected Non-Alzheimer Pathophysiology
T+  Tau positive measured by CSF P-tau or tau PET
TMT  Trail Making Test
TREM2  Triggering Receptor Expressed on Myeloid Cells 2
T-tau  Total Tau Protein
YKL-40  Chitinase-3-like Protein-1
Articles in the thesis

**Paper I:**


**Paper II:**

*Biomarker profiling beyond amyloid and tau – CSF markers, hippocampal atrophy and memory change in cognitively unimpaired older adults.* Idland AV, Sala-Llonch R, Watne LO, Brækhus A, Hansson O, Blennow K, Zetterberg H, Sørensen Ø, Walhovd KB, Wyller TB, Fjell AM. (Submitted manuscript)

**Paper III:**

1 Introduction

This thesis is about Alzheimer’s disease-associated biomarkers in cognitively normal older adults and delirium. I will therefore start with an introduction to Alzheimer’s disease, its neuropathology and pathophysiology, and different Alzheimer’s disease-associated biomarkers. Next, I will introduce the reader to the concepts “delirium” and “cognitively normal older adults”, describe cognitive changes and neuropathological changes occurring in these groups, and explain the relevance of studying Alzheimer’s disease-associated biomarkers in these groups.

1.1 Alzheimer’s disease

1.1.1 Introduction

Alzheimer’s disease (AD) is a neurodegenerative disease characterized by the accumulation of extracellular deposits of abnormally folded amyloid-β (Aβ) peptides (amyloid plaques) and intraneuronal inclusions (neurofibrillary tangles [NFTs]) mainly composed of hyperphosphorylated forms of the microtubule-stabilizing protein tau (1). The disease was first described by Alois Alzheimer in 1906 in a patient with clinical symptoms of dementia (2). AD is the most common cause of dementia (3), which is a syndrome characterized by progressive and chronic cognitive decline that interferes with independence in everyday activities and that is not due to delirium or a major psychiatric disorder, e.g. depression (4, 5). 62 % of all dementia cases are attributable to AD, whereas vascular dementia accounts for 17 %, mixed dementia for 10 %, Lewy body dementia (DLB) for 4 %, frontotemporal dementia (FTD) for 2 %, and Parkinson’s disease (PD) dementia for 2 % of all dementia cases (3). The most characteristic clinical symptom of AD dementia is amnestic cognitive impairment, i.e. problems including learning and recall of newly learned information (5), typically including progressive problems with episodic memory, i.e. the memory of personal experiences occurring at a specific time and place (6-8). Non-amnestic presentations with deficits in word finding, spatial cognition, or executive dysfunction are also common (5). Aging is the major risk factor for AD, and consequently the prevalence of AD dementia will increase as the world’s population ages. Estimates suggest that 50 million people worldwide have dementia (9) increasing dramatically to approx. 132 million in 2050 (10), meaning that AD will account for 82 million dementia
cases in 2050. This will represent a massive societal and economic challenge, yet there is no cure or disease-modifying treatment for AD. Studies targeting AD neuropathology specifically, e.g. Aβ, have failed to show clinical benefit so far (11-14).

Biomarkers can be used to assess neuropathologies in living individuals, and over the two last decades, research criteria for AD have integrated biomarkers of amyloid plaques and NFTs in the diagnostic process (5, 15-17). Earlier stages of AD gained recognition in the 1990s (18). The term mild cognitive impairment (MCI) was introduced and denotes an intermediate state between cognitively normal and dementia, where cognition is impaired, but not to a degree interfering with independence in activities of daily living (19). Research has shown that both individuals with MCI and cognitively normal individuals can have NFTs and amyloid plaques in their brains (12, 18, 20, 21). Accordingly, AD is thought to have a preclinical phase where neuropathology progress over years before cognitive symptoms develop, and since 2010 research diagnostic criteria have incorporated one or more preclinical AD phases (more about this in section 1.1.5), with biomarkers of amyloid plaques and NFTs as the evidence of AD neuropathology in cognitively normal adults (1, 17, 22). Furthermore, treatment of AD is assumed to be most effective in the preclinical phase, before neurodegeneration is too severe (23). Therefore, current avenues of AD research focus on individuals without dementia who have biomarker signs of AD neuropathology and on biomarkers of other neuropathologies that may assist with early diagnosis of AD or with predicting clinical progression. Yet, the implications of such brain pathologies in individuals without dementia are not established (24-28).

1.1.2 Neuropathology

The core neuropathological features of AD are amyloid plaques and NFTs. Neurodegeneration and neuroinflammation are other common features. In addition, cerebral amyloid angiopathy, dystrophic neurites, neuropil threads, granulovacuolar degeneration, and Hirano bodies frequently coexist. Mixed pathology with a combination of AD pathologies and other pathologies, e.g. vascular pathology and Lewy bodies, is also frequent in older individuals with AD dementia (29).
**Amyloid pathology**

Amyloid plaques are extracellular deposits of abnormally folded Aβ peptides with 40 and 42 amino acids (Aβ40 and Aβ42, respectively) (6, 30). Aβ is constitutively secreted from cells under normal physiological conditions (31), and brain Aβ is normally degraded by enzymes and cleared through different mechanisms (32). The Aβ peptides result from sequential cleavage of the transmembrane protein amyloid precursor protein (APP) by β- and γ-secretases. Aβ can accumulate to form soluble oligomers (consisting of 2 to 12 peptides) (30). Oligomers can further aggregate to form insoluble fibrils which successively can amass into plaques (33). There are two kinds of amyloid plaques, diffuse plaques and dense-core plaques. Diffuse plaques are poorly marginated extracellular amyloid deposits, whereas dense-core plaques (also called neuritic plaques) contain a core of fibrillar Aβ that is surrounded by dystrophic neurites, reactive astrocytes and activated microglial cells. Dense-core plaques are used for the pathological diagnosis of AD, while diffuse plaques are commonly found in cognitively intact individuals, and therefore not used for the diagnosis of AD (30). The main constituent of amyloid plaques in AD is the Aβ42 isoform (34). The deposition pattern of amyloid is less predictable than the spatiotemporal progression pattern of NFTs; however, the general pattern is that amyloid mainly accumulates in the neocortex, and only later progresses to involve the allocortex (including medial temporal lobe [MTL] areas like the entorhinal cortex and hippocampus), the basal ganglia, brain stem nuclei, and the cerebellum (30, 35, 36).

**Tau pathology/neurofibrillary tangles**

Tau is a microtubule-associated protein in the axon of neurons that is essential for microtubule stabilization and axonal transport (37, 38). Phosphorylation of tau is believed to cause it to self-aggregate (38), and NFTs are intraneuronal aggregates that are mainly composed of helical filaments of hyperphosphorylated tau (30, 39). In AD, tau is moved to the somatodendritic compartment where hyperphosphorylation of tau results in misfolding and aggregation of the protein into NFTs (30). Hyperphosphorylated tau is also found in neurophil threads (aggregates in dendrites or axons) and in the dystrophic neurites surrounding amyloid plaques (30). Typical tau pathology in AD starts in the entorhinal cortex and hippocampus in the MTL (Braak stages I-III) before it goes on to affect the associative areas of the neocortex (Braak stages IV-VI). The primary sensory, motor and visual cortices are not affected until the last stage (30, 35).
Neurodegeneration

Neurodegeneration is the progressive loss of structure and function of neurons that ultimately results in death of neurons, and it is a common feature of neurodegenerative diseases. Neurodegeneration in AD is typically seen as symmetric cortical atrophy. MTL structures like the entorhinal cortex and hippocampus are affected very early in AD, and the posterior cingulate gyrus and the adjacent precuneus are also affected in an early disease stage (40-42). The atrophy further extends to the rest of the cortex in a temporal-parietal-frontal course (41-47), a pattern mirroring the sequence of NFT accumulation (35). The visual, motor, and sensory cortices are typically spared from atrophy until late disease stages. Neuronal death contributes to atrophy. The neuronal loss affects the same regions and cortical layers as NFTs, and it is shown to be severe in MTL regions (48, 49). However, the neuronal loss is more extensive than the number of NFTs (50), suggesting that there is one mechanism for neuronal death of tangle-bearing neurons and another mechanism affecting tangle-free neurons. Damage and loss of synapses also contributes to atrophy, affecting the same regions and cortical layers as NFTs and neuronal loss (30, 51, 52). As synaptic loss is shown to exceed the neuronal loss, it is though that synaptic loss happens before neuronal loss. Consequently, synaptic loss is also the best anatomical correlate of cognitive performance in AD dementia (53, 54).

Neuroinflammation

Neuroinflammation is the immune-related response of the central nervous system (CNS), including astrocytic and microglial reactions, to alterations in the environment. Neuroinflammation represent a common feature of several neurodegenerative diseases, including PD, FTD, and AD, and the inflammation may promote neurodegeneration and progression of these diseases (55). Glial cells in amyloid plaques were first discovered by Alois Alzheimer (56), and we now know that both reactive astrocytes and activated microglia surround amyloid plaques in brains of patients with AD (57-59). Further, several studies using immunohistochemistry, mRNA-measurements, and other methods, have found abnormal expression of different inflammatory mediators, e.g. cytokines, in the AD brain (60). Moreover, in vivo studies of neuroinflammation, using positron emission tomography (PET)-ligands binding to activated microglia, have shown activated microglia in frontal, temporal, parietal, and occipital association cortical regions, the cingulate cortex, and also in the striatum in patients with AD (61-63), corresponding to regions with increased amyloid load.
1.1.3 Pathophysiology

The amyloid cascade hypothesis

The pathogenesis of AD remains unclear, and several hypotheses regarding AD pathophysiology have been proposed. Nonetheless, the amyloid cascade hypothesis has been the dominant hypothesis in AD research for the past decades (11). The hypothesis suggests that accumulation of Aβ in the brain is the primary pathological process, initiated by an imbalance between production and clearance of Aβ (64). Aβ is in turn thought to cause formation of NFTs and subsequent neuronal dysfunction and neurodegeneration (11). Evidence supporting a central role for Aβ is the fact that all familial AD mutations increase the production of toxic forms of Aβ (11, 65). Moreover, individuals with trisomy 21 (Down’s syndrome) have 3 copies of APP, causing lifelong overproduction of Aβ, and they are shown to have AD neuropathology in the brain already in their teens (11, 66). Further, a beneficial APP mutation, the Icelandic mutation A673T, is shown to reduce Aβ production and protect against sporadic AD dementia and cognitive decline in individuals without AD dementia (67). Furthermore, apolipoprotein E (ApoE) is involved in clearance of Aβ, and the APOE-ε4 allele is shown to be a strong genetic risk factor for sporadic AD (11, 68). Although fibrillar amyloid within plaques was initially presumed to cause Aβ-associated toxicity, soluble Aβ oligomers are now thought to be most toxic. Numerous studies have shown that Aβ oligomers may cause injury to synapses and neuronal processes of brain neurons and induce other AD pathologies, including tau pathology (69-72). Plaques may represent a protective “reservoir” sequestering Aβ oligomers until it reaches a saturation limit, after which oligomers diffuse into the environment (11). The temporal relationship between AD pathologies has been further explored in biomarker studies. Such studies have suggested that amyloid pathology, measured by amyloid PET or cerebrospinal fluid (CSF) Aβ42, starts many years before clinical symptoms develop, and predates change in biomarkers of tau pathology (CSF tau and tau PET) which in turn predates change in biomarkers of neurodegeneration and synaptic dysfunction (structural magnetic resonance imaging [MRI] and fluorodeoxyglucose PET), further supporting the amyloid cascade hypothesis (6, 73).

On the other hand, accumulation of evidence over the past years has provoked a dispute about the validity of the amyloid cascade hypothesis (74-76), and it has been argued that the hypothesis relies on the unconfirmed assumption that amyloid pathology causes AD dementia
Findings that contradict the hypothesis are the poor correlation between amyloid plaque burden and cognitive impairment, the finding of abundant Aβ deposits in cognitively normal older adults, and the numerous unsuccessful clinical trials of anti-amyloid treatments (11-14). Yet, Selkoe et al. have presented counterarguments to these findings based on emerging studies, e.g. that amyloid deposition is an early event causing downstream pathological changes that ultimately leads to cognitive impairment, explaining the poor correlation between amyloid burden and degree of cognitive impairment, that anti-amyloid agents have not been effective because the clinical trials have included patients in late stages of AD, and that individuals without dementia who show abundant amyloid plaques post–mortem have simply not been rigorously tested before death (11, 74). Later versions of the amyloid cascade hypothesis have also integrated other important pathological processes like neuroinflammation and oxidative stress as factors mediating the formation of NFTs and neurodegeneration (11), and the hypothesis still remains the dominant explanatory model of AD. Yet, the dispute about the amyloid cascade hypothesis highlights the need of more studies of the consequences of biomarker signs of amyloid pathology and other neuropathologies in individuals without cognitive impairment.

Other pathophysiological theories

As the literature suggests the causality of AD dementia is likely more complex than the amyloid cascade hypothesis can explain (77), and other pathophysiological mechanisms are also suggested to play the main role in AD pathology, e.g. tau pathology and neuroinflammation. The tau hypothesis proposes that tau pathology occurs before amyloid pathology, and that tau pathology is the primary cause of AD-related neurodegeneration (78). Both neuropathological and biomarker studies support that tau pathology can present prior to amyloid pathology (79, 80), indicating that these pathologies arise independently. Furthermore, tau pathology is more closely linked to cognitive impairment than amyloid pathology (81, 82). Yet, the mechanisms by which tau pathology causes neurodegeneration are not well understood (83).

Increased microglial activation has also been shown to be present early in AD (63, 84), and this has led to a model proposing that microglial activation is a very early event in the disease, perhaps starting even before amyloid deposition (58). Soluble Aβ oligomers and other misfolded and aggregated proteins are recognized by microglia and astrocytes via binding to pattern recognition receptors, and this induces an innate immune response characterized by
release of inflammatory mediators (58, 85). Genome-wide association studies finding that
triggering receptor expressed on myeloid cells 2 (TREM2), a microglial surface receptor, and
also several other genes involved in the innate immune system are associated with sporadic AD
dementia, supports a causal role of neuroinflammation in AD pathogenesis (58, 86, 87). There
is an ongoing debate about whether the inflammation is protective, detrimental, or perhaps both,
and this may depend on disease stage. Nevertheless, the inflammatory process is ultimately
thought to cause functional and structural damage to neurons contributing to disease pathology.
Resulting neuronal damage can in turn activate microglia and possibly create a vicious cycle of
amyloid deposition, tangle formation, neuronal damage, and neuronal death (58, 85).

1.1.4 Alzheimer’s disease-associated biomarkers

A biomarker is an objectively measured indicator of a normal or pathological biological process
or a pharmacological treatment response (88). The measure can be a physiological,
biochemical, or an anatomic parameter (89). Biomarkers can be used as a diagnostic tool,
allowing earlier and more specific identification of pathology, predict disease progression,
depth our understanding of pathogenesis, guide selection of patients with evidence of disease
pathologies to treatment trials, and act as surrogate endpoints in clinical trials (90-92). In AD,
the most established biomarkers can be divided into biochemical CSF biomarkers and imaging
biomarkers. In the next sections, I will mainly focus on the biomarkers used in the papers of
this thesis, but other AD-associated biomarkers will also be briefly mentioned.

Biochemical markers

Biochemical biomarkers for AD have mainly been sought and measured in blood and CSF.
Three waves have been described in the search for biofluid AD biomarkers (93). The first wave
was the discovery and validation of the core AD CSF biomarkers Aβ42, total tau (T-tau) and
phosphorylated tau (P-tau) (Figure 1), giving rise to the CSF AD profile (low levels of Aβ42,
and high levels of T-tau and P-tau). The second wave was the search for additional CSF
biomarkers for other aspects of AD pathophysiology, e.g. neurofilament light (NFL) reflecting
neuroaxonal injury, chitinase-3-like protein-1 (YKL-40) reflecting neuroinflammation, and
fatty acid-binding protein 3 (FABP3) reflecting neuronal damage (Figure 1). The third wave is
the search and development of blood based biomarkers (see examples below).
CSF is a body fluid of which 25% occupies the ventricles of the brain and the rest bathes the brain and the spinal cord, and the intracranial volume ranges from 140 to 270 mL. Production of CSF happens mainly in the choroid plexus located in the ventricles, and further diffuses into the basal cisterns, the subarachnoid space, and the spinal cord. CSF turnover is high (~0.4 mL/min), and the entire volume is replaced around 4 times daily. Some important physiological roles of CSF are to remove metabolic by-products produced by neurons and glial cells, and circulation of biologically active bodies throughout the brain (94). For biomarker discovery in brain disorders, an obvious advantage of CSF is the proximity to the brain parenchyma because brain proteins are secreted from the brain’s interstitial fluid space and into CSF. CSF can be easily collected by lumbar puncture, a diagnostic procedure where a needle is inserted into the subarachnoid space.

Core Alzheimer’s disease cerebrospinal fluid biomarkers

**Aβ42.** Aβ42 is produced by sequential cleavage of APP by two enzymes, β-site APP-cleaving enzyme 1, also called β-secretase, and presenilin complex, also called γ-secretase. CSF Aβ42 is shown to be low in AD (92, 95-97), which may be a result from aggregation of Aβ42 into amyloid plaques, with lower amounts of Aβ42 remaining to be secreted to the CSF (98). CSF Aβ42 concentrations have been shown to correlate inversely with postmortem plaque counts (99, 100) and with plaque load on amyloid PET scans (96, 101, 102), proving that CSF Aβ42 reflects in vivo amyloid pathology. Interestingly, discordancy with low CSF Aβ42 but negative amyloid PET has primarily been found in cognitively normal older adults and patients with early AD, suggesting that CSF Aβ42 may be an earlier biomarker of brain amyloid pathology than amyloid PET (103, 104). Accordingly, studies have shown a good diagnostic accuracy of CSF Aβ42 in MCI and early AD dementia (105, 106). However, abnormal CSF Aβ42 is not specific for AD dementia, as a substantial proportion of patients with other forms of dementia, e.g. vascular dementia, FTD, and DLB, also have abnormal CSF Aβ42, especially among the oldest adults (107, 108).

**Tau.** The tau protein is abundant in neuronal axons in the CNS, and six isoforms of the protein are expressed in the human brain (37, 38). Phosphorylation is a common post-translational modification of tau, and NFTs are composed of abnormally hyperphosphorylated tau proteins (39). CSF T-tau increases in AD (92, 95-97), and the label “total” means that all six isoforms are measured irrespective of phosphorylation state (96). CSF T-tau is suggested to reflect the intensity of axonal degeneration (109), and high levels have been associated with faster disease
progression in MCI and AD dementia (110-113). CSF T-tau increase is not specific to AD, shown by an increase in CSF T-tau levels with traumatic brain damage (114), after stroke (115), and by very high levels in Creutzfeldt-Jacob disease (116). CSF P-tau concentrations, however, do not increase after stroke (115, 117). CSF P-tau levels increase in AD dementia (95-97), and reflect tangle pathology, supported by a correlation between CSF P-tau levels and neocortical NFT counts (118). A recent study has found high levels of CSF P-tau in preclinical AD, despite normal tau PET scans (119), suggesting that CSF P-tau levels increase before tau aggregates are identified on PET scans. Nonetheless, abnormal CSF P-tau is also common in non-AD dementias (107, 108).

![Figure 1](image-url)  
**Figure 1.** CSF core Alzheimer's disease biomarkers and novel biomarkers reflecting Alzheimer's disease-related neuropathological processes. Copyright 2019 by Henrik Zetterberg. Adapted with permission.

**Novel CSF biomarkers reflecting pathological processes beyond amyloid and tau**

Novel biomarkers explored in the second wave in the search for biofluid AD biomarkers are sought to give insight to different pathological processes associated with AD (120), such as neuroinflammation. Novel biomarkers are also needed to complement the core AD biomarkers, e.g. it is necessary to find better biomarkers of AD progression (92, 120, 121). A wide range of
biomarker candidates have been evaluated (91, 95, 122), including three of the biomarkers in the articles of this thesis, namely NFL, YKL-40, and FABP3. Some other promising CSF biomarker candidates are synaptic proteins like neurogranin, synaptotagmin-1, and synaptosomal-associated protein 25 (96, 123), and the microglial activation biomarker sTREM2 (123). In the next section I will elaborate on the three biomarkers included in the papers of this thesis.

NFL. Neurofilaments are cytoskeletal components of neuronal axons and they are composed of four subunits, known as NFL (neurofilament light), NFM (neurofilament middle), NFH (neurofilament heavy), and α-internexin or peripherin (124). Neurofilaments are important for axonal radial growth and nerve conduction velocities, and may also have a role in synaptic function (124, 125). NFL is expressed in neurons both in the central and the peripheral nervous system (126), and damage of CNS neurons will release NFL into the extracellular compartment resulting in increased CSF NFL levels (124, 125). Consequently, CSF NFL is regarded a neurodegeneration biomarker reflecting axonal injury, axonal loss, and neuronal death (124, 125). In accordance with its role as a neurodegeneration biomarker, NFL has rather consistently been shown to correlate positively with CSF T-tau and P-tau (112, 113, 127-129), whereas associations between NFL and CSF Aβ42 have been weak negative or non-existent (112, 113, 128, 129). CSF NFL levels are shown to increase acutely in stroke (130, 131) and traumatic brain injury (114). CSF NFL concentrations are also elevated in different neurodegenerative disorders including AD (95, 132-137), with highest levels in FTD and amyotrophic lateral sclerosis (135). Thus, NFL is not a specific biomarker for the axonal injury occurring in AD (136). CSF NFL has been shown to predict faster brain atrophy, including hippocampal atrophy, and faster cognitive decline in MCI patients (129), and to predict faster progression in AD and other neurodegenerative disorders (135, 138). Hence, NFL is suggested to be a general biomarker for disease intensity and progression (133, 139). NFL may also be used as a biomarker of treatment response (133, 140, 141), or to rule out neurodegeneration (123). Studies also suggest that NFL is a preclinical marker of neurodegenerative disorders (142, 143), with levels increasing parallel to protein deposition in the brain (143).

YKL-40. YKL-40 is a glycoprotein mainly expressed by astrocytes in the human brain (144-147). The functions of YKL-40 remain unclear, however, astrocytes expressing YKL-40 are found in proximity to activated microglia (144, 148) and astrocytic YKL-40 transcription may
be induced by cytokines (interleukin-1β and tumor necrosis factor-α) released by macrophages/microglia (148), suggesting that YKL-40 is involved in neuroinflammation. CSF YKL-40 is, therefore, considered a biomarker of neuroinflammation (149, 150). Expression of YKL-40 and increased CSF YKL-40 levels are found in both acute and chronic neurological diseases (144, 145, 149, 150), so the biomarker is not specific to AD. In AD, astrocyte YKL-40 expression has been shown in the proximity to amyloid plaques (145, 147), confirming the protein’s involvement in the neuroinflammatory response to amyloid pathology. Furthermore, studies have shown significant correlations between CSF YKL-40 and the core CSF AD biomarkers (147, 151-155), most consistently positive correlations with T-tau and P-tau (150), suggesting that CSF YKL-40 levels are related to neurofibrillary neurodegeneration. Accordingly, CSF YKL-40 levels are higher in patients with AD dementia compared to controls (95, 147, 156), and concentrations have been shown to increase over time in patients with MCI and AD dementia (157, 158). CSF YKL-40 levels may be increased already in the preclinical phase of AD (159, 160), and are also suggested to predict AD progression (157, 160).

**FABP3.** Fatty acid-binding protein 3, also called heart-type FABP, is an intracellular fatty acid transport protein expressed in neurons and a wide range of other tissues, including the myocardium (161-164). In the human brain, FABP3 is expressed in many different areas with highest expression in the pons (164). Because FABP3 is a cytosolic protein, damage to a cell will cause release into the extracellular environment, thus CSF FABP3 is considered a biomarker of neuronal damage (165, 166). CSF FABP3 levels have consistently been shown to be elevated in AD dementia (95, 167, 168), and in MCI patients who progress to AD dementia (169-171). Significant correlations between CSF FABP3 and the core CSF AD biomarkers have been most consistently reported for T-tau and P-tau (170-174), and CSF FABP3 is associated with longitudinal atrophy of AD-vulnerable neuroanatomic regions (175), suggesting a role for CSF FABP3 in neurofibrillary degeneration. However, FABP3 is not a specific biomarker for AD. FABP3 concentrations have been shown to increase in blood within hours after acute stroke (176, 177), within days in CSF after subarachnoid hemorrhage (178), and to be elevated in CSF in various neurodegenerative diseases (168, 172, 174, 179, 180), being especially high in Creutzfeldt–Jakob disease in which neuronal degeneration is very rapid (165, 168). Furthermore, CSF FABP3 levels are shown to be elevated in preclinical AD (181, 182), and to predict progression to different dementia subtypes (169, 171, 183, 184).
Blood biomarkers

The third wave in the search for biofluid AD biomarkers explored blood based biomarkers. Blood is a more accessible biofluid than CSF, and a blood biomarker would therefore be preferable for use in tests in primary care, and for repeated sampling in longitudinal assessments. Blood biomarkers for AD have, however, been difficult to develop for several reasons, e.g. very small amounts of a biomarker crosses the blood-brain barrier and dilution in blood results in an even lower concentration, and expression of the biomarker in peripheral tissues can make it hard to detect the contribution from CNS (93, 185). Despite this, some promising blood-based biomarkers have been identified (93), e.g. NFL which correlates well with CSF NFL (133, 186-188). The situation has been less clear for Aβ proteins and tau biomarkers (93, 189). However, ultrasensitive assays has advanced the field during the past few years, and recent studies have shown promising results also for amyloid and tau biomarkers in blood (190, 191). It is likely that new techniques like ultrasensitive measurement techniques, neuron-enriched exosome preparations, and microRNA will offer new opportunities for blood based biomarkers in the future (92, 93, 192, 193).

Imaging biomarkers

The most commonly used imaging biomarkers for AD include structural brain imaging and molecular imaging.

Structural (i.e. anatomical) imaging is most commonly obtained using MRI. MTL structures are important for episodic memory (194, 195), and atrophy of MTL structures like the hippocampi, entorhinal cortex, parahippocampal gyrus, and amygdala is classical in AD dementia (Figure 2) (196). For clinical purposes MTL atrophy is often graded by visual assessment (197). Further, hippocampal atrophy is considered one of the major AD biomarkers (198). Several studies have shown that hippocampal atrophy assessed by MRI is associated with post-mortem neurodegenerative pathology including Braak staging (199-202), with a clinical diagnosis of AD dementia (203-205), and with severity of cognitive symptoms and episodic memory deficits in MCI and AD dementia (206-211). Reported hippocampal atrophy rates vary across reports, depending on the samples used and how atrophy is measured, but has been reported to be as high as 4.66 % annually in patients with AD dementia compared to 1.41 % in controls (212, 213). Hippocampal atrophy can also predict progression from MCI to AD dementia (214, 215). However, hippocampal atrophy is not a specific biomarker for AD and is
also present in other forms of dementia (216-218), e.g. cerebral age-related TAR DNA-binding protein 43 (TDP-43) and sclerosis which is often misdiagnosed as AD dementia because of the presence of progressive memory impairment (219). In addition to hippocampal atrophy, the atrophy of different subfields of the hippocampus is an evolving biomarker, and especially the CA1 field in the head of the hippocampus is a strong imaging biomarker candidate for AD (43). It is also possible to measure atrophy across the entire cortex. Studies using these methods have shown that the cortical atrophy pattern follows the Braak staging of NFT pathology, and an “AD signature” consisting of a composite of thickness estimates derived from regions impacted in AD dementia is an often used biomarker (43, 46). Other studied atrophy biomarkers for AD are entorhinal cortex atrophy, and atrophy of various subcortical structures like amygdala (43).

Figure 2. Hippocampus atrophy in Alzheimer’s disease (A) and healthy control (B). From “Multimodal imaging in Alzheimer’s disease: validity and usefulness for early detection” by Teipel et al., 2015, The Lancet Neurology, Volume 14, p.1037-1053. Copyright [2015] by Elsevier Ltd. Reprinted with permission.

The most well established molecular imaging AD biomarkers are $^{18}$F-fluorodeoxyglucose ($^{18}$F-FDG) PET and amyloid PET, and both techniques are clinically available (197). FDG PET assesses regional brain metabolism of glucose, and brain hypometabolism is a biomarker of synaptic dysfunction and neurodegeneration (197, 220). Amyloid PET is a biomarker of amyloid pathology using amyloid tracers binding to fibrillary amyloid, and the technique visualizes the regional cerebral $\beta$ deposition (197, 221, 222). In recent years, PET tracers for tau have been developed allowing visualization of tau pathology in the brain, and tau PET will probably have implications for the future (222). The role of other imaging techniques like functional MRI and diffusion tensor imaging in AD is also being investigated (197).
1.1.5 Biomarkers in criteria for Alzheimer’s disease

Historically, the first AD diagnostic criteria were published in 1984 by the Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association (223). The criteria were solely clinical, and a diagnosis of “probable AD” was made clinically in patients with dementia, whereas a diagnosis of “definite AD” required a neuropathological examination after death. Over the past decades, biomarkers have received an increasing role in new research diagnostic criteria for AD. In 2007, the International Working Group (IWG) for New Research Criteria for the Diagnosis of AD published the first research criteria integrating CSF Aβ42, T-tau and P-tau and other AD biomarkers in the diagnostic process (17). These criteria, and also later research criteria, have suggested using biomarkers to diagnose AD in patients with MCI or dementia (5, 15, 16). The term “prodromal AD” has been used to describe patients with MCI and positive core AD biomarkers (16).

In 2010, Jack et al. put forward a hypothetical model of the temporal evolution of the major AD biomarkers, suggesting that AD biomarkers become abnormal many years before cognitive symptoms arise. The model suggests that amyloid pathology biomarkers become abnormal first, CSF tau next, and last MRI and FDG-PET biomarkers (73, 224). In line with this model, the IWG group defined the first criteria for preclinical AD in 2010, incorporating the core AD biomarkers for diagnosing AD in cognitively normal adults (225). Later, somewhat different criteria have been suggested (22, 23, 226). Some criteria have defined preclinical AD based on only a positive amyloid biomarker, whereas other have required the combination of a positive amyloid biomarker and a positive biomarker of tauopathy or neuronal injury.

The Alzheimer’s continuum

The latest addition to AD research criteria was published by the National Institute on Aging—Alzheimer’s Association (NIA-AA) in 2018 (1). The research framework defines AD solely based on biomarkers reflecting core AD pathologies, whereas cognitive symptoms are used only to stage severity of the disease. The diagnosis is based on the amyloid/tau/neurodegeneration (A/T/N) classification scheme for AD biomarkers (227), where A includes amyloid PET, CSF Aβ42, and CSF Aβ42/Aβ40 ratio, T includes CSF P-tau and tau PET, and N includes structural MRI, FDG PET, and CSF T-tau. The ATN system allows for inclusion of other biomarkers to the three categories and for addition of other categories in the future. Three main biomarker categories are defined 1) normal AD biomarkers (A-T-N-,
negative biomarkers in all categories), 2) Alzheimer’s continuum where an amyloid biomarker is positive (A+), and 3) non-AD pathologic change with a negative amyloid biomarker (A-) but positive tau (T+) and/or neurodegeneration biomarkers (N+). The Alzheimer’s continuum can be further divided into three categories: 1) Alzheimer’s pathologic change (A+T-N-), 2) Alzheimer’s disease (A+T+N- or A+T+N+), and 3) Alzheimer’s and concomitant suspected non Alzheimer’s pathologic change (A+T-N+). Further, based on cognitive symptoms, research participants can be divided into the stages “cognitively unimpaired”, “mild cognitive impairment”, and “dementia”. For research participants within the Alzheimer’s continuum, a numeric staging system of cognitive symptoms ranging from 1-6 is suggested. It is important to note that these criteria represent a research framework and are not intended for use in clinical practice. In fact, the use of core AD biomarkers to diagnose AD in individuals without dementia is still controversial, and Cochrane reports conclude that the predictive value of individual biomarkers is low and do not recommend using the biomarkers for this purpose (24-28).

The term AD has traditionally been used interchangeably to describe the typical multidomain amnestic dementia syndrome and AD neuropathological changes. However, a major implication of the new NIA-AA research framework is that the definition of AD has been separated from the clinical syndrome to become an entirely biological construct, emphasizing that AD is characterized by neuropathological change of amyloid plaques and NFTs. As stated in the framework paper itself, up to 60% of cognitively normal individuals over the age of 80 years have AD neuropathological changes, but many of them will never develop cognitive symptoms (1). Therefore, it is a matter of debate whether it is meaningful to talk about AD without clinical symptoms. This debate emphasizes the need of more studies of the implications of core AD pathologies and other neuropathologies in individuals without dementia.

In this thesis I have used the term AD to denote Alzheimer’s disease, the neurodegenerative disorder characterized by amyloid plaques and NFTs in the brain, whereas I have referred to the clinical dementia syndrome (with or without neuropathological or biomarker verification) as AD dementia. Still, as the main basis for the term AD is studies of patients with an AD dementia, the terms AD and AD dementia are not always easily distinguishable. In such cases I have only written AD. I have used the terms “preclinical AD” and “prodromal AD” to describe, respectively, cognitively normal individuals and MCI patients with biomarkers signs of AD neuropathology. I have also used the A/T/N-system to describe biomarker abnormalities in individuals without dementia.
1.2 Cognitively normal older adults

1.2.1 What is normal cognition?

Cognition is a higher cortical function that includes the mental process of acquiring knowledge, understanding and perceiving through experiences, thought, and senses. Cognitive ability can be divided into different cognitive domains, including memory, language, processing speed, attention, visuospatial abilities, and executive functions (228).

Cognitively unimpaired is in the NIAA-AA research framework defined as “cognitive performance within the expected range for that individual based on all available information” (1). It is further stated that this information can include clinical judgement and/or cognitive test results, and the test results may be based on comparison to normative data. Normative data are often adjusted for age and education, and may also be adjusted for other factors like sex and occupation. Subjective cognitive decline and/or subtle decline on serial testing is considered acceptable. The definition of cognitively normal is unavoidably also dependent of the definition of MCI. Cognitive impairment in MCI is defined by a cognitive performance below the range expected for that individual (1). Objective cognitive impairment in MCI is often defined as cognitive performance more than 1.5 standard deviations (SD) below age- and education adjusted normative data (15, 19, 229), although there is no consensus, and the prevalence of MCI among community-dwelling adults without dementia has been shown to range from 4 % to 70 % depending on which criteria that are used (230). Further, there is no consensus on which or how many cognitive tests that should be used neither to classify cognitively normal nor MCI (230-234). Accordingly, the definition of cognitively normal/unimpaired varies between studies (235). Ideally, a consensus diagnosis based on all available information including an interview of the participant and an informant and serial cognitive assessment should be used to determine if a person is cognitively normal (232).

Cognitively normal adults are shown to exhibit some of the same features as patients with AD dementia, e.g. episodic memory decline (236), amyloid pathology (13), and hippocampal atrophy (213), although to a smaller extent. Details about these features and the relationships between them in cognitively normal adults will be elaborated on in the next sections.
1.2.2 Cognitive changes in cognitively normal adults

Memory, processing speed, and executive function are cognitive domains known to decline with normal aging (228). Decline of episodic memory starts in the 60s, or maybe even in the 20s or 30s, and is shown to accelerate with increasing age (Figure 3) (237, 238). However, progressive episodic memory impairment is also a cardinal symptom of AD dementia (5, 17), and episodic memory impairment is more profound and shows higher rates of decline than in normal aging many years before a patient is diagnosed with AD dementia (236, 239-241). Consequently, studies of age-related changes, including cognition in normal aging, can be complicated if participants are misdiagnosed as cognitively normal (misclassification bias) (228, 242). Furthermore, studies of cognition in normal aging can be affected by selection bias, cohort bias, attrition bias, and practice effects (228, 242). Despite these limitations, the literature consistently shows that changes in cognition, including episodic memory, occur with normal aging (228, 237, 243, 244). Nevertheless, studies have shown substantial inter-individual differences in memory trajectories in aging, with subgroups of older individuals showing well-preserved memory, often referred to as “successful cognitive aging” (245-248).

Figure 3. Memory trajectories in normal cognitive aging. Means and standard errors of the cross-sectional and three-occasion longitudinal data and estimates of quasi-longitudinal relations in the memory domain. The quasi-longitudinal trajectories are portrayed as originating at the first longitudinal occasion, and extending over an interval equal to the average longitudinal interval. Quasi-longitudinal values are only reported for the Time 1 and Time 3 occasions to minimize clutter in the figures. From “Trajectories of normal cognitive aging” by Salthouse, 2019, Progress in Neurobiology, Volume 34, p.17-24. Copyright [2019] by American Psychological Association. Adapted and reprinted with permission.
1.2.3 **Brain atrophy in cognitively normal adults**

A large body of evidence show that the gross brain volume decreases with around 0.5% annually in cognitively normal older adults from the age of 60 (249-251). Significant atrophy has been shown across almost the entire cerebral cortex in cognitively normal individuals, but atrophy rates vary substantially across regions (252, 253). The frontal and temporal lobes appear especially susceptible to age-related atrophy (252-258). Most studies of cortical thickness and volume have reported a monotone, more or less linear trajectory of atrophy with age, although accelerated atrophy has also been shown in some regions, e.g. the entorhinal cortex (258-263). Hippocampal atrophy is also rather consistently shown to accelerate with higher age, but before the age of 60 years hippocampal volumes are relatively stable (Figure 4) (258, 264-269). Hippocampal atrophy is also a pathologic feature in patients with AD dementia, but the hippocampal atrophy in AD dementia has been shown to be of greater magnitude and accelerate faster than in cognitively normal older adults (213, 269). Furthermore, age-related atrophy of the hippocampus, and other brain regions, has also been shown to occur in individuals without amyloid pathology, and in other healthy groups with low risk of AD (260, 270), demonstrating that preclinical AD is not the only factor driving the hippocampal atrophy in aging.

![Image](image-url)

Different brain structure characteristics have been associated with cognitive performance in cognitively normal older adults. The most consistent findings are the relationship between frontal brain measures and executive function, and the relationship between MTL/hippocampal brain measures and memory (271, 272). Longitudinal studies support a relationship between hippocampal atrophy and episodic memory decline in cognitively normal older adults (273-278). Further, cognitively normal individuals with biomarker signs of amyloid pathology in combination with abnormal hippocampal atrophy show a steeper memory decline than those with abnormality of only one of these biomarkers (279-282). However, hippocampal atrophy is also associated with episodic memory decline in older adults with very low risk of AD, suggesting that this relationship is not specific for preclinical AD (270).

1.2.4 **Amyloid and tau pathology in cognitively normal adults**

A large proportion of cognitively normal individuals is shown to exhibit pathological levels of both amyloid and tau pathology (12, 283). A commonly used example is 101 year old Sister Mary from the “Nun Study” who performed very well on cognitive tests despite autopsy findings of amyloid plaques and NFTs that satisfied neuropathological criteria for AD (284). This is not unique (12, 20, 21, 285-288), and up to 40% of individuals who were cognitively normal on the last assessment before death reached neuropathological criteria for AD (12). Further, entorhinal cortex tau pathology (Braak stage I) is nearly ever-present in cognitively normal 80-year olds at autopsy (82, 288). These findings have also been confirmed in vivo, with up to 40% of cognitively normal individuals having positive amyloid PET (13, 289), and up to 18% being classified as Braak stage III/IV on tau PET (290, 291), depending on the age of the individuals. Other biomarker studies further support these findings, showing that the prevalence of cognitively normal individuals with normal AD biomarkers (A-T-N-) declines from the age of 50 years, while the prevalence of abnormal biomarkers increase with age (283, 292, 293). At the age of 70 years, around 30% of cognitively normal individuals are A+, around 25% are T+, and around 5% are A+T+N+, increasing further to around 20% A+T+N+ at 80 years (283, 292-295).

Notably, some cognitively normal older adults exhibit tau pathology, but have sparse or absent amyloid pathology. The neuropathological finding of NFTs in the absence of amyloid plaques called primary age-related tauopathy (PART) is common in older individuals (296), and it is suggested to be a separate pathologic entity. Biomarker studies suggest that at least 13% of
cognitively normal older adults show an A-T+ biomarker profile (283, 297), and tau deposition outside of the MTL on tau PET is shown to be common in amyloid negative cognitively normal individuals (298).

Longitudinal studies with repeated biomarker measures show that CSF Aβ42 decrease in cognitively normal adults (158, 208, 299, 300), however the decrease may be confined to amyloid negative individuals (158). Further, amyloid slowly accumulates on amyloid PET over time in both A+ and A- cognitively normal adults, with highest accumulation rates shown in A+ individuals (301-306). Moreover, CSF T-tau and P-tau levels increase over time in cognitively normal older adults (158, 208, 300), and the P-tau increase may be limited to A+ individuals (158). The few studies on longitudinal tau PET in cognitively normal older adults show tau accumulation rates up to 3 % per year that are possibly highest in amyloid positive individuals, but lower than in patients with AD dementia (307-310).

**Relationships between core AD biomarkers, brain atrophy, and cognition in cognitively normal adults**

As explained above, many of the same changes in brain and cognition are common to aging and AD dementia. Because aging itself is the major risk factor for sporadic AD dementia, and in order to understand why the aging brain is vulnerable to AD dementia, it is important to understand the role of brain changes in cognitively normal adults. We need a better understanding of the relationship of different neuropathological processes to brain and cognitive changes in cognitively normal adults. Importantly, a better understanding of the relationship between the core AD biomarkers and brain and cognitive changes in cognitively normal adults is essential to understand their role in aging in general, not confined to AD dementia. One central question is whether the core AD biomarkers inherently reflect pathological and detrimental processes in the brain, or whether some reflect processes that may be common for normal aging and AD dementia. Answering this question will have implications if cognitively normal adults with biomarker signs of AD pathology are to be included in clinical trials.

There are many studies on the interrelationships between core AD biomarkers, brain atrophy, and cognition in cognitively normal older adults, however the relationships are far from understood, and more research is needed. Following, some examples of relationships between these features will be given.
Core AD pathology and cognition. The most studied relationship is between amyloid pathology and cognition. Meta-analyses have shown poorer cognitive performance and cognitive decline in several cognitive domains including episodic memory in amyloid positive vs amyloid negative cognitively normal older adults (311-314). The differences are, however, small, and it is uncertain whether they are clinically significant. Neuropathological studies suggest a stronger relationship between the burden of neocortical NFTs and cognition than between amyloid plaques and cognitive performance (82), and this has been supported by some biomarker studies in cognitively normal individuals (310, 315-322). CSF P-tau and tau PET measures have been associated with cognitive impairment and decline in different cognitive domains, including memory, in cognitively normal individuals (290, 316, 317, 319-321, 323-326), although a large study did not find an association between P-tau and cognition (327). Several studies of cognitively normal individuals suggest that a combination of abnormal amyloid and tau biomarkers, or the synergy between them, may have a stronger relationship to cognition than each single biomarker (297, 315, 318, 319, 325, 328-331). Biomarkers of AD pathology have also been associated with progression from cognitively normal to MCI and AD dementia (282, 301, 310, 315, 330, 332-339).

Core AD pathology and brain atrophy. Many studies have examined the relationship between core AD biomarkers and brain atrophy in cognitively normal older adults. Both CSF Aβ42 and amyloid load on PET have been associated with atrophy in AD-vulnerable regions, including the hippocampus (302, 340-350). Results are, however, not consistent (320, 322, 340, 351-355), and biomarkers of amyloid pathology have also been associated with atrophy in regions not vulnerable in AD (341, 356). CSF T-tau and P-tau have also been associated with atrophy in hippocampus and other AD-vulnerable regions (322, 340, 357-360), whereas other studies have not found such relationships (343, 354, 356, 361). Tau PET results so far have indicated a relationship between MTL tau deposition and MTL atrophy (307, 320, 341, 362). Some reports suggest that only individuals with evidence of both amyloid and tau pathology, and not individuals with amyloid pathology only, show MTL atrophy (305, 363, 364), and tau pathology may mediate the effects of amyloid pathology on brain atrophy (310, 365). Further, the association between amyloid pathology on memory has been suggested to be mediated by brain atrophy (207, 352, 366, 367).

Relationships between amyloid and tau pathology. Cross-sectional studies of CSF Aβ42 and Tau proteins have suggested weak to moderate negative correlations (112, 363, 368), although
no correlation has also been shown (297, 315, 322). In longitudinal studies, an increase in CSF P-tau has been associated with a decrease in CSF Aβ42 and an increase in amyloid accumulation on PET (322, 360). The recent development of tau PET has made it possible to assess spatial relationships between amyloid and tau. Most studies have found positive associations between amyloid and tau deposition (325, 341, 369-373), and studies indicate that more amyloid deposition regardless of its localization is associated with greater tau deposition in temporal brain regions (341, 370-372), notably in the entorhinal cortex (372). However, a longitudinal study has shown higher tau accumulation rates in amyloid positive than negative individuals in widespread cortical areas including neocortical regions (308), and this finding has been supported by a cross-sectional study (374). A recent longitudinal PET study showed that amyloid accumulation was associated with subsequent tau disposition and accumulation, but tau accumulation was not associated with later amyloid deposition, hence supporting the amyloid cascade hypothesis (310).

1.2.5 Other neuropathological changes in cognitively normal adults

Neuropathologies other than amyloid and tau pathology are also common with advancing age, and the brains of older persons often contain a mixture of pathologies (375). Autopsy studies show that patients with dementia in old age have multiple pathologies (376-379), and research suggest that the neuropathological substrates of cognitive impairment in older individuals are likely multifactorial (380, 381). Cognitively normal adults can have autopsy findings of hippocampal sclerosis (382), TDP-43 proteinopathy (383), α-synucleinopathy (287, 288, 384), argyrophilic grain disease (288), and cerebrovascular pathologies (287, 288, 384). These pathologies may represent preclinical pathology of one or more non-AD neurodegenerative disorders, e.g. FTD, DLB, and PD, and may contribute to neurodegeneration in both A+ and A- cognitively normal adults. Aging is associated with neuroinflammation (385, 386), and activation of microglia with age is also evident in cognitively normal adults (387-389). Evidence suggest that neuroinflammation contributes to neurodegeneration (55). Many of the same neuropathological changes are evident in aging and dementia, but the role of different neuropathologies alone and in combination in cognitively normal adults is not established. Therefore, we need a better understanding of the relationships between various neuropathologies, including neuroinflammation and neurodegeneration, and of the
consequences of different neuropathologies independently and in combination, in cognitively normal adults. Biomarker studies can provide important insights into this matter.

**Suspected Non-AD Pathophysiology**

Biomarker studies show that about 15% of cognitively normal 65 year olds are N+, increasing to around 50% at the age of 80 (283, 292, 293). Roughly 25% of cognitively normal adults aged > 65 years show a biomarker profile with abnormal neurodegeneration-biomarkers, but normal Aβ-biomarkers (A-N+) (226, 297), and this biomarker profile is sometimes called Suspected Non-Alzheimer Pathophysiology (SNAP) (226, 390, 391). It is shown that individuals with SNAP are usually older than A-N- individuals (391, 392) and that pathologies underlying SNAP are heterogeneous (391). Some studies have found similar MTL atrophy rates in SNAP and A-T- (282, 305), whereas others have found greater brain atrophy rates in SNAP (393). Likewise, some studies have shown cognitive decline in cognitively normal adults with SNAP compared to A-N- (279, 294, 394, 395), whereas others have not (282, 297). Further, some cognitively normal individuals with SNAP are shown to become amyloid positive (i.e. convert from A-N+ to A+N+) later (282, 305, 355, 390), and some progress to dementia (330, 390, 396), even to AD dementia (330), highlighting the heterogeneity of the SNAP group.

**YKL-40, NFL and FABP3 in cognitively normal adults**

The role of CSF NFL, YKL-40, and FABP3 in AD and other neurodegenerative diseases is still being explored (120). When papers I and II in this thesis were planned in 2015, most studies of these three novel CSF biomarkers in cognitively normal older adults were in the form of control groups for patient groups with neurological diseases. Still, not a lot studies assessing the role of these biomarkers specifically in cognitively normal older adults exist, but concentrations of all three biomarkers are shown to increase with increasing age in cognitively normal adults (397-399). CSF NFL, YKL-40, and FABP3 are shown to have non-existent or weak relationships to CSF Aβ42 (112, 142, 154, 171, 397, 400), whereas moderate to strong positive correlations have been found between the three novel biomarkers and CSF T-tau and P-tau (112, 142, 154, 171, 397, 401). Some studies have found that concentrations of the three biomarkers are higher in A+ than A- cognitively normal adults (112, 182). Before paper I of this thesis was published, no studies had examined the relationship between any of the three novel biomarkers and MTL atrophy in cognitively normal adults, and still no studies on the
relationship of YKL-40 and FABP3 to MTL atrophy exist. Further, only a few studies have examined the relationship of the three novel biomarkers to cognitive decline in cognitively normal adults, and in some of these studies high levels of CSF NFL, YKL-40, and FABP3 have been associated with cognitive decline or development of cognitive impairment (112, 142, 181, 402, 403). More results from studies of these three biomarkers in cognitively normal adults will be discussed in section 5.1.2.

1.2.6 Resilience against neuropathology

As evident from the above sections, there are inter-individual variations in the degree of brain changes and in cognitive trajectories among cognitively normal older adults. “Reserve” is a concept trying to explain the difference between observed cognitive performance and the cognitive performance expected in an individual with a given degree of age- or disease-related brain changes (404). The proposed nomenclature defines two reserve concepts, namely “cognitive reserve” and “brain reserve”. Cognitive reserve is an active model of reserve, where individual differences in the ability to adapt new, or compensatory, cognitive processes determine how well a person can cope with brain changes. These processes can be innate, but can also be influenced by lifetime exposures like education, physical exercise, social engagement, and occupation. Cognitive reserve is a theoretical construct, but it has been measured using different proxies including sociobehavioral proxies (e.g. education, occupational complexity, and IQ), and by functional imaging looking at activation of brain networks (404, 405). As an example, some studies suggest that cognitively normal individuals with higher measures of cognitive reserve (e.g. IQ) can tolerate more AD pathology before they develop cognitive impairment (406-408). Brain reserve is a passive model of reserve, and it includes all anatomical and structural aspects of the brain, except neuropathology. Brain reserve comprehends the number of neurons and synapses, and it has been measured both in vivo and postmortem using proxies like whole-brain measures, intracranial volume, gray matter volume and cortical thickness. The concept is that individuals with a greater brain reserve have more neurons and synapses to loose before cognitive symptoms emerge (404).
1.3 Delirium

1.3.1 Introduction

Delirium is characterized by an acute disturbance in attention, awareness, and cognition (e.g. memory, orientation, and perception) that represents a change from the person’s habitual status, and the symptoms usually fluctuate during the day. The syndrome is by definition caused by a medical condition, substance withdrawal/intoxication, or multiple etiologies, and the symptoms should not be better explained by a neurocognitive disorder, e.g. dementia (4). Delirium is common in hospitalized patients, and prevalence and incidence rates vary depending on the population studied, with highest delirium rates found in older patients and in intensive care and postoperative settings (409, 410). About 50 % of patients in intensive care units (411), up to 50 % of patients with hip fracture (412, 413), and 20 % of patients in general medical wards (410, 414, 415) experience delirium. Risk factors predisposing a patient to delirium have been identified, and delirium can be triggered by a wide range of medical conditions (more details in section 1.3.3). The pathophysiology of delirium is still not understood, which limits our ability to design therapeutic interventions, and no established pharmacological prevention or treatment measures for delirium exists (416-419). Multicomponent non-pharmacological interventions are still the mainstay for both prevention and treatment of delirium (420, 421), although they are mainly documented to be effective for prevention of delirium (419, 422, 423). Multiple approaches are recommended, including orientation, adequate hydration and nutrition, early mobilization, pain management, hearing and vision optimization, and sleep enhancement (419). Delirium is associated with many adverse outcomes, including longer hospital stay, increased risk of institutionalization, mortality, and future cognitive decline and dementia (414, 424, 425).

1.3.2 Delirium diagnosis

Delirium is a clinical diagnosis, and no diagnostic test or biomarker for delirium exist. The most commonly used diagnostic criteria in delirium research are the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) (4), and International Classification of Diseases, 10th edition (ICD-10) (426). Both criteria comprise acute and fluctuating change in awareness, attention, and cognition triggered by disease, but the ICD-10 criteria also includes the features psychomotor disturbance and sleep disturbance.
Studies suggest that only 20-30% of patients with delirium are recognized (410, 427, 428). Many tools have been developed for recognition and diagnosing delirium (419, 429), yet there is no consensus regarding which tests should be used to assess the cognitive domains impaired in delirium (430). The Confusion Assessment Method (CAM) (431) is a widely used and good delirium instrument (429, 432, 433). A more newly developed delirium screening tool is the 4AT (434), which now widely used internationally, and its sensitivity was recently shown to be better than the CAM (435). In any case, it is recommended that a positive screening test for delirium should be followed by a definitive diagnosis by a trained and experienced clinician who has examined and tested specific diagnostic features of delirium, such as attention, memory, and orientation (419). Many blood biomarkers have been explored in the search for an objective diagnostic tool for delirium, yet none have been found to be clinically useful (436, 437). Electroencephalography studies have shown promising results the past years, but awaits further evaluation. (438, 439).

Differentiation of delirium and chronic cognitive impairment/dementia can be difficult, especially when delirium occurs in patients with dementia, also referred to as delirium superimposed on dementia (430, 440, 441). The main differences in diagnostic criteria are that the cognitive impairment is acute in delirium as opposed to progressive and chronic in dementia, and that delirium is triggered by a medical condition (4). Information from relatives or caregivers in order to determine prior cognitive function and whether the cognitive function is acutely changed, accompanied by a thorough medical workup for common causes of delirium, are important approaches to help disentangle the conditions. Notably, DLB can be particularly challenging to discriminate from delirium, as symptoms in DLB fluctuate substantially and are similar to delirium, e.g. visual hallucinations and attentional deficits. Further, attention deficit is a core delirium symptom, and the ability to focus and shift attention is fairly preserved until moderate to severe dementia stages. Arousal is also affected in delirium, but is maintained in advanced dementia stages (430). Furthermore, delirium usually, but not always (442), resolves, and sometimes a definitive delirium diagnosis can only be made when the patient’s cognitive function is seen to improve.

1.3.3 Risk factors for delirium

The risk factors of delirium are traditionally divided into predisposing factors and precipitating factors. An explanation model of delirium suggests that the development of delirium depends
on the inter-relationship between the patient’s vulnerability (predisposing factors) and the exposure to noxious insults (precipitating factors) (443). A highly vulnerable patient can develop delirium after only a minor insult, whereas a less vulnerable patient requires a more noxious insult to develop delirium. It is common that a patient has both several predisposing factors and multiple factors triggering delirium. Some of the most consistently reported predisposing factors are dementia/chronic cognitive impairment, older age, comorbidity, and functional impairment (409, 412, 444-446). A wide range of medical conditions can trigger delirium, however, common precipitating factors are infections, illness severity scored using Acute Physiology and Chronic Health Evaluation II (APACHE II) (447), surgery, trauma, and drugs (particularly anticholinergic, psychoactive, and sedative drugs) (409, 412, 444, 446). Biomarker predictors for delirium have also been assessed (448, 449), and some inflammatory biomarkers may have potential (448). However, there are few studies and most have assessed only the association between the biomarker and delirium occurrence, thus more research is needed.

1.3.4 Pathophysiology

The pathophysiology of delirium is probably very complex, and as delirium has a multifactorial etiology, it is unlikely that one common pathophysiological pathway exists. Many pathophysiological hypotheses for delirium have been proposed through the years, such as the neuroinflammatory hypothesis (450, 451), neuroendocrine hypothesis (451, 452), the oxidative stress hypothesis (451, 453), the neurotransmitter hypothesis (451, 454, 455), and the network disconnectivity hypothesis (456), each supported by various degrees of scientific evidence. Infections are common triggers of delirium, and several biomarker studies, and also a post-mortem study, have supported a role of inflammation in delirium pathophysiology (448, 449, 457). The neuroinflammatory hypothesis suggests that a peripheral inflammatory process can signal across the blood-brain barrier, and induce a neuroinflammatory state in the brain, which ultimately causes delirium. An aberrant response to acute or chronic stress, mediated by high glucocorticoid levels, may cause delirium (452, 458, 459). This is suggested by the neuroendocrine hypothesis, and the hypothesis has received some support from biomarker studies that have found an association between elevated cortisol in blood and CSF and delirium (448, 460). Oxidative stress is known to cause damage to cells, proteins, and DNA, and the oxidative stress hypothesis proposes that inadequate oxidative metabolism in the brain leads to neurotransmitter abnormalities that results in delirium symptoms (453, 459). Because
neurotransmitters are the ultimate signal transmitters in the brain, they are likely to be involved in delirium. The neurotransmitter hypothesis suggest that neurotransmitter changes (excess or deficiencies) of certain neurotransmitters can cause the symptoms of delirium. The changes are thought to include excess of dopamine, noradrenaline, and glutamate, and deficiency of acetylcholine (451, 459). The expression of behavioral changes in delirium may depend on the relative neurotransmitter alterations, e.g. acetylcholine has a key role in attention (455). Last, the network disconnectivity hypothesis points out that the brain is organized into networks of interacting brain regions. The hypothesis poses that brain network disturbances and changes in connectivity patterns between brain networks results in the symptoms of delirium (456, 461, 462).

Maldonado has written several papers on delirium pathophysiology (451, 459, 463), and his latest paper proposes a new delirium hypothesis that integrates all of the above mentioned hypothesis, namely the system integration failure hypothesis (459). This hypothesis point to areas where the different pathophysiological hypotheses intersect, e.g. that brain inflammation may induce cholinergic hypoactivity. It further suggests that a combination of alterations in neurotransmitters and network disconnectivity is the ultimate disturbance that gives rise to behavioral and cognitive symptoms in delirium. The clinical delirium phenotype depends on the variable contribution of different pathophysiological mechanisms, and this interplay is also affected by the patient’s baseline physiological characteristics.

1.3.5 The interrelationship between delirium and dementia

Epidemiological interrelationship

Many studies report that delirium affects more than 50 % of hospitalized patients with dementia (464), and pre-existing cognitive impairment or dementia is found to increase the risk of delirium up to five times (465). A longitudinal population-based study, Vantaa 85+, showed that for every point lost on the global cognitive test the Mini Mental Status Examination (MMSE), delirium risk increased 5 % (466). The relationship between poorer preoperative performance on cognitive tests and higher risk of postoperative delirium has also been shown in populations without dementia (467-469). Conversely, there is also consistent evidence that delirium increases the risk of dementia and long-term cognitive decline, and a review report odds ratios ranging from 6 to 41 (465). In a study of 3 large population-based cohort studies,
subjects who experienced delirium declined 0.37 points more annually on MMSE compared to those who did not (425). Accordingly, delirium has been associated with increased rates of cognitive decline in both patients with and without dementia (465, 470-472).

**Pathophysiological interrelationship**

The nature of the interrelationship between delirium and dementia remains unclear, and many possible links have been proposed (465). Delirium may only be a marker of the brain’s vulnerability to dementia, by unmasking underlying preclinical or unrecognized dementia-related neuropathology. In line with this, the system integration failure hypothesis suggests that if the brain is already neurophysiologically deranged, the precipitant physiological insult may be less severe before the overall system is disrupted, thereby resulting in symptoms of delirium (459). Further, it is possible that the precipitating factor of delirium, e.g. inflammation, rather than delirium itself, causes neuronal death and hence dementia. Moreover, delirium may also independently, or through interactions with dementia-related neuropathological processes, be involved in pathological processes causing permanent neuronal loss (465).

Studies of the relationship between delirium and dementia-related neurobiological processes including investigations of fluid and neuroimaging biomarkers, post-mortem brains, and animal models could improve our understanding of delirium pathophysiology and the biological mechanisms underlying the interrelationship between delirium and dementia. In the past few years, an increasing number of such studies have been conducted, although the literature on this topic is still scarce. Several neuroimaging studies have been performed in delirium, but a recent review points out that study methods and populations are heterogeneous. Nevertheless, results indicate that brain atrophy, a higher burden of white matter lesions, and loss of white matter integrity may be associated with delirium (473). Further, a few studies have assessed neuronal injury biomarkers in delirium. Some studies have found that neuronal injury biomarkers known to be increased in dementia are also increased in delirium (474), examples are S100β (448, 475, 476) and NFL (477, 478), suggesting chronic neurodegeneration and/or acute neuronal injury in patients with delirium. A few studies have also assessed the relationship between AD biomarkers and delirium, one of these is paper III in this thesis. When paper III was written, only two studies on the relationship between the core AD CSF biomarkers and delirium were published, and they showed conflicting results (479, 480). Further discussion of the relationship between delirium and AD pathologies will therefore be given in section 5.1.3.
A large post-mortem study including almost 1000 brains assessed the association between a history of delirium and the burden of NFTs, amyloid plaques, vascular lesions, and Lewy bodies (425). None of the pathologies were significantly different in individuals with and without a history of delirium. However, the total burden of dementia-related pathologies were found to account for -0.39 points decline in MMSE score annually, similar to the cognitive decline attributable to delirium alone (-0.37 points annually). In addition, the combination of dementia pathology and delirium accounted for additional -0.16 points annually, suggesting that delirium in the presence of dementia-related neuropathologies may involve detrimental pathophysiological processes.

Some rodent models have studied mice with underlying vulnerability, e.g. older age and neurodegeneration, in combination with minor insults intended to trigger delirium-like behavior. One interesting model with neurodegenerative pathology (ME7 prion disease) has been used to test the effect of different inflammatory insults on delirium-like behavior. Studies have shown that inflammation causes acute behavioral changes in mice with underlying prion disease pathology, but not in normal animals, neither in animals with prion disease pathology not exposed to an inflammatory insult (466, 481-483), suggesting that delirium may unmask an underlying neurodegenerative process. Systemic inflammation has in animals with neurodegenerative pathology also been shown to acutely induce CNS inflammation (481-484) and neurodegeneration (484), and accelerate clinical progression of the neurodegenerative disease (482).


2 Aims of the thesis

The overall aim of this thesis was to increase our knowledge about the role of different neuropathological processes in cognitively normal older adults and delirium using AD-associated CSF biomarkers reflecting core AD pathology and other neuropathological processes.

We therefore wanted to 1) examine relationships between different biomarkers reflecting various processes in the brain (papers I and II), 2) test whether subgroups of older adults with similar biomarker profiles could be detected using clustering analyses (paper II), and 3) examine the relationship of biomarkers and also biomarker-based subgroups to brain and cognitive changes known to be present both in aging and AD dementia (i.e. hippocampal atrophy and memory decline) (papers I and II) in cognitively normal older adults. We also wanted to examine whether the core AD biomarkers are related to delirium in patients with and without dementia (paper III).
3 Materials and Methods

3.1 Cohorts and study designs

The articles in this thesis are based on two different cohorts. Papers I and II are based on a cohort of patients who underwent elective surgery in spinal anesthesia. I was in charge of this project from start of inclusion and through 3 year follow-up. Paper III is based on a cohort of patients who underwent hip fracture surgery in spinal anesthesia, recruited by my supervisor, Leiv Otto Watne.

3.1.1 Elective surgery cohort

We recruited patients scheduled for elective surgery in spinal anesthesia turning 65 years or older the year of inclusion from February 2012 to June 2013. The patients underwent gynecological (genital prolapse) or urological (benign prostate hyperplasia, prostate cancer, or bladder tumor/cancer) surgery at Oslo University Hospital, or orthopedic surgery (knee or hip replacement) at Diakonhjemmet Hospital in Oslo. Exclusion criteria were dementia, PD, previous stroke with sequela, or other CNS disorders likely to affect cognition. We performed a clinical assessment including cognitive testing of the participants prior to surgery. Blood samples and CSF samples were obtained by the anesthesia team in conjunction with spinal anesthesia. Brain MRIs were taken during the first months after surgery (Figure 5).

The cohort was originally planned to result in a reference sample for CSF biomarkers and brain MRI measures in older adults. This reference sample should include only those individuals from the original cohort who based upon cognitive testing at the time of recruitment and through a five-year annual follow-up were free from dementia, MCI, and other brain disorders. At 2-year follow-up, collaboration with the Center for Lifespan Changes in Brain and Cognition at the Department of Psychology at the University of Oslo was established. This collaboration received grants for a project named “New biomarkers for early detection of Alzheimer’s disease”. A sub-aim of this study was to test new CSF biomarkers for cerebral and cognitive changes in cognitively normal older adults, and papers I and II in this thesis are a result of this collaboration. The collaboration has resulted in even more measures of biomarkers, e.g. brain MRI at three more time-points, and in extended time of follow-up (7-year follow-up started in September 2019) (Figure 5).
At baseline, participants were excluded from follow up if they had poor test results, if we did not obtain a CSF sample or a brain MRI at baseline, or if they were unpractical to reach (lived/moved too far from Oslo). During follow up, participants who have developed dementia have been excluded from further follow up. Others have discontinued follow up at different time points due to various reasons, mainly because they did not want to attend further follow up assessments, were unpractical to reach (lived/moved too far from Oslo), or died. Some participants have skipped a follow-up appointment, e.g. because they did not have time, but have later continued follow-ups.

3.1.2 Hip fracture cohort

The cohort of patients with hip fracture were recruited from Oslo Orthogeriatric Trial: a randomized, controlled, single-blind trial comparing orthogeriatric care in the acute geriatric ward with usual orthopedic care for hip fracture patients (413, 485). All patients acutely admitted to the Ullevaal Clinic of Oslo University Hospital with a hip fracture from September 2009 to January 2012 were assessed for eligibility. Patients were excluded if they were moribund or if the hip-fracture was caused by a high-energy trauma. In addition to assessments during the index stay, patients underwent assessment four and twelve months after surgery, as the primary outcome of the study was cognitive performance four months after surgery. Delirium incidence during the index hospital stay was a secondary outcome, and there was no difference in delirium rates between intervention and control group.

A predefined secondary aim of the study was to collect CSF samples in order to study pathophysiological mechanisms in delirium. Therefore, CSF was collected in conjunction with spinal anesthesia in as many participants as possible.
Figure 5. Assessments and follow-up of the elective surgery cohort.

Figure 6. Assessments and follow-up of the hip fracture cohort.
3.2 Clinical assessments

In both cohorts, a wide range of data has been collected. I will only describe the assessments used in the articles included in this thesis.

3.2.1 Elective surgery cohort

A clinical assessment including cognitive tests was performed at baseline and annually thereafter. Demographic information was recorded at baseline. The participants’ medical history and medication use have been registered at all assessments. Baseline assessments were performed by me when I was a medical student or by a medical doctor. Follow-up assessments have been performed by medical students. All have been trained to perform the test battery.

Cognitive assessment

Participants were assessed with a multi-domain battery of cognitive tests before surgery, comprising MMSE (486), Clock Drawing Test (487), Word List Memory Task from Consortium to Establish a Registry for Alzheimer’s Disease (CERAD) (488), Trail Making Test (TMT) A and B (489), Kendrick Object Learning Test (490), and verbal fluency (FAS test and Animal Naming) (491). The same test battery has been administered annually since baseline, although Kendrick Object Learning Test has been excluded from the test battery at 3, 6 and 7 year follow-ups.

MMSE is a cognitive instrument used to assess global cognition. The score ranges from 0 to 30; a higher score represents better cognitive functioning, and it covers several cognitive domains including orientation, language, attention, memory, and visuospatial abilities (486). Clock Drawing Test is a cognitive screening instrument, and it is sensitive to visuospatial and executive function (487), and we used a scoring system ranging from 0 (worst score) to 5 (best score). The Word List Memory Task from CERAD assesses verbal episodic memory (488). First, immediate recall (i.e. Word List Learning) is assessed by presenting 10 words printed on different cards. The subject is asked to read out load and memorize the words, and then asked to recall as many words as possible. The words are presented twice again, each time in a different order, and the number of words recalled is registered after each trial. The maximum score is 10 words x 3 = 30 words. Next, delayed recall (i.e. Word List Recall) is assessed after about five minutes delay. The subject is asked to recall as many of the 10 words as possible,
giving a maximum score of 10. Last, recognition (i.e. Word List Recognition) is assessed immediately after delayed recall. The subject is presented a list of 20 words, 10 words from the list previously presented and 10 distractor words. The subject is asked to identify the 10 words from the list previously presented, and the maximum scores are 10 correct yes-responses and 10 correct no-responses. TMT A and B assess processing speed and executive functions (489). The score is the number of seconds it takes to complete the task. Kendrick Object Learning Test is a visual test of episodic memory (490). The subject is presented four boards with 10, 15, 20, and 25 pictures, and is allowed to look at the boards for 30, 45, 60, and 75 seconds, respectively. The maximum score is 70 (total number of pictures). Verbal fluency tested with the letters F, A and S assesses phonemic fluency (492). Subjects are asked to say words beginning with a letter for one minute. The score is the total number of words named for the three letters. Semantic fluency is assessed in a similar way, only that the subject is asked to say words (beginning with any letter) from a category e.g. animals (Animal Naming) or pieces of clothing (492, 493).

3.2.2 Hip fracture cohort

Delirium assessments were performed daily preoperatively and for five days postoperatively in all patients, and daily until hospital discharge in patients with delirium (Figure 6). A diagnosis of dementia at admission was based upon consensus between two experienced medical doctors.

Delirium screening

Delirium was assessed by the study physician (Leiv Otto Watne) or one of the two study nurses. They used the CAM (431), which is a diagnostic algorithm consisting of four items: 1) acute onset and fluctuating course, 2) inattention, 3) disorganized thinking, and 4) altered level of consciousness. The diagnosis of delirium requires the presence of item 1 and 2 and either 3 or 4. CAM scores were based on an interview with the patient, including tests of cognition, attention and alertness (digit span test, orientation and delayed recall), information from close relatives and nurses, and review of hospital records from the last 24 hours. Delirium assessments were done regularly only Monday through Friday, however staff members were interviewed and hospital records were reviewed every Monday, in order to discover potential episodes of delirium occurring during weekends.
Cognitive status

Relatives or health professionals (e.g. from nursing homes) were interviewed regarding the patients pre-fracture status. Pre-fracture cognitive status was assessed using the Informant Questionnaire on Cognitive Decline in the Elderly short form (IQCODE-SF) (494). A person who has known the patients for 10 years or longer is asked to fill out the questionnaire which covers cognitive change in 16 items. Each item is ranged on a five point scale, where 1 is “much better than 10 years ago”, 3 is “no change over the last 10 years”, and 5 is “much worse than 10 years ago”. The score is given as a mean measure, and higher scores represent greater cognitive decline. IQCODE-SF was scored during the hospital stay (pre-fracture cognitive status), and at follow-up four and twelve months after hospital discharge. A cognitive test battery consisting of MMSE (486), Clock Drawing Test (487), and Word List Learning and Recall from CERAD (488) was also performed at both follow-up visits. The Clinical Dementia Rating (495, 496) was used as a measure of severity of dementia. The scale consists of six domains of cognitive and functional performance, each rated 0 to 3, adding up to a sum score of 0 (normal) to 18 (severe dementia). The “sum of boxes” scoring was used (497) and based on the all available information, and scores were obtained at baseline (pre-fracture score) and both follow-up visits. For the diagnosis of dementia at baseline, one specialist in old age psychiatry (Knut Engedal) and one specialist in geriatric medicine (Torgeir Bruun Wyller) each assessed whether the hip fracture patients fulfilled the ICD-10 criteria for dementia (426). They had access to all the above mentioned information on cognitive status (except delirium status) from the index stay and one-year follow-up, and also other relevant information extracted from clinical records e.g. previous dementia diagnoses, cognitive test results and results from questionnaires on independence in everyday activities recorded prior to fracture, during the index stay, and/or at one-year follow-up. The inter-rater agreement was very good (kappa 0.87), and disagreements were resolved through discussion.

Clinical and functional status

Previous and current diagnoses were registered from medical records. Independence in everyday activities was recorded at baseline (pre-fracture status) and both follow-ups using two questionnaires 1) Barthel Index of Activities of Daily Living (498) which assesses the patient's dependency in the basic activities of daily living, such as personal hygiene, toilet hygiene, self-feeding and functional mobility. The score ranges from 0 to 20, with lower scores indicating increased dependency, and 2) Nottingham Extended Activities of Daily Living Scale (499),
assessing extended activities of daily living, such as participation in social activities, hobbies, moving in the community, and household activities. The score ranges from 0 to 66, with lower scores indicating increased dependency. The American Society of Anaesthesiologists’ classification of Physical Health (ASA) score (500) before hip fracture surgery was registered from anesthesiology records. ASA score grades preoperative health of a patient in five classes where I is a completely healthy fit patient and V is a moribund patient. APACHE II (447) score on admission was calculated based on information extracted from medical records, though without information on hematocrit and arterial blood gasses. APACHE II score measures severity of disease and increasing scores are associated with hospital death.

3.3 Biological samples

Blood samples were taken preoperatively, and CSF was collected in conjunction to spinal anesthesia in both cohorts.

Collection and handling of biological samples at baseline

In the elective surgery cohort, an ethylenediamine tetraacetic acid (EDTA) tube was one of the collected blood tubes. The EDTA tube was frozen at minus 20°C as soon as possible, and moved to a minus 80°C freezer after it was completely frozen (and within 76 hours).

CSF was collected into polypropylene tubes by the anesthesia team in both cohorts. Up to 4 mL was collected in the hip fracture cohort, whereas up to 10 mL was collected in the elective surgery cohort. In both cohorts, centrifugation of CSF samples for 10 minutes was done as soon as possible, within 4 hours for the hip fracture cohort and 2 hours for the elective surgery cohort. After centrifugation, the supernatant was transferred into polypropylene storage tubes with screw caps and stored at minus 80°C.

ApoE genotyping

EDTA tubes were thawed and 0.5 mL blood from each patient was transferred into polypropylene tubes, frozen, and sent on dry ice to the Clinical Neurochemistry Laboratory at Sahlgrenska University Hospital, Mölndal, Sweden, for ApoE genotyping. Blood samples were genotyped for APOE (gene map locus 19q13.2) using TaqMan Allelic Discrimination
technology (Applied Biosystems, Carlsbad, CA, USA). Genotypes were obtained for the two SNPs that are used to unambiguously define the ε2, ε3, and ε4 alleles (rs7412 and rs429358).

**CSF analyses**

CSF samples were sent on dry ice to the Clinical Neurochemistry Laboratory at Sahlgrenska University Hospital, Mölndal, Sweden. Aliquots of CSF taken directly from the freezer were sent from the elective surgery cohort. CSF from hip fracture patients had been thawed, aliquoted one more time in polypropylene tubes, and frozen again before they were sent.

Aβ42, T-tau and P-tau181 concentrations were determined using INNOTEST enzyme-linked immunosorbent assays (Fujirebio, Ghent, Belgium). NFL concentrations were measures using a commercial ELISA (UmanDiagnostics, Umeå, Sweden), YKL-40 concentrations using a commercial ELISA (R&D systems, Minneapolis, MN), and FABP3 concentrations using an immunoassay with electrochemiluminescence detection (MSD® Human FABP3 kit, Meso Scale Discovery, Gaithersburg, MD, USA). Analyses were performed by board-certified laboratory technicians who knew which cohort the participants belonged to, but who were masked to other clinical data including delirium status. Intra-assay coefficients of variation were 9-13%. The lower limit of detection was 50 pg/mL for NFL, 4880 pg/mL for YKL-40, and 0.103 pg/mL for FABP3.

**3.4 MRI**

MRI scans have been performed every second year in the elective surgery cohort, at baseline, 2 year, 4 year, and 6 year follow-up. Not all participants have had an MRI taken. The most common reasons for no MRI are contraindications (pacemaker etc.), claustrophobia, or that the participant refused for other reasons.

**MRI acquisition and processing**

T1-weighted MPRAGE 3D images were acquired with a 1.5 T Siemens Avanto scanner using a 12-channel head coil (TR=2400 ms, TE=3.79 ms, Field of View=240 mm, slice thickness=1.20 mm, pixel size = 1.25x1.25 mm). The same MRI scanner and T1 weighted sequence was used at baseline and at each follow up. Brain measures were obtained using the freely available semi-automated brain image morphometric software package FreeSurfer, which allows
measurement of neuroanatomic volumes, cortical thickness, surface area, and cortical gyrification of regions of interest throughout the brain (501, 502). We used FreeSurfer version 5.3 (Paper I) and 6.0 (Paper II), and its longitudinal processing scheme (https://surfer.nmr.mgh.harvard.edu), which is designed to minimize bias, increase statistical power, and increase reliability. The longitudinal processing scheme creates a subject-specific intermediate template space and image from all available time points (503), thus approximating the subject’s anatomy, and processing steps of each time point is then initialized with the template (504, 505). The image is processed through several steps in order to obtain the brain measures, and manual image editing is allowed after each stage to ensure quality control. For paper I, Sala-Llonch inspected reconstructed surfaces and volumes from individual and longitudinal processing steps, e.g. skull stripping, and gray-white matter segmentation, and performed manual corrections when necessary. For paper II, manual corrections were deemed unnecessary, and processing was repeated for the MRIs from time point 1 and 2 and also included the MRIs from time point 3 and 4. The segmentation algorithm assigns labels to all brain regions of each individual image, based on an atlas comprising probabilistic information on the location of structures (502). The atlas is obtained from a training set of images that have been accurately manually labeled (502). Segmentation of subcortical structures has been performed in our cohort, giving volumes for e.g. cerebral white matter, cerebral cortex, volumes of the ventricles, thalamus, caudate, putamen, pallidum, hippocampus, amygdala, and white matter hypointensities. Cortical thickness has also been measured (402). In the articles of this thesis, only hippocampal volume measures have been used.

3.5 Selection of participants for different articles

Paper I

The article was written after 2-year follow-up of the elective surgery cohort was completed. Only participants with available CSF data and/or brain MRI(s) were included. We selected only cognitively normal participants by first excluding participants who had been offered referral to the hospital for further cognitive assessment. Further, selection of cognitively normal participants was based on test results from the last available cognitive assessment in the study (baseline, 1-year, or 2-year follow-up), using results from all cognitive tests (11 test scores). First, we included all participants with MMSE score $\geq 27$, and next, for participants with MMSE score $< 27$, only those with none or one other abnormal test scores(s) were included. An
abnormal score was defined as a score below 4 on Clock Drawing Test (487) and as a score < -1.5 SD from the mean according to the following norms adjusted for age, sex, and/or education for the remaining tests: Word List Learning and Recall (488), Word List Recognition (506), Trail Making Tests (507), Kendrick Object Learning Test (508), FAS-test (493), Animal Naming (509). We also excluded participants with CSF NFL levels > 4000 pg/mL (i.e. more than ±3 SD from the mean value). This selection resulted in a sample of 144 cognitively normal participants with CSF analyses and/or MRI at baseline (sample A in the paper) and 88 participants with available CSF NFL analyses and MRI at both baseline and 2-year follow-up (sample B in the paper). Generalized Additive Mixed Models (GAMMs) were performed in subsamples of sample A. For the association between age and hippocampal volume, all individuals with MRI available from at least one time point were included (n=123), whereas for the association between age and hippocampal volume in NFL+ and NFL-, a valid NFL value was also required (n=108). All other statistical analyses were performed in a sample B. Some analyses were repeated in additional subsamples, as described in the article and its supplementary flow-chart. Most importantly, we created sub-samples with very low risk of AD based on APOE status (no ε4 alleles), amyloid status (amyloid negative based on different cut-offs for CSF Aβ42 levels), memory change (stable or improved delayed recall score on Word List Recall at 2-year follow-up compared to baseline), and P-tau status (P-tau negative defined as CSF P-tau < 60 pg/mL).

Paper II

The article was written at the end of 6-year follow-up in the elective surgery cohort. We selected only participants with CSF data available for all biomarkers used in the article. Further, we performed a review of all neurological diagnoses and MRI findings in the sample, and we excluded participants with diagnoses/lesions that we found likely to affect cognition or measures of hippocampal volume (details in article supplementary). Participants were excluded either from baseline, or from time of debut of the disease/lesion. Importantly, all participants who had received a diagnosis of dementia or MCI during follow up, had a cognitive impairment according to hospital medical records, had developed other neurodegenerative diseases during follow up, and participants who based on a cognitive assessment in the study had been offered referral to the hospital for further cognitive assessment were excluded from baseline. Last, from the remaining sample, selection of participants cognitively normal at baseline was based on the following procedure: 1) we included all participants with MMSE ≥ 28, if also Clock Drawing Test score was above 4, and 2) we included all participants with CSF NFL levels < 4000 pg/mL.
Test score was $\geq 4$ (487), and Word List Recall score was $\geq -1.5$ SD from the mean according to age, sex, and education adjusted norms (510, 511), and 2) we included participants with MMSE $< 28$, if Clock Drawing Test score was $\geq 4$, and also test scores for Word List Recall score, TMT A, TMT B, FAS-test, and Animal Naming were $\geq -1.5$ SD from the mean according to norms (TMTs (510, 512), FAS-test (513), Animal Naming (510, 513)). Our selection resulted in 99 participants.

**Paper III**

This article was written after all baseline examinations and follow-ups of the hip fracture cohort were finished. All patients in the cohort with available CSF data for all biomarkers used in the article and available delirium status were selected (n=129).

**Cut-off values for biomarker positivity**

In paper I, NFL positivity was defined using a median split (NFL+ if NFL levels $> 902$ pg/ml), whereas $\text{A}\beta 42+$ was defined as $\text{A}\beta 42 < 550$ pg/mL (514), and P-tau+ as $\text{P-tau} \geq 60$ pg/mL (515). In paper II, we classified participants into AT groups according to the NIA-AA criteria (1). The criteria for amyloid positivity (A+) was $\text{A}\beta 42 < 530$ pg/mL and for tau positivity (T+) $\text{P-tau} > 60$ pg/mL (515). In paper III, $\text{A}\beta 42+$ was defined as $\text{A}\beta 42 < 530$ pg/mL, T-tau+ as $> 350$ pg/mL, and P-tau+ as $\text{P-tau} \geq 60$ pg/mL (515).

**3.6 Statistical analysis**

In this section I will describe the most important statistical methods used in the articles of the thesis. Please see the respective papers for a description of all statistical methods used in the articles.

Tests of significance, correlations, linear and logistic regressions, and calculations of Cohen’s D were performed using SPSS versions 21-25 (IBM, Armonk NY). Clustering analyses were run in Matlab (MathWorks Inc). GAMMs were run in the PING data portal (http://pingstudy.ucsd.edu/welcome.html) (paper I) (516), and in R (https://www.r-project.org) using Rstudio (www.rstudio.com) IDE and using the package “mgcv” (papers I and II) (517). The significance level was set at $p < 0.05$ for all analyses.
Comparisons (Paper I-III)

We used paired-samples t-test to determine if the hippocampal volume change was significantly different from zero in paper I. In paper II, we used Cohen’s D to compare the relative contributions of the different biomarkers to the clustering. Cohen’s D is the difference of mean biomarker concentration between groups divided by the pooled standard deviation of these groups weighted for group size. Effect sizes > .80 are considered as large, > .50 as medium and > .20 as small. In paper III, Mann-Whitney U-test was used for comparing independent continuous variables due to non-normal distribution of the data. Chi-Square or Fisher's exact tests were used for comparison of categorical variables between groups.

Correlations (papers I and II)

Correlation is a measure of the relationship between two continuous variables. A positive correlation coefficient means that the value of one variable increases as the value of the other variable increases. A negative coefficient means that the value of one variable increases as the value of other variable decreases. In paper I, Pearson’s correlations were used to test relationships between age, CSF biomarkers, and hippocampal atrophy rate. In paper II, correlations were tested using bivariate Spearman correlations. Correlations between CSF biomarker levels were adjusted for age using partial Spearman correlations.

Clustering analyses (paper II)

Clustering analysis is used to identify natural groupings of similar objects from a data set. It divides objects into groups in such a way that objects in the same clusters have more similar characteristics, and objects in different clusters are more distinct. Clustering can be achieved using different algorithms. We used agglomerative hierarchical clustering where the two closest clusters or objects are successively merged until only one common cluster remains. The first step of hierarchical clustering is to calculate the distance between data points using a similarity measure. We calculated the distance between variables using Spearman correlation to account for non-normal distribution of the data. The next step is to choose a linkage function which determines how the distance between two clusters should be calculated. The linkage function then links pairs of objects or clusters to each other until all the objects in the data set are linked together in a hierarchical tree of clusters, a dendrogram. We used the “Ward's linkage” where each step merges the two clusters that results in the least increase in total within-cluster variance.
after merging. To remove the effect of age from all biomarkers, we computed independent linear regressions of each biomarker against age, and the residuals of these regressions were used for the clustering analysis. We ran two different cluster analyses. Cluster analysis 1: We used cluster analysis to establish clusters of CSF biomarkers with shared behavior across participants. In this analysis, all the available CSF biomarkers were used as variables and the participants as observations. The purpose of this was to see which CSF biomarkers that tended to go together across different number of clusters. Cluster analysis 2: The purpose of the analyses was to identify subgroups of participants with similar CSF biomarker profiles. In this analysis, participants were used as variables and the biomarkers as observations. Thus, in cluster analysis 1, we tested which biomarkers that clustered together (the CSF biomarkers were the variables), while in cluster analysis 2 we tested which participants that clustered together (the participants were the variables). One advantage of the hierarchical clustering is that it results in a dendrogram that makes it easy to visualize how the clusters are formed. Accordingly, the number of biomarker groups in paper II was established by inspection of the hierarchical distribution of the dendrogram. Clustering analyses were performed by the postdoctoral researcher on the project, R. Sala-Llonch.

**Linear regression (Paper I)**

Simple linear regression is used to evaluate the relationship between a continuous predictor variable and a continuous outcome variable, whereas multiple linear regression allows for two or more predictor variables (continuous and/or categorical). We used linear regression to test whether CSF NFL levels predicted hippocampal atrophy rate. The first regression model included NFL and age. Next, we added Aβ42 and P-tau separately to the model in conjunction with NFL and age. As we wanted to assess the contributions of each biomarker, several biomarkers were entered in the same model even though we expected some of them to correlate. The final regression model was also tested for robustness by adding the potential confounders white matter hypointensities and sex. Sensitivity analyses excluding outliers (defined as studentized deleted residuals > ± 2) were run for all regression models. The analyses were also repeated in low-risk groups.
Logistic regression (Paper III)

Logistic regression is used to assess the relationship between one or more predictor variables (categorical or continuous) and a categorical outcome variable, e.g. a dichotomous variable. In paper III, we used logistic regression to adjust for potential confounders of the relationship between CSF biomarkers and delirium in patients without dementia. We adjusted for the potential confounding factors age, gender, and IQCODE-SF score using logistic regression with delirium as the outcome variable. We built separate regression models for each of the CSF biomarkers/biomarker-ratios, forced all the potential confounding factors into the model, and checked if the association between the biomarker/biomarker-ratio was upheld. Sensitivity analyses excluding outliers (defined as standard residuals $> \pm 3$ or Cooks distance $> 1$) were run for all regression models.

Generalized Additive Mixed Models (papers I and II)

GAMMs is a statistical model expanding on the properties of generalized additive models. Being an additive model, the model is able to model non-linear curves. The optimal shape of the non-linearity is determined automatically, thus relationships of any degree of complexity can be modelled without specification of the basic shape of the curve (e.g. linear or quadratic). Being a mixed model, the model can handle the dependencies between repeated measures from the same individuals, also if time spacing is uneven. Further, mixed models allow the use of all available data, such that data from subjects with missing data points in follow-up can be used. For more mathematical details on GAMMs, please see Wood 2006 (517).

In paper I, GAMMs were used to confirm and illustrate findings from correlations and multiple linear regression analyses. The first model tested the association between age and hippocampal volumes over time adjusted for sex. The second model also tested the relationship between age and hippocampal volume, adjusted for sex and P-tau, however, the sample was divided into NFL+ and NFL- by a median split. In paper II, GAMMs were used to assess whether biomarker-based subgroups showed different trajectories of hippocampal volume or memory over time. Memory score from Word List Recall for up to seven time points was used as outcome variable, biomarker group as factor, participant-specific time since baseline as covariate, and we included a time x biomarker group interaction term. Sex and baseline age were included as covariates of no interest. Random intercept was included. The same analyses were run with hippocampal volume for up to four time points as outcome variable. The same variables and covariates as
for the memory analyses were included. In addition, estimated total intracranial volume was included as an additional covariate of no interest. GAMM fits are typically evaluated and inspected based on p- and F-values, edf (effective degrees of freedom) as a measure of the complexity of the curve, as well as by inspecting the plotted graphs. GAMM analyses were carried out by my co-supervisor prof. A.M. Fjell.

3.7 Ethical considerations

The studies were conducted in accordance with the Declaration of Helsinki and approved by the Regional Committee for Ethics in Medical Research in Norway (REK 2009/450 and REK 2011/2052). All participants in the elective surgery cohort and also cognitively intact hip fracture patients provided written informed consent. Hip fracture patients with reduced ability to give informed consent received a simplified information leaflet, and a next of kin was given the full written information. Those who were unable to give a valid informed consent were included based on assent from the next of kin.
4 Main results

In this section the main results from the papers included in this thesis are summarized. Please see the respective papers for a complete description of the results.

Paper I

The main result of this paper was that high CSF NFL levels predicted higher hippocampal atrophy rate in cognitively normal older adults, also after adjusting for age, Aβ42, and P-tau in regression models. The results from the regression analyses were also confirmed by GAMM showing more hippocampal volume change in NFL+ than NFL- individuals (Figure 7). The relationship between NFL and hippocampal atrophy was also tested in several subgroups with very low risk of AD, and the main result was upheld in all groups, including a group of Aβ42 and P-tau negative individuals without APOE ε4 alleles and no memory change over 2 years. We also found that NFL correlated positively with age and P-tau, but not with Aβ42.

![Figure 7. Relationship between age and hippocampal volume in NFL+ and NFL- participants. Estimated group slopes with 95 % confidence intervals are displayed. Abbreviations: NFL, neurofilament light. From “CSF neurofilament light levels predict hippocampal atrophy in cognitively healthy older adults.” By Idland et al., 2017, Neurobiol Aging. 2017 Jan;49:138-144. Copyright [2016] by Elsevier Inc. Reprint permission not required.](image-url)
There were several important results in this paper on cognitively normal older adults. First, CSF T-tau, P-tau, YKL-40, NFL and FABP3 were all positively correlated, whereas Aβ42 did not correlate with any of the other CSF biomarkers. Accordingly, clustering analysis of biomarkers showed that the novel CSF biomarkers NFL, FABP3, and YKL-40 clustered with T-tau and P-tau, while Aβ42 was separated out in an independent cluster (Figure 8).

Second, clustering analyses of participants identified two main biomarker profiles, where one biomarker group had more abnormal levels of all biomarkers (group 2) compared to the other biomarker group (group 1) (Figure 9). The group with more pathological biomarkers was further split in one group characterized by higher T-tau, P-tau and FABP3 levels (group 2.2) and one group characterized by lower Aβ42, higher NFL, and tendencies to higher YKL-40 levels (group 2.1) (Figure 9). Similarly, the group with less pathological biomarker levels was split into one group characterized by higher T-tau, P-tau and FABP3 levels (group 1.1), and one group with lower Aβ42 and slightly higher NFL levels (group 1.2) (Figure 9).
Figure 9. Hierarchical clustering of participants. Left panel: Subject-wise correlation matrix and dendrogram of the groups at the different levels. Right panel: Mean Z scores of each variable within each group. The z-scores are calculated for the current sample, yielding a sample sum of 0 and a standard deviation of 1, thus the groups tend to approximately mirror each other around the y=0 axis when the group sizes are similar.

Third, hippocampal volume trajectories did not differ between any of the clustering-based biomarker groups nor the AT groups.

Forth, the group with highest levels of T-tau, P-tau, and FABP3 (group 2.2) showed more memory decline over 6.8 years compared to the group with lowest levels of the same biomarkers (group 1.2) (Figure 10). No other differences in memory trajectories were seen between cluster-based biomarker groups. A+T+ and A-T+ groups did not show significantly different changes in hippocampal volume or memory compared to a normal AD biomarker group (A-T-), although there was a tendency for the A-T+ group to show more memory decline over time.
Figure 10. Longitudinal change in memory function across biomarker groups. Upper panel: GAMM-fitted change slope in memory score for Group 1 and Group 2 across time. Lower panel: Memory slopes for each of the groups in the 4-cluster solution.

Paper III

The main result of this paper was that the core AD biomarkers (low CSF Aβ42, high CSF T-tau, low Aβ42/T-tau, low Aβ42/P-tau) were significantly associated with delirium in patients without dementia (Figure 11), also in adjusted analyses. There was a tendency toward an association between high CSF P-tau and delirium in patients without dementia. We found no associations between the CSF biomarkers Aβ42, T-tau, and P-tau and delirium in patients with dementia.
Figure 11. Cerebrospinal fluid (CSF) Aβ42, T-tau and Aβ42/T-tau ratio in patients without dementia with and without delirium. There were 16 patients with delirium and 49 without. P-values were calculated using Mann-Whitney U-tests. From “Preclinical Amyloid-β and Axonal Degeneration Pathology in Delirium.” By Idland et al., 2017, Journal of Alzheimer’s Disease. Volume 55, p. 371-379. Copyright [2017] by IOS Press. Reprinted and with permission.

Patients were also divided into four groups according to amyloid and/or tau pathology (Aβ42- and P-tau-, Aβ42- and P-tau+, Aβ42+ and P-tau-, Aβ42+ and P-tau+). In patients without dementia, we found that a significantly greater proportion of the patients with both amyloid and tau pathology developed delirium compared to the other three groups (Figure 12). In patients with dementia, there were no difference in proportion of patients with delirium between the biomarker groups. It must however be noted, that these analyses were unadjusted.
Figure 12. Delirium occurrence according to combinations of amyloid and tau pathology. Percent of patients who developed delirium in conjunction to the hip fracture according to grouping by positive and negative Aβ42 and P-tau values in hip fracture patients without dementia. Subjects were classified as Aβ42+ (< 530 ng/L) or Aβ42- (≥ 530 ng/L), and P-tau+ (≥ 60 ng/L) or P-tau- (< 60 ng/L). P-values were calculated using Chi-Square and Fisher’s Exact Tests. From “Preclinical Amyloid-β and Axonal Degeneration Pathology in Delirium.” By Idland et al., 2017, Journal of Alzheimer's Disease. Volume 55, p.371-379, Supplementary Material. Copyright [2017] by IOS Press. Adapted and reprinted and with permission.
5 Discussion

5.1 Discussion of main findings

Summary of main findings

The main aim of this thesis was to increase our knowledge about the role of different neuropathological processes in cognitively normal older adults and delirium using AD-associated CSF biomarkers reflecting core AD pathology and other neuropathological processes. We therefore examined relationships between different biomarkers reflecting various neuropathological processes in cognitively normal older adults. We found that the novel biomarkers NFL, YKL-40, and FABP3 correlated with and clustered with T-tau and P-tau, but not with Aβ42. Further, we tested whether subgroups of older adults with similar biomarker profiles could be detected using clustering analyses. We found one group with more abnormal levels of all biomarkers compared to the other biomarker group. The group with more abnormal biomarkers was further split in one group characterized by higher T-tau, P-tau, and FABP3 and one group characterized by lower Aβ42, higher NFL, and slightly higher YKL-40. Similarly, the group with most normal biomarkers was split into one group characterized by higher T-tau, P-tau, and FABP3, and one group characterized by lower Aβ42 and slightly higher NFL. Furthermore, we examined the relationship of biomarkers and biomarker-subgroups to brain and cognitive changes known to be present both in aging and AD dementia (i.e. hippocampal atrophy and memory decline). We found that high CSF NFL levels predicted higher hippocampal atrophy rate in cognitively normal older adults, independently of AD pathology, and also in individuals with very low risk of AD. We also found that the clustering-based biomarker subgroup of cognitively normal older adults characterized by highest T-tau, P-tau, and FABP3 showed more memory decline compared to the group with lowest levels of the same biomarkers. Conversely, the clustering-based biomarker subgroup characterized by lowest Aβ42, highest NFL and slightly elevated YKL-40 and a group with biomarker defined AD (A+T+) did not show more memory decline compared to groups with less pathological biomarkers. Moreover, we examined whether the core AD biomarkers were related to delirium in patients with and without dementia. We found that the core AD biomarkers were significantly associated with delirium in patients without dementia, and also that the proportion of patients with delirium was significantly higher in the A+T+ group compared to all other AT groups.
5.1.1 Consequences of amyloid and tau pathology in older adults without dementia

AD has a long-lasting preclinical phase, and although still controversial (77), biomarkers of amyloid pathology are believed to be the first detectable sign of AD. Accordingly, the recent NIA-AA Research Framework defines Alzheimer’s pathologic change based on a positive amyloid biomarker (A+) alone (1), and regardless of the clinical stage. Amyloid positivity has, however, only small effects on cognition in cognitively normal older adults (311-314), and there is substantial evidence that Aβ deposition has to be accompanied with tauopathy in order to have considerable impact on cognition (297, 329, 331, 518). In the NIA-AA Research Framework, AD is defined as positivity of both markers of brain Aβ deposition (A+) and markers of tauopathy (T+) (1). However, a substantial number of cognitively normal older adults show biomarker evidence of Aβ deposition and/or tauopathy (12, 283), but the implications of this brain pathology are not established.

Papers II and III in this thesis reported on cognitive consequences of amyloid and tau pathology in individuals without dementia. In paper III, we found a significant association between the core AD biomarkers Aβ42 and T-tau and delirium in patients without dementia, whereas only a trend was seen for P-tau. The Aβ42/T-tau and Aβ42/P-tau ratios were also significantly associated with delirium in patients without dementia. Although P-tau was not significantly associated with delirium, a 38% of the A+ patients were also T+, and the proportion of patients developing delirium was significantly higher in A+T+ individuals than in A+T- individuals (56% vs 13%), suggesting a role of tauopathy in the presence of Aβ deposition in delirium. This finding parallels previous research showing steeper cognitive decline in A+T+ cognitively normal individuals compared to A+T- individuals (297). As the association between Aβ deposition and delirium was found in patients without dementia, the findings in paper III suggests that Aβ deposition has cognitive consequences also in the preclinical stage. The results remained significant also after adjusting for cognitive functioning, further supporting this implication, and also raising the question of whether delirium is an early symptom of AD. It must however be noted that it is likely that some of the patients had MCI, and thus had prodromal, and not preclinical, AD. Research on consequences of core AD pathology in MCI patients also parallels our findings, showing an association between each of the core AD biomarkers and disease progression (95). Furthermore, a combination of CSF Aβ42 and P-tau is found to be a better predictor of disease progression in MCI than CSF Aβ42 alone (515, 519).
The association between core AD pathologies and more rapidly developing and transient cognitive impairments is less studied. In addition to the delirium studies that will be discussed in detail in section 5.1.3, core AD biomarkers have been assessed in relation to postoperative cognitive dysfunction (POCD) (520-522). POCD is a usually transient cognitive impairment arising days to weeks after anaesthesia and surgery, and it is detected using neuropsychological tests (523). Studies of individuals without dementia who have been neuropsychologically tested prior to surgery and also at different intervals postoperatively have shown that the CSF Aβ42/T-tau ratio is associated with POCD at 1 week as well as 3 to 6 months postoperatively (520, 521), and that CSF Aβ42 is associated with POCD 3 months postoperatively (522). These findings correspond with our findings in delirium. Similarly, mice with brain Aβ deposition are shown to develop cognitive impairment after surgery, whereas wild type mice of the same age do not (524). Only one small study has assessed the association between CSF P-tau and POCD (522), and in accordance with our findings in delirium, there was no significant association. The finding that AD pathology is associated with both delirium and POCD in individuals without dementia suggests that AD neuropathology represents a vulnerability factor, predisposing the individual to acute and subacute cognitive decline. Research has shown that higher cognitive reserves may protect against delirium and POCD (525, 526). Accordingly, one may speculate that individuals with preclinical AD have cognitive and/or brain reserves that allow them to compensate for the pathology, but when exposed to an insult (precipitating factor), they are no longer able to compensate, and hence develop cognitive impairment. Thus, delirium and POCD may be early manifestations of AD.

In the above discussion I have shown how paper III adds to the existing literature suggesting that core AD pathologies have cognitive consequences also in patients without dementia. Particularly, paper III was the first paper to convincingly show that core AD biomarkers were associated with acute cognitive impairment (i.e. delirium). Next, I will discuss the role of amyloid and tau pathology in cognitively normal older adults.

As opposed to paper III, in paper II, we did not find that the combination of biomarker signs of Aβ deposition and tauopathy had cognitive consequences when assessed in cognitively normal older adults. That is, A+T+ individuals did not show more memory decline over 6.8 years compared to A-T- individuals. Neither, the clustering-based biomarker subgroup with biomarker signs of Aβ deposition and also slight tauopathy (group 2.1) showed more memory decline compared to a subgroup with less pathological biomarkers. This finding differs from
several previous studies reporting that cognitively normal individuals with both amyloid and tau pathology show accelerated cognitive decline compared to those with one or none of these pathologies (297, 319, 325, 329). We can only speculate why we did not find such an association. One explanation may be that our selection of cognitively normal older adults based on several cognitive tests including an episodic memory test, and our exclusion of participants who during follow up developed cognitive impairment, may have excluded the individuals with cognitive consequences of core AD pathology who likely contributed to the findings in other studies. Supporting this speculation, a meta-analysis assessing the risk of clinical progression in different stages of preclinical AD found heterogeneity in prevalence of biomarker positivity and risk of progression (527), and suggested that the reason could be that methods for definition of cognitively normal individuals differed between studies. As cognitively normal individuals show steeper memory decline if Aβ deposition is accompanied with pathological hippocampal volumes (279, 282), we have also speculated that the key to understand why some biomarker groups with Aβ deposition and/or tauopathy only showed age-expected memory decline is that these participants did not show higher than age-expected hippocampal atrophy. Furthermore, the degree of neurodegeneration, rather than its presence, has been shown to play a role in cognitive decline in amyloid positive individuals (112, 113, 279, 281, 319). Therefore, the slight tauopathy in our clustering-based subgroup may not be sufficient to cause memory decline.

Nevertheless, the findings in paper II underscore the fact that older adults can have good cognitive function for their age, and show age-expected changes in memory function over several years, despite biomarker profiles indicating Aβ deposition and tauopathy. Our findings suggest that core AD pathology does not always lead to accelerated hippocampal atrophy or memory decline in older adults. However, there is possibility that more experimental cognitive test could be more sensitive than the usual clinical instruments used to assess the participants in the current study. Another explanation, with support in previous literature, is that these individuals have higher cognitive and/or brain reserves enabling them tolerate more AD pathology before they develop cognitive symptoms (406-408), thus it is possible that their rate of memory decline may accelerate later. Research on such factors that promote resilience to pathology may give cues on possible therapeutic targets for prevention of development of dementia in individuals with preclinical AD (518). One interesting question is whether these individuals will eventually develop AD dementia. A recent meta-analysis reported 38 % risk of progression in A+T+ individuals over a period up to 10 years, however, this also shows that more than half did not progress (527). Post-mortem studies show that 20-40 % of individuals
who were cognitively normal on the last assessment before death reached neuropathological
criteria for AD (12, 287), and we will never know if these individuals would eventually develop
AD dementia if they lived long enough. The inconsistencies of existing studies on clinical
consequences of preclinical AD also elucidates the need for studies with very extensive follow-
up of cognitively normal adults with evidence of core AD pathology using sensitive cognitive
measures. Such studies will be necessary to determine if and when anti-amyloid treatments
should be started in cognitively normal adults if they become available, in order to avoid
treating persons who would never experience symptoms of preclinical AD pathology.
Additionally, in paper II we reported that 32 % of the cognitively normal older adults did not
have core AD pathology, and there is evidence that some individuals may never develop AD
pathology (20, 21, 283, 287). Future studies on such groups of very low-risk older adults may
give cues on how to prevent development of AD pathology (528).

Prevalences of tauopathy and neurodegeneration increase with increasing age (283), and these
pathologies can be found also in the absence of amyloid pathology, e.g. as PART (autopsy
based) (296) or SNAP (biomarker based) (390). We found that the clustering-based biomarker
subgroup with the highest levels of P-tau, T-tau, and FABP3 (i.e. evidence of most tauopathy
and neuronal damage) showed more memory decline than the clustering-based biomarker
subgroup with the lowest levels of the same biomarkers. Accordingly, the A-T+ group also
showed tendencies toward more memory decline compared to the A-T- group. Our finding
indicates that the degree of tauopathy and neuronal damage plays a role in cognitive decline
independently of amyloid pathology, which is in line with studies on PART (529) and amyloid
negative cognitively normal individuals (279, 394). Furthermore, this finding parallels studies
showing that higher levels of neurodegeneration biomarkers are associated with greater
cognitive decline in amyloid biomarker positive individuals (112, 113, 279, 281, 319), and
studies showing an association between neurodegeneration biomarkers and cognitive decline
also after adjusting for the degree of amyloid pathology (142, 316, 317, 319, 320, 530),
suggesting that a high degree of neurodegeneration is associated with cognitive decline
irrespective of the pathology causing it. The association between high T-tau and delirium in
paper III is also in accordance with this suggestion, although this association was not adjusted
for Aβ42. It must however be noted that the Aβ42 concentration in the clustering-based
subgroup characterized by highest T-tau, P-tau, and FABP3 was also close to the pathological
threshold. Thus, we cannot exclude that the combination of subthreshold amyloid positivity and
a high degree of tauopathy and neurodegeneration was responsible for more memory decline in this group.

In paper II, we found no differences between any of the clustering-based biomarker subgroups or the AT groups in hippocampal volume trajectories. As summarized in section 1.2.4, reports on the relationship between biomarkers of core AD pathology and hippocampal atrophy are not consistent, although many find that more pathological levels are related to hippocampal atrophy (302, 322, 340, 345, 349, 357, 360). It has also been reported that A+T+ cognitively normal individuals show more hippocampal volume decline than A-T- and A+T- (305). As for the lack of more than age-related memory decline in A+T+ individuals, we could speculate the lack of association between biomarker groups and hippocampal atrophy may due to selection criteria for our sample. Lack of power to detect differences is also a possibility, however studies with lower power than ours have found associations between AD biomarker groups and MTL atrophy in cognitively normal older adults (366, 531, 532). It should still be mentioned that in a subsample in paper I, we found that low A\(\beta\)42 predicted higher hippocampal atrophy rate in cognitively normal older adults.

In the discussion above I have shown how findings in paper II relate to previous literature on the consequences of amyloid and tau pathology in cognitively normal older adults. Notably, our findings suggest that older adults may uphold age-expected cognitive function and hippocampal integrity even when harboring AD pathologies. Our findings also suggested that a high degree of tauopathy and neuronal damage is associated with more memory decline in cognitively normal older adults.

5.1.2 Novel biomarkers in cognitively normal older adults

Many novel biomarkers can be measured in CSF, however this discussion will primarily focus on the novel CSF biomarkers examined in the articles of this thesis, namely YKL-40, NFL, and FABP3. These three biomarkers have been reported to be elevated in AD dementia (95), and also in diverse acute and chronic neurological diseases (114, 136, 144, 149, 164, 168). Furthermore, the biomarkers have been associated with brain atrophy in AD (YKL-40, NFL, FABP3) (113, 160, 175, 533) and other neurodegenerative diseases (YKL-40 and NFL) (534-537). All three biomarkers have also been associated with disease progression in AD (112, 113, 157, 160, 170) and other neurodegenerative diseases (133, 135, 184, 536, 538-541). Their role in aging is, however, less studied.
In paper I, we found that high CSF NFL levels predicted a higher hippocampal atrophy rate in cognitively normal older adults, also after adjusting for age, Aβ42, and P-tau. When paper I was written, the literature on the relationship between NFL and atrophy was scarce. Only one study had assessed NFL in relation to longitudinal volume change, showing that high NFL predicted more hippocampal atrophy in a cohort of MCI patients (129). Paper I was the first to assess the relationship between NFL and hippocampal volumes in a cognitively normal cohort. In line with our findings, several later studies have found a relationship between high NFL (measured in blood or CSF) and more hippocampal atrophy in cognitively normal adults (542-544). This relationship has also been found in patients with AD dementia and FTD (545, 546). NFL has also been associated with other brain measures of atrophy (e.g. cortical thickness) and neurodegeneration (e.g. FDG-PET) in cognitively normal adults (401, 533, 542, 543, 547-549), supporting its role as a general neurodegeneration biomarker.

We also examined the relationship between NFL and hippocampal atrophy in several subgroups unlikely to have preclinical AD, and higher NFL was associated with higher atrophy rates in all groups. Consequently, one implication of our results is that NFL likely indicates AD-independent neurodegeneration. This is in agreement with studies showing a relationship between NFL and hippocampal atrophy and other measures of neurodegeneration in A-individuals (113, 533, 549), and in non-AD dementia (535, 546, 550). Moreover, NFL increase has been shown to correspond with onset and progression of different proteopathic lesions (Aβ deposition, α-synucleinopathy, tauopathy) in an animal study (143). NFL concentrations are also shown to increase after acute brain damage (114, 131), and, equal to our study, other studies have found that higher NFL levels are associated with faster rates of atrophy (545, 550), suggesting that concentrations of NFL reflect the intensity of axonal injury. As suggested by the association between NFL and hippocampal atrophy, recent studies have found an association between high NFL and poorer memory performance and more memory decline in cognitively normal adults (403, 544, 551, 552). Furthermore, high NFL levels have been associated with cognitive decline in MCI patients (112, 129, 542, 545) and AD dementia (112, 545), and with disease progression of different neurodegenerative diseases (133, 135, 184, 536, 538). Altogether, NFL appear to be a not disease-specific neurodegeneration and progression biomarker. Our results further suggest that NFL also reflects age-expected neurodegeneration and that common neurodegenerative processes are ongoing in aging and disease.
In paper II, we found that the novel biomarkers NFL, YKL-40, and FABP3 correlated with and clustered with T-tau and P-tau, but not with Aβ42. Such a divide between Aβ42 and other CSF biomarkers, including FABP3 and P-tau, in clustering analyses has also been shown in a cohort comprising cognitively normal individuals and patients with MCI and AD dementia (181), and non-existent or weak relationships between each of the three novel biomarkers and Aβ42 have previously been shown in cognitively normal adults (112, 142, 154, 171, 397, 400). The divide suggests that axonal degeneration, neuroinflammation, and neuronal injury are processes related to tauopathy and neurodegeneration, and not to Aβ deposition, in cognitively normal older adults. In agreement with our study, positive correlations between the Tau biomarkers and YKL-40, NFL, and FABP3 have been found in previous studies of cognitively normal adults (112, 142, 171, 401). The association between YKL-40 and Tau has been shown in both A+ and A- cognitively normal individuals (154), suggesting that the link between neurodegeneration and neuroinflammation is independent of amyloid deposition. Accordingly, YKL-40 and NFL have been shown to be elevated in T+ and N+ groups compared to A+T-N- or A-T-N- groups across cognitive stages (153, 543). In keeping with previous studies (112, 401), we also found that NFL, FABP3, and YKL-40 were all positively correlated, further supporting a link between neuroinflammation and neurodegeneration. This was the first study to assess the relationship of FABP3 to YKL-40 and NFL in a cognitively normal cohort, however positive correlations had previously been reported across cognitively normal, MCI and AD dementia (181, 553).

In paper II, we also tested whether subgroups of older adults with similar biomarker profiles could be detected using clustering analyses. The analysis identified one group with more abnormal biomarkers compared to the other group. The most abnormal group was further split into one group characterized by Aβ deposition (low Aβ42), axonal degeneration (high NFL), and slight inflammation (elevated YKL-40) and one group characterized by tauopathy (high T-tau and P-tau) and neuronal damage (high FABP3). Similarly, the most normal group was split into one group characterized by slight tauopathy and higher FABP3, and one group with lower Aβ42 and slightly higher NFL, paralleling the split in the most abnormal group, although biomarker levels were more normal in both these subgroups. Previous clustering analyses of cognitively normal adults had only included the CSF biomarkers Aβ42 and Tau (554, 555), thus we did not know how the six biomarkers would contribute to the clustering. Intriguingly, we found that all six CSF biomarkers differed between at least two groups with a relatively large effect size, suggesting that also the three novel biomarkers contributed to the subgrouping of
individuals. Yet, a recently published study clustered cognitively normal individuals and individuals with subjective memory complaints, MCI, and AD dementia using CSF Aβ42, T-tau, P-tau, NFL, and YKL-40 (556). In agreement with findings in paper II, this study found two clusters characterized by neurodegenerative processes, but not Aβ deposition. 25% of the cognitively normal individuals were found in one of these two subgroups, whereas 75% were found in the remaining three clusters characterized by more normal biomarker levels. The largest proportion (35%) of cognitively normal individuals was found in a cluster characterized by high Aβ42, low T-tau, and low YKL-40. Contrary, 62% of patients with AD dementia were included in the two neurodegeneration groups, suggesting that neurodegeneration, independent of Aβ deposition, is essential for the development of cognitive impairment. In this study all biomarkers except Aβ42 contributed to the clustering of individuals, and YKL-40 contributed prominently to classification of one of the neurodegeneration clusters.

Multiple neuropathologies are common in older adults, and they may have additive and/or synergistic effects on brain atrophy and cognition. YKL-40, NFL, and FABP3 have been found to interact with core AD biomarkers in prediction of hippocampal atrophy and other brain measures (113, 151, 155, 175, 401, 533, 549), and previous studies have suggested that combinations of these novel biomarkers and core AD biomarkers could be useful for prediction of disease progression (112, 120, 147, 170, 184, 536). Hence, in paper II, we also assessed if the cluster-based biomarker profiles could predict hippocampal atrophy and memory decline. As discussed above, neither the clustering-based biomarker profiles, nor the AT groups, showed more than age-related hippocampal atrophy. However, the clustering-based biomarker group characterized by highest concentrations of FABP3, T-tau, and P-tau showed worse memory decline, whereas the AT-classification was not associated with memory decline. Thus, subgrouping of participants based on a wide range of biomarkers improved prediction of memory decline compared to the canonical AT-classification. This finding suggest that it is meaningful to use multiple and novel biomarkers for characterization of brain states in cognitively normal older adults, and that clustering studies may identify biomarker profiles relevant for clinical outcomes, and thus clinical trials.

In the discussion above I have shown how findings in papers I and II relate to findings in other studies of NFL, YKL-40, and FABP3. In agreement with later studies, findings in paper I indicate that NFL is related to brain atrophy independently of AD neuropathology and reflects the intensity of axonal injury. In paper II, we found that the three novel biomarkers are related
to tauopathy and neurodegeneration, and not amyloid pathology, which is in accordance with previous literature. Finally, this was the first study clustering cognitively normal individuals based on biomarkers other than Aβ42 and Tau. We found that the three novel biomarkers contributed to classification of cognitively normal older adults, and the resulting subgrouping of participants predicted subsequent memory decline better than the traditional AT-classification.

5.1.3 The interrelationship between AD pathologies and delirium

Dementia is an established and strong risk factor for delirium (412), and AD is the most common cause of dementia (3). Low cognitive performance scores and MCI in patients without dementia have also been associated with increased risk of delirium (467-469, 557, 558). In order to understand the link between cognitive impairment and delirium, it is relevant to study the interrelationship between AD pathologies and delirium.

Before paper III, only two studies had examined the relationship between CSF core AD biomarkers and delirium (479, 480). Similar to our study, Witlox et al. had examined this relationship in a cohort of hip fracture patients without dementia (480). However, they found no association between the biomarkers and delirium. Xie et al. had examined the relationship between AD core biomarkers and delirium in patients without dementia undergoing surgery in spinal anesthesia (479). They found no association between individual AD biomarker concentrations and delirium, but when they divided the participants into quartiles based on Ab40/T-tau and Ab42/T-tau ratios, they found a higher delirium incidence in the lowest quartile. Hence, results from previous studies were conflicting. In paper III, we found that biomarker levels of Aβ42 and T-tau, and also Aβ42/T-tau and Aβ42/P-tau ratios, were associated with delirium in patients without dementia. Our study was therefore the first to show a convincing association between CSF core AD biomarkers and delirium. It was also the first to assess the relationship between these biomarkers and delirium in patients with dementia, but we found no association between core AD biomarkers and delirium in this subgroup. Explaining the conflicting results in our study and the study of Witlox et al. is not straightforward, but we have suggested that more normal AD biomarkers concentrations in their study than ours indicate that AD neuropathology was less prevalent in their study. Results from the only later study on CSF core AD biomarkers and delirium is in keeping with our findings (559). Cunningham et al. found that Aβ42 concentrations, but not T-tau or P-tau, were associated with
delirium in an elective surgery cohort without dementia. In conclusion, studies of the relationship between CSF core AD biomarker and delirium have provided mixed results, but they suggest a link between amyloid pathology and delirium.

The relationship between AD pathology and delirium has also been assessed using other AD-associated biomarkers. Amyloid biomarkers in blood have been associated with delirium (560, 561), but amyloid biomarkers in blood are less reliable biomarkers of brain Aβ deposition than CSF measures (at least until recently (190, 191)). Further, the relationship between APOEε4 positivity, an AD genetic risk factor associated with higher prevalences of amyloid and tau positivity (13, 293), and delirium has been assessed in several studies. The studies have yielded mixed results, but a meta-analysis from 2016 found no association between the APOEε4 allele and the occurrence of delirium (562). Moreover, AD-associated brain atrophy measures like AD-signature cortical thickness and hippocampal volumes have been associated with delirium incidence, duration or severity in some studies (563-566), but not all (567). Hippocampal microstructural abnormalities have also been associated with delirium incidence and severity (568). Moreover, a study of 3 large population-based cohort studies found no difference in post-mortem AD pathologies between individuals with and without a history of delirium (425).

In summary, our study and some other existing studies suggest that AD pathology is associated with delirium and may be a vulnerability marker. This has been shown in populations without dementia, indicating that delirium may unmask preclinical or prodromal AD. Other studies did however not find a relationship between AD pathology and delirium, and larger studies on different populations are needed. If future studies proves that AD pathology is a risk factor for delirium, this would implicate that AD biomarkers may be used to identify patients at risk of delirium and that disease-modifying therapies for AD will decrease delirium rates if/when they become available.

Delirium is associated with accelerated cognitive decline in patients with AD dementia and in patients without dementia (470-472). It is possible that delirium is itself (or its precipitating factor) accelerates AD neuropathology or initiates other pathological changes leading to cognitive decline and dementia. Another possibility is that delirium is only a marker of the degree of AD pathology, and that it is the AD pathology that causes later cognitive decline. The only study on longitudinal cognitive decline in relation to dementia pathologies and history of delirium found that delirium alone, and also the interaction between pathological burden and delirium, accounted for some of the cognitive decline (425), suggesting that delirium itself may
cause neuropathological changes leading to cognitive decline. Additionally, a recent study showed a sharper postoperative rise in plasma NFL in patients who developed delirium compared to those who did not (478). This finding suggests that delirium is associated with acute neuroaxonal injury, and it is likely that such neuroaxonal injury has cognitive consequences. The relationship between delirium and later cognitive decline is, however, far from understood, and longitudinal studies with biomarkers and repeated cognitive assessments are needed. If future research shows that delirium triggers pathological processes leading to dementia, prevention and treatment of delirium will be important means for preventing dementia.

5.2 Methodological considerations

5.2.1 Study designs

The hip fracture cohort and the elective surgery cohort both represent convenience samples, and a limitation is therefore that results may not be generalized to the general populations we seek to study, that is all hospitalized patients and all cognitively normal adults.

The cognitively normal older adults studied are highly educated and likely to be healthier than the general population of cognitively normal older adults. They may therefore have less brain pathologies and memory decline than the general population of cognitively normal older adults. Furthermore, participants who died or declined further follow-up may have been less healthy, which may have led to an underestimation of progression of hippocampal atrophy and memory decline in our sample. Delirium pathophysiology is likely to differ according to the factors triggering delirium. Therefore, the interrelationship between AD pathology and delirium may not be the same in other populations, e.g. younger populations or populations with more noxious insults like sepsis or multi-trauma triggering delirium.

CSF sampling from patients who are undergoing a lumbar puncture anyway may, however, be the most feasible way to obtain CSF from older adults for biomarker studies. Actually, when we performed lumbar punctures to obtain CSF samples from the elective surgery cohort at 4 year follow-up, we only obtained CSF from 38 of the 105 cognitively tested participants. One cause was that antithrombotic medication was an exclusion criteria for lumbar puncture, but this was not an issue when CSF was obtained in conjunction to spinal anesthesia. Also, several participants did not want to undergo a lumbar puncture for the study's purpose only. Therefore,
the collection of CSF in conjunction to spinal anesthesia may have resulted in inclusion of cognitively normal older adults that are usually excluded from studies on CSF biomarkers, thereby increasing the generalizability of our study. Individuals using anticoagulants are for instance excluded from lumbar puncture in the Alzheimer's Disease Neuroimaging Initiative (ADNI) (569).

CSF biomarkers concentrations from only one time point was used in both cohorts. This is a limitation in the hip fracture cohort in particular, because we do not know the temporal relationship between biomarker changes and delirium. Some patients had ongoing delirium at the time of CSF sampling, and although the core AD biomarker concentrations likely represents the patient’s brain state before the hip fracture, we cannot rule out that acute changes in these biomarkers occurred before, during, or after delirium. Longitudinal samples would also be informative in the elective surgery cohort, but the CSF biomarkers in the papers in this thesis have not yet been analyzed in the samples from 4 year follow-up.

In the elective surgery cohort, postoperative delirium assessments were planned, but regrettably not performed due to lack of resources. Since an acute preoperative change in biomarker concentrations is less likely in this population, it would have been an advantage to assess the relationship between core AD biomarkers and delirium also in this cohort. A delirium assessment in this cohort would also enable studies of brain imaging measures in relation to delirium and studies on longitudinal cognitive decline in relation to AD neuropathology and delirium.

### 5.2.2 Diagnoses and selection of participants

**Delirium.** Diagnosing delirium can be challenging. In order to differentiate delirium and dementia, it is essential to determine whether the cognitive impairment represents a change from baseline. Certain information can be hard to obtain as information from someone who knows the patient is often essential. Further, dichotomization of “delirium” and “no delirium” is an oversimplification as delirium in reality is not a binary phenomenon, but represent a continuum of cognitive symptoms. Accordingly, the term “subsyndromal delirium” has been used to describe a state along this continuum that falls in-between “delirium” and “no delirium” (570), but it is not established if a patient with subsyndromal delirium should be considered a case or a control. Furthermore, delirium symptoms by definition fluctuate, and accordingly may be missed.
Our study tried to overcome these challenges in diagnosing delirium in several ways. We obtained information from persons who knew the patients, reducing the likelihood of falsely diagnosing a patient with cognitive symptoms due to dementia with delirium. Review of hospital records was also performed to identify delirium symptoms that were not present during the daily delirium assessments, e.g. only during the night. A limitation was, however, that delirium was not assessed regularly during the weekends, and therefore some delirium cases may erroneously have been classified as a control. Yet, we believe that well-trained personnel interviewing hospital staff and reviewing of case notes from the weekend compensated this in part. All patients with available CSF were included in paper III in order to obtain the highest possible statistical power. Exclusion of patients with subsyndromal delirium could have been done, but this would have reduced our sample size substantially. Hence, subsyndromal delirium was included and classified as “no delirium” in paper III.

**Dementia.** Ideally, the diagnosis of dementia is based on an assessment performed before the patient is hospitalized, but this was not feasible in acutely admitted hip fracture patients. It is also known that a large proportion of dementia cases are not recognized before hospitalization (441), so classification of dementia based on an existing pre-fracture dementia diagnosis would not be suitable. We therefore believe that our consensus diagnosis of dementia was the best possible approach in our cohort. The very good agreement between the two raters of dementia, is also a strength of our study.

**Cognitively normal.** As deliberated on in section 1.2.1, there is no consensus for the definition of “cognitively normal”, and definitions differ between studies. Ideally, a consensus diagnosis based on all available information should have been used to determine whether study participants were cognitively normal at each time point of the follow up (232). Consensus diagnoses are unfortunately not yet available in the elective surgery cohort due to lack of resources. We have therefore selected cognitively normal older adults based on results on cognitive tests. Our goal was to study brain changes in cognitively normal older adults only, and we therefore considered individuals referred to further cognitive testing at baseline or later, and individuals who were diagnosed with cognitive impairment during follow up, as not cognitively normal. This increased the likelihood of studying brain changes common in aging, but rendered us unable to accurately study the predictive value of the CSF biomarkers in preclinical AD. It would also have been difficult to include participants referred to cognitive testing and participants who developed MCI and dementia because they were regrettably, by
study design, excluded from further follow-up. This was due to the fact that the initial purpose of the study was to obtain a cognitively normal cohort.

The definition of “cognitively normal” differed between papers I and II. Criteria in paper I were less strict than in paper II, e.g. for individuals with MMSE score < 27, we allowed one abnormal test score out of the remaining 10 test scores in paper I, whereas individuals with MMSE score < 28 had to have normal scores on all other tests in order to be considered cognitively normal in paper II. Further, in paper I, all participants with MMSE score ≥ 27 were considered cognitively normal, whereas several normal test scores were necessary in paper II. Because MCI can be diagnosed based on only one abnormal test score, it can be argued that some of the individuals in our sample may have MCI. Moreover, a definition of cognitive normality based on MMSE alone may overlook individuals with subtle cognitive impairments. We compensated for this in paper I by reproducing results in a subgroup with stable memory function for two years and in paper II by stricter criteria for the definition of “cognitively normal”. Another difference is that test scores from the last available testing was used for the definition of “cognitively normal” in paper I, whereas in paper II we used only baseline test scores. The rationale for this was that we wanted to assess memory change over time in paper II, and therefore a definition based on the last available cognitive assessment could have affected this outcome in an unpredictable way. It may be argued that some of the individuals in paper II may have developed cognitive impairment during follow up, but we believe that our exclusion of participants who were diagnosed with cognitive impairment during the 6 year follow up largely prevented this. It must also be noted that different norms were used to define abnormality on different cognitive test scores in papers I and II. The rationale for this was that we used the same norms as the Memory Clinic at Oslo University Hospital, and the preferred norms of the clinic were changed in the time between writing papers I and II, e.g. the new norms for the CERAD neuropsychological battery published during this period were included (510).

The selection of several other subgroups with low risk of AD in paper I was a novel approach, strengthening the suggestion the association between NFL and hippocampal atrophy was independent of AD.

5.2.3 Choice of CSF analyses and progression measures

CSF measures. Analyses of CSF Aβ42, T-tau, and P-tau was pre-planned in the study in the elective surgery cohort. Analyses of the core AD biomarkers in the hip fracture cohort, and
analyses of the novel biomarkers NFL, YKL-40, and FABP3 in the elective surgery cohort, were performed after recommendations from our collaborators at the Clinical Neurochemistry Laboratory at Sahlgrenska University Hospital. Analyses of CSF samples in this expert laboratory and that all samples were analyzed in one round represents strengths of our studies. We used different cut-off values for Aβ42 in paper I compared to II and III, however the lowest (i.e. less strict) cut-off was suggested by our laboratory, and not many cognitively normal subjects fell within the range between the different cut-offs as the distribution of CSF Aβ42 was bimodal.

**Hippocampal volume change.** Longitudinal change of hippocampal volumes was chosen as an outcome variable in papers I and II. The rationale for this choice was that we wanted to study the relationship between AD-associated biomarkers and brain changes known to be present both in aging and AD, and although hippocampal atrophy is a hallmark for AD, it is also one of the brain areas showing the largest decline in aging and in cognitively normal adults at very low risk of AD (213, 264, 270). Longitudinal measurement of brain volumes represents a strength of our study. Another strength is that the same MRI scanner and T1 weighted sequence was used at baseline and follow up.

**Memory change.** Longitudinal change of memory was chosen as an outcome variable in paper II. The rationale for this choice was that we wanted to study the relationship between AD-associated biomarkers and cognitive changes known to be present both in aging and AD. Similar to hippocampal atrophy, episodic memory decline is a cardinal symptom of AD (5, 17), but it is also the form of memory that shows the largest degree of decline in normal aging (7, 8, 571). We chose to use the Word List Recall score as our memory measure because delayed recall tests may be most sensitive to memory decline (572), and this was the only delayed recall test available in the elective surgery cohort. The Word List Memory Task is part of the CERAD neuropsychological battery which is widely used and well-validated (573). However, a more challenging episodic memory test, e.g. the California Verbal Learning Test-II (574), would possibly be more sensitive to slight memory changes in our cohort of cognitively normal older adults. However, since the initial purpose in of this study was to obtain a cognitively normal control cohort, we used the same tests as the memory clinic of our hospital. Construction of a composite memory score using test scores from all Word List Memory Task items and the Kendrick Object Learning Test may also be more sensitive to memory change (575, 576), and we plan to explore this approach in future work. Nevertheless, since slight memory changes in
cognitively normal older adults may only be detected by repeated measurements, the longitudinal follow up with up to seven assessments represents a major strength of our study.

### 5.2.4 Statistical issues

Due to the explorative designs and convenience samples, power calculations for the relationships studied in this thesis were not performed before study enrolment. Consequently, there is a possibility that lack of associations are due to type II errors.

Subgrouping of hip fracture patients according to delirium status and dementia status resulted in some small subgroups. For example, only 10 hip fracture patients with dementia did not develop delirium. Differences in biomarker concentrations between patients with and without delirium were also smaller in patients with dementia due to more pathological biomarker concentrations. Therefore, it is possible that larger studies will be able to detect a difference also in patients without dementia. Due to a limited statistical power, we could only adjust for some confounding factors, and it is possible that adjustment for other potential confounding factors, e.g. APACHE II, would have affected our results. Yet, the hip fracture cohort is among the largest cohorts on CSF biomarkers in delirium (449), and the sample size enabled us to study the association between CSF biomarkers and delirium separately in patients with and without dementia.

The sample size in the papers on cognitively normal older adults is similar to many other studies, but smaller than in the ADNI and some aging studies. I have speculated that the lack of difference in hippocampal atrophy between biomarker groups in paper II may be due to lack of power. However, it must be noted that the long follow up interval and the many longitudinal tests yield an increase in power that is much larger than a similar increase in the number of participants would have done. Furthermore, a major advantage of GAMM in our setting is that the power could be increased by taking advantage of all longitudinal and cross-sectional observations.

For paper II, we discussed how to report differences in biomarkers concentrations between clusters with a statistician (Ø. Sørensen). Since the clusters were already made based on differences in biomarker levels, tests of significance would not be valid. We therefore chose to report the relative contributions of the different biomarkers to the clustering using Cohen’s D.
5.3 Ethical considerations

Informed consent is a challenge in studies of patients with cognitive impairments, including delirium, as many of these individuals are incapable of understanding what the consent entails. However, a strict requirement of informed consent would make it impossible to improve our knowledge about diseases prevalent in cognitively impaired individuals. Consequently, the Norwegian Health Research Act states that “competence to give consent may cease to apply wholly or partly if the patient, on account of a physical or mental disorder, senile dementia or mental retardation, is clearly incapable of understanding what the consent entails.” (577).

Further, it is stated that the study may only be done if “a) potential risks or disadvantages for the person are insignificant, b) the individual involved should is not averse to it, and c) there is reason to assume that the results of the research may be of use to the person concerned or other people with the same age-specific disorder, disease, injury or condition”. The ethical committee considered sampling of CSF in conjunction to spinal anesthesia to be in conformity with these requirements. I do, however, believe that CSF sampling by lumbar puncture may be too invasive to be in agreement with the law. Furthermore, repeated CSF sampling in delirium would likely be adverse for the patient, e.g. contribute to maintenance of delirium, and repeated biomarker measures is therefore only ethical and feasible for biomarkers that can be obtained by minimally invasive procedures, e.g. blood sampling.

Assessment of $APOE$ genotype and CSF concentrations of Aβ42, T-tau and P-tau, and screening with brain MRIs, in presumably cognitively normal individuals introduces ethical problems. The odds ratio for AD dementia is ~12 in $APOEe4$ homozygotes and ~3 in $APOEe4$ heterozygotes compared to non-E4 carriers (578). Disclosure of an individual’s $APOEe4$ status could therefore possibly cause emotional difficulties if the person was $APOEe4$ positive, however a study testing this hypothesis found only mild and brief psychological problems in $APOEe4$ positive individuals (579). Nevertheless, when planning the study, we decided that the study participants should not be informed about their $APOE$ status, but if they requested this information we would refer them to genetic counselling. Furthermore, although the positive predictive value of abnormal core AD biomarkers for later dementia in cognitively normal adults is not established, disclosure of biomarker status introduces similar ethical issues (580). We have, therefore, chosen not to inform study participants about their biomarker status unless they have asked. Finally, MRI scans have revealed tumors, cerebral infarctions, chronic vascular changes, and brain atrophy in several study participants. We have chosen to only
inform participants about findings if may have a clinical consequence, and participants have been referred for further evaluations if necessary. Similarly, participants with poor scores on cognitive tests have been referred to further cognitive testing in our hospital.
6 Conclusion and Implications

This work has provided information that increases our knowledge about the role of different neuropathological processes in cognitively normal older adults and delirium.

The novel biomarkers NFL, YKL-40, and FABP3 were linked to T-tau and P-tau, but not to Aβ42, in cognitively normal adults. These findings suggest that the neuropathological processes neuronal damage, axonal damage, and neuroinflammation are accompanying tauopathy and neurodegeneration, and not Aβ deposition.

Concentrations of NFL, YKL-40, and FABP3 differed between at least two clustering-based groups with a relatively large effect size, suggesting that the three novel biomarkers contribute to subgrouping of cognitively normal older adults.

High NFL levels predicted higher hippocampal atrophy rate in cognitively normal older adults, also when adjusting for AD pathology and in subgroups unlikely to have preclinical AD. This finding suggests a role of NFL as a biomarker of AD-independent neurodegeneration and further suggested that common neurodegenerative processes are ongoing in aging and disease. Moreover, the clustering-based subgroup characterized by highest levels of T-tau, P-tau, and FABP3 showed more memory decline, suggesting that a high degree of neurodegeneration and tauopathy is associated with greater memory decline in cognitively normal older adults. Conversely, the biomarker groups based on only Aβ42 and P-tau showed no differences in memory trajectories. Thus, the clustering-based classification of the participants improved prediction of memory decline compared to the canonical AT-classification. This finding further suggests that it is meaningful to use the three novel biomarkers for characterization of brain states in cognitively normal older adults, and that clustering studies may identify biomarker profiles relevant for clinical trials. Also, several subgroups with abnormal biomarker levels did not show more than age-expected hippocampal volume and memory decline, suggesting that older adults may uphold age-expected cognitive function and hippocampal integrity even when harboring brain pathologies.

In patients without dementia, the core AD biomarkers (low CSF Aβ42, high CSF T-tau, low Aβ42/T-tau, low Aβ42/P-tau) were significantly associated with delirium, and the proportion of patients with delirium was significantly higher in A+T+ individuals than those with only one or no abnormal biomarkers. Contrary, there was no association between core AD biomarkers
and delirium in patients with dementia. These findings suggest that AD pathologies may underlie the interrelationship between delirium and dementia. Furthermore, the results indicate that AD neuropathology has clinical consequences (i.e. delirium) in patients without dementia, and raise the question of whether delirium is an early symptom of AD.

In summary, we have shown that CSF biomarkers previously associated with AD are also associated with the clinical outcomes delirium, hippocampal atrophy, and memory decline in individuals without dementia. Although much more research is needed, the understanding of neuropathological processes in delirium and cognitively normal older adults with and without core AD pathology may eventually have implications for the development of drugs that can prevent and/or treat delirium, AD, and other neurodegenerative diseases by acting upon specific neuropathological processes involved.
7 Future perspectives

This work has shown that novel AD-associated CSF biomarkers can predict brain and cognitive changes in cognitively normal older adults. The role of these biomarkers should therefore be assessed further in larger studies, using also other measures of brain and cognitive changes, and over longer periods of follow-up. One interesting approach would be to assess whether the relationships of novel biomarkers to brain and cognitive changes differs between A+ and A-, or between A+T+ and A-T-, cognitively normal individuals. Future studies should also assess whether similar relationships between biomarkers and outcomes are present in patients with MCI and neurodegenerative diseases.

Further, the role of other novel biomarkers like neurogranin, tau PET, and blood biomarkers should be assessed, and in order to assess the individual role of the neuropathological process reflected by the biomarker, analyses must be adjusted for levels of biomarkers reflecting other neuropathological processes. Post-mortem studies will also be important to determine the individual and combined added values of different biomarkers. Envisioning that validated assays for biomarkers reflecting many other neuropathological processes common in aging and age-related neurodegenerative disorders will become available, we will be able to more accurately classify cognitively normal adults according to their brain states, and further improve our understanding of which neuropathological processes that have clinical consequences.

Larger studies can be accomplished by merging different cohorts of cognitively normal adults, but large multicenter studies using the same study protocol would be a better design. Population based cohorts with follow-up across the life-span including individuals irrespective of their cognitive status, would produce the most generalizable results, however such studies requires large resources. Additionally, international efforts are needed to determine a uniform definition of “cognitively normal”.

In our elective surgery cohort, we are planning to continue annual follow-ups, assess the role of serum NFL and CSF neurogranin, assess longitudinal biomarker changes in blood and CSF, assess other brain measures like cortical thickness, different subcortical volumes, and functional MRI measures, and use other cognitive outcome measures like a composite memory score. The cohort is currently also included in larger international cohorts of cognitively normal individuals (581, 582).
This work has suggested that AD neuropathologies may underlie the interrelationship between delirium and dementia. Future work should assess this in larger and different populations, and also using other biomarkers of dementia pathology, e.g. imaging biomarkers and other fluid biomarkers associated with neurodegenerative diseases. Especially, validated blood biomarkers should be assessed, e.g. NFL and also amyloid biomarkers measured using ultrasensitive measurement techniques, and repeated measurements before, during, and after delirium would distinguish risk biomarkers from biomarkers fluctuating over the course of delirium. As for studies of cognitively normal older adults, the individual contribution of neuropathological processes should be assessed by adjusting for other important biomarkers. Furthermore, in addition to assessing whether the association between biomarkers and delirium differs between patients with and without delirium, it will also be interesting to assess if it differs between patients with and without biomarker signs of AD neuropathy.

Large-scale multicenter cohorts including different patient populations are needed to assure generalizability of the findings. Longitudinal studies are needed to determine whether delirium itself causes brain changes leading to cognitive decline. Such longitudinal studies should study populations where a comprehensive cognitive assessment and biomarker assessment can be performed before they experience a potentially delirium triggering insult. Next, delirium occurrence and severity should be detected, ideally for all delirium episodes occurring during the study period, and the study participants should be followed with repeated cognitive assessments for a long period of time. Analyses should assess whether delirium, biomarkers of different neuropathologies, or a combination of the two is associated with cognitive decline. Further and detailed recommendations for future research on the relationship between delirium and AD has also recently been proposed by Fong et al (474).

In the hip fracture cohort, we have already assessed the relationship between NFL and delirium (477), and we are planning to also assess whether YKL-40, FABP3, and neurogranin are associated with delirium.
References

75. Castellani RJ, Smith MA. Compounding artefacts with uncertainty, and an amyloid cascade hypothesis that is 'too big to fail'. J Pathol. 2011;224(2):147-52.


108


112


Wong CL, Holroyd-Leduc J, Simel DL, Straus SE. Does this patient have delirium?: value of bedside instruments. JAMA. 2010;304(7):779-86.


Papers I-III
CSF neurofilament light levels predict hippocampal atrophy in cognitively healthy older adults

Ane-Victoria Idland\textsuperscript{a,b,*}, Roser Sala-Llonch\textsuperscript{b}, Tom Borza\textsuperscript{c}, Leiv Otto Watne\textsuperscript{a,d}, Torgeir Bruun Wyller\textsuperscript{a,c}, Anne Brækhuse\textsuperscript{e,f,g}, Henrik Zetterberg\textsuperscript{h,i}, Kaj Blennow\textsuperscript{h}, Kristine Beate Walhovd\textsuperscript{b}, Anders Martin Fjell\textsuperscript{b}

\textsuperscript{a} Oslo Delirium Research Group, Department of Geriatric Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway
\textsuperscript{b} Research Group for Lifespan Changes in Brain and Cognition, Department of Psychology, University of Oslo, Oslo, Norway
\textsuperscript{c} Centre for Old Age Psychiatric Research, Inlandet Hospital Trust, Hamar, Norway
\textsuperscript{d} Department of Basic Medical Sciences, University of Oslo, Oslo, Norway
\textsuperscript{e} Department of Geriatric Medicine, Oslo University Hospital, Oslo, Norway
\textsuperscript{f} Norwegian National Advisory Unit on Ageing and Health, Vestfold Hospital Trust, Tønsberg, Norway
\textsuperscript{g} Department of Neurology, Oslo University Hospital, Oslo, Norway
\textsuperscript{h} Department of Psychiatry and Neurochemistry, Clinical Neurochemistry Lab, Institute of Neuroscience and Physiology, The Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden
\textsuperscript{i} Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK

A R T I C L E   I N F O

Article history:
Received 24 May 2016
Received in revised form 3 September 2016
Accepted 21 September 2016
Available online 28 September 2016

Keywords:
Cerebrospinal fluid
Magnetic resonance imaging
Normal aging
Neurofilament light
Hippocampal atrophy rate

A B S T R A C T

Cerebrospinal fluid (CSF) neurofilament light (NFL) is a marker of axonal degeneration. We tested whether CSF NFL levels predict hippocampal atrophy rate in cognitively healthy older adults independently of the established CSF Alzheimer’s disease (AD) biomarkers, β-amyloid (A\textsubscript{β} 1–42), and phosphorylated tau (P-tau). We included 144 participants in a 2-year longitudinal study with baseline CSF measures and 2 magnetic resonance images. Eighty-eight participants had full data available. A subgroup of 36 participants with very low AD risk was also studied. NFL predicted hippocampal atrophy rate independently of age, β-amyloid 1–42, and P-tau. Including NFL, P-tau, and age in the same model, higher NFL and lower P-tau predicted higher hippocampal atrophy (\(R^2 = 0.20\), NFL: \(\beta = -0.34; p = 0.003\); P-tau: \(\beta = 0.27; p = 0.009\)). The results were upheld in the participants with very low AD risk. NFL predicted neurodegeneration in older adults with very low AD probability. We suggest that factors previously shown to be important for brain degeneration in mild cognitive impairment may also impact changes in normal aging, demonstrating that NFL is likely to indicate AD-independent, age-expected neurodegeneration.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Hippocampal atrophy rates are higher in patients with Alzheimer’s disease (AD) than in cognitively healthy older adults (Barnes et al., 2009). However, hippocampal atrophy is known to accelerate from midlife onward also in persons with low AD risk (Fjell et al., 2013), and hippocampus is one of the brain areas with highest atrophy rate in aging (Fjell et al., 2013), reported to be around 1% annually (Fjell et al., 2013; Fraser et al., 2015). Thus, identification of biomarkers predicting hippocampal atrophy is critical for understanding brain changes both in normal aging and early AD. Interestingly, a recent study showed that cerebrospinal fluid (CSF) neurofilament light (NFL) subunit levels predicted hippocampal atrophy in mild cognitive impairment (MCI) patients (Zetterberg et al., 2016), indicating that CSF NFL could be a progression marker in AD.

Neurofilaments are important cytoskeletal components of neuronal axons, and CSF NFL levels are believed to reflect axonal degeneration (Petzold, 2005; Zetterberg et al., 2006). CSF NFL levels are associated with age (Rosengren et al., 1996; Skillback et al., 2014; Vagberg et al., 2015), white-matter (WM) lesions (Sjogren et al., 2001), AD (Olsson et al., 2016; Petzold et al., 2007; Skillback et al., 2014; Zetterberg et al., 2016), and other neurodegenerative diseases (Backstrom et al., 2015; Petzold et al., 2007; Skillback et al., 2014; Steinacker et al., 2016; Teunissen et al., 2005). Previous studies have indicated a relationship between high CSF NFL and...
lower brain volume in frontotemporal lobe dementia (Scherling et al., 2014), and nondenominated older adults (Bjerke et al., 2014; Vagberg et al., 2015), but studies are mostly cross-sectional and results have not been consistent (Bendlin et al., 2012; Khalil et al., 2013). The relationship between CSF NFL levels and hippocampal atrophy in cognitively healthy older adults has never been tested but is critical for understanding whether NFL is a general or disease-specific atrophy marker. Thus, the objective of this study was to test whether CSF NFL levels predict hippocampal atrophy rate in cognitively healthy older adults independently of the established CSF AD biomarkers β-amyloid 1–42 (Aβ 42) and phosphorylated tau (P-tau) (Blenow et al., 2010).

2. Methods

2.1. Participants

We recruited patients scheduled for elective gynecological (genital prolapse), urological (benign prostate hyperplasia, prostate cancer, or bladder tumor/cancer), or orthopedic (knee or hip replacement) surgery in spinal anesthesia turning 65 years or older the year of inclusion. Dementia, previous stroke with sequelia, Parkinson’s disease, and other neurodegenerative diseases likely to affect cognition were exclusion criteria. Participants were assessed with a multidomain battery of cognitive tests before surgery, comprising the Mini–Mental Status Examination (MMSE) (Folstein et al., 1975), Clock Drawing Test (Shulman, 2000), Kendrick Object Learning Test (Kendrick et al., 1979), and verbal fluency (The Controlled Word Association Test, with the letters F, A and S, and Animal Naming) (Spreen and Strauss, 1991), giving 11 test scores. Blood and CSF samples were collected by the anesthesiologist in conjunction with spinal anesthesia, and participants underwent magnetic resonance imaging (MRI) after surgery. The mean time between CSF sampling and MRI at baseline was 8 weeks (standard deviation [SD] range: 6 [2–20 to 24]). Participants underwent a second MRI and were tested with the same battery of cognitive tests at 2-year follow-up (mean time between MRIs 2.2 years, SD range: 0.3 [1.6–2.9]; see Table 1).

We selected participants as shown in Supplementary Fig. 1. CSF was available from baseline only, while the majority of the participants had 2 MRI scans. Only participants with CSF data and/or brain MRI(s) were included. We selected only cognitively healthy participants based on the following procedure: First, participants offered referral to cognitive assessment were excluded. Next, we included all participants with MMSE score ≥27. Finally, for participants with MMSE score <27, only those with none or one other abnormal test score(s) when last tested were included. Abnormal score was defined as more than 1.5 SD below the mean normal value for age, sex, and educational level. Four participants with CSF NFL levels >4000 pg/mL (i.e., more than ±3 SD from the mean value) were excluded. This resulted in 144 participants with CSF analyses or MRI at baseline (sample A) and 88 participants with CSF NFL analyses and MRI at both time points (sample B). After further screening of sample B, some participants with additional conditions (details Supplementary Table 1) were excluded, resulting in sample C. From samples B and C, we created subgroups with very low risk of AD by excluding participants in a hierarchical manner: (1) no apolipoprotein E (APOE) 4 alleles (samples D and H); (2) also Aβ 42 >550 pg/mL (Mulder et al., 2010) and stable or improved delayed recall score on Word List Memory Task at 2-year follow-up compared to baseline (samples E and I); and (3) (i) also P-tau <60 pg/mL (samples F and J) and (ii) also Aβ 42 >650 pg/mL (samples G and K). It is possible to effectively define low-risk group based on APOE status only, but we created additional low-risk groups to further reduce AD risk by including Aβ 42 and memory function as further criteria. Several cutoff values for CSF Aβ 42 levels are described in the literature, ranging from 500 to 650 pg/mL (Fagan et al., 2009; Mulder et al., 2010; Niemantsverdriet et al., 2016; Zwan et al., 2016). We used 550 pg/mL as our cutoff, and in addition, we increased the cutoff to 650 pg/mL for 1 subgroup to be more conservative.

2.2. Ethical considerations

The study was conducted in accordance with the Declaration of Helsinki and approved by the Regional Committee for Ethics in Medical Research in Norway (REK, 2011/2052). All participants provided written informed consent.

2.3. MRI acquisition and processing

T1-weighted MPRAGE 3D images were acquired with a 1.5 T Siemens Avanto scanner using a 12-channel head coil (repetition time = 2400 ms, echo time = 3.79 ms, field of view = 240 mm, slice thickness = 1.2 mm, pixel size = 1.25 × 1.25 mm). Images were processed with FreeSurfer (version 5.3) and its specific longitudinal stream (https://surfer.nmr.mgh.harvard.edu). For each MRI, the FreeSurfer pipeline performs a set of automated procedures for the cortical reconstruction and volumetric segmentation, documented elsewhere (Dale et al., 1999; Fischl et al., 2002). We used hippocampal volume measures and WM hypointensities estimations obtained from the automated segmentation. More specifically, the FS-segmentation algorithm assigns labels to all the brain

---

**Table 1 Demographics, CSF biomarkers, and hippocampal measures**

<table>
<thead>
<tr>
<th>Variable</th>
<th>All participants (sample A, n = 144)</th>
<th>Participants with MRI at both time points and CSF NFL analyses (sample B, n = 88)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at baseline, y</td>
<td>73 (6), 64–91</td>
<td>73 (6), 64–89</td>
</tr>
<tr>
<td>Sex, male</td>
<td>68 (47)</td>
<td>43 (49)</td>
</tr>
<tr>
<td>Education, y</td>
<td>14 (4), 7–23</td>
<td>15 (3), 8–23</td>
</tr>
<tr>
<td>Hypertension</td>
<td>61 (42)</td>
<td>29 (33)</td>
</tr>
<tr>
<td>MMSE score, baseline</td>
<td>29 (12), 25–30</td>
<td>29 (13), 25–30</td>
</tr>
<tr>
<td>MMSE score, 2-y follow-up</td>
<td>29 (14), 21–30³</td>
<td>29 (12), 24–30</td>
</tr>
<tr>
<td>APOE genotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3/E2</td>
<td>12 (9)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>E3/E3</td>
<td>68 (53)</td>
<td>44 (53)</td>
</tr>
<tr>
<td>E4/E2</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>E4/E3</td>
<td>44 (34)</td>
<td>31 (37)</td>
</tr>
<tr>
<td>E4/E4</td>
<td>4 (3)</td>
<td>3 (4)</td>
</tr>
<tr>
<td>CSF Aβ 42, pg/mL</td>
<td>718 (208), 275–1179⁴</td>
<td>724 (203), 275–1175</td>
</tr>
<tr>
<td>CSF P-tau, pg/mL</td>
<td>60 (20), 25–115⁴</td>
<td>61 (19), 26–110</td>
</tr>
<tr>
<td>CSF NFL pg/mL</td>
<td>1163 (507), 487–312³</td>
<td>1141 (558), 510–312</td>
</tr>
<tr>
<td>Aβ 42 &lt;(-550 pg/mL)</td>
<td>34 (26)⁵</td>
<td>24 (27)</td>
</tr>
<tr>
<td>Months between MRIs</td>
<td>26 (3), 19–35</td>
<td></td>
</tr>
<tr>
<td>Hippocampal volume,</td>
<td>3505 (396), 2373–4544</td>
<td></td>
</tr>
<tr>
<td>baseline, mm³</td>
<td>3464 (407), 2425–4514</td>
<td></td>
</tr>
<tr>
<td>2-y follow-up, mm³</td>
<td>2.64 (5.3), 2.00–14.4</td>
<td></td>
</tr>
<tr>
<td>Hippocampal volume,</td>
<td>0.55 (1.08), 0.00–2.14</td>
<td></td>
</tr>
</tbody>
</table>

Values are n (%) and mean (SD), range. Key: Aβ 42, β-amyloid 1–42; APOE, apolipoprotein E; CSF, cerebrospinal fluid; MMSE, Mini–Mental Status Examination; MRI, magnetic resonance imaging; NFL, neuromyelitis optica; P-tau, phosphorylated tau.

³ Based on information from the participant and patient records.

⁴ n = 115.
⁵ n = 129 and n = 83, respectively.
⁶ n = 130.
⁷ n = 128.
regions of each individual scan, based on an available probabilistic atlas obtained from a training set of subjects which has been accurately manually labeled (Fischl et al., 2002). Both the hippocampal volume and the WM hypointensities are defined from this available atlas. Hippocampal volume was not normalized by estimated intracranial volume because the main analyses were done on rate of atrophy, where normalizations are not recommended. WM hypointensities appear as dark WM on the T1-weighted image and are obtained from the overall sum of regions within the WM with T1-intensity values within a certain range defined from the probabilistic atlas. This measure is related to WM lesions but is considered less sensitive than WM hyperintensities based on T2 or FLAIR images. The FreeSurfer longitudinal stream includes methods designed to minimize the bias to any time point in a participant and which lead to increased statistical power, better separation of groups based on atrophy, and higher reproducibility. These include the generation of a procedure-specific intermediate template followed by a principal norm of each time point to this template (Jovicich et al., 2013; Reuter et al., 2012). For both the individual and longitudinal processing steps, reconstructed surfaces and volumes were visually inspected and manually corrected when necessary.

2.4. APOE genotyping

Blood samples were genotyped for APOE (gene map locus 19q13.2) using TaqMan Allelic Discrimination technology (Applied Biosystems, Carlsbad, CA, USA). Genotypes were obtained for the 2 SNPs that are used to unambiguously define the ε2, ε3, and ε4 alleles (rs7412 and rs429358).

2.5. CSF collection and analyses

CSF was collected in polypropylene tubes, centrifuged at room temperature for 10 minutes, the supernatant aliquoted into polypropylene tubes, and frozen at –80 °C pending analyses. Mean time from CSF sampling to freezing was 83 minutes (SD [range]: 21 [30–129]; median) from CSF sampling to freezing was 83 minutes. Analyses were performed by board-certified laboratory technicians masked to clinical data. Intra-assay coefficients of variation were 9%–13% and the lower limit of detection for NFL was 50 pg/ml. The ELISA method for CSF Aβ 42 has been fully validated analytically (Vanderstichele et al., 2000) and also validated against the Joint Committee for Traceability in Laboratory Medicine approved mass spectrometry Reference Measurement Procedure for CSF Aβ 42 (Leinenbach et al., 2014) and show high consistency in results over time and between batches when adhering to strict laboratory analytical procedures (Palmqvist et al., 2014). CSF T-tau and P-tau levels were strongly correlated (r = 0.96, p < 0.001); thus, we only used CSF P-tau in the main statistical analyses.

2.6. Statistical analysis

We calculated hippocampal atrophy rate as the annual percent change in hippocampal volume (average of both hemispheres), normalized by the average volume across time points and divided by years between scans. We also calculated the average WM-hypointensities volume across time points for use as a control variable. We tested associations between CSF biomarkers, age, and hippocampal atrophy rate using SPSS (version 22). Generalized additive mixed models (GAMMs) implemented in R (www.r-project.org) using the package ‘mgcv’ (Wood, 2006) was used to derive the age function for hippocampal atrophy and for the relationship between hippocampal change and CSF NFL levels, taking advantage of all longitudinal and cross-sectional observations, run through the PING data portal (http://pingstudys.ucsd.edu/welcome.html) (Bartsch et al., 2014). Akaike information criterion (AIC) (Akaike, 1974) was used to guide model selection and help guard against overfitting. For analyses including CSF biomarkers and not MRI measures, we used age at the day of CSF sampling. For analyses including MRI measures, we used the age at the day of baseline MRI. Significance was set at p < 0.05.

We tested correlations between CSF biomarkers, age and hippocampal atrophy rate using Pearson correlations, and hippocampal volume change using paired samples t-test. We performed multiple linear regression analyses to test associations between age, CSF biomarkers, and hippocampal atrophy rate. Regressions were performed in several steps. The first model included NFL and age as predictors of hippocampal atrophy rate. Next, we tested the predictive power of Aβ 42 and P-tau levels separately in conjunction with NFL in the model. The resulting regression model was tested for stability by including sex and WM hypointensities separately. All analyses were done in the main sample (sample B). The most important analyses were also repeated in sample C (details Supplementary Table 2). We also tested the final regression model within the very low AD risk subgroups (samples D–K). Sensitivity analyses were performed with and without outliers (defined as studentized deleted residuals > ±2) for all regression models. Although we expected correlations between the explanatory variables, we chose not to use data reduction methods, such as principal component or cluster analysis, to be able to evaluate the contributions from each biomarker separately. Finally, we ran mediation analyses in sample B using the SPSS macro INDIRECT (Preacher and Hayes, 2008). Mediation is present if the relationship between the predictor variable and the dependent variable (c) attenuates when accounting for a third variable (the mediator) (c’). The % reduction was calculated as (c – c’)/c. The significance of the indirect effect (a x b) was tested using bootstrapped confidence intervals. Standardized coefficients were obtained using z-scores.

3. Results

3.1. CSF biomarkers, hippocampal volume, and demographic factors

Demographics and CSF biomarker and MRI characteristics are shown in Table 1 and Supplementary Table 3. NFL levels correlated positively with age (r = 0.45, p < 0.001), while P-tau (r = 0.09, p = 0.41) and Aβ 42 levels (r = 0.05, p = 0.67) did not correlate with age. NFL levels correlated positively with P-tau levels (r = 0.23, p = 0.03), but not with Aβ 42 levels (r = 0.07, p = 0.53).

3.2. CSF NFL levels and hippocampal atrophy rate

CSF NFL levels were negatively correlated with baseline hippocampal volume, averaged bilaterally, (r = −0.25, p = 0.02). Hippocampal atrophy, measured as the difference in hippocampal volumes between baseline and follow-up, was significantly different from zero (mean [SD], range: −40.49 mm³ [74.71], −24.66 to −56.32, t = 5.09, p < 0.001), and the mean annual atrophy rate was −0.55%. Age correlated with higher atrophy rate (r = −0.26, p = 0.01), indicating accelerated atrophy with increasing age. This relationship was confirmed with GAMM for the full sample, as illustrated in Fig. 1. We ran a multiple regression analysis using NFL
and age as predictors of hippocampal atrophy rate. Higher NFL levels predicted higher hippocampal atrophy rate (p = 0.02; see Table 2). Age was not a significant predictor in this model. Next, Aβ 42 level was also introduced as a possible predictor and did not predict hippocampal atrophy rate independently of NFL, while NFL was still significant (Table 2). The last step included P-tau as a predictor together with age and NFL, and we obtained a model with higher NFL levels and lower P-tau levels predicting higher hippocampal atrophy rate (Table 2) independently of age. Regression analyses results were unchanged when excluding 5–6 outliers per analysis. Substitution of P-tau with T-tau in this last step gave the same results (Supplementary Table 4), while only NFL was a significant predictor when P-tau was substituted with Aβ 42/P-tau ratio (data not shown).

The relationship between NFL and hippocampal volume was also tested with GAMM to take advantage of all data points, obtaining an optimal fit based on both cross-sectional and longitudinal information. The sample was divided into NFL+ and NFL− by a median split, and the relationship between hippocampal volume and age was plotted in each group, with sex as a covariate. AIC for the model was 2587 and NFL status yielded a highly significant contribution (t = −2.96, p < 0.005). Removing NFL increased AIC to 2595, indicating a worse fit. Adding P-tau as a covariate did not improve the model fit (AIC = 2588), and P-tau did not contribute significantly (t = −0.18, p = 0.85) while NFL still did (t = −2.66, p < 0.01). Thus, the initial model was preferred and plotted in Fig. 2.

3.3. Adjusting for effect of WM hypointensities and sex on hippocampal atrophy rate

Because vascular brain pathology may affect the relationship between NFL and hippocampal atrophy (Sjögren et al., 2001), we entered WM hypointensities into the regression model including age, NFL, and P-tau levels as predictors of hippocampal atrophy rate. NFL and P-tau levels were still significant predictors of hippocampal atrophy rate, whereas WM hypointensities were not predictive (data not shown). Results were unchanged after exclusion of 6 outliers. We adjusted for sex in the same way as for WM hypointensities, and NFL and P-tau levels were the only significant predictors of hippocampal atrophy rate (data not shown). Sex was not a significant predictor; however after exclusion of 6 outliers, sex was also a significant predictor (higher atrophy rates in males).

3.4. CSF NFL levels and hippocampal atrophy in low-risk subgroups

The most important analyses were repeated in sample C (exclusions after further screening of sample B). NFL did not correlate with age in this sample (r = −0.18, p = 0.13); however, when excluding 4 statistical outliers (studentized deleted residuals > ±2), the correlation was significant (r = −0.36, p = 0.002). All other correlation results were unchanged from sample B, and hippocampal volume change was significantly different from zero. In linear regression analyses, NFL was the only significant predictor (borderline significant in analyses adjusted for sex; Supplementary Table 5); however, after excluding 4–5 outliers per analysis, all results were unchanged from sample B, except that Aβ 42 was also a significant predictor of hippocampal atrophy rate (Supplementary Table 6). We further applied our final regression model including age, NFL, and P-tau levels as predictors of hippocampal atrophy rate in the very low AD risk subgroups from sample B (samples D–G). In the first subgroup, participants without APOE4 alleles, higher NFL levels and lower P-tau levels predicted higher hippocampal atrophy rate independent of age as in the full sample (Supplementary Table 7). The

![Fig. 1. Relationship between age and hippocampal volume. Adjusted for sex. The graph shows mean slope with 95% confidence interval. Data points from participants with MRI are displayed, including within-person changes for those with MRI at both time points. Abbreviation: MRI, magnetic resonance imaging.](Image)

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Multiple linear regression with hippocampal atrophy rate as dependent variable (full sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent variables</td>
<td>R²</td>
</tr>
<tr>
<td>Age</td>
<td>0.13</td>
</tr>
<tr>
<td>NFL</td>
<td>−0.001</td>
</tr>
<tr>
<td>Age</td>
<td>0.16</td>
</tr>
<tr>
<td>NFL</td>
<td>−0.001</td>
</tr>
<tr>
<td>Aβ 42</td>
<td>0.001</td>
</tr>
<tr>
<td>Age</td>
<td>0.20</td>
</tr>
<tr>
<td>NFL</td>
<td>−0.001</td>
</tr>
<tr>
<td>P-tau</td>
<td>0.016</td>
</tr>
</tbody>
</table>

Sample B (n = 88).
Key: Aβ 42, β-amyloid 1–42; CI, confidence interval; NFL, neurofilament light; P-tau, phosphorylated tau.

![Fig. 2. Relationship between age and hippocampal volume in NFL+ and NFL− participants. Adjusted for sex. NFL+ (≥902 pg/mL) and NFL− (<902 pg/mL) participants are defined by a median split. Estimated group slopes with 95% confidence intervals are displayed. Data points from participants with MRI are displayed, including within-person changes for those with MRI at both time points. Abbreviations: MRI, magnetic resonance imaging; NFL, neurofilament light.](Image)
results were unchanged when also excluding Aβ 42 positive participants and those with declining memory function (Supplementary Table 7). Furthermore, exclusion of participants with P-tau levels > 60 pg/mL, increase of the Aβ 42 cutoff from 550 to 650 pg/mL, and also exclusion of 2–3 outliers per analysis did not alter the results (data not shown). Results were the same in the very low AD risk subgroups from sample C (samples H–K).}

3.5. Mediation analyses

We tested the mediating (indirect) effect of NFL on the relationship between age and hippocampal atrophy rate (Fig. 3). NFL was a significant mediator, with confidence interval of –0.24 to −0.01, and accounted for 36% of the age effect on hippocampal atrophy rate. In our model, the total effect of age on hippocampal atrophy was β = −0.23 equal to the sum of the direct effect of age (β = −0.15) and the indirect effect through the relationship with NFL (β = −0.08).

4. Discussion

High CSF NFL levels predicted higher hippocampal atrophy rate in cognitively healthy older adults. Although previous studies have demonstrated this in samples of high-risk participants, that is, MCI patients (Zetterberg et al., 2016), here we show that the relationship was replicated in a sample with very low AD risk and that NFL predicted hippocampal atrophy independently of the established AD CSF biomarkers Aβ 42 and P-tau. This suggests that CSF NFL may be an important marker of neurodegeneration both in normal aging and in age-related neurodegenerative diseases.

The only previous study assessing CSF NFL in relation to longitudinal volume change in older adults found that higher NFL levels were associated with deterioration in whole-brain, ventricular and hippocampal volume in MCI patients (Zetterberg et al., 2016). However, cross-sectional studies in nondemented adults have been more inconsistent. One study found that CSF NFL correlated with ventricular size, but not with sulcal atrophy (Bjerke et al., 2014), a second study found a correlation between brain parenchymal fraction and CSF NFL that did not survive adjustment for age (Vagberg et al., 2015), whereas a third study found no relationship between baseline CSF NFL levels and gray-matter volumes 3.5 years later (Bendlin et al., 2012). In frontotemporal dementia, higher CSF NFL is associated with lower gray and WM volumes, including in the temporal lobe (Scherling et al., 2014), whereas findings in multiple sclerosis (MS) and related disorders are less straightforward (Eikelenboom et al., 2003; Khalil et al., 2013). Thus, previous literature on the association between CSF NFL and brain volumes is scarce and inconsistent, but the only longitudinal study in older adults is in line with our findings (Zetterberg et al., 2016), showing that high CSF NFL predicts more hippocampal atrophy in MCI patients. The present study takes these results further by showing that the NFL-atrophy association is likely not caused by AD-specific mechanisms but is important also in AD-independent, age-expected hippocampal decline.

Neurofilaments are abundant in neuronal axons where they are essential for axon radial growth (Petzold, 2005) but are also found in soma and dendrites of neurons (Trojanowski et al., 1986). NFL is expressed in neurons in both the central and peripheral nervous system (Trojanowski et al., 1986), including hippocampus (Vickers et al., 1994). Thus, following neuronal damage, NFL is believed to be released into the extracellular compartment resulting in increased CSF NFL levels (Petzold, 2005). Age is associated with increasing CSF NFL levels in several studies (Khalil et al., 2013; Rosengren et al., 1996; Skillback et al., 2014; Steinacker et al., 2016; Vagberg et al., 2015), suggesting that CSF NFL levels increase with normal aging. Interestingly, we found that NFL levels could explain more than one-third of the age-related increase in hippocampal atrophy rates. As accelerated decline of the hippocampus also in normal aging is observed independently of AD-related pathology (Fjell et al., 2013), this is an important finding. Thus, our results indicate that CSF NFL levels reflect processes characterizing normal aging.

There has recently been increasing focus on amyloid-independent neurodegeneration in aging, often referred to as suspected non-Alzheimer pathology, making it important to map out correlates of atrophy also in AD-typical areas in Αβ 42–negative older adults (Jack et al., 2016). Thus, we created a subgroup with very low AD risk (only Αβ 42–negative participants), in which our finding was upheld. This bolsters that AD brain pathology is not a confounder of the relationship between CSF NFL and hippocampal atrophy rate and suggests that CSF NFL most likely reflects neurodegeneration processes in normal aging. Furthermore, in another subgroup analysis, CSF NFL predicted hippocampal atrophy rate after exclusion of participants with additional risk conditions, supporting that CSF NFL likely reflects normal aging processes. Previous studies suggest that CSF NFL may reflect the rate of ongoing neurodegeneration. High CSF NFL levels are seen days after a bout in amateur boxing (Zetterberg et al., 2006), with subsequent decrease during the next months, CSF NFL levels are highest in MS patients with an ongoing relapse (Malnemström et al., 2003), and high CSF NFL levels are associated with progression of neurodegenerative diseases (Backstrom et al., 2015; Skillback et al., 2014; Steinacker et al., 2016; Zetterberg et al., 2016). Accordingly, CSF NFL levels are higher in the rapidly progressing neurodegenerative disease amyotrophic lateral sclerosis than in AD which progresses more slowly (Steinacker et al., 2016), and MCI patients have CSF NFL levels intermediate between those of AD patients and controls (Zetterberg et al., 2016). Thus, CSF NFL may reflect that similar neurodegenerative processes are ongoing in both normal aging and diseases, and the CSF NFL levels may reflect the progression rate of the processes.
The etiologies of neuronal damage and neurodegeneration, and thus high NFL levels, can be manifold. Cerebrovascular pathology, including stroke (Norgren et al., 2003) and WM lesions (Sjogren et al., 2001), has been associated with elevated CSF NFL levels. Because clinically silent cerebrovascular pathology is prevalent in older adults without dementia (Ikram et al., 2008; Vermeer et al., 2002), cerebrovascular pathology may be one cause of elevated NFL levels in our study. Associations between WM lesions and hippocampal atrophy have been shown previously, although not consistently (Appelman et al., 2009). It is still unknown whether this represents a causal link or is due to shared risk factors (Appelman et al., 2009). Because vascular brain pathology may affect the relationship between NFL and hippocampal atrophy, we adjusted our final regression model for WM hypointensities. NFL was still a significant predictor of hippocampal atrophy rate, indicating that NFL predicts hippocampal atrophy rate independently of cerebrovascular pathology, although it cannot be ruled out that more sensitive measures of WM lesions could yield other results. Unexpectedly (Tosun et al., 2010), in the final model, higher P-tau levels predicted lower hippocampal atrophy rates. However, P-tau was not significantly related to atrophy in GAMM, neither when no covariates but P-tau were included in the regression model. We can thus not exclude the possibility that the unexpected relationship with hippocampal atrophy is due to shared variance with the other covariates NFL and age. One explanation for the finding may be that our study could have excluded individuals with high CSF P-tau levels and high hippocampal atrophy rates, as they are more likely to have dementia or cognitive impairment. Thus, this result should be interpreted with caution. Because there were correlations among the biomarkers, data reduction methods such as principal component or cluster analysis could have been used to optimize classification accuracy. In this study, this was not done because we aimed to evaluate the contributions from the different biomarkers separately. However, this would be an important step for future studies to develop optimal combinations of variables in terms of classification accuracy.

Several limitations should be addressed. The main limitation is that although the likelihood of confounding by symptomatic AD is low, we cannot rule out the possibility that symptomatic neurodegenerative pathology of other etiologies may in part account for some of the relationship between CSF NFL and hippocampal atrophy. Furthermore, although participants were followed for 2 years, we cannot be sure that they do not develop neurodegenerative diseases later. A second limitation is the method used to measure WM lesion load. We obtained a measure based on the automated labeling of the T1 signal, but we believe that this measure could be more accurate if other modalities, designed specifically to evaluate WM, such as T2 or FLAIR were available. Finally, a third limitation is the fact that our study had a mainly exploratory aim, in which it was difficult to address the issue of multiple comparisons. However, we believe that our results were consistently supported by the different analyses undertaken.

5. Conclusion

CSF NFL predicted neurodegeneration in older adults with very low probability of AD. The present results suggest that factors previously shown to be important for brain degeneration in MCI may also impact brain changes in normal aging, demonstrating that NFL is likely to be a marker of AD-independent, age-expectated neurodegeneration. Future research needs to explore the predictive value of this biomarker with regard to onset and progression of prodromal AD. For this purpose, even longer follow-ups than the current 2-year interval are necessary.

Disclosure statement

Dr Watne has given a lecture on delirium for Lilly. Dr Bruun Wyller has given lectures on delirium for Pfizer, Roche, AstraZeneca, and Nycomed. Dr Blennow has served on Advisory Boards for IBL International and Roche Diagnostics. Dr Walhovd has given a lecture on lifespan changes in brain and cognition for Shire International GmbH (2015) and has served in an expert group for ILSI Europe, for both of which honoraria were paid. The other authors report no conflicts of interest.

Acknowledgements

The authors would like to thank Knut Engedal, MD, PhD (University of Oslo and Vestfold Hospital Trust) and Anette Hylen Ranhoff, MD, PhD (Diakonhjemmet Hospital and Kavli Research Center for Geriatrics and Dementia) for general support; Gry Torsater Dahl, MD (Diakonhjemmet Hospital), Randi Steen, MD (Oslo University Hospital), Arne Myklebust, MD (Oslo University Hospital), and Dagfinn Tore Kollerv, MD (Oslo University Hospital) for collection of data; and Asta Kristine Håberg, MD, PhD (St Olav’s University Hospital) for planning the baseline MRI protocol. The authors would also like to thank the study participants and the contributions of the Department of Gynecology, the Department of Urology, and the Department of Anesthesiology at Oslo University Hospital; the Department of Orthopedic Surgery and the Department of Anesthesiology at Diakonhjemmet Hospital in Oslo, Norway.

The study was mainly funded by the Medical Student Research Program at the University of Oslo and The National Association for Public Health’s dementia research program, Norway. Furthermore, they have received funding from Inlandet Hospital Trust (grant number: 150201), the Knut and Alice Wallenberg Foundation, the Swedish Research Council (grant numbers: 2013-3546 and K2013-61X-14002-13-5), and the Torsten Söderberg Foundation at the Royal Swedish Academy of Sciences. The funding sources had no role in the study design; in the collection, analysis, and interpretation of the data; in the writing of the report; or in the decision to submit the article for publication.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.neurobiolaging.2016.09.012.

References


A.-V. Idland et al. / Neurobiology of Aging 49 (2017) 138–144
Figure e-1. Sample selection

Supplementary Data

Aβ42 = cerebrospinal fluid β-amyloid 1-42. AD = Alzheimer’s disease. APOE = Apolipoprotein E. CSF = cerebrospinal fluid. NFL = cerebrospinal fluid neurofilament light chain. MRI = magnetic resonance imaging. P-tau = cerebrospinal fluid phosphorylated tau.
Table e-1. Cognitively healthy participants who have reported central nervous system related conditions (ICD-10 codes) (World Health Organization, 1992)

<table>
<thead>
<tr>
<th>Inflammatory diseases of the central nervous system (G00–G09)</th>
<th>MRI at both time points and CSF NFL analyses (n=12)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Three participants had had meningitis before inclusion in the study. One has epilepsy as a possible sequela.</td>
<td>2</td>
</tr>
<tr>
<td>- One participant had had encephalitis without any reported sequela many years before inclusion in the study.</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Systemic atrophies primarily affecting the central nervous system (G10–G13)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- None</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extrapyramidal and movement disorders (G20–G26, except essential tremor [G25.0])</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- None</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other degenerative diseases of the nervous system (G30–G32)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- None</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Epilepsy, episodic and paroxysmal disorders (G40–G41)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- One participant has epilepsy with symptom debut before inclusion in the study. This participant has also had meningitis.</td>
<td>1</td>
</tr>
<tr>
<td>- One participant developed epilepsy after inclusion in the study.</td>
<td>1</td>
</tr>
<tr>
<td>- Two participants had epilepsy during childhood/youth (one possibly due to head trauma), but had not had a seizure in decades before the inclusion in the study.</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cerebrovascular, episodic and paroxysmal disorders (G45–G46)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Four participants had had transient ischemic attack before inclusion in the study.</td>
<td>1</td>
</tr>
<tr>
<td>- Two participants had had transient global amnesia before inclusion in the study.</td>
<td>2</td>
</tr>
<tr>
<td>- One participant has had transient global amnesia after inclusion in the study.</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cerebral palsy and other paralytic syndromes (G80–G83)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- None</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other disorders of the nervous system (G90–G99)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- None</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Viral infections of the nervous system (A80–A89)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- None</td>
<td>-</td>
</tr>
</tbody>
</table>

---

*Inflammatory diseases of the central nervous system (G00–G09)

- Three participants had had meningitis before inclusion in the study. One has epilepsy as a possible sequela.
- One participant had had encephalitis without any reported sequela many years before inclusion in the study.

*Systemic atrophies primarily affecting the central nervous system (G10–G13)

- None

*Extrapyramidal and movement disorders (G20–G26, except essential tremor [G25.0])

- None

*Other degenerative diseases of the nervous system (G30–G32)

- None

*Epilepsy, episodic and paroxysmal disorders (G40–G41)

- One participant has epilepsy with symptom debut before inclusion in the study. This participant has also had meningitis.
- One participant developed epilepsy after inclusion in the study.
- Two participants had epilepsy during childhood/youth (one possibly due to head trauma), but had not had a seizure in decades before the inclusion in the study.

*Cerebrovascular, episodic and paroxysmal disorders (G45–G46)

- Four participants had had transient ischemic attack before inclusion in the study.
- Two participants had had transient global amnesia before inclusion in the study.
- One participant has had transient global amnesia after inclusion in the study.

*Cerebral palsy and other paralytic syndromes (G80–G83)

- None

*Other disorders of the nervous system (G90–G99)

- None

*Viral infections of the nervous system (A80–A89)

- None
<table>
<thead>
<tr>
<th>Disease</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>All participants (n=22)^a</td>
<td>MRI at both time points and CSF NFL analyses (n=12)^a</td>
</tr>
<tr>
<td>Benign neoplasm of meninges, parts of central nervous system, pituitary gland, craniopharyngeal duct or pineal gland (D32-D33 and D35.2-D35.4)</td>
<td>0</td>
</tr>
<tr>
<td>- One participant has a pituitary gland tumor. The tumor has been asymptomatic for decades.</td>
<td></td>
</tr>
<tr>
<td>Malignant neoplasm of meninges, brain, spinal cord, cranial nerves and other parts of central nervous system (C70-72 and C79)</td>
<td>-</td>
</tr>
<tr>
<td>- None</td>
<td></td>
</tr>
<tr>
<td>Neoplasm of uncertain or unknown behaviour of meninges, central nervous system, pituitary gland, craniopharyngeal duct or pineal gland (D42-43 and D44.3-D44.5)</td>
<td>-</td>
</tr>
<tr>
<td>- None</td>
<td></td>
</tr>
<tr>
<td>Head injury (S00–S09,T00.0, T02.0, T03.0, T04.0, and T06.0)</td>
<td>4</td>
</tr>
<tr>
<td>- Three participants had had head trauma with skull fracture many years before inclusion in the study.</td>
<td></td>
</tr>
<tr>
<td>- One participant had had head trauma with epilepsy as a possible sequela.</td>
<td></td>
</tr>
<tr>
<td>Cerebrovascular diseases (I60–I69)</td>
<td>1</td>
</tr>
<tr>
<td>- Four participants had had stroke without sequela before inclusion in the study.</td>
<td></td>
</tr>
</tbody>
</table>

^aTwo participants have two conditions each: One has had meningitis and has epilepsy. One has had a head trauma with epilepsy as a possible sequela. MRI = magnetic resonance imaging. CSF = cerebrospinal fluid. NFL = neurofilament light. Participants are excluded in the subgroup analysis with results shown in table e-2 (n=12).
Table e-2. Statistical analyses performed in the different samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Criteria</th>
<th>Analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Baseline CSF biomarker measures or MRI (n=144)</td>
<td>- Demographics</td>
</tr>
</tbody>
</table>
| B      | CSF NFL, Aβ42, P-tau, and T-tau measures and MRI at both time points available (n=88) | - Demographics  
- Mean time from CSF sampling to freezing  
- Mean time from CSF sampling to MRI at baseline  
- Mean time between MRIs  
- Correlations between age and CSF biomarkers.  
- Correlations between CSF biomarkers  
- Correlation between hippocampus volume at baseline and CSF NFL  
- Correlation between age and hippocampal atrophy rate  
- Paired sample T-test on hippocampal volume change  
- Stepwise linear regression with age, CSF NFL, Aβ42, P-tau, T-tau, Aβ42/P-tau ratio, WMHypointensities and sex as predictors of hippocampal atrophy rate  
- Mediation analyses |
| C      | Exclusions after further screening of sample B (n=76) | - Correlations between age and CSF biomarkers.  
- Correlations between CSF biomarkers  
- Correlation between hippocampus volume at baseline and CSF NFL  
- Correlation between age and hippocampal atrophy rate  
- Paired sample T-test on hippocampal volume change  
- Stepwise linear regression with age, CSF NFL, Aβ42, P-tau, T-tau, Aβ42/P-tau ratio, WMHypointensities and sex as predictors of hippocampal atrophy rate |
| D-K    | Very low AD risk subgroups of sample B and C (see Figure e-1 for details) | Linear regression with final model including age, NFL, and P-tau as predictors of hippocampal atrophy rate |
# Table e-3. Cognitive test results at baseline and two-year follow-up.

<table>
<thead>
<tr>
<th>Participants with MRI at both time points and CSF NFL analyses (n=88)</th>
<th>Baseline</th>
<th>Two-year follow-up</th>
<th>Range of score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock drawing test</td>
<td>4.8 (.5), 3 to 5</td>
<td>4.8 (.4), 3 to 5</td>
<td>0-5</td>
</tr>
<tr>
<td>Immediate recall&lt;sup&gt;a&lt;/sup&gt;</td>
<td>21(3), 11 to 27</td>
<td>22 (3), 14 to 30</td>
<td>0-30</td>
</tr>
<tr>
<td>Delayed recall&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.8 (1.8), 3 to 10</td>
<td>7.7 (1.7), 3 to 10</td>
<td>0-10</td>
</tr>
<tr>
<td>Correctly recognized yes&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>9.4 (.8), 7 to 10</td>
<td>9.8 (.5), 7 to 10</td>
<td>0-10</td>
</tr>
<tr>
<td>Correctly identified no&lt;sup&gt;ac&lt;/sup&gt;</td>
<td>10.0 (.2), 9 to 10</td>
<td>9.9 (.3), 8 to 10</td>
<td>0-10</td>
</tr>
<tr>
<td>Kendrick OLT</td>
<td>44 (7), 26 to 63</td>
<td>44 (8), 22 to 61</td>
<td>0-70</td>
</tr>
<tr>
<td>Verbal fluency (FAS)</td>
<td>44 (11), 19 to 81&lt;sup&gt;d&lt;/sup&gt;</td>
<td>47 (13), 17 to 83</td>
<td>Unlimited</td>
</tr>
<tr>
<td>Animal naming</td>
<td>21 (6), 5 to 32</td>
<td>23 (6), 7 to 38</td>
<td>Unlimited</td>
</tr>
<tr>
<td>TMT A</td>
<td>50 (17), 26 to 120</td>
<td>46 (17), 19 to 101</td>
<td>Unlimited&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>TMT B</td>
<td>118 (61), 34 to 466</td>
<td>114 (53), 31 to 270&lt;sup&gt;f&lt;/sup&gt;</td>
<td>Unlimited&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Values are mean (SD), range. <sup>a</sup>From Word List Memory Task, <sup>b</sup>Number of 10 words presented in the Word List Memory Task correctly recognized, <sup>c</sup>The number of 10 distractor words correctly identified, <sup>d</sup>n=87. Missing in one participant, <sup>e</sup>Usually stopped if not finished within 300 seconds, <sup>f</sup>n=87. One participant did not finish TMT B. TMT = Trail Making Test; OLT = Object Learning Test.
<table>
<thead>
<tr>
<th>Independent variables</th>
<th>$R^2$</th>
<th>B</th>
<th>95% CI</th>
<th>$\beta$</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.18</td>
<td>-.026</td>
<td>-.065 to .013</td>
<td>-.14</td>
<td>.20</td>
</tr>
<tr>
<td>NFL</td>
<td>-.001</td>
<td>-.001 to -.0002</td>
<td>-.34</td>
<td>.004</td>
<td></td>
</tr>
<tr>
<td>T-tau</td>
<td>.002</td>
<td>.0003 to .003</td>
<td>.24</td>
<td>.02</td>
<td></td>
</tr>
</tbody>
</table>

Sample B (n=88). CI = confidence interval. T-tau = cerebrospinal fluid total tau. NFL = cerebrospinal fluid neurofilament light.
Table e-5. Multiple linear regression with hippocampal atrophy rate as dependent variable (Sample C).

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>R²</th>
<th>B</th>
<th>95% CI</th>
<th>β</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.09</td>
<td>-.009</td>
<td>-.049 to .032</td>
<td>-.05</td>
<td>.67</td>
</tr>
<tr>
<td>NFL</td>
<td></td>
<td>-.0005</td>
<td>-.001 to -.00006</td>
<td>-.28</td>
<td>.03</td>
</tr>
<tr>
<td>Age</td>
<td>.13</td>
<td>-.01</td>
<td>-.05 to .029</td>
<td>-.06</td>
<td>.60</td>
</tr>
<tr>
<td>NFL</td>
<td></td>
<td>-.001</td>
<td>-.001 to -.00008</td>
<td>-.29</td>
<td>.02</td>
</tr>
<tr>
<td>Aβ42</td>
<td></td>
<td>.001</td>
<td>-.0002 to .002</td>
<td>.19</td>
<td>.09</td>
</tr>
<tr>
<td>Age</td>
<td>.13</td>
<td>-.011</td>
<td>-.051 to .029</td>
<td>-.07</td>
<td>.59</td>
</tr>
<tr>
<td>NFL</td>
<td></td>
<td>-.001</td>
<td>-.001 to -.0001</td>
<td>-.32</td>
<td>.01</td>
</tr>
<tr>
<td>P-tau</td>
<td></td>
<td>.011</td>
<td>-.001 to .024</td>
<td>.20</td>
<td>.08</td>
</tr>
</tbody>
</table>

Sample C (n=76). CI= confidence interval. Aβ42= cerebrospinal fluid β-amyloid 1-42. P-tau = cerebrospinal fluid phosphorylated tau. NFL = cerebrospinal fluid neurofilament light.
Table e-6. Multiple linear regression with hippocampal atrophy rate as dependent variable (excluding outliers from sample C).

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>R²</th>
<th>B</th>
<th>95 % CI</th>
<th>β</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.21</td>
<td>-.005</td>
<td>-0.04 to .03</td>
<td>-.04</td>
<td>.77</td>
</tr>
<tr>
<td>NFL</td>
<td></td>
<td>-.001</td>
<td>-0.001 to -.0003</td>
<td>-.43</td>
<td>.001</td>
</tr>
<tr>
<td>Age&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.27</td>
<td>-.015</td>
<td>-0.05 to .02</td>
<td>-.10</td>
<td>.40</td>
</tr>
<tr>
<td>NFL</td>
<td></td>
<td>-.001</td>
<td>-0.001 to -.0003</td>
<td>-.43</td>
<td>.001</td>
</tr>
<tr>
<td>Aβ42</td>
<td></td>
<td>.001</td>
<td>.0001 to .002</td>
<td>.24</td>
<td>.02</td>
</tr>
<tr>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.31</td>
<td>-.016</td>
<td>-0.049 to .017</td>
<td>-.11</td>
<td>.35</td>
</tr>
<tr>
<td>NFL</td>
<td></td>
<td>-.001</td>
<td>-0.001 to -.0004</td>
<td>-.48</td>
<td>.0002</td>
</tr>
<tr>
<td>P-tau</td>
<td></td>
<td>.016</td>
<td>.006 to .026</td>
<td>.32</td>
<td>.003</td>
</tr>
</tbody>
</table>

CI = confidence interval. Aβ42 = cerebrospinal fluid β-amyloid 1-42. P-tau = cerebrospinal fluid phosphorylated tau. NFL = cerebrospinal fluid neurofilament light. <sup>a</sup>n=71. <sup>b</sup>n=72.
Table e-7. Multiple linear regression with hippocampal atrophy rate as dependent variable (low AD risk subsamples)

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>R²</th>
<th>B</th>
<th>95% CI</th>
<th>β</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.41</td>
<td>-.001</td>
<td>-.053 to .05</td>
<td>-.01</td>
<td>.96</td>
</tr>
<tr>
<td>NFL</td>
<td>-.001</td>
<td>.023</td>
<td>-.022 to .069</td>
<td>.15</td>
<td>.30</td>
</tr>
<tr>
<td>P-tau</td>
<td>.029</td>
<td>-.001</td>
<td>-.002 to -.001</td>
<td>-.72</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

<sup>a</sup>APOE-negative participants, sample D (n=48).<sup>b</sup>Very low AD risk subsample, sample E (n=36). CI = confidence interval. P-tau = cerebrospinal fluid phosphorylated tau. NFL = cerebrospinal fluid neurofilament light.
eReferences

Biomarker profiling beyond amyloid and tau – CSF markers, hippocampal atrophy and memory change in cognitively unimpaired older adults

Ane-Victoria Idland, Roser Sala-Llonch, Leiv Otto Watne, Anne Brækhus, Oskar Hansson, Kaj Blennow, Henrik Zetterberg, Øystein Sørensen, Kristine B. Walhovd, Torgeir Bruun Wyller, Anders Martin Fjell

a Center for Lifespan Changes in Brain and Cognition, Department of Psychology, University of Oslo, Norway.
b Oslo Delirium Research Group, Department of Geriatric Medicine, Institute of Clinical Medicine, University of Oslo, Norway.
c Institute of Basic Medical Sciences, University of Oslo, Norway.
d Department of Biomedicine, University of Barcelona, Spain.
e Department of Geriatric Medicine, Oslo University Hospital, Norway.
f Norwegian National Advisory Unit on Ageing and Health, Vestfold Hospital Trust, Norway.
g Department of Neurology, Oslo University Hospital, Norway.
h Department of Clinical Sciences, Lund University, Lund, Sweden.
i Memory Clinic, Skåne University Hospital, Lund, Sweden.
j Department of Psychiatry and Neurochemistry, The Sahlgrenska Academy at University of Gothenburg, Sweden.
k Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Sweden.
l Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, UK.
m UK Dementia Research Institute at UCL, London, UK.
n Department of Radiology and Nuclear Medicine, Oslo University Hospital, Norway.
Abstract
Brain changes commonly occurring in aging can be indexed by biomarkers. We used cluster analysis to identify sub-groups of individuals with different biomarker profiles. CSF from cognitively unimpaired individuals (n=99, 64-93 years) was analyzed for concentrations of Aβ42, phosphorylated tau (P-tau), and total tau (T-tau), and biomarkers for neuroinflammation (YKL-40), neuronal (FABP3) and axonal damage (NFL). Hippocampal volume and memory were assessed across multiple follow-up examinations covering up to 6.8 years. Clustering revealed one group (39%) with more pathological concentrations of all biomarkers, which could further be divided in one group (20%) characterized by tauopathy and high FABP3 and one (19%) by brain β-amyloidosis, high NFL, and slightly higher YKL-40. The clustering approach clearly outperformed classification based on Aβ42-P-tau alone in prediction of memory decline, with the individuals with most tauopathy and FABP3 showing worse development of memory over 6.8 years, but not more hippocampal volume change. The results demonstrate that older adults can be classified based on biomarkers beyond amyloid and tau, with improved prediction of memory decline.
1. Introduction

Biomarkers play an increasingly important role in research on age-related neurological conditions and diseases. Numerous studies have consistently shown a marked decrease in cerebrospinal fluid (CSF) concentration of $\beta$-amyloid 1-42 (A$\beta$42) together with increased total tau (T-tau) and phosphorylated tau (P-tau) in Alzheimer’s disease (AD) dementia and mild cognitive impairment cases showing progression to AD (Olsson, et al., 2016). The National Institute on Aging – Alzheimer’s Association’s (NIA-AA) new research framework defines AD solely based on biomarkers reflecting the core pathologies of AD, while clinical symptoms only are used for staging of the disease (Jack, et al., 2018). These criteria rely on the amyloid/tau/neurodegeneration (A/T/N) classification scheme for the core AD biomarkers (Jack, et al., 2016a), wherein CSF A$\beta$42 reflects brain $\beta$-amyloidosis (Strozyk, et al., 2003,Tapiola, et al., 2009), P-tau reflects tau pathology/neurofibrillary tangles (Buerger, et al., 2006,Tapiola, et al., 2009), and T-tau reflects neurodegeneration (Hesse, et al., 2001,Ost, et al., 2006,Zetterberg, et al., 2006). CSF concentrations of A$\beta$42, T-tau, and P-tau are also known to change even in cognitively well-functioning older adults (Jansen, et al., 2015,Toledo, et al., 2015). Accordingly, cognitively unimpaired older adults with abnormal amyloid and tau biomarkers are defined as having preclinical AD based on the NIA-AA criteria.

However, higher age comes with a multitude of different conditions and changes in the brain in addition to amyloid deposition and tau, and many proteins related to these processes can be measured in CSF, such as neuroinflammation and different aspects of neurodegeneration. Chitinase-3-like protein 1 (YKL-40) is mainly expressed in astrocytes in the human brain (Bonneh-Barkay, et al., 2010), and YKL-40 expressing astrocytes are found close to activated microglia (Bonneh-Barkay, et al., 2010). Consequently, CSF YKL-40 is believed to be biomarker of neuroinflammation (Baldacci, et al., 2017,Dhiman, et al., 2019). Fatty acid binding protein 3 (FABP3) is expressed in neurons of the brain (Pelsers, et al., 2004), where it is involved in transport of fatty acids. FABP3 is found in the cytosol, and it is released following cellular injury, thus CSF FABP3 is considered a biomarker of neuronal damage (Dhiman, et al., 2019,Pelsers, et al., 2004). Neurofilament light chain protein (NFL) is a cytoskeletal component of neuronal axons (Khalil, et al., 2018). NFL is released from neuronal axons in response to neuronal damage, and CSF NFL is believed to reflect axonal degeneration (Dhiman, et al., 2019,Khalil, et al., 2018). CSF YKL-40, FABP3, and NFL are not disease-specific biomarkers, and all of them have been found to be increased in both acute and chronic brain diseases (Baldacci, et al., 2017,Bonneh-Barkay, et al., 2010,Bridel, et al., 2019,Olsson, et al.,
Although all the above mentioned CSF biomarkers reflect at least partly separate brain pathological processes, some processes may also be interrelated, which can be reflected by relationships between biomarkers.

Relationships between the core AD CSF biomarkers and emerging CSF biomarkers like YKL-40, FABP3, and NFL in cognitively unimpaired older adults have mainly been studied by assessing correlations between the individual biomarkers. Our knowledge about how CSF biomarkers may cluster is, however, limited. Further, acknowledging that most age-related brain changes are the result of a number of different processes that probably vary across individuals, it is a major task to be able to group older adults according to their brain states. Clustering analyses can be used to identify subgroups with multiple co-occurring biomarker features. Unfortunately, beyond Aβ42 and Tau (Nettiksimmons, et al., 2010, Racine, et al., 2016), we do not know whether clustering analyses can be used to classify cognitively unimpaired older adults in meaningful subgroups characterized by partly different and partly overlapping brain pathology.

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Related pathological process</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF Aβ42</td>
<td>Amyloid deposition</td>
</tr>
<tr>
<td>CSF FABP3</td>
<td>Neuronal damage</td>
</tr>
<tr>
<td>CSF phosphorylated tau</td>
<td>Tau phosphorylation/tangle formation</td>
</tr>
<tr>
<td>CSF total tau</td>
<td>Altered tau metabolism/neurodegeneration</td>
</tr>
<tr>
<td>CSF YKL-40</td>
<td>Neuroinflammation</td>
</tr>
<tr>
<td>CSF NFL</td>
<td>Axonal damage/neurodegeneration</td>
</tr>
</tbody>
</table>

CSF=cerebrospinal fluid. Aβ42= β-amyloid 1-42. FABP3= fatty acid binding protein 3. YKL-40= Chitinase-3-like protein 1. NFL = neurofilament light.
To address these questions, we first performed correlation and clustering analyses to assess relationships between established and emerging CSF biomarkers (Table 1) in order to examine how biomarkers for different brain states are related at different superordinate levels. Secondly, we used the CSF biomarkers to identify participants with similar biomarker profiles using a blind, data-driven clustering approach across participants. The rationale was to test whether subgroups of older adults could be detected based on biomarker profiles. Third, we assessed whether these subgroups showed different trajectories of memory and hippocampal volume change across multiple follow-up examinations distributed over an interval up to 6.8 years. The performance of the clustering approach in prediction of memory decline and hippocampal volume change over time was compared to the NIA-AA classification based on amyloid and tau.

2. Materials and methods

2.1 Participants
We recruited patients scheduled for elective gynecological (genital prolapse), urological (benign prostate hyperplasia, prostate cancer, or bladder tumor/cancer) or orthopedic (knee or hip replacement) surgery in spinal anesthesia turning 65 years or older the year of inclusion. Dementia, previous stroke with sequela, Parkinson’s disease and other neurodegenerative diseases likely to affect cognition were exclusion criteria at baseline. Participants were assessed with a multi-domain battery of cognitive tests before surgery, including the Mini Mental Status Examination (MMSE) (Folstein, et al., 1975), Clock Drawing Test (Shulman, 2000), Word List Memory Task (Morris, et al., 1989), Trail Making Test A (TMTA) and B (TMTB) (Reitan, 1955), and verbal fluency (FAS test and Animal Naming) (Spreen and Strauss, 1991). CSF samples were collected by the anesthesiologist in conjunction with the spinal anesthesia. The participants also underwent magnetic resonance imaging (MRI) after surgery. The mean time between CSF sampling and MRI at baseline was 8 weeks (standard deviation [SD] [range]: 6 [-20 to 24]). Participants were tested with the same battery of cognitive tests annually and with MRIs biennially for up to 6.8 years.

From all recruited participants, we selected only participants with available CSF data available for all biomarkers (Aβ42, T-tau, P-tau, YKL-40, FABP3, and NFL). Further, we performed a review of all neurological diagnoses and MRI findings at baseline or occurring through follow-up in the cohort. We excluded participants with diagnoses/lesions that we found likely to affect
cognition or measures of hippocampal volume (details in Supplementary Material). Notably, we excluded all participants who had received a diagnosis of dementia or mild cognitive impairment during follow up, had a cognitive impairment according to hospital medical records, had developed other neurodegenerative diseases during follow up, and participants who, based on a cognitive assessment in the study, had been offered referral to the hospital for further cognitive assessment. Last, from the remaining sample, selection of participants cognitively normal at baseline was based on the following procedure: 1) we included all participants with MMSE ≥ 28, if also Clock Drawing Test score was ≥ 4, and Word List Recall score was ≥ -1.5 SD from the mean according to age, sex, and education adjusted norms, and 2) we included participants with MMSE < 28, if Clock Drawing Test score was ≥ 4, and also test scores for Word List Recall score, TMT A, TMT B, FAS-test, and Animal Naming were ≥ -1.5 SD from the mean according to norms. Our selection resulted in 99 participants with CSF analyses available for all biomarkers, of which 99 had been cognitively assessed one or more times (55 participants with seven cognitive assessments), and 85 had one or more MRIs (33 participants with four MRIs). Demographics, cognitive test results and CSF biomarker characteristics are shown in Table 2. 73 participants were excluded (see reasons in Figure S1 and Supplementary text). Excluded participants had poorer performance on MMSE and Word List Recall compared to the included participants (Table S1). Age, years of education, sex distribution, and CSF biomarker characteristics did not differ.
### Table 2. Demographics

<table>
<thead>
<tr>
<th></th>
<th>Cognitively unimpaired older adults (n=99)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at baseline, years</td>
<td>72 (68 to 78)</td>
</tr>
<tr>
<td>Sex, male</td>
<td>50 (51)</td>
</tr>
<tr>
<td>Education, years</td>
<td>14 (12 to 17)</td>
</tr>
<tr>
<td>MMSE score, baseline</td>
<td>29 (28 to 30)</td>
</tr>
<tr>
<td>CERAD, Word List recall score</td>
<td>6 (5 to 8)</td>
</tr>
<tr>
<td>CSF Aβ42, pg/mL</td>
<td>731 (512 to 866)</td>
</tr>
<tr>
<td>CSF FABP3, pg/mL</td>
<td>4.56 (3.36 to 5.93)</td>
</tr>
<tr>
<td>CSF P-tau, pg/mL</td>
<td>59 (46 to 75)</td>
</tr>
<tr>
<td>CSF T-tau, pg/mL</td>
<td>347 (272 to 486)</td>
</tr>
<tr>
<td>CSF YKL-40, pg/mL</td>
<td>225210 (175208 to 280877)</td>
</tr>
<tr>
<td>CSF NFL, pg/mL</td>
<td>1026 (794 to 1482)</td>
</tr>
</tbody>
</table>

Values are median (IQR) and N (%). MMSE = Mini Mental Status Examination. CERAD= Consortium to Establish a Registry for Alzheimer’s disease. CSF = cerebrospinal fluid. Aβ42 = β-amyloid 1-42. FABP3 = fatty acid binding protein 3. P-tau = phosphorylated tau. T-tau = total tau. YKL-40 = Chitinase-3-like protein 1. NFL = neurofilament light.

#### 2.2 Ethical considerations

The study was conducted in accordance with the Declaration of Helsinki and approved by the Regional Committee for Ethics in Medical Research in Norway (REK 2011/2052). All participants provided written informed consent.

#### 2.3 Magnetic resonance imaging acquisition and processing

T1-weighted MPRAGE 3D images were acquired with a 1.5 T Siemens Avanto scanner using a 12-channel head coil (TR=2400 ms, TE=3.79ms, Field of View=240mm, slice thickness=1.20mm, pixel size = 1.25x1.25mm). The same scanner was used at baseline and all follow ups. Images were processed with FreeSurfer (version 6.0) and its specific longitudinal stream (https://surfer.nmr.mgh.harvard.edu). For each MRI, the FreeSurfer pipeline performs a set of automated procedures for the cortical reconstruction and volumetric segmentation, documented elsewhere (Dale, et al., 1999, Fischl, et al., 2002). More specifically, the segmentation algorithm
assigns labels to all the brain regions of each individual scan, based on an available probabilistic atlas obtained from a training set of participants which has been accurately manually labeled (Fischl, et al., 2002). The hippocampal volume is defined from this available atlas. The FreeSurfer longitudinal stream includes methods designed to minimize the bias to any time point in a participant and which lead to increased statistical power, better separation of groups based on atrophy, and higher reproducibility. These include the generation of a subject-specific intermediate template followed by a projection of each time point to this template (Jovicich, et al., 2013, Reuter, et al., 2012).

2.4 CSF collection and analyses
CSF was collected in polypropylene tubes, centrifuged at room temperature for 10 minutes, the supernatant aliquoted into polypropylene tubes, and frozen at -80 °C pending analyses. Samples were sent on dry ice to the Clinical Neurochemistry Laboratory at Sahlgrenska University Hospital, Mölndal, Sweden, for analyses. CSF Aβ42, total tau (T-tau) and P-tau concentrations were measured using INNOTEST enzyme-linked immunosorbent assays (Fujirebio, Ghent, Belgium), CSF NFL concentrations using a commercial ELISA (UmanDiagnostics, Umeå, Sweden), YKL-40 concentrations using a commercial ELISA (R&D systems, Minneapolis, MN), and FABP3 concentrations using an immunoassay with electrochemiluminescence detection (MSD® Human FABP3 kit, Meso Scale Discovery, Gaithersburg, MD, USA). Analyses were performed by board-certified laboratory technicians masked to clinical data. Intra-assay coefficients of variation were 9-13%. All participants had detectable levels of all biomarkers.

2.5 Statistical analysis
We tested correlations between baseline age and CSF biomarkers and between the CSF biomarkers using bivariate and partial Spearman correlations in SPSS (version 25). These analyses were undertaken to describe the structure of the data, not to test specific hypotheses.

Clustering analyses were performed in Matlab (MathWorks Inc). Clustering analysis is used to identify natural groupings of similar variables from a data set. First, we calculated the distance between variables using Spearman correlation, to account for non-normal distribution of the data. We then used the ‘ward’ or inner squared distance as a linkage function to group the variables into clusters. The variables were reordered with the optimal leaf order and a hierarchical dendrogram was used to represent the clusters at the different levels. To remove
the effect of age from all biomarkers, we computed independent linear regressions of each biomarker against age, and the residuals of these regressions were used for the clustering analysis. The clustering approach was completely data-driven. As the relationships between several of the included biomarkers are not known a priori, we did not impose restrictions of the clustering algorithms.

We performed two cluster analyses:

**Cluster analysis 1**: We used cluster analysis to establish clusters of CSF biomarkers with shared behavior across participants. In this analysis, all the available CSF biomarkers were used as variables and the participants as observations. The purpose of this was to see which CSF biomarkers that tended to go together across different number of clusters.

**Cluster analysis 2**: In this analysis, we used the CSF biomarkers to identify sub-groups of participants. Thus, we ran the cluster analysis to define groups of participants with the same profiles of CSF biomarkers. Here, participants were used as variables and the biomarker concentrations as observations. Thus, in cluster analysis 1, we tested which biomarkers that clustered together (the CSF biomarkers were the variables), while in cluster analysis 2 we tested which participants that clustered together (the participants were the variables).

The number of clusters was established by inspection of the hierarchical distribution of the dendrogram, and we explored different numbers of clusters. In order to characterize the biomarker profile for each clustering-based subgroup of participants, differences between the subgroups for each biomarker were quantified by calculating Cohen’s d (the pairwise difference in mean biomarker values between groups divided by the pooled standard deviation weighted for group size). This was done to map the relative contributions of the different biomarkers in the grouping of participants. According to established rules of thumbs, we considered effect sizes ≥ .80 as large, ≥ .50 as medium and ≥ .20 as small.

As an alternative to the clustering approach, we classified participants into biomarker groups according to the NIA-AA criteria (Jack, et al., 2016a). The criteria for amyloid positivity (A+) was Aβ42 <530 pg/mL and for tau positivity (T+) P-tau > 60 pg/mL according to established cut-offs (Hansson, et al., 2006). T-tau was not used for classification of neurodegeneration (N+), because of a very strong correlation between T-tau and P-tau (r=0.96, p < 0.001).
Finally, we tested for intercept or slope differences in memory and hippocampal volume over time as a function of biomarker group by use of Generalized Additive Mixed Models (GAMM) run in R (https://www.r-project.org) using Rstudio (www.rstudio.com) IDE. GAMM uses the package “mgcv” (Wood, 2006). Memory score from Word List Recall, for up to seven time points covering up to 6.8 years was used as outcome variable, biomarker group (“cluster”) as factor, participant-specific time since baseline as covariate, and we included a time x biomarker group interaction term. Sex and baseline age were included as covariates of no interest. Random intercept was included. Separate analyses were run including number of memory test sessions completed as a proxy to control for practice effects. The same analyses were run for hippocampal volume across time, covering up to 6.81 years. The same variables and covariates as for the memory analyses were included. In addition, estimated total intracranial volume was included as an additional covariate of no interest. A major advantage of GAMM in the present setting is that relationships of any degree of complexity can be modelled without specification of the basic shape of the relationship, and GAMM is thus especially well-suited to map trajectories of neurocognitive variables which can be assumed to be non-linear and where the basic form of the curve is not known (Fjell, et al., 2010). This means that if the trajectories of a given measure are compared across groups of participants, GAMM will detect possible slope differences around inflection points. GAMM fits are typically evaluated and inspected based on p- and F-values, edf (effective degrees of freedom) as a measure of the complexity of the curve, as well as by inspecting the plotted graphs. We also used the package “simr” in R to calculate how many annual examinations would be required to detect differences in memory change between the biomarker groups with 80% power for our sample with the given effect sizes (Green and MacLeod, 2016).

3. Results

3.1 CSF biomarker correlations

CSF Aβ42 did not correlate with age or any of the other CSF biomarkers. CSF T-tau, P-tau, YKL-40, FABP3, and NFL were all positively correlated with age (Table 3). Correlations between CSF biomarkers were therefore adjusted for age. T-tau, P-tau, YKL-40, NFL and FABP3 were all positively correlated (Table 3). Such a correlation pattern between the biomarkers suggested that it could be possible to identify higher-order structures in the data, i.e. clusters of biomarkers.
Table 3. CSF biomarker correlations

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Aβ42</th>
<th>FABP3</th>
<th>P-tau</th>
<th>T-tau</th>
<th>YKL-40</th>
<th>NFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aβ42</td>
<td>-.01</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.96</td>
</tr>
<tr>
<td>FABP3</td>
<td></td>
<td>.26</td>
<td>.15</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-tau</td>
<td>.28</td>
<td>.10</td>
<td>.79</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-tau</td>
<td>.29</td>
<td>.05</td>
<td>.79</td>
<td>.96</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YKL-40</td>
<td>.36</td>
<td>.01</td>
<td>.44</td>
<td>.62</td>
<td>.67</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>NFL</td>
<td>.47</td>
<td>.02</td>
<td>.52</td>
<td>.32</td>
<td>.33</td>
<td>.35</td>
<td>-</td>
</tr>
</tbody>
</table>

Numbers represent Spearman’s Rho and p-values for the first and second line, respectively. Correlations between the CSF biomarkers are adjusted for age by partial correlations. **Bold** indicates p < .05. CSF = cerebrospinal fluid. Aβ42 = β-amyloid 1-42. FABP3 = fatty acid binding protein 3. P-tau = phosphorylated tau. T-tau = total tau. YKL-40 = Chitinase-3-like protein 1. NFL = neurofilament light.

3.2 Cluster analysis 1: Clusters of CSF biomarkers

The cluster analysis yielded different levels of separation of the CSF biomarker clusters (Figure 1). At Level 1, one cluster was formed by Aβ42 and a second cluster by the remaining CSF biomarkers. At Level 2, the second group from Level 1 was further subdivided into one cluster consisting of FABP3, T-tau, P-tau and YKL-40, and one cluster formed by NFL. At Level 3, the cluster formed by FABP3, T-tau, P-tau and YKL-40 was split into two clusters, one consisting of YKL-40 only and one with the remaining CSF biomarkers (FABP3, T-tau and P-tau). At the final level, the tau biomarkers were separated from FABP3.
3.3 Cluster analysis 2: Subgroups of participants identified by CSF biomarkers

We ran the cluster analysis to identify subgroups of participants with similar biomarker characteristics. We found that the participants could be divided in two groups, Group 1 (n = 60) and Group 2 (n = 39), respectively (Figure 2). To map out the relative contributions of each biomarker to the grouping, Cohen’s d was calculated for each biomarker (Table 4). Group 2 participants were characterized by more pathological biomarker results for all biomarkers, with Cohen’s d > .80 – considered a large effect size - for all except NFL, where Cohen’s d was > .20 (small effect size). These differences should be interpreted as descriptions of the pattern of biomarker differences most contributing to the grouping. Group 2 exceeded pathological thresholds for P-tau (> 60 pg/mL), and T-tau (> 350 pg/mL) (Hansson, et al., 2006), and the mean Aβ42 concentration was so close to the pathological threshold of <530 pg/mL (Hansson, et al., 2006) that it should be considered abnormal. Group 1 had normal values for all core AD biomarkers (Aβ42, P-tau, T-tau). There were no
differences in age, sex, years of education, or cancer morbidity between Group 1 and 2 (Table S2).

**Figure 2.** Hierarchical clustering of participants. Left panel: Subject-wise correlation matrix and dendrogram of the groups at the different levels. Right panel: Mean Z scores of each variable within each group. The z-scores are calculated for the current sample, yielding a sample sum of 0 and a standard deviation of 1, thus the groups tend to approximately mirror each other around the y=0 axis when the group sizes are similar.
Table 4. Differences in CSF biomarkers between two subgroups based on cluster analysis

<table>
<thead>
<tr>
<th></th>
<th>Group 1 (n=60)</th>
<th>Group 2 (n=39)</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Aβ42, pg/mL</td>
<td>812</td>
<td>150</td>
<td>538</td>
</tr>
<tr>
<td>FABP3, pg/mL</td>
<td>4.31</td>
<td>1.57</td>
<td>5.93</td>
</tr>
<tr>
<td>P-tau, pg/mL</td>
<td>53</td>
<td>15</td>
<td>76</td>
</tr>
<tr>
<td>T-tau, pg/mL</td>
<td>317</td>
<td>104</td>
<td>499</td>
</tr>
<tr>
<td>YKL-40, pg/mL</td>
<td>209512</td>
<td>67132</td>
<td>277101</td>
</tr>
<tr>
<td>NFL, pg/mL</td>
<td>1155</td>
<td>678</td>
<td>1491</td>
</tr>
</tbody>
</table>


In a second level analysis, each group was further divided in two smaller groups, Group 1.1 (n = 24), Group 1.2 (n = 36), Group 2.1 (n = 20), and Group 2.2 (n = 19). The group of participants with generally more pathological biomarker values (Group 2) was split in one subgroup (Group 2.2) with more pathological values of tau (T-tau, P-tau) and FABP3, and one (Group 2.1) with more pathological values of Aβ42 and NFL and tendencies to higher levels of YKL-40 (Cohen’s d = .28) (Table 5). Both Group 2.1 and Group 2.2 participants exceeded pathological thresholds for P-tau and T-tau referred above. Only Group 2.1 had an Aβ42 concentration satisfying usual criteria for amyloid positivity, although the concentration in Group 2.2 was also close to the pathological threshold. Group 2.1 participants also had mean NFL values of 1831 pg/mL, close to an established cut-off value of 1850 for this age-range (Yilmaz, et al., 2017).

The participants in groups 1.1 and 1.2 had less pathological biomarker values than the participants in 2.1 and 2.2, but could still be differentiated. Group 1.1 had more pathological levels of FABP3 and tau, while Group 1.2 had more pathological levels of Aβ42 and slightly higher levels of NFL (Cohen’s d = .37) (Table 5). There were no differences in YKL-40 between Group 1.1 and 1.2. Group 1.1 had mean tau levels above the pathological thresholds. These findings parallel the results from the comparison between Group 2.1 and 2.2, with higher
tau- and FABP characterizing one group and more pathological levels of Aβ42 characterizing the other, although it must be noted that Group 1.2 still was not close to amyloid positivity.

**Table 5.** Differences between four subgroups based on clustering analysis

<table>
<thead>
<tr>
<th></th>
<th>Group 1.1 (n=24)</th>
<th>Group 1.2 (n=36)</th>
<th>d</th>
<th>Group 2.1 (n=20)</th>
<th>Group 2.2 (n=19)</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>SD</td>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Aβ42</td>
<td>888</td>
<td>136</td>
<td>762</td>
<td>139</td>
<td>0.91</td>
<td>472</td>
</tr>
<tr>
<td>FABP3</td>
<td>5.45</td>
<td>1.59</td>
<td>3.54</td>
<td>0.99</td>
<td>1.52</td>
<td>5.31</td>
</tr>
<tr>
<td>P-tau</td>
<td>63</td>
<td>14</td>
<td>47</td>
<td>13</td>
<td>1.21</td>
<td>67</td>
</tr>
<tr>
<td>T-tau</td>
<td>379</td>
<td>88</td>
<td>275</td>
<td>93</td>
<td>1.15</td>
<td>431</td>
</tr>
<tr>
<td>YKL-40</td>
<td>205901</td>
<td>80065</td>
<td>211920</td>
<td>58047</td>
<td>0.09</td>
<td>289103</td>
</tr>
<tr>
<td>NFL</td>
<td>1004</td>
<td>317</td>
<td>1256</td>
<td>827</td>
<td>0.37</td>
<td>1831</td>
</tr>
</tbody>
</table>

Means are based on raw data. CSF concentrations of biomarkers are measured in pg/mL. d = Cohen’s d. M=mean. SD=standard deviation. Aβ42= β-amyloid 1-42. FABP3= fatty acid binding protein 3. P-tau = phosphorylated tau. T-tau = total tau. YKL-40= Chitinase-3-like protein 1. NFL = neurofilament light. **Bold** indicates Cohen’s d ≥ .50, *italics* indicates Cohen’s d ≥ .20.

Group 2.1 was significantly older than groups 1.1 and 2.2, otherwise there were no differences in age, sex, years of education, or cancer morbidity between the four clustering-based subgroups (Table S3).

### 3.4 Differences in hippocampal atrophy between the biomarker subgroups

In the full sample, GAMM with time since baseline (interval) as predictor, including random effects for intercept, showed that hippocampal volume was significantly reduced over time in a slightly accelerated fashion (edf = 2.2, F = 168.5, p < 2e-16). Significant atrophy was seen for both groups in the two-cluster solution (Group 1: edf = 1.8, F = 109.7, p < 2e-16; Group 2: edf = 1.1, F = 132.8, p < 2e-16). Directly comparing hippocampal change over time between the groups from the two-cluster solution, there were no significant differences in hippocampal volume loss over time (F = 1.0, p = .32), and no offset difference (p = .63) (see Figure 3). We repeated the analyses, comparing volume change pairwise between the groups from the four-cluster solution, finding no significant effects (all p’s ≥ .20).
3.5 Memory differences between the biomarker subgroups

In the full sample, GAMM with time since baseline (interval) as predictor, including random effects for intercept, controlling for baseline age and sex, showed that memory scores, measured as number of words recalled, followed an inverted U-shaped trajectory over the 6.8 year interval (edf = 2.9, F = 11.65, p = 7.22e-7), see Figure 4. The initial increase is likely due to practice effects. Thus, we re-ran the analyses, also controlling for practice effects using number of test sessions completed. This showed a linear negative effect of interval on memory score (edf = 1.0, F = 12.06, p = .0006), and a positive but gradually reduced effect of number of test sessions as a proxy for practice effects (edf = 2.0, F = 21.24, p = 6.55e-8).
Figure 4. Longitudinal change in memory function in the full sample
Upper panel: Memory scores over time in the full sample. Left bottom panel: Estimated memory performance over time corrected for practice effects. Bottom right panel: Estimated practice effects plotted over time.

After having established the trend for the change in memory scores over time in the total sample, we tested whether memory differed between biomarker groups in terms of intercept or slope, covarying for baseline age and sex (Figure 5, Table 6). Comparing Group 1 and Group 2, we found a significant difference in slope (edf = 1.0, F = 4.7, p = .030) if practice effects as indexed by number of follow up test sessions were included as covariate, and a trend if not (p = .098). Plotting the results showed more memory decline in the group (Group 2) with the more pathological biomarkers.
Pairwise comparisons between the memory trajectories from each of the groups in the four-cluster solution showed significantly more memory decline for Group 2.2 compared to Group 1.2. (p = .014). Group 2.2 was the group with the highest levels of tau and FABP3, while Group 1.2 was the group with the lowest levels of the same biomarkers. No other differences between groups in memory trajectories were seen. The proportion of participants followed up with cognitive assessment did not differ between the four cluster-based biomarker groups at any of the seven time points (p=0.76).
Table 6. Change in memory as a function of biomarker group

<table>
<thead>
<tr>
<th></th>
<th>edf</th>
<th>F</th>
<th>P slope</th>
<th>P intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2 group model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>2.8</td>
<td>11.12</td>
<td>2.37e-06</td>
<td></td>
</tr>
<tr>
<td>Group 2 vs 1</td>
<td>1.3</td>
<td>3.65</td>
<td>0.098¹</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>4 group model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1.1 vs 1.2</td>
<td>1.0</td>
<td>2.78</td>
<td>0.10</td>
<td>0.74</td>
</tr>
<tr>
<td>Group 1.2 vs 2.1</td>
<td>1.0</td>
<td>2.20</td>
<td>0.14</td>
<td>0.13</td>
</tr>
<tr>
<td>Group 1.2 vs 2.2</td>
<td>1.0</td>
<td>6.05</td>
<td>0.01</td>
<td>0.16</td>
</tr>
<tr>
<td>Group 2.1 vs 2.2</td>
<td>1.0</td>
<td>0.76</td>
<td>0.38</td>
<td>0.27</td>
</tr>
</tbody>
</table>

GAMMs were run to test effects of biomarker group on changes in memory performance over time. Changes in memory performance were tested against time and then it was tested whether the effect of time on memory differed between biomarker groups. Baseline age and sex were used as covariates of no interest. Edf: effective degrees of freedom (a measure of deviation from linearity). ¹ p < .030 if practice effects were corrected for.

3.6 Memory and hippocampus changes in NIA-AA defined biomarker groups

As an alternative to the clustering approach, we classified participants as AD (A+/T+, n = 19), A-T+ (n=27), and normal AD biomarkers (A-/T-, n = 32), based on cut-offs defined above. We tested if the groups A+T+ or A-T+ showed different changes in hippocampal volume or memory over 6.8 years compared to the A-T- group. For neither hippocampal volume (edf = 1.0, F = 0.59, p = .44, n = 130 observations) nor memory (edf = 1.0, F = .04, p = .84, n = 278 observations) were significant slope differences seen between A+T+ vs. the A-T- group. There was a tendency for A-T+ participants to show more memory decline over time compared to the A-T- group (edf = 1, F = 2.9, p = .089). No effect of A-T+ on hippocampal volume change was found (p = .16). To assess whether the lack of difference in memory trajectories between the AD group based on biomarkers as described by NIA-AA and the normal AD biomarker group was due to too short follow up interval or too small sample, we ran power simulations based on the observed effects (see Supplemental Information for details). These simulations showed than even if we follow the participants over 15 years, we would not find a significant difference in memory slope between the AD group defined by the NIA-AA criteria vs. the normal AD biomarker group (12% power at α = .05 with 15 time points spanning 15 years) (Figure 6). This shows that it is unlikely that the biomarker defined AD group and the normal AD biomarker
group will experience different changes in memory function over the next 15 years. We also tested how increasing the number of participants included in the analyses would affect the power to detect slope differences between the groups. These simulations demonstrated that with a sample size of 1050 participants, power to detect an effect was no more than 13%. Thus, while the clustering approach was able to define subgroups of participants with different biomarker profiles showing significant differences in memory slope, using the NIA-AA criterion for AD, we were not able to detect differences, and this is highly unlikely to be due to short follow-up interval or a limited sample size.

![Figure 6](image)

**Figure 6** Memory slope in AD vs no-AD
Participants were divided in an “AD” group based on amyloid and tau (A+/T+) and a no-AD group (A-/T-). Left panel shows the estimated differences in memory trajectories over 15 years. Middle panel shows how power to detect a slope difference between AD and no-AD increases as a function of number of follow-ups. Right panel shows how power increase as a function of sample size.

### 4. Discussion

This study on biomarkers in cognitively unimpaired older adults has three key findings. First, clustering analysis of biomarkers showed that the novel CSF biomarkers NFL, FABP3, and YKL-40 clustered with T-tau and P-tau, whereas Aβ42 was separated out in an independent cluster. Second, clustering analyses of participants identified two main biomarker profiles, where one biomarker group had more abnormal levels of all biomarkers compared to the other biomarker group. At the 4-cluster level, the group with more pathological biomarkers was split in one group characterized by tauopathy and FABP3 and one group by brain β-amyloidosis, NFL, and YKL-40. Third, the group with tauopathy and FABP3 showed worse development of memory function over 6.8 years compared to a group with less pathological biomarker levels. The clustering-based classification of the participants yielded better predictions of memory decline across the subsequent seven years than a canonical classification based on amyloid and tau.
Relationships between CSF biomarkers

We assessed the relationship between the established AD biomarkers Aβ42, reflecting amyloid deposition, T-tau, reflecting neuro-axonal degeneration, and P-tau, reflecting phosphorylation state of tau and possibly neurofibrillary tangle pathology, and the novel biomarkers YKL-40, reflecting neuro-inflammation, FABP3, reflecting neuronal damage, and NFL, reflecting axonal injury. Interestingly, the cluster analyses revealed a principal divide between Aβ42 and the rest of the biomarkers. While T-tau, P-tau, YKL-40, FABP3, and NFL clustered together at the highest level, Aβ42 was separated in a single cluster. Thus, in cognitively unimpaired older adults, the less established CSF biomarkers clustered with P- and T-tau, while showing no relationships to Aβ42. Such a divide between Aβ42 and other CSF biomarkers including FABP3 and P-tau has also been shown using clustering analyses in a cohort with individuals from the entire AD cognitive continuum (Harari, et al., 2014). Further, lack of relationship between Aβ42 and the Tau-proteins in cognitively unimpaired adults has also been found in previous studies (Roe, et al., 2013,Xiong, et al., 2016), although weak to moderate negative correlations have also been reported (Bos, et al., 2019,Pettigrew, et al., 2015,Soldan, et al., 2019). Non-existent relationships between Aβ42 and YKL-40 (Olsson, et al., 2013a), NFL (Bruno, et al., 2012), and FABP3 (Olsson, et al., 2013b) in cognitively unimpaired individuals are also in agreement with previous studies, but weak negative (Alcolea, et al., 2015a,Bos, et al., 2019) and weak positive (Alcolea, et al., 2015a,Kern, et al., 2019) correlations have also been shown in some larger studies. The cluster analysis results show that these less established biomarkers cluster with the Tau proteins, independently of amyloid pathology.

The very high correlation between T-tau and P-tau was expected, being consistent with previous studies of cognitively unimpaired adults (Blennow, et al., 1995,Bos, et al., 2019,Soldan, et al., 2019). The strength of this correlation (r = .96) suggests that these two CSF tau-markers are statistically collinear in cognitively unimpaired older adults. The neuronal injury biomarkers, NFL and FABP3, were positively correlated both with each other and also with T-tau and P-tau, supporting their role as neurodegeneration biomarkers even in cognitively unimpaired older adults. This finding is in agreement with previous studies of cognitively unimpaired adults showing positive correlations between Tau-proteins and NFL (Bos, et al., 2019,Kern, et al., 2019,Melah, et al., 2016), and FABP3 (Olsson, et al., 2013b), respectively. NFL was the first biomarker to separate from the cluster with Tau biomarkers, supporting the hypothesis that CSF NFL provide information on neurodegeneration that is at least partly different from CSF T-tau (Mattsson, et al., 2016). Interestingly, the neuroinflammation biomarker YKL-40 was also
positively correlated with all the neurodegeneration biomarkers, indicating a link between neuroinflammation and neurodegeneration in aging. Previous data showing positive correlations between YKL-40 and T-tau and P-tau in cognitively unimpaired adults (Alcolea, et al., 2015a,Bos, et al., 2019,Melah, et al., 2016,Olsson, et al., 2013a), and in neurodegenerative diseases (Craig-Schapiro, et al., 2010,Hall, et al., 2018,Nordengen, et al., 2019,Wennstrom, et al., 2015), and studies finding associations between elevated YKL-40 and white matter degeneration (Racine, et al., 2019), brain atrophy (Alcolea, et al., 2015b,Alcolea, et al., 2017,Janelidze, et al., 2018,Swanson, et al., 2016), and cognitive function (Bos, et al., 2019,Janelidze, et al., 2018,Sala-Llonch, et al., 2017), provide further support for this link both in aging and neurodegenerative diseases. The association between YKL-40 and Tau has been shown in both A+ and A- cognitively normal individuals (Alcolea, et al., 2015a), suggesting that the link between neurodegeneration and neuroinflammation is independent of amyloid deposition. Positive correlations between YKL-40 and NFL (Bos, et al., 2019,Melah, et al., 2016) in cognitively normal adults have also been reported previously, whereas this is, to our knowledge, the first study to explore the relationship of FABP3 to YKL-40 and NFL in a cognitively unimpaired population. Positive correlations of FABP3 with YKL-40 and NFL have, however, previously been reported in populations including patients (Bjerke, et al., 2011,Harari, et al., 2014).

Grouping of participants based on biomarker profiles
Clustering analyses, based on objective biomarker measures and blind to any clinical evaluation, revealed one group, consisting of 39% of the total sample, with more abnormal concentrations of all biomarkers. Although normative data are not available for all the biomarkers, this group had pathological levels of Aβ42, T-tau, and P-tau according to previously established criteria (Hansson, et al., 2006). Further clustering analyses separated a group of participants with non-pathological biomarker concentrations (Group 1.2), consisting of 36% of the sample. Studies show that there may be a proportion of older adults that never develops amyloid or neurodegenerative pathology (Jack, et al., 2014,Khachaturian, et al., 2004), and future studies on such groups of very low-risk older adults may give cues on how to prevent development of various brain pathologies (Vemuri, 2018). However, the proportion of individuals without pathology may be over-estimated in our sample, as we did not include cases with clinical diagnoses. Also, we cannot exclude that some of the participants with a non-pathological biomarker profile may develop pathological biomarkers later.
The three remaining groups represented different patterns of increased or pathological biomarker values. Group 1.1 was the least pathological of these, showing evidence for slight tauopathy. The most pathological group in the two-cluster solution was divided in one group with more pathological values of Tau (T-tau, P-tau) and FABP3 (Group 2.2), and one with more pathological values of Aβ42 and NFL, and tendencies to higher levels of neuroinflammation (YKL-40) (Group 2.1). Although the mean tau levels in Group 2.1 were clearly lower than in Group 2.2, they still exceeded pathological thresholds. Accordingly, Group 2.1 participants on average satisfied the NIA-AA criterion for AD (Jack, et al., 2018). The participants in Group 2.1 also had mean NFL values close to an established cut-off value of 1850, and slightly higher concentrations of YKL-40, suggesting ongoing neuroinflammation and axonal degeneration in addition to brain β-amyloidosis and tau pathology. Emerging evidence suggests that neuroinflammation in concert with AD neuropathology may contribute to the development of clinical symptoms (Craig-Schapiro, et al., 2010, Heneka, et al., 2015, Merluzzi, et al., 2018), possibly through contributing to neurodegeneration (Alcolea, et al., 2015b, Heneka, et al., 2015, Janelidze, et al., 2018).

The tauopathy found in three of the four identified groups may partly be age-related (Crary, et al., 2014, Lowe, et al., 2018), such as in primary age-related tauopathy (PART), although preclinical phases of other tauopathies cannot be excluded (Arendt, et al., 2016). As Tau was measured in CSF, we could not assess the patho-anatomical location, i.e. whether it is spread outside the medial temporal lobe. Group 2.2 was characterized by a neurodegeneration biomarker pattern with tauopathy and elevated FABP3, and was the group with highest levels of Tau and FABP3. Such a pattern of neurodegeneration can represent normative, age-expected brain changes. Group 2.2 may also, according to some systems, be classified as SNAP (Jack, et al., 2016b), in which e.g. clinically silent cerebral microvascular disease, hippocampal sclerosis or aging-expected processes could be responsible for the neurodegeneration. It should also be noted that this group had Aβ42 levels in a grey zone around the pathological threshold for amyloid positivity.

The present results suggest that patients can be divided into subgroups based on their biomarker profiles also beyond amyloid and tau. This yields more extensive information about patients than what can be obtained by using biomarkers in isolation. Interestingly, all the six CSF biomarkers differed between at least two of the groups with a relatively large effect size, suggesting that all have contributed to the clustering results. We did not attempt to cluster the
participants based on a subset of the CSF biomarker to test if any was redundant. Except p-tau and t-tau, which are statistically almost collinear, it seems that inclusion of all the biomarkers contributes to the different biomarker profiles of the subgroups. Although further studies are needed before firm conclusions can be reached, clustering may be a promising approach to identify patients with various biomarker profiles for clinical trials, intervention studies, and in the clinic to improve diagnosis and prognosis.

This biomarker-based grouping of the participants suggests that linear staging of CSF biomarkers, where the biomarkers become abnormal at different times in an ordered sequence, does not apply to the present data. It is possible that a well-defined clinical endpoint can follow a fixed chain of events in an orderly fashion. In cognitively unimpaired older adults, however, biomarkers do not seem to adhere to a fixed linear staging. Unfortunately, we do not have longitudinal data on the biomarkers in combination with different clinical endpoints, which would be necessary for proper staging of the biomarkers.

**Biomarker profiles in prediction of hippocampal volume change and memory decline**

The full sample showed an inverted U-shaped trajectory of memory scores over the 6.8 year interval since the baseline testing. As practice effects are well-known to increase performance on memory tests in longitudinal studies (Ronnlund, et al., 2005), we attempted to tear apart real change in memory from practice-induced inflation of the scores. This analysis showed a linear decline in the corrected scores, accompanied by a decelerating increase due to repeated test exposure. Testing the difference in memory trajectories between the group with the most pathological biomarkers versus the group with the least pathological biomarkers, the first group showed a slightly worse memory development over the examination interval. Examining this pattern in more detail in the four-cluster solution, the group with most tauopathy and highest FABP3 showed significantly more memory decline compared to the group with normal biomarker levels. This indicates that a high degree of neuronal damage is the biomarker feature most predictive of memory decline in cognitively unimpaired older adults. Earlier clustering studies of cognitively normal adults have also found that subgroups with more neurodegeneration or tauopathy show greater rates of cognitive decline, e.g. greater memory decline in a subgroup mainly characterized by lower Aβ42 and higher P-tau (Racine, et al., 2016), and greater global cognitive decline in a subgroup characterized by more brain atrophy, more white matter hyperintensities, lower Aβ42, higher T-tau, and higher P-tau (Nettiksimmons, et al., 2010). FABP3, YKL-40, or NFL have never been used for clustering of
cognitively unimpaired adults. Yet, a recent study clustering individuals from the entire cognitive continuum using CSF NFL, YKL-40, and the core AD biomarkers reported that a subgroup characterized by high T-tau and P-tau, but not Aβ42 levels, included almost 50% of all patients with mild cognitive impairment and AD dementia, respectively (Toschi, et al., 2019). High levels of each of the three less established biomarkers have, however, been associated with cognitive decline or development of cognitive impairment in cognitively normal adults (Bos, et al., 2019,Harari, et al., 2014,Kern, et al., 2019,Sala-Llonch, et al., 2017).

Moreover, Group 2.2 may have brain β-amyloidosis. The predictive value of brain β-amyloidosis for later cognitive or clinical symptoms is however controversial (Morris, et al., 2018). Although brain β-amyloidosis may be a risk factor for cognitive decline (Hedden, et al., 2013), up to 40% of cognitively unimpaired older adults have brain β-amyloidosis (Jansen, et al., 2015), depending on the age of the participants. Likely, amyloidosis has to be accompanied with neurodegeneration or tau pathology in order to result in dementia and cognitive decline (Burnham, et al., 2016,Desikan, et al., 2012,Merluzzi, et al., 2018,Soldan, et al., 2019). Therefore, we cannot exclude that the combination of grey zone amyloid positivity and neuronal damage is responsible for more memory decline in Group 2.2. However, the A-T+ group showed a tendency toward more memory decline compared to the group with normal AD biomarkers, suggesting that the association between neurodegeneration and memory decline in this sample is independent of amyloid pathology. Furthermore, Group 2.1 with lowest Aβ42 (i.e. most amyloid pathology), and also tauopathy and axonal degeneration, and Group 1.1 with slight tauopathy, showed only age-expected changes in memory performance over time. This finding also underscores the fact that older adults can have good cognitive function for their age, and show age-expected changes in memory function over several years, despite biomarker profiles indicating amyloid, tau and/or neurodegeneration pathology.

Interestingly, comparing memory change between the group with AD according to the NIA-AA A+/T+ criterion with the group with normal AD biomarkers (A-/T-), we did not observe any difference. This finding differs from several previous studies reporting that cognitively normal individuals with both amyloid and tau pathology show accelerated cognitive decline compared to those with one or none of these pathologies (Desikan, et al., 2012,Soldan, et al., 2019). Actually, simulations showed that even if we follow the participants for 15 years after baseline, or increase the sample size to above 1000 participants, it is unlikely that we would see a difference in memory slope between the biomarker defined AD group and the non-AD group.
Thus, while the clustering approach was able to define subgroups of participants with different biomarker profiles that showed more memory decline over time, no differences in memory outcome was seen using a simple AD vs non-AD dichotomy based on Aβ42 and P-tau alone. These results demonstrated that the clustering approach, taking advantage of multiple biomarkers beyond amyloid and tau, clearly outperformed the NIA-AA AD biomarker classification system in prediction of memory decline.

There were no differences between any of the clustering-based biomarker groups or the NIA-AA defined groups in hippocampal volume trajectories, suggesting that none of these biomarker profiles were associated with more than age-related hippocampal atrophy. Others have, however, reported that A+T+ cognitively normal individuals show accelerated hippocampal atrophy compared to A-T- and A-T+ individuals (Gordon, et al., 2016). The relationship between novel biomarkers and medial temporal lobe atrophy is less studied, but higher NFL (Mattsson, et al., 2016, Pereira, et al., 2017), YKL-40 (Alcolea, et al., 2015b, Swanson, et al., 2016), and FABP3 (Desikan, et al., 2013) have been associated with medial temporal lobe atrophy in populations including both cognitively unimpaired and impaired individuals. We have previously shown an association between higher NFL levels and higher hippocampal atrophy rates in individuals from the same cohort (Idland, et al., 2017), and one study found no relationship between YKL-40 and hippocampal volume in cognitively unimpaired adults (Melah, et al., 2016). However, the relationship between medial temporal lobe atrophy and FABP3 has never been assessed in a cognitively unimpaired population. Previous research has shown that cognitively normal individuals with amyloid pathology show steeper memory decline if they also have pathological hippocampal volumes (Bilgel, et al., 2018, Burnham, et al., 2016). Accordingly, we speculate that the key to understand why some biomarker groups with amyloid and/or tau positivity only showed age-expected memory decline is that these participants did not show higher than age-expected hippocampal atrophy. Although followed for up to 6.8 years, there were no differences between any of the biomarker groups in hippocampal volume trajectories. Differences would have been expected if the sample also had included participants showing cognitive impairment such as in Alzheimer’s dementia. We propose that older adults may uphold age-expected cognitive function for many years, even when harboring pathological biomarker profiles, as long as hippocampal atrophy is within the age-expected range.
**Strengths and limitations**

The inclusion of cognitively unimpaired adults only, constitutes both a strength and a weakness, causing the sample probably to be more homogenous than the general population, which likely affected both the clustering of biomarkers and the clustering of participants. Thus, the conclusions drawn are valid for cognitively unimpaired older adults only – the clusters may be different if examined in populations of participants with mild cognitive impairment or dementia. On the other hand, the sample consisted of surgical patients including some patients who had cancer surgery, which also may reduce the generalizability of the results, although the consequences of this is difficult to assess. Nevertheless, we did not find any differences in cancer morbidity between the biomarker groups. Strengths of our work include measurement of both established and novel biomarkers, and multiple longitudinal measures of hippocampal volume and memory over 6.8 years. We also used a data-driven method, rather than defined cut-off values, to assess relationships between biomarkers.

**5. Conclusion**

Here we show that CSF biomarkers of AD pathophysiology can be grouped in superordinate clusters, and that Aβ42 is the biomarker with the least connections to other established as well as more novel CSF biomarkers. Using a large collection of CSF biomarkers enabled us to identify subgroups of participants with different biomarker profiles. This clustering-based grouping of participants outperformed biomarker profiling based on the NIA-AA AD classification system in predicting memory change over 6.8 years. The analyses of changes in memory function further showed that older adults may uphold age-expected cognitive function and hippocampal integrity even when harboring abnormal biomarker profiles, such as tauopathy, underscoring the complex relationship between cognitive function, maintenance, resilience and brain health in aging (Stern, et al., 2018). Understanding the conditions for maintained cognitive function in aging despite various types of brain changes will be a major task for future research.
Acknowledgements

The authors would like to thank Knut Engedal, MD, PhD (University of Oslo and Vestfold Hospital Trust) and Anette Hylen Ranhoff, MD, PhD (Diakonhjemmet Hospital and Kavli Research Center for Geriatrics and Dementia) for general support, Gry Torsæter Dahl, MD (Diakonhjemmet Hospital), Randi Støen, MD (Oslo University Hospital), Arne Myklebust, MD (Oslo University Hospital) and Dagfinn Tore Kollerøs, MD (Oslo University Hospital) for collection of data, and Asta Kristine Håberg MD, PhD (St Olav's University Hospital) for planning the baseline MRI protocol. The authors would also like to thank the study participants and the contributions of the Department of Gynecology, the Department of Urology, and the Department of Anesthesiology at Oslo University Hospital, the Department of Orthopedic Surgery and the Department of Anesthesiology at Diakonhjemmet Hospital in Oslo, Norway.
References


Figure S1

172 participants tested at baseline

Excluded from baseline
- Diagnosed dementia or mild cognitive impairment - 5
- Cognitive impairment according to medical records - 2
- Referred/offered referral for cognitive assessment - 4
- Parkinson's disease - 2
- Other neurological disorders - 5
- Not cognitively unimpaired according to defined criteria - 38
- Missing CSF data - 17

99 participants with available CSF data for all biomarkers
- Number of cognitive assessments:
  - 7: n=55
  - 6: n=11
  - 5: n=4
  - 4: n=6
  - 3: n=6
  - 2: n=5
  - 1: n=12
- Number of magnetic resonance imagings:
  - 4: n=33
  - 3: n=23
  - 2: n=17
  - 1: n=12

CSF = cerebrospinal fluid.
**Table S1. Demographics**

<table>
<thead>
<tr>
<th></th>
<th>Included participants (n=99)</th>
<th>Excluded participants (n=73)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at baseline, years</td>
<td>72 (68 to 78)</td>
<td>74 (70 to 81)</td>
<td>0.11</td>
</tr>
<tr>
<td>Sex, male</td>
<td>50 (51)</td>
<td>35 (48)</td>
<td>0.74</td>
</tr>
<tr>
<td>Education, years</td>
<td>14 (12 to 17)</td>
<td>13 (10 to 17)</td>
<td>0.22</td>
</tr>
<tr>
<td>MMSE score, baseline</td>
<td>29 (28 to 30)</td>
<td>29 (28 to 30)</td>
<td><strong>0.03</strong></td>
</tr>
<tr>
<td>CERAD, Word List Recall score</td>
<td>6 (5 to 8)</td>
<td>5.5 (4 to 7)</td>
<td><strong>&lt;.001</strong></td>
</tr>
<tr>
<td>CSF Aβ42, pg/mL</td>
<td>731 (512 to 866)</td>
<td>763 (524 to 852)</td>
<td>0.98</td>
</tr>
<tr>
<td>CSF FABP3, pg/mL</td>
<td>4.56 (3.36 to 5.93)</td>
<td>4.54 (3.23 to 6.18)</td>
<td>0.98</td>
</tr>
<tr>
<td>CSF P-tau, pg/mL</td>
<td>59 (46 to 75)</td>
<td>55 (45 to 68)</td>
<td>0.31</td>
</tr>
<tr>
<td>CSF T-tau, pg/mL</td>
<td>347 (272 to 486)</td>
<td>359 (267 to 427)</td>
<td>0.80</td>
</tr>
<tr>
<td>CSF YKL-40, pg/mL</td>
<td>225210 (175208 to 280877)</td>
<td>198629 (157963 to 266920)</td>
<td>0.13</td>
</tr>
<tr>
<td>CSF NFL, pg/mL</td>
<td>1026 (794 to 1482)</td>
<td>1189 (848 to 1618)</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Values are median (IQR) and N (%). aN=72, bN=51, cN=49. MMSE = Mini Mental Status Examination. CERAD = Consortium to Establish a Registry for Alzheimer’s disease. CSF = cerebrospinal fluid. Aβ42 = β-amyloid 1-42. FABP3 = fatty acid binding protein 3. P-tau = phosphorylated tau. T-tau = total tau. YKL-40 = Chitinase-3-like protein 1. NFL = neurofilament light. P-values are calculated using Mann-Whitney U-test and Chi-Square Test. **Bold** indicates p < .05.
Inclusion/exclusion criteria based on neurological diagnoses and magnetic resonance imaging findings at baseline or occurring through follow-up in the cohort:

Excluded from baseline

- Diagnosed dementia or mild cognitive impairment.
- Cognitive impairment according to medical records.
- Referred/offered referral for cognitive assessment based on a cognitive assessment in the study.
- Parkinson’s disease.
- Epilepsy at baseline or developed later.

Excluded from longitudinal analyses from the time of debut

- Brain infarction on magnetic resonance imaging affecting cortical areas in cerebrum or hippocampi.
- Stroke affecting cortical areas, hippocampi, or with unknown localization.
- Previous meningitis or encephalitis with sequelae.
- Head trauma with permanent neurological sequela or known structural brain abnormality.
- Brain parenchymal tumors.
- Transitoric global amnesia.
- Neuroborreliosis.

Included

- Previous meningitis or encephalitis without sequelae (n=3).
- Transient ischemic attack (n=7).
- Epilepsy at younger age, but not at baseline in the study or later, and not using any antiepileptic medications (n=2).
- Head trauma without neurological sequela or known structural brain abnormality (n=5, including one participant with stroke not affecting cortical areas or hippocampi and one who has had epilepsy at younger age)
- Brain infarctions localized subcortically (but not affecting hippocampi), in the brain stem, or the cerebellum (n=4).
- Stroke not affecting cortical areas or hippocampi (n=3)
- Intracranial tumors not affecting the brain parenchyma, e.g. meningeomas (n=4).

Only three participants were included at baseline (in clustering analyses) and excluded from longitudinal analyses from the time of debut of disease/lesion. These included one participant with neuroborreliosis (excluded from a time point eight months before debut of symptoms), one participant with a craniopharyngioma (excluded from the time when the tumor was evident on MRI), one participant with a cortical brain infarction (excluded from the time of debut).
Table S2. Differences between two subgroups based on clustering analysis

<table>
<thead>
<tr>
<th></th>
<th>Group 1 (n=60)</th>
<th>Group 2 (n=39)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at baseline, years</td>
<td>71 (67 to 77)</td>
<td>75 (69 to 80)</td>
<td>0.10</td>
</tr>
<tr>
<td>Sex, male</td>
<td>32 (53)</td>
<td>18 (46)</td>
<td>0.49</td>
</tr>
<tr>
<td>Education, years</td>
<td>14 (12 to 17)</td>
<td>12 (11 to 16)</td>
<td>0.10</td>
</tr>
<tr>
<td>Cancer at baseline</td>
<td>17 (28)</td>
<td>11 (28)</td>
<td>0.99</td>
</tr>
<tr>
<td>Metastatic cancer at baseline</td>
<td>3 (5)</td>
<td>5 (13)</td>
<td>-</td>
</tr>
</tbody>
</table>

Values are median (IQR) and N (%). P-values are calculated using Mann-Whitney U-test and Chi-Square Test.
Table S3. Differences between four subgroups based on clustering analysis

<table>
<thead>
<tr>
<th></th>
<th>Group 1.1 (n=24)</th>
<th>Group 1.2 (n=36)</th>
<th>Group 2.1 (n=20)</th>
<th>Group 2.2 (n=19)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at baseline, years</td>
<td>70 (67 to 74)*</td>
<td>72 (67 to 78)</td>
<td>77 (70 to 83)</td>
<td>75 (68 to 77)</td>
<td>0.047</td>
</tr>
<tr>
<td>Sex, male</td>
<td>12 (50)</td>
<td>20 (56)</td>
<td>10 (50)</td>
<td>8 (42)</td>
<td>0.82</td>
</tr>
<tr>
<td>Education, years</td>
<td>15.5 (13 to 17)</td>
<td>14 (12 to 16)</td>
<td>12 (11 to 16)</td>
<td>13 (10 to 16)</td>
<td>0.25</td>
</tr>
<tr>
<td>Cancer at baseline</td>
<td>6 (25)</td>
<td>11 (31)</td>
<td>6 (30)</td>
<td>5 (26)</td>
<td>0.96</td>
</tr>
<tr>
<td>Metastatic cancer at baseline</td>
<td>0 (0)</td>
<td>3 (8)</td>
<td>2 (10)</td>
<td>3 (16)</td>
<td>-</td>
</tr>
</tbody>
</table>

Values are median (IQR) and N (%). P-values are calculated using Kruskal-Wallis Test, post-hoc Mann-Whitney U test, and Chi-Square Test. *P <0.05 compared to 2.1.
Power Analysis

Background

This document shows the results of power simulations for the interaction between AD classification and the effect of time on the number of words recalled.

We have a dataframe `NBM_data` with the following variables of interest. The R data type is shown in parentheses:

- Tiord_utsatt: Number of words recalled, between 0 and 10 (`numeric`).
- Interval_mem: Time since baseline in years (`numeric`).
- AD_OF: AD classification, 0 or 1 (`order.factor`).
- BL_Age: Age at baseline in years (`numeric`).
- Sex (`factor`).
- ID: subject identifier (`factor`).

We define the following generalized additive mixed model. The response is the number of words recalled, and it has a smooth term for time since baseline (`ti(Interval_mem)`).

Since AD_OF is an ordered factor, the term `ti(Interval_mem, by=AD_OF)` shows how the smooth term for interval differs between subjects classified with AD compared to subjects not classified with AD. The argument `random = list(ID = ~1)` defines a random intercept per subject ID.

```r
library(mgcv)
tiord_model_AD <- Tiord_utsatt ~ ti(Interval_mem) + AD_OF + ti(Interval_mem, by=AD_OF) + BL_Age + Sex
fit_tiord_model_AD <- gamm(tiord_model_AD, data = NBM_data, random = list(ID = ~1))
```

Looking at the model output, we see that there is a significant effect of Interval_mem on Tiord_utsatt, but the interaction between Interval_mem and AD_OF has p-value 0.841. That is, we cannot conclude that there is a difference between the trajectories of subjects classified with AD compared to subjects not classified with AD.

```r
summary(fit_tiord_model_AD$gam)

## Family: gaussian
## Link function: identity
##
## Formula:
## Tiord_utsatt ~ ti(Interval_mem) + AD_OF + ti(Interval_mem, by = AD_OF) + BL_Age + Sex
##
## Parametric coefficients:
##                      Estimate  Std. Error   t value Pr(>|t|)
## (Intercept)      17.9202048  2.18547077  8.20000  9.80e-15 ***
## AD_OF.L          -0.3372612  0.31331246  -1.07573   0.28339
## BL_Age           -0.1486367  0.02931244  -5.07129  7.37e-07 ***
## SexMale          -0.1419497  0.41429321  -0.34282   0.73161
## ---
```

---
Studying the shape of the fits, we see the following:

- There is a strong positive effect of time the first 3 or 4 years, until it starts to decline (left part of plot below).
- Subjects classified with AD seem to have a slightly weaker increase and stronger decline compared to subject with AD (right part of plot below), when considering the point estimate (solid line). However, the confidence intervals are wide, and include both a strong negative interaction and a strong positive interaction. We thus cannot rule out any of these scenarios based on this model and data.
Increasing the number of timepoints

The interaction effects points in the same direction as the solution with 4 cluster, i.e., that the pathological cases have a weaker increase with time. We can therefore hypothesize that the true effect of AD is on the order of magnitude seen by the point estimate above, and perform a power analysis to investigate over how many timepoints we would have to follow the subjects in order to obtain sufficient statistical power to conclude that there is a difference, at a 0.05 significance level.

To perform the power analysis, we extrapolate beyond the time intervals given in our dataset. The plot below shows how the smooth terms diverge between subject with and without AD.

The simulation goes as follows:

- For each subject, we add a given number of new time intervals at one-year intervals beyond the values in the current data, until the subject is 90 years old. Then the following steps are repeated in a Monte Carlo simulation:
  - Randomly sample an intercept term per subject, with mean zero and standard deviation given by the model fit.
  - Randomly sample a response vector (Tiord_utsatt) from the normal distribution with mean given by the model prediction for the subject’s data and standard deviation given by $\sigma_e$ from the model fit, and add to this the random intercept sampled in the previous step. The response was truncated to lie between 0 and 10.
Fit a model to the randomly sampled data, and record the p-value of the interaction term \( t_i(\text{Interval}_{\text{mem}}, \text{by} = \text{AD}_{\text{OF}}) \).

We start the simulations at 7 timepoints per subject, which is the number of timepoints sampled for most of the subjects. Then additional timepoints are added in increments of 1, up to 15.

The resulting power curve is shown below. It shows that the power increases with the number of timepoints, but even with 15 timepoints we only have about 12% power to detect the interaction, at a 5% significance level.

Increasing the number of subjects

An alternative way of increasing the power would be by increasing the number of subjects, keeping the number of timepoints at the current level. To this end, we performed the following steps in a Monte Carlo simulation:

- New data with a given number of unique IDs were generated by randomly sampling (with replacement) values of Sex, AD_OF, and BL_Age from the columns of the data.
- A random intercept was sampled for each new observation.
- Each new sampled observation was expanded to have observations at 7 timepoints, from 0 to 6 years after baseline.
- A random response vector Tiord_utsatt was generated by sampling from the normal distribution with mean given by the model prediction for the subject’s data and
standard deviation given by $\sigma_e$ from the model fit and adding this to the random intercept for the subject. The response was truncated to lie between 0 and 10.

- A model was fit to the data and the p-value for the interaction term $\text{ti}$(Interval_mem, by = AD_OF) was saved.

Starting at 50 subjects, we simulate an increasing number in increments of 200, up to 1,050 subjects.

The simulated power curve is shown below. Currently our dataset has 51 unique subjects. As the power curve shows, increasing this number to 1,050 would still leave us with an estimated power of no more than 13%.