DEPRESSION & COGNITION IN THE ELDERLY: NEUROIMAGING PERSPECTIVE

Aleksandra Lebedeva

Stockholm 2017
All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by E-Print AB 2017
© Aleksandra Lebedeva, 2017
Depression & Cognition in the Elderly: Neuroimaging Perspective
THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Aleksandra Lebedeva, MD

Principal Supervisor:
Professor Dag Aarsland
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Neurogeriatrics

Co-supervisor(s):
Docent Eric Westman
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Clinical Geriatrics

Professor Lars-Olof Wahlund
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Clinical Geriatrics

Professor Bengt Winblad
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Neurogeriatrics

Opponent:
Professor Stefan Klöppel
University of Freiburg
University Medical Center Freiburg
Center for Geriatric Medicine and Gerontology

Examination Board:
Docent Linda Knutsson
Lund University
Department of Medical Radiation Physics

Professor Albert Leentjens
Maastricht University
Department of Psychiatry and Neuropsychology

Docent Ingvar Karlsson
Göteborg University
Department of Psychiatry and Neurochemistry
Dedicated to all patients and their families
ABSTRACT

This thesis examines the relationship between depression and brain structure in the elderly with (Study I, III) and without (Study II, IV) cognitive impairment (Alzheimer’s disease and mild cognitive impairment). Individuals from four independent cohorts were included. Participants had either a depressive episode (Study II, III) or depressive symptoms, as measured with different depression scales (Study I, IV). Studies I and II have cross-sectional design, and studies III and IV are longitudinal. Main outcomes were cortical thickness of the brain and volumes of different structures (hippocampus, ventral diencephalon, including hypothalamus and corpus callosum), or atrophy rate of the thickness and volumes (Study IV).

We found in all the cohorts that depressive symptoms were associated with cortical thinning in the same region – the left temporoparietal junction. Depression-related thinning was observed in three cohorts (Studies I, IV) in superior temporal cortex and temporal pole. In two non-demented cohorts (Studies II, IV) angular cortex was also involved in depression. Longitudinal analysis revealed that thinning in these regions is secondary to depressive symptoms (study IV). In two cohorts (Study I, II) fusiform cortex was involved in depression. In study IV, we also were able to assess thinning which developed in parallel with depressive symptoms. It covered medial superior frontal cortex and lingual cortex.

The number of depressive episodes was associated with cortical thinning in the left temporal pole in women (Study II) and reduced volume of the right ventral diencephalon in both – men and women (Study III).

We have found moderating effect of gender on the relationship between cortical thickness and depression onset. Women with late-onset depression (>65 years) but not men had the widespread thinning in the prefrontal cortex compared to early-onset depressed.

The volume of the right hippocampus and thickness of the superior frontal cortex were positively associated with a level of global cognition measured with the mini-mental state examination (MMSE) This effect was more pronounced in the subgroup of late-onset depressed (Study II).

The volume of the right ventral diencephalon was associated with cognitive decline (MCI or dementia diagnosis) one year later in the elderly with a depressive episode (study III). Adding baseline MMSE to the classifier increased its accuracy.

Total and phosphorylated tau were associated with cortical thinning in the cluster covering right posterior cingulate cortex and precuneus and cluster covering right parahippocampal and fusiform gyri in the AD patients with depressive symptoms from the KI cohort (Study I). No association has been found in non-depressed AD patients.

Higher baseline saliva cortisol levels in non-demented individuals (Study IV) were associated with widespread cortical atrophy in temporal, prefrontal and parietal cortex bilaterally and the right hippocampus, independently of age and MMSE.

To sum-up, depression was associated with thinning (Studies I, II) and subsequent atrophy (Study IV) in the superior temporal, supramarginal, temporal pole, lingual, fusiform and parahippocampal cortex. Cortical thinning in the superior frontal and lingual regions developed in parallel or prior to the depressive symptoms. The afore-mentioned regions are involved in social perception (processing of the information about others, experience positive emotions related to other people and building an integrative picture of another person), and are among the first to be impaired in Alzheimer’s disease. Elevated cortisol explained atrophy in these and a number of other regions, including the hippocampus, suggesting that depression and Alzheimer’s disease may be connected via cortisol-related brain damage. Depression-related atrophy in the ventral diencephalon leads to
impaired cognitive performance. Assessment of cognitive function during the depressive episode, combined with brain structural measurements may have a prognostic value. Future studies should evaluate if a detailed neurocognitive assessment of elderly patients during the depressive episode would help to identify those at high risk of dementia. It is also important to test if stress-reduction interventions in individuals at-risk of Alzheimer’s disease would be effective in its prevention.
LIST OF SCIENTIFIC PAPERS


CONTENTS

1 Introduction .................................................................................................................. 9
  1.1 History and state-of-the-art of depression-dementia studies .................................. 9
  1.2 Depression .......................................................................................................... 9
    1.2.1 Types of depression ....................................................................................... 9
    1.2.2 Epidemiology ............................................................................................. 10
    1.2.3 Cognitive impairment in depression ............................................................ 10
    1.2.4 Pathophysiology (Theories of depression) ............................................... 11
    1.2.5 Potential connection between the depression theories ............................ 13
  1.3 Dementia ............................................................................................................. 14
    1.3.1 Causes of dementia ...................................................................................... 14
    1.3.2 Alzheimer’s disease ...................................................................................... 14
    1.3.3 Types of AD ................................................................................................ 14
    1.3.4 Epidemiology ............................................................................................. 15
    1.3.5 Stages of AD and a clinical picture ............................................................. 15
    1.3.6 Pathophysiology of AD ............................................................................ 16
    1.3.7 Potential connections between the AD theories ....................................... 18
  1.4 Shared mechanisms of depression and dementia (are they pieces of the
    same puzzle?) ....................................................................................................... 19
    1.4.1 The Glucocorticoid Cascade Hypothesis .................................................. 19
    1.4.2 “Age-by-disease interaction hypothesis of late-life depression” .......... 19
    1.4.3 Brain aging ............................................................................................... 20
    1.4.4 Affective and cognitive functions in the brain ........................................ 20
  1.5 Neuroimaging ..................................................................................................... 22
    1.5.2 Neuroimaging studies of depression ....................................................... 25
    1.5.3 Neuroimaging studies of AD ................................................................. 27

2 Aims ......................................................................................................................... 29

3 Methods ................................................................................................................... 30
  3.1 Ethical approval ................................................................................................. 30
  3.2 Study participants ............................................................................................. 30
  3.3 Clinical assessment ........................................................................................... 31
    3.3.1 Definition and measurement of depressive symptoms and
         depression diagnosis .................................................................................. 31
    3.3.2 Diagnosis of AD and MCI ....................................................................... 33
  3.4 Neuropsychological assessment ........................................................................ 33
  3.5 Laboratory data .................................................................................................. 34
    3.5.1 CSF ........................................................................................................... 34
    3.5.2 Cortisol ..................................................................................................... 35
  3.6 MRI methods ..................................................................................................... 35
    3.6.1 Acquisition parameters ............................................................................. 35
    3.6.2 Freesurfer analyses .................................................................................... 37
  3.7 Statistical methods ............................................................................................ 38
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-HT</td>
<td>5-hydroxtryptamine</td>
</tr>
<tr>
<td>5-HTT</td>
<td>5-hydroxtryptamine transporter</td>
</tr>
<tr>
<td>a-beta</td>
<td>Amyloid-beta</td>
</tr>
<tr>
<td>ACC</td>
<td>Anterior-cingulate cortex</td>
</tr>
<tr>
<td>ACTH</td>
<td>Adrenocorticotropic hormone</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>APOE</td>
<td>Apolipoprotein E (gene)</td>
</tr>
<tr>
<td>ApoE</td>
<td>Apolipoprotein E</td>
</tr>
<tr>
<td>APP</td>
<td>Amyloid precursor protein</td>
</tr>
<tr>
<td>CSF</td>
<td>Cerebral spinal fluid</td>
</tr>
<tr>
<td>DLPFC</td>
<td>Dorsolateral prefrontal cortex</td>
</tr>
<tr>
<td>DMN</td>
<td>Default-mode network</td>
</tr>
<tr>
<td>Dp</td>
<td>Depressive symptoms</td>
</tr>
<tr>
<td>DSM-IV</td>
<td>Diagnostic and Statistical Manual of mental disorders IV</td>
</tr>
<tr>
<td>EOD</td>
<td>Early-onset depression</td>
</tr>
<tr>
<td>FF</td>
<td>Fusiform cortex</td>
</tr>
<tr>
<td>fMRI</td>
<td>Functional MRI</td>
</tr>
<tr>
<td>GCH</td>
<td>Glucocorticoid-cascade theory</td>
</tr>
<tr>
<td>GLM</td>
<td>General linear model</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HPA</td>
<td>Hypothalamic–pituitary–adrenal axis</td>
</tr>
<tr>
<td>ICD-10</td>
<td>International classification of disease, tenth revision</td>
</tr>
<tr>
<td>IDO</td>
<td>Indole amine 2,3-dioxygenase</td>
</tr>
<tr>
<td>ICV</td>
<td>Intracranial volume</td>
</tr>
<tr>
<td>LC</td>
<td>Lingual cortex</td>
</tr>
<tr>
<td>LLD</td>
<td>Late-life depression</td>
</tr>
<tr>
<td>LOD</td>
<td>Late-onset depression</td>
</tr>
<tr>
<td>LTP</td>
<td>Long-term potentiation</td>
</tr>
<tr>
<td>MCI</td>
<td>Mild cognitive impairment</td>
</tr>
<tr>
<td>MD</td>
<td>Major Depression</td>
</tr>
<tr>
<td>MDD</td>
<td>Major depressive disorder</td>
</tr>
<tr>
<td>ML</td>
<td>Machine learning</td>
</tr>
<tr>
<td>MM</td>
<td>Multivariate methods</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Medical Term</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>MTL</td>
<td>Medial-temporal lobe</td>
</tr>
<tr>
<td>PC</td>
<td>Parahippocampal cortex</td>
</tr>
<tr>
<td>Pes</td>
<td>Parameter estimates</td>
</tr>
<tr>
<td>PET</td>
<td>Positron-emission tomography</td>
</tr>
<tr>
<td>PET</td>
<td>Positron-emission imaging</td>
</tr>
<tr>
<td>PIB</td>
<td>Pittsburg compound</td>
</tr>
<tr>
<td>rf</td>
<td>Radiofrequency</td>
</tr>
<tr>
<td>RF</td>
<td>Random forest</td>
</tr>
<tr>
<td>SCI</td>
<td>Subjective cognitive impairment</td>
</tr>
<tr>
<td>SFC</td>
<td>Superior frontal cortex</td>
</tr>
<tr>
<td>SMG</td>
<td>Supramarginal cortex</td>
</tr>
<tr>
<td>SSRI</td>
<td>Serotonin-reuptake inhibitors</td>
</tr>
<tr>
<td>STC</td>
<td>Superior temporal cortex</td>
</tr>
<tr>
<td>TPJ</td>
<td>Temporoparietal junction</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>BDNF</td>
<td>Brain-derived neurotrophic factor</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

A large body of evidence suggests that depression is a risk factor for dementia, including dementia due to Alzheimer’s disease. The purpose of this thesis was to increase understanding of how depression and AD are connected. We hope that the new evidence and ideas in this field of research will lead to improvement of preventive and treatment strategies against depression and cognitive impairment; and will ultimately lead to increased life expectancy and better quality of life.

1.1 HISTORY AND STATE-OF-THE-ART OF DEPRESSION-DEMENTIA STUDIES

Already in the 19th-century it was noticed, that affective disorders can lead to cognitive impairment (1). However, in the middle of the 20th-century research was mostly aiming to differentiate pseudodementia (cognitive impairment, which resolved with depression) in geriatric depression from early stages of dementia (2–4). The logic behind this aim is clear – correct diagnosis should lead to the prescription of correct treatment, assuming that depression and dementia are two distinct pathologies. However, already back then, judging, from the description of the study groups, it was clear that depression and dementia overlap - depressed patients develop dementia and dementia is often comorbid with depression. Moreover, depression and dementia were found to have similar results in tests designed to differentiate them. For instance, it was suggested that unsuppression of cortisol in Dexamethasone Suppression Test is highly specific for depression (5), but later it was also shown to be valid for dementia and schizophrenia (6,7).

Modern studies suggest that the relationship between depression and dementia is complex. Depression may be considered as a risk factor for MCI and AD dementia (8,9). Duration of depression and number of depressive episodes may negatively affect memory-related brain structures (hippocampus and entorhinal cortex) (10,11). Alternatively, depression may be an early symptom of AD – one which may in fact outpace cognitive impairment (12).

1.2 DEPRESSION

The term depression is relevant to the diagnosis of Major Depressive Disorder (MDD) but also is an umbrella term encompassing various conditions mainly characterized by low mood and loss of interest. Other symptoms include appetite or weight loss or gain; insomnia or hypersomnia; agitation or retardation; loss of energy or fatigue (DSM-IV); loss of confidence or self-esteem (ICD-10); worthlessness or guilt; reduced concentration or indecisiveness; as well as thoughts of suicide or suicide attempt.

1.2.1 Types of depression

Depression is considered a “multifactorial and heterogeneous disorder that varies in terms of symptom severity, psychiatric comorbidity, and clinical course, including recurrence and response to treatment” (13). It can be classified by:
- The main cause - primary (unknown cause) or secondary (to some other disease or trauma)
- Severity – major or minor/subclinical level/subsyndromal
- Age of onset - childhood, adolescent, mid-life, geriatric. The terms early onset and late-onset depression are common.
- Pathological affect – unipolar vs bipolar
- Comorbidity with other psychiatric (anxiety) neurologic (stroke, AD) or somatic (diabetes, hypothyroidism) diseases
- Number of depressive episodes – first-episode or recurrent
- Stage – depressive episode vs remission
- Treatment resistance

1.2.2 Epidemiology
One of every five individuals experiences a depressive episode at least once in his/her life (14). A European survey concluded that people with a 12-month history of mental or emotional discomforts had a 36% chance of developing major depression (MD) (13). Depression is expected to be the single leading cause of disease burden by 2020 (16). It is associated with increased morbidity and mortality (17). High comorbidity between somatic diseases and depression might be a result of accelerated aging processes (18). The most replicable psychosocial risk factors for development of depression are related to chronic or acute stress (19). Among them are loneliness (divorce or widowhood) negative life events (illness, loss of close relatives or friends), unemployment and low income, childhood adversities, low social support and low education (20). Individuals with a history of childhood trauma have twice the risk of developing MD (21). Furthermore, patients with MD and a history of childhood trauma show higher symptom severity, a poorer course and less efficacious treatment effect than patients with MD without childhood trauma (22).

1.2.3 Cognitive impairment in depression
Many studies have shown that acute depression is associated with impairment in cognitive functioning, including attention, executive functions and episodic memory (23,24). In attempt to specify cognitive impairment associated with depression, a term pseudodementia was created. It was later discovered that pseudodementia is associated with persistent cognitive impairment (25). Even in euthymic/remitted state patients demonstrate cognitive deficits (26), suggesting that deficits cannot be explained purely by the depressive state. Importantly, those, diagnosed with the first-episode depression had significantly less cognitive deficits compared to patients with recurrent depression (27). Patients with previous depressive episodes and late-onset depressed have the worst cognitive performance (28). Elevated cortisol levels correlated with greater impairment in hippocampus-dependent cognitive functions in depressed individuals (29). Cognitive deficit is such a prominent feature of depression that it has been considered as a core component of depression (30). Yet, there is still no consensus criteria of cognitive impairment in depression (31).
1.2.4 Pathophysiology (Theories of depression)

1.2.4.1 The serotonin deficiency theory

The serotonin deficiency theory suggests that depression is a result of a reduced serotonin (5-HT) transmission in the brain (32). The transmission may be reduced due to insufficient levels of serotonin in the brain or as a result of serotonin receptors dysfunction. A large amount of evidence has accumulated in favor of the involvement of the serotonergic system in depression. Earlier studies have shown that depression is associated with reduced levels of serotonin and serotonin metabolites in cerebrospinal fluid (CSF) (33,34) CSF serotonin levels were shown to reflect serotonin levels in the cerebral cortex (35). Remitted depressed with low tryptophan levels had higher risk of new episodes (36). Moreover, restoration of the serotonin levels and transmission seems to be the main mechanism behind all modern antidepressant treatments (37).

Earlier studies reported that individuals with two short alleles of the serotonin transporter have higher risks of developing depression (38). Recent meta-analysis did not find any significant interaction between the 5-HTT polymorphism and vulnerability to stress (39). Studies suggest that increasing deficit in central serotonin via tryptophan deplletion can be concomitant with negative biases in remitted depressed (40,41). Conversely, antidepressant treatment may elevate the processing of positive emotional information in healthy individuals and patients with the depressive episode (42). However, antidepressant (including serotonin reuptake inhibitors - SSRI) treatment does not improve mood within hours (despite of increase in plasma levels), sometime weeks are needed to achieve some clinical effect, suggesting that other mechanisms may be involved. Role of the serotonin in neuroplastic processes has been shown. SSRIs were effective in reducing depression phenotype (also reducing corticosterone and elevating brain-derived neurotrophic factor (BDNF) levels) in mice in favorable (enriched) environment, but were associated with worsening of depression and related markers in the stressful environment (43). This might work through the ability of serotonin to prevent stress-induced blockage of neurogenesis in the hippocampus and dentate gyrus (where serotonin receptors are highly expressed) (44), (45,46). Hippocampal neurogenesis after restoration of serotonin transmission can explain a decrease in cortisol levels in depressed after remission since it might also restore the functional ability of the hippocampus to block cortisol secretion (negative feedback) (47). Importantly, hippocampal neurogenesis is required for SSRI antidepressive effect (48).

1.2.4.2 Hypothalamic–pituitary–adrenal axis dysregulation

Many studies have shown that depression is associated with the hypothalamic–pituitary–adrenal (HPA) axis dysregulation (49,50). The HPA axis is controlled by a corticotropin-releasing hormone produced in the hypothalamus, which stimulate secretion of adrenocorticotropic hormone (ACTH) in the pituitary gland. In turn, ACTH stimulates secretion of cortisol from adrenal glands. Cortisol is the main glucocorticoid hormone in humans. It influences gene expression and protein synthesis in most tissues, including neurons. Increased levels of cortisol in blood signal to the hippocampus and hypothalamus,
which normally leads to decrease in ACTH and subsequent decrease in cortisol levels (negative feedback) (Figure 1). The earliest studies had shown that depressed individuals have cortisol non-suppression by the dexamethasone test, which indicates impaired negative feedback in HPA axis. Later it has been shown that adverse events, especially in early childhood lead to higher risks of depression later in life and are associated with slower hippocampal growth and impairment of the negative feedback in HPA axis. This was not surprising, given the role of the hippocampus in cortisol suppression (47). Elevated levels of cortisol have been described in depressed compared with controls (50), in MD patients during the episode compared with remitted (51). Hypertrophy of adrenal glands has also been demonstrated in MD patients (52).

The damaging effect of glucocorticoids under adverse conditions has been termed "allostatic load", referring to the body’s “cost” of adaptation to adverse conditions (53). "Adaptation in the face of potentially stressful challenges involves activation of neural, neuroendocrine and neuroendocrine-immune mechanisms”, a process referred to as "allostasis" or "stability through change" by Sterling and Eyer (54). Allostasis is an important mechanism of maintaining homeostasis. When such mechanisms are engaged only rarely, they are beneficial for survival. McEwen suggests that the following factors may affect one’s ability to handle stressful situations: 1) subjective interpretation of the situation as a threat (based on individual experience) 2) health (diseases or genetic predispositions) (53). Studies in rodents show that there might be a genetic predisposition towards HPA hyper-reactivity (55). Human studies show that HPA hyper-reactivity in mothers predicts HPA hyper-reactivity in daughters (56), which may have both - genetic and environmental causes. Epigenetic studies have shown that stress reactivity profile of an individual is programmed by hormonal (incl. HPA) maternal effects (57). Taken together, these findings suggest that the HPA axis dysregulation in depression is “rather the manifestation of persistent neurobiological abnormalities that predispose to depression” than a consequence of clinical depression (58).

1.2.4.3 Inflammation

Depression is associated with a chronic inflammatory process (59,60). Meta-analysis based on 29 publications showed that elevated pro-inflammatory biomarkers are significantly associated with depression (61). Moreover, depression and inflammatory diseases often co-occur (62). Evidence suggests that inflammation is a causal factor for the development of depression (63). Exposure to immunomodulatory agents increases the risk of developing depression (64) when inhibition of inflammatory pathways can improve mood (65). Evidence regarding the depression-inflammation connection led to the development of the cytokine hypothesis, according to which, environmental stress induces the production of cytokines (interleukins, tumor necrosis factor- α and interferon- α and γ). These in turn activates indole amine 2,3-dioxygenase (IDO), subsequently utilizing tryptophan to the IDO pathway (a part of the kynurenine pathway) and decreasing the availability of tryptophan for serotonin synthesis (66), (67) (Figure 1).
While investigating the causes of chronic inflammation in depression, researchers have found several possible sources, including: sleep deprivation, obesity, lack of exercise, vitamin D deficiency, atopic and autoimmune disorders, infections and stress (68,69). Based on the connection between depression, stress and inflammation, social signal transduction theory of depression was formulated. According to this theory, social threat and adversity up-regulate components of the immune system involved in inflammation. The key mediators of this response - pro-inflammatory cytokines, can provoke behavioral changes (sickness behavior), which, actually, describes depressive symptoms: sadness, anhedonia, fatigue, psychomotor retardation, and social withdrawal (70). However, given that 95% of the population has latent/chronic viral and bacterial infection, it seems unlikely that inflammation, which follows acute cortisol elevation (which has immunosuppression effect), activates without any antigen stimulation (71).

Importantly, inflammation is a normal physiological reaction. Various factors listed above, probably lead to a deficiency in adaptive immunity and subsequent chronic inflammation. Deficiency in adaptive immunity has been associated with impaired social behavior across various species (flies, fish, rodents, humans), suggesting the co-evolutionary link between social/aggregation behavior and an efficient anti-pathogen response (72).

### 1.2.5 Potential connection between the depression theories

Genetic predisposition, concomitant with epigenetic programming and environmental stress lead to HPA system modification biasing toward hyper-reactivity and leading to structural and functional changes in the fronto-limbic network associated with specific "low trust/negative bias" behavioral patterns and impaired cognitive functions (73,74). Chronic
cortisol elevation also leads to impairment of immune response with subsequent decrease of adaptive immunity and chronic low-grade inflammation. Tryptophan is extensively utilized by the kynurenine pathway, developing tryptophan depletion, a resulting deficit of tryptophan for serotonin synthesis, consequent serotonin deficiency in the brain, which in turn, is clinically expressed as depression (66).

1.3 DEMENTIA

Dementia is a clinical syndrome characterized by progressive loss of cognitive functions. It also can be defined as an “acquired organic mental disorder with loss of intellectual abilities of sufficient severity to interfere with social or occupational functioning. The dysfunction involves many domains, like memory, behavior, personality, judgment, attention, spatial relations, language, abstract thought, and other executive functions” (MeSH).

A diagnosis of dementia requires a decline in both memory and reasoning that leads to deterioration from the previous level of functioning (31). Symptom duration should be evident for at least six months.

1.3.1 Causes of dementia

Possible causes of dementia are: Alzheimer’s disease (AD), vascular pathology, Lewy body disease, Frontotemporal dementia, Creutzfeldt-Jakob disease, Huntington disease, Wernicke-Korsakoff Syndrome, various brain traumas, infections (neurosyphilis, HIV), B12 vitamin deficiency, and toxicities of various types.

Interestingly, even in non-AD dementias, cognitive decline correlates with the amount of AD pathology (75), suggesting that AD pathology might be responsible for the cognitive component of decline observed across other neurodegenerative diseases. This might be explained by the regional distribution of the pathological process since in AD it involves predominantly medial temporal and other regions of the cerebral cortex, which play a major role in cognitive functions.

1.3.2 Alzheimer's disease

The most common cause of dementia in the elderly is Alzheimer's disease. It is a “chronic progressive neurodegenerative disease, characterized by impairment of memory, judgment, attention span, and problem-solving skills, followed by severe apraxia and a global loss of cognitive abilities” (76). The condition’s pathology is marked by severe neuronal loss, senile plaques, and neurofibrillary tangles (see below). No cure exists for Alzheimer's disease. Current treatment can only temporarily slow the worsening of symptoms.

1.3.3 Types of AD

AD can be classified by:
Onset - Early onset before 65 (familial/genetic AD) vs late onset (sporadic)
Stage – preclinical, subjective cognitive impairment (SCI) – in some patients, mild cognitive impairment (MCI), mild, moderate, severe/advanced
Level of diagnostic confidence - possible (atypical), probable (typical), definite (typical + histological confirmation)
Phenotype - typical (amnestic with hippocampal atrophy) vs atypical (posterior/frontal atrophy/logopenic)
Preclinical states - asymptomatic (at risk) vs presymptomatic (autosomal-dominant mutation)

1.3.4 Epidemiology
Approximately 10% of people 65 years or older have dementia, and among this 60 % have AD. AD risk factor can be divided into modifiable - physical inactivity, depression, hypertension, hypercholesterolemia, obesity, diabetes, low education, alcohol intake, infections and non-modifiable - age, gender, APOE4 allele, Down syndrome risk factors. Risk factors can initiate a pathogenic cascade of events in the brain starting in mid-life or even earlier.

1.3.5 Stages of AD and a clinical picture
The slow and progressive natural course of AD can be conceptualized as passing through three stages. Initially, there is a latent stage in which the pathogenesis of AD slowly develops but there are no overt symptoms. Some patients may have subjective complains without objective impairment measured with the neuropsychological tests (SCI). Then, there is a prodromal stage in which mild impairments, especially in memory, begin to emerge but are not sufficient to interfere with normal daily functioning - mild cognitive impairment (MCI) (77). Finally, the disease eventually progresses to a clinical stage in which the symptoms worsen and ultimately fulfill clinical criteria for dementia (72).

The classic clinical symptoms of the disease are an amnesic type of memory impairment, visuospatial deficits, and deterioration of language. Behavioral symptoms (depression, agitation) are also common, especially in the early stages of the disease (79). There is a progression from loss of complex instrumental day-to-day activities, to loss of the basic activities of daily living (80).

The prevalence of depressive symptoms in AD is around 20% to 45% (81,82). Younger age, comorbidities, bereavement, greater impairment in daily activities and previous depressive episodes are risk factors for depression in AD (83,84). Depressive symptoms in AD are associated with higher rates of cognitive decline (85). Depression due to AD can be diagnosed when all criteria of dementia of Alzheimer type are fulfilled, and three (or more) typical depressive symptoms have been detected during the same 2-week period and represent a perturbation from the previous physiological activity (82).
1.3.6 Pathophysiology of AD

Classical hallmarks of the neuropathology of AD are extracellular Aβ plaques, intracellular neurofibrillary tangles, neuronal loss and brain atrophy especially pronounced in the fronto-limbic area, including the hippocampus.

1.3.6.1 Amyloid

Amyloidosis is a condition in which soluble plasma proteins transform into insoluble, fibrillar form. A-beta deposition expands to the regions that receive neuronal projections from regions already exhibiting a-beta (from neocortex to deep brain structures). According to the amyloid cascade hypothesis, amyloid deposition is a primary event in AD (86). A-beta is a 39–43 amino acid peptide produced by a proteolysis of the a-beta precursor protein (APP). The lengths associated with Alzheimer's disease are 40 and 42 amino acids long. These longer forms have higher aggregation properties. Amyloid precursor protein is a source of A-beta. APP can be cleaved by two alternative pathways – non-amyloidogenic by a combined action of a- and y-secretase, and amyloidogenic by a combined action of b- and y-secretase (around 10% of APP) (87), which happens after the complex is internalized into endosome. Presenilin is a sub-unite of a y-secretase, which cuts the APP. Its mutation leads to increase in the ratio of a-beta 42 produced compared to a-beta 40 and is the most common genetic cause of AD. Proteolysis, astrocytic degradation and drainage via CSF and blood are the main cleavage pathways for a-beta (88). Factors that trigger amyloid aggregation, include: temperature rise, acidosis, hyperosmosis (89) and excessive neuronal activity (90). Importantly, SSRI has been shown to reduce a-beta plaque formation by 78% (91).

It is still under debate how exactly a-beta causes neuronal damage. Two views are dominating in the literature. A classical view is that a-beta plaques are toxic to neurons. They disturb neuronal connections. This leads to the functional impairment, neuroinflammation, and apoptosis. An alternative view is that a-beta oligomers are associated with neuronal damage (92). Amyloid antibodies can cleave plaques, however there is no evidence that reduction of extracellular amyloid depositions decrease neuronal damage and slows progression of cognitive impairment. Nevertheless, many clinical trials attempt to treat AD patients with amyloid antibodies, and, as could be expected, with no positive outcome (93). Even if aggregation causes damage to the cells, such aggregated state signals a final stage and offers no potential treatment target. From another hand, rodents that were genetically mutated to produce a-beta oligomers and no plaques still developed AD phenotype (94). An alternative model which produced both – oligomers and plaques did not develop more severe cognitive impairment compared with oligomer model. A-beta oligomers were associated with inflammation and neuronal loss, but not A-beta plaques (95). One of the mechanisms is that oligomers damage neuronal membrane including the membrane receptors (96) which might lead to the calcium influx and cell death (97). Also small soluble oligomers inhibit hippocampal long-term potentiation, which impairs the recall of a complex learned behavior in rats (98). The other mechanism is that a-beta oligomers alter neuronal insulin receptors (99). Also intracellular a-beta oligomers might disrupt the functioning of neuronal
organelles. In particular accumulation of a-beta oligomers in the mitochondrial membrane can activate apoptosis (100).

It is still not known how A-beta forms a plaque. It is possible that release to the extracellular space and plaque formation may occur after the neuronal death. Interestingly, it has been shown that activation of 5-HT4 receptors decrease levels of beta-amyloid peptides, enhance non-amyloidogenic pathway and increases neuronal survival (101).

Apolipoprotein E (ApoE) is essential for a-beta formation (102). Individuals with ApoE4 allele have increased risks of developing AD. Interestingly, global distribution of APOE variants in the world is different, with a significantly higher frequency of APOE4 in pigmy and aborigine cultures of hunter-gatherers with a lack of stable food supply, suggesting that APOE4 allele predisposes to more conservative lipid metabolism (103). Remarkably, most patients with familial renal amyloidosis are heterozygous for mutations in the genes for apolipoprotein AI, apolipoprotein AII, or fibrinogen A alpha-chain (104).

The physiological role of Amyloid It is suggested that a-beta oligomers are involved in LTP (long-term potentiation) blockage (105). Recent evidence suggests the role of A-beta in locale brain immunity. Interestingly, amyloid plaque burden was not associated with the severity of the cognitive impairment in AD (106).

1.3.6.2 Tau protein

Aggregates of hyperphosphorylated tau protein-- neurofibrillary tangles-- are the primary markers of AD pathology. It has been shown that elevated CSF tau levels reflect the degree of neuronal loss (107,108). CSF levels of total tau (measured with antibodies unspecific to forms of the tau protein) is a sensitive, but not specific marker of AD (109). Later studies suggested that phosphorylated tau --may be more specific in discrimination with other dementias (110). Neurofibrillary tangles and elevated CSF tau can be found it many other conditions associated with neuronal loss (traumatic brain injury, other neurodegenerative diseases, encephalitis, brain tumors). However the causes and mechanisms of tau hyperphosphorylation are not well understood. The physiological function of Tau protein is to stabilize cytoskeletal microtubules predominantly in neurons. Tau will be discussed further in the section 1.5.3 “Neuroimaging studies in AD”.

1.3.6.3 Inflammation

Evidence suggests that there is an association between inflammation and AD. Acute inflammatory responses are often beneficial in clearing foreign material from the brain, but the chronic inflammation present in AD may be pathogenic as it can cause neurotoxicity by damaging neighboring cells and is even thought to promote the accumulation of plaque pathology (111,112). Presence of inflammatory molecules in AD brain has been described in multiple studies (113,114).

The source of the inflammation and its causal relation to the established AD mechanisms are not known. Some researchers suggest that forms of AD pathology, like degenerating tissues
and a-beta plaques, trigger inflammatory responses (115). However, other studies suggest that astrocytosis is an early event in AD, even preceding amyloid plaque deposition (116). And amyloid deposition is not associated with neuronal loss in the absence of inflammation (117). Moreover, prevention of the amyloid deposition in mice does not restore neuronal connections if inflammation persists (118). Multitracer PET studies of pre-symptomatic patients with familial AD also reveal that astrocytosis temporally precedes amyloid deposition (119,120). Also, the fact that the main cause of systemic amyloidosis is inflammation suggests that this mechanism deserves special attention in brain amyloidosis. To note, brain disease accompanied with an inflammatory response, like prion disease, HIV, and brain syphilis are also associated with subsequent amyloid deposition. Other evidence is that infectious diseases increase the risk of subsequent AD (121,122).

The "pathogen hypothesis" of AD suggests that amyloid is an anti-microbial peptide, active against most common neurotropic bacteria (123).

Some studies have shown that non-steroid anti-inflammatory treatment slows down AD progression and also reduce risks of AD development (124,125). However, other studies did not show any benefit anti-inflammatory drugs (126). In case, if “pathogen hypothesis” is true, anti-inflammatory treatment is not an appropriate approach.

1.3.6.4 HPA dysregulation

Studies have demonstrated a bidirectional relationship between neurodegeneration and HPA axis dysregulation- From one side, cortisol has a damaging effect on the aging brain, indeed, a clear aging effect of itself (127–131). It has even been proposed that HPA-axis dysfunction is central to the development of AD (132). Indeed, hypothalamic dysfunction can explain the overlap in symptoms between depression and AD (mood, appetite, sleep, memory, autonomic). Consistent with our finding, several previous imaging studies have shown structural and functional abnormalities in the hypothalamus in MCI, preclinical AD and AD compared with control groups (133–136). For instance, Hall et al. have demonstrated reduced basal forebrain and hypothalamus volumes in preclinical AD, and interestingly the combination of reduced forebrain and hippocampal volumes was associated with more rapid cognitive decline. (137). On the other hand, aging is associated with cortisol non-suppression. This is also found in AD and can be explained with age- and AD- related hippocampal atrophy and subsequent impairment of cortisol down-regulation.

1.3.7 Potential connections between the AD theories

AD pathology may develop due to i) increased a-beta production (in genetically predisposed individuals or triggered by inflammation or both), and toxic effect of a-beta oligomers on neurons ii) microglia-related a-beta degradation, e.g. may be disrupted by a parallel inflammatory process (138) (which might be triggered by cortisol elevation and immunosuppression), iii) changed local physical conditions that leads to accelerating a-beta aggregation and damage of the fronto-limbic structures. Neurons could be additionally compromised by toxic levels of cortisol, which may provoke excessive neuronal activity,
leading to amyloid hyperproduction and subsequent blockage of the LTP (which may in fact be a protective mechanism against the excitotoxicity). A frequent occurrence of this process may lead to the structural damage of the brain. Hippocampal atrophy will lead to impaired negative feedback of the HPA axis, chronic cortisol elevation, chronic inflammation, excessive stimulation of the DMN structures and a-beta deposition in a vicarious circle.

1.4 SHARED MECHANISMS OF DEPRESSION AND DEMENTIA (ARE THEY PIECES OF THE SAME PUZZLE?)

Biological factor observed in both – depression and AD are: HPA dysfunction, inflammation and brain atrophy. All these factors are associated with accelerated aging. The term aging reflects the process of getting older. Within a more biological framework, “aging reflects the process in which a variety of stressors are no longer adequately counteracted by the body’s protective functions” (139).

1.4.1 The Glucocorticoid Cascade Hypothesis

HPA axis affects all the organs and tissues including the brain. In the 1930's, Hans Selye discovered that chronic stress can shorten the lifespan (140). In 1986 Robert Sapolsky with his colleagues formulated the “Glucocorticoid cascade hypothesis” (GCH) which suggests that stress is “cumulatively damaging aging tissue” (128). They also found that aged male rats had prolonged corticosterone secretion after stress-exposure. Cortisol not only damage aging tissues, it has an aging effect on the brain and the body itself (131). Based on the endocrinologic abnormalities major depressive disorder has been considered a syndrome of “premature aging” (141).

1.4.2 “Age-by-disease interaction hypothesis of late-life depression”

Brandon Chad McKinney with colleagues has reported that many of the genes involved in aging-regulation are also involved in most common neurodegenerative and neuropsychiatric disorders (142). They have proposed the “age-by-disease interaction hypothesis” which suggests that late-life depression, schizophrenia and AD may be associated with epigenetic changes which are normally occur during aging. In other words, “the biologic processes disrupted in aging overlap to a significant degree with those recruited in a number of brain diseases including LLD” (143). The genes involved in both – aging and disorders are mainly responsible for inflammation, oxidative stress responses, mitochondrial function, synaptic and calcium regulation.

Another research teams have supported the “age-by-disease interaction hypothesis” analyzing 410 GWAS studies (144). The data analysis has shown shared aging pathways in the most common age-related diseases (neurodegenerative, vascular, metabolic, cancer and other). Authors claim that they have provided the first direct evidence that conserved pathways of aging simultaneously influence multiple age-related diseases in humans.
Another interesting marker of aging is a leukocyte telomere length. It gets shorter with age, in depressed (145) and in AD (146).

1.4.3 Brain aging

Franke et al. created a predictive model of aging based on structural brain measurements of healthy individuals 19-86 years (147). They observed that applying the model to people with mild AD resulted in overestimation of their age on 10 years on average (147). Next, the framework (BrainAGE) was utilized to predict conversion from MCI to AD (148). With accuracy rates of up to 81%, BrainAGE outperformed cognitive tests and CSF biomarkers. In 2014 Koutouleris et al. using BrainAGE have shown that the brain of the patients with schizophrenia, major depression, borderline personality disorder, and individuals in at-risk mental states for psychosis age faster than controls. Group-level analyses showed that BrainAGE was highest in schizophrenia (+5.5 y) group, followed by major depression (+4.0), borderline personality disorder (+3.1), and the at-risk mental states (+1.7) groups (149). BrainAGE predicted patient status as well, suggesting that the pattern of brain changes is similar between aging and brain disorders. The process of accelerated aging might explain why the same fronto-limbic regions are impaired in AD, depression, and schizophrenia.

Together these lines of evidence, suggest that depression and AD may be connected via biological processes, related to accelerated aging.

1.4.4 Affective and cognitive functions in the brain

James Papez coined the term “limbic system” as a set of paleocortical and deep brain structures playing central role in emotional behavior (150). Limbic structures are the cingulate gyrus, the parahippocampal gyrus, the hippocampal formation, the amygdala and the hypothalamus with mammillary bodies (151). Large number and histological polymorphism of the structures involved in emotion regulation can be explained by a complexity of emotion processing. Emotions depend upon the integrated activity of neural networks that modulate arousal, autonomic function, motor control, and somatosensation. Indeed, despite the phenomenological differences, affective/motivational and cognitive aspects of behavior are closely intertwined. So, it is not surprising that atrophy within the crucial components of the “emotional circuit”, such as the basal forebrain, temporal and prefrontal cortex, seen in Alzheimer’s disease is associated with a severe impairment in memory and other cognitive functions (133,134,137). This remarkable overlap between brain regions involved in emotions and cognition, memory, in particular, is quite reasonable from the evolutionary perspective. The main purpose of memory is to obtain, keep and retrieve information which is required for survival. Stimuli relevant for survival always have emotional valence and accompanied by a reaction of the autonomous system and switch to the specific behavioral repertoire. For instance, the hippocampus is a crucial region for novelty detection, spatial orientation, and spatial memory (152). But also hippocampal atrophy is the most replicable finding in depression (153,154). Hippocampus can be subdivided into anterior and posterior parts.
Anterior part is more active in response to close stimuli, emotionally salient stimuli and strongly connected with amygdala and emotional learning. Posterior part is more active in response to new places, open spaces; focus on background and spatial memory. This functional segregation makes a lot of sense, since “the ability to identify quickly in the environment emotionally salient information, including danger and reward, and to form rapid and appropriate behavioral responses is critical to survival” (Darwin 1872/1965), but it is also necessary to memorize a territory and a path towards or from such important objects. The role of the hippocampus in memorizing location of important objects is supported by the fact, that food-storing species have a larger hippocampus than non-storing species (155). In humans, hippocampus might have even more sub-divisions (156). Hippocampus and parahippocampal gyrus in humans are also involved in learning and retrieval social information (157).

The role of the temporal lobe in assessment of the significance of the stimuli has been shown in lesion studies. After bilateral temporal lobectomy, Rhesus monkeys did not develop any clear cognitive or motor impairment, however, developed "psychic blindness" or emotional agnosia, which was expressed in emotional indifference of the subjects (158). By "psychic blindness" authors mean forms of behavior which seem to indicate that the “ability to recognize and detect the meaning of objects on the basis of visual criteria alone is either lost or seriously disturbed, although the animal exhibits no, or at least no gross, defect in the ability to discriminate visually. The bilateral temporal monkey shows a strong tendency to approach animate and inanimate objects without hesitation. This tendency appears even in the presence of objects which previously called forth avoidance reactions, extreme excitement and other forms of emotional response" (159).

The other region involved in both depression and AD is the anterior cingulate cortex (ACC). Not surprisingly it also can be subdivided into “affective” and “cognitive” subregions. The ventral ACC (affective subdivision) consists of the sub- and pregenual anterior cingulate gyri. Ventral ACC is activated in response to pain (and physical and psychological). It is a key node of the salience network (160). It is antagonistic with the executive network. Caudal and dorsal ACC represents the cognitive subdivision. ACC plays an important role in the integration of cognitive and emotional data.

The presence of affective and cognitive subdivisions within brain structures can be explained by a suggestion that both sets of functions had to be present at each stage of human brain evolution in order to adapt to the changing environment.

Prefrontal cortex plays a key role in higher cognitive functions. In particular, DLPFC plays an important role in executive functions, including speech, but also in cognitive control. Medial and lateral orbitofrontal cortices are the key nodes in reward and punishment networks in human and other primates (161).

Damage to a specific brain region may lead to a loss of function, but normally, in order to execute a complex goal-oriented behavior, coordinated work of multiple regions and
structures is required. Studies have identified several functional networks in the brain, linked
to a particular cognitive ability. A network can be defined as a set of interacting brain regions
known to have activity highly correlated with each other and distinct from other networks in
the brain (162).

According to the two-stream hypothesis, there are two major perceptional (visual and audial)
streams in the neocortex – dorsal and ventral (163). The ventral stream is processing visual
information about the object – "what" stream. The path is: the thalamus -> primary and
associative visual cortex (or auditory) -> posterior, central and anterior inferior temporal
cortex. Damage to the ventral stream can cause inability to recognize faces/speech or
interpret facial expression/intonation. The dorsal stream is recognition of the direction and
orientation in space – “where” stream.

Another well-established system is a salience network. It consists of the anterior insula,
dorsal anterior cingulate cortex, the amygdala, the ventral striatum, and the substantia nigra
(164). It is involved in various complex functions, including social behavior.

Recently it was proposed that there are two attention networks – dorsal (intraparietal sulcus
and frontal eye field), which is involved in the top-down selection of stimuli, and ventral
(dorsolateral prefrontal cortex - DLPFC and temporoparietal junction - TPJ) which is
involved in detection of behaviorally relevant stimuli (new/salient/unexpected). Attention
system is strongly connected with the dorsal stream, by redirecting it to the salient stimuli
(165).

DLPFC and TPJ may also form cognitive control network (strongly connected with
attention), which is crucial for goal-oriented behavior. Together with amygdala networks
(perception (overlaps with a ventral stream), affiliation and aversion), cognitive control
network is involved in social behavior (166).

Cognitive control network plays role in attention to the external world. The reciprocal
function has the default-mode network (DMN), with its internally directed mentation
involving long-term memory. It is activated when a person is not focused on the outside
world and the brain is at wakeful rest (daydreaming, free thinking about past/future/self/others). The most active DMN regions are precuneus, medial superior frontal
cortex, and inferior parietal cortex. DMN suppression supports certain types of goal-directed

cognitive processes (167).

1.5 NEUROIMAGING

Imaging techniques are potentially powerful methods for identifying endophenotypes that are
associated with, or are indicative of, a vulnerability to certain pathological conditions. They
can be used to discriminate between the diseases, to predict treatment effectiveness,
prognosis, and dynamics of the pathological process including its response to treatment.

Neuroimaging techniques can be classified into:
- Structural [Magnetic Resonance Imaging (MRI) (T1 - gray matter, T2, DTI - white matter,), Computer Tomography (CT) (can be used for stroke, fracture or tumor identification)]

- Functional [involving radioactive tracers: Positron-emission tomography (PET) (PIB-amyloid binding, FTG-glucose, astrocytosis and tau markers), SPECT; MRI-based: fMRI (resting-state or during task), Arterial-spin labeling], and many others

T1 brain MRI is an optimal approach to study structure of the gray matter. There are various ways to measure gray matter structure: using cortical thickness (mm), using surface area (mm2), gray matter volume (mm3). Fractional anisotropy and mean diffusivity can be used to measure a direction of the water molecules in the brain.

**The basic principles of MR physics**

MRI utilizes the resonance of protons (hydrogen) to generate signal, which is then processed to generate images. The direction of the vectors representing a magnetic moment of protons is randomly arranged when not exposed to a magnetic field. In presence of an external (primary) magnetic field (B0) the protons align along the direction of the magnetic field (longitudinal magnetization). Larger field strength (measured in Tesla, T) results in a larger magnetization (larger fraction of protons gets excited) and therefore superior signal-to-noise ratio. Then protons are excited by a radiofrequency (rf) pulse (MHz), generating another magnetic field (B1). As a result magnetization vector is tipped onto towards the transverse plane (transverse magnetization). When rf is switched off, the excited protons emit energy and return to their original state (relaxation). Relaxation in the longitudinal axis (parallel to the B0) is called T1 relaxation, in transverse axis – T2 relaxation. T1 and T2 relaxations vary depending on a tissue structure.

**What are we measuring?**

Current resolution of most MRI machines does not allow to identify what components of the gray matter contribute most to the atrophy and on what stages. Changes in the cortical thickness or gray matter volumes may reflect many different processes. Glia composes around 50% of the cortex it is important to understand what contribution it makes on a macrostructural level. Important to note, that brain structure is not as rigid as was suggested – it can change its volume or thickness just after months of stimulation, like it was shown after psychotherapy or learning new skills (168,169). This indicates that gray matter structure may represent long-term consequences of functional activity of the region, however it is not known yet, whether all regions of the cerebral cortex equally susceptible to modification.

1.5.1.1  **MRI data analysis**

**General Linear Modeling**

In the general linear modeling (GLM) a linear combination of predictors is used to explain the outcome. The formula is: $y=X*\beta$, where $y$ is the outcome variable (cortical thicknesses at each vertex), $X$ is a set of predictors (gender, age, depression status), and $\beta$ is the vector.
of parameter estimates. The assumption here is that residuals are independent and are normally distributed.

**Multivariate methods**

Biological systems are complex. Thus, cannot be efficiently described using one or two variables. Modern analytical techniques in neuroscience allow extracting a large amount of data (hundred thousands of voxels or vertices in brain MRI, SNPs in GWAS, proteins in spectrometry). However, the most widely used statistical methods are univariate. Therefore, they are not always optimal for dealing with high complexity of biological data and neuroimaging features in particular. In contrast, multivariate methods (MM), can handle large datasets, allowing exploring intrinsic and sometimes hidden patterns in the data (which might carry more useful information than one isolated measurement), make effective classification and predictive models. Another important property is ability to identify variables that contribute most to overall variance of the data or variance of interest. Multivariate techniques can be useful for both – mechanistic understanding of the biological process and for solving practical classification and prediction problems. Classification can be defined as a general task in data analysis that needs the construction of a classifier that assigns a class label to instances described by a set of attributes. MM can be classified as:

- **Unsupervised** (principal component analysis, factor analysis, clustering) – uncover unobserved factors, can be used for outlier detection
- **Supervised** (random forest, support vector machines, partial least squares) – uses a set of predictors and observed outcome to build an algorithm able to predict the non-observed outcome (classification/regression).

MM is a general term which covers all methods of analysis of the multivariate data when machine learning (ML) is a more narrow term, referring to the predictive application of the MM (in contrast with conventional statistical methods, where population inference is emphasized). Also conventional statistical methods are usually combined with a hypothesis-driven approach when ML is more suitable for the data-driven analysis.

The main steps of the multivariate data analysis include:

- Choice of the method(s) and model parameters (depends on the aims, sample size, if the sample is balanced or not, computational resources)
- Preprocessing of the data (scaling/centering/dimension reduction/outlier detection – depends on the data and assumptions of the method)
- Model derivation and interpretation (training, assessing the parameters of the model)
- Model validation/assessment of its generalizability (cross-validation/out-of-bag error estimation/testing set - preferable)

**Random Forest**

Random forest is an ensemble learning method for classification, regression, and clustering. The idea of the ensemble learning methods is to generate many classifiers and aggregate their
results. The rationale behind it is that combination of learning models increases the classification accuracy. Two main methods are boosting (170) and bagging (171). In boosting, successive trees give extra weight to points incorrectly predicted by earlier predictors. In the end, a weighted vote is taken for prediction. In bagging, successive trees do not depend on earlier trees — each is independently constructed using a bootstrap sample of the data set. In the end, a simple majority vote is taken for prediction.

In 2001 Breiman proposed random forest (RF). In contrast to standard decision trees (where each node is split using the best split among all variables), in the random forest, many trees trained on a randomly drawn subset of features are ultimately combined together by a majority vote. This strategy has been shown to produce highly accurate prediction models and to be robust against overfitting (172). Other advantages of the RF are: it’s non-parametric (not sensitive to outliers), fast, and has few tuning parameters – number of trees and the number of variables in the random subset at each note.

1.5.2 Neuroimaging studies of depression

It is challenging to study brain correlates of psychiatric disorders, due to etiological heterogeneity, technological limitations, the confounding effects of medication, and non-disease related inter-individual variation in brain morphology and function (173). Nevertheless, some regional changes were replicated in many neuroimaging studies of depression.

Structural findings

Cross-sectional studies

Hippocampal atrophy is one of the most replicable structural findings in depression (153,154,174). It correlates with the number of depressive episodes and is a common finding in recurrent depression, especially in elderly populations. However it is not as common in the first-onset MDD, suggesting that structural changes in the hippocampus are secondary to the disease progression possibly reflecting progressive impairment of the HPA system. Hippocampal volume is larger in remitted compared with non-remitted MDD, supporting the role of the hippocampal neurogenesis in SSRI efficacy (48).

Ventricular expansion is also a common finding in depression (153), providing additional evidence of the relationship between depression and neurodegeneration.

White matter pathology is a common finding in elderly MDD cohorts. Most probably it is secondary to a cerebrovascular disease and associated inflammation (173).

Increased volume of the pituitary gland was observed in two of three studies, supporting the role of endocrine abnormalities in depression (153).

Study assessing the whole brain structure in first-episode depressed demonstrated reduction of the GM and cortical thickness in the temporal, insular and parietal GM (175).
In late-life depression, a decrease in GM was observed in the parahippocampal cortex, hippocampus and amygdala, prefrontal and subcallosal cortex. Increased GM volume was observed in occipital area (176). In AD, depressive symptoms were associated with atrophy across all the cortical surfaces except for occipital (177).

**Longitudinal studies**

There are not many longitudinal imaging studies in depression (178). The study, using hippocampus as a ROI, have found that patients with late-onset depression (>60 years) exhibited faster hippocampal volume loss then early-onset depressed (179). Depressive symptoms were associated with white matter volume reduction in AD patients (180).

Even fewer studies assessed conducted a longitudinal assessment the whole brain or whole cerebral cortex. One study showed gray matter reduction in patients with depressive episode in most cortical regions after three-year follow-up period (superior, medial orbitofrontal, superior temporal, fusiform and parahippocampal) and the hippocampus (181). The other two studies have found cross-sectional changes in the cerebral gray matter (temporal and insular cortex), but no longitudinal changes (182,183). The discrepancy can be explained by the proportion of remitted patients at the follow-up assessment, which was less in the study, where significant changes were observed. The authors concluded, that depression-related GM changes are state-dependent and may resolve after remission (183).

Not much evidence regarding the influence of pre-existing structural changes on depressive symptoms (178). WML were associated with subsequent depressive symptoms (184).

**Brain circuits**

Major depression is associated with alterations in large-scale brain networks (185). Depression was associated with reduced connectivity within the frontoparietal network, reduced connectivity between frontoparietal network regions and parietal regions of the dorsal attention network involved in external focus. Depression was also associated with increased connectivity within the DMN, and h between control systems and regions of the DNM, or with DMN non-suppression during cognitive tasks (186). Abnormalities within the DMN were reported for the first-episode drug-naive patients (187). Successful antidepressant treatment is accompanied with increased fMRI activity in the DLPFC and VLPFC (cognitive control and attention), and decreased activity in the hippocampus, cingulum, and precuneus (DMN) (188).

Reward and non-reward (punishment) systems were found to be impaired in depression (161) (189). In particular decrease in functional connectivity of the “reward” medial orbitofrontal cortex with memory-related regions (MTL) and increase in functional connectivity of the non-reward or "punishment" lateral orbitofrontal region with precuneus (sense of self) and angular cortex (social perception). Amygdala "social" networks, perception network, in particular, are impaired in LLD, including LLD with mild cognitive impairment (MCI) (190). (166) .
**PIB PET studies**

Patients with mild cognitive impairment (MCI) and depressive symptoms had more amyloid deposition and converted to Alzheimer’s disease (AD) faster compared with patients without depressive symptoms (191,192). AD patients with a lifetime history of depression had higher levels of both plaque and tangle formation in the hippocampus than brains of patients with AD without a history of depression (193). Importantly, A-beta positive individuals had significantly higher risks of developing depressive symptoms 4.5 years later compared with A-beta negative (194).

1.5.3 **Neuroimaging studies of AD**

Neuroimaging techniques have gained a pivotal role in the diagnostics of AD. They are used as biomarkers of both pathology and progression of AD and to differentiate it from other diseases.

**Structural findings**

Brain atrophy, measured with T1-weighted sequences, can support a clinical diagnosis of AD and assess disease progression. Structural MRI (medial temporal lobe atrophy) is a biomarker of neuronal injury in AD (195) and in MCI due to AD (196).

**Atrophy progression**

Already on the stage of subjective memory impairment (SCI), some patients demonstrate a structural pattern of the AD dementia (197). Such patients have significant cortical thinning in medial and lateral temporal cortex, temporoparietal junction, inferior parietal cortex, cingulum and ventrolateral prefrontal cortex compared with healthy-like SCI. Another study directly compared cortical thickness between SCI and controls. SCI had a thinner entorhinal cortex and no differences in the hippocampal volume (198). SCI with thinner cortex in the temporal and parietal regions had a higher risk of developing AD (199).

Widespread atrophy (longitudinal data) throughout the cortex (most of the prefrontal, temporal and parietal regions) was detected already at the stage of MCI when comparing progressive MCI with clinically stable MCI (200,201) (Figure 2).

AD dementia is characterized by severe atrophy of all the cortical regions, even primary motor and sensory areas (202).

Prominent brain atrophy in AD, even on the stage of MCI, allows using brain structure measurements to discriminate between AD and controls, build predictive models for MCI-to-AD conversion, and even discriminate from other dementias. The accuracy of AD-control discrimination is usually quite high about 80-96% for the testing set (203,204). The accuracy of prediction of the MCI-to-AD conversion is about 75-81% (148,204,205). The combination
of the MRI and multivariate methods was successfully implemented in discrimination between AD and fronto-temporal dementia (203) and AD with dementia with Lewy bodies (206). Individual prediction values (scores) derived from the classification (207) can be used in clinical practice in order to improve AD diagnostics, prediction, and assessment of disease progression.

Functional

Like in depression, AD patients demonstrated non-suppression of the DMN during the cognitive tasks (208,209).

The most consistent in fMRI studies of AD is probably a decrease in the hippocampal activity during the encoding of new information (210,211).

PET

Amyloid imaging

Distribution of the amyloid plaques in the living brain can be assessed using Pittsburgh Compound-B (PIB) tracer (212). Already in early days of multimodal PIB-PET + fMRI studies, it was observed that regional distribution of the amyloid plaques overlaps with the DMN (213). It was suggested that default activity may predispose cortical regions to amyloid deposition. This observation is in line with animal studies, which shows that neuronal activity is an important factor of amyloid secretion to the extracellular space (where it forms a plaque and can be detected by PET).

Amyloid deposition thought to be the earliest marker of AD pathology (214), but later studies revealed that astrocytosis may be an earlier event (116,119).

Tau imaging

Recently developed Tau tracers (THK5317, THK5351, AV-1451, and PBB3) can shed light on the regional distribution of the tau protein (215) and its spatial and temporal relationship with other AD-related processes.

AD patients had higher tau deposition in the medial temporal and cingulate cortex compared with controls (216). In other study, MCI patients had higher tau deposition only in the entorhinal cortex, when AD patients had higher tau deposition in most cortical regions compared with controls (217).
2 AIMS

The general aim of the project was to investigate structural brain changes associated with depression in elderly people with different levels of cognitive functioning using brain MRI. Additional aims were to study underlying pathophysiological processes (markers of AD pathology and cortisol levels), and whether MRI can predict future cognitive decline in depressed elderly.

Study I
To investigate structural changes of the brain cortex associated with depressive symptoms in elderly people with AD and their association with CSF markers of AD pathology

Study II
To examine whether depression is associated with cortical thinning and reduced hippocampal volumes in nondemented elderly

Study III
To assess if structural brain MRI in elderly MDD patients can predict mild cognitive impairment (MCI) or dementia 1 year prior to the diagnosis

Study IV
To evaluate the rate of atrophy in the cerebral cortex, hippocampi, and ventral diencephalon over a 4-year period in middle-aged and older individuals with depressive symptoms, and to test if depression-related brain atrophy is associated with cortisol levels.
3 METHODS

3.1 ETHICAL APPROVAL

All studies were conducted according to the Declaration of Helsinki and subsequent revision. The study I was approved by the Regional ethics committee in Stockholm for human studies (KI) and by ethical review boards in each participating center (ADNI). The studies II and III were approved by Regional Committee of Medical Research Ethics and Privacy and Data Protection Officer at Oslo University Hospital. Study IV was approved by the Regional ethics committee in Umea.

3.2 STUDY PARTICIPANTS

Information about the study participants is summarized in table 1.

The participants of the study I were AD patients recruited from the Memory clinic at University Hospital, Karolinska Huddinge, Sweden (KI cohort) and from multiple sites in the US (ADNI cohort). ADNI was established in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public–private partnership. The purpose of ADNI was to test whether MRI, PET and other biological markers are useful in clinical trials of MCI and early AD.

The participants of the study II and III were recruited from the “Prognosis of Depression in the Elderly” (PRODE) cohort. PRODE is an observational Norwegian multicenter prospective study. Nine departments of old-age psychiatry across Norway participated in the study. The collection of data was performed by health professionals working in these departments. Standardized assessment scales were used. Assessors received standardized training prior to the study period and recurrent biannual training during the study period in order to secure reliable data.

The participants in the study IV were drawn from the Betula prospective cohort study on memory, health, and aging that was started in 1988 (218). Betula is a study of healthy individuals that have been randomly selected from the population in a town of about 120,000 inhabitants. Currently, Betula includes six samples (S1–S6) and about 4500 individuals in total. Many of the participants have been re-examined every 5 years. At each test wave (T1 1988–1990; T2 1993–1995; T3 1998–2001; T4 2003–2005; T5 2008–2010; T6 2013–2014), health examinations were performed along with extensive cognitive testing. At T5 and T6, structural and functional MRI was performed on a randomly recruited sub-sample of Betula participants from samples 1 and 3, and on an additional new sample of participants in the age range between 25 and 80 years, stratified by age and gender (n=376) (219).
3.3 CLINICAL ASSESSMENT

Participants of the studies underwent clinical assessment, including:
Demographic: Age, gender, education, occupation, family status
Medical examination - Interview performed by a physician with the patient and an informant, physical and neurological examination, vascular risk profile, information on depression and antidepressant use, Routine blood chemistry, CSF (KI and ADNI), Brain MRI.
For the PRODE cohort instrumental activities of daily living (IADL) and general health status using the General Medical Health Rating scale (GMHR) were also measured. Clinical data was collected using semi-structured interviews with the patients and caregivers and case notes. All the health-care professional involved in the data collection underwent a standardized training prior to the study period and again twice a year during the study period in order to secure reliable data (220).
For the Betula cohort health assessment consisted of anthropometric measurements, blood-pressure measurement, and information gathering regarding drug use and medical history over the last 5 years. Dementia was investigated via routine clinical methods.

3.3.1 Definition and measurement of depressive symptoms and depression diagnosis

**KI cohort**
The Cornell Scale for Depression in Dementia (CSDD) was used to assess depressive symptoms in the KI cohort. Depression was defined as a CSDD score of six or more (221). The scale has 19-items; scores are ranging from 0 to 38. The CSDD was completed by a licensed geriatrician or psychiatrist, or an experienced nurse after training. The CSDD is designed for the assessment of depression in older people with dementia who can at least communicate basic needs. The CSDD differentiates between the diagnostic categories and severity of depression. The scale was tested for validity, reliability and sensitivity in different settings (community, hospital, home care) (222). Scores are determined by a combination of prior observation and two interviews: 20 minutes with the career and 10 minutes with the patient.

**ADNI cohort**

Subjects with subsyndromal depression were defined with a score of 1–5 on the Geriatric Depression Scale (GDS) and non-depressed subjects with a score of 0. The ‘memory’ item, asking “if the subject feels he/ she has more memory problems than most”, was excluded due to its association with cognitive performance. The GDS created by Yesavage, et al. has been tested and used extensively with the older population. The GDS Long Form has 30 yes/no questions in reference to how they felt over the past week (223). A Short Form of the scale consists of 15 questions, which were selected from the long version, based on their discriminative ability. The Short Form is preferable for physically ill and mildly—to--moderately demented patients due to the short attention spans and fatigue. In a validation study comparing the Long and Short Forms of the GDS for self-rating of symptoms of depression, both were successful in differentiating depressed from non-depressed adults with a high correlation (r = .84, p < .001) (224). GDS was performed by a study/site physician.

**PRODE cohort**

A depressive episode was diagnosed based on the International Classification of Diseases (ICD-10) criteria. Diagnostic criteria for depression ICD-10 include a list of ten depressive symptoms:

Key symptoms are: persistent sadness or low mood; and/or loss of interests or pleasure, fatigue or low energy at least one of these, most days, most of the time for at least 2 weeks

Associated symptoms are: disturbed sleep, poor concentration or indecisiveness, low self-confidence, poor or increased appetite, suicidal thoughts or acts, agitation or slowing of movements, guilt or self-blame

The severity of depressive symptoms was also assessed with several depression scales: Montgomery-Asberg Depression Scale, Cornell Scale for Depression in Dementia, Hamilton Depression Scale, Hospital Anxiety and Depression Scale.

**Betula cohort**
Center for Epidemiologic Studies Depression Scale (CES-D 10) (225) is a 10-item self-report scale designed to measure depressive symptomatology in the general population. Responses capture the frequency of feelings and behaviors over the past 7 days and are rated on a 4-point scale ranging from 0 (rarely or none of the time) to 3 (most or all of the time) on 10 items. This brief screening tool for depressive symptoms was taken from the longer well-validated 20-item Center for Epidemiological Studies Depression Scale (226), which has a cut-off > 16 for major depression. There is a high agreement between cut-off score of 10 on the CES-D 10 and cut-off score of 16 on the CES-D 20 (225,227). Thus, in this study, a depression was defined as a CES-10 score ≥ 10. The assessment was performed by a trained research nurse (228).

3.3.2 Diagnosis of AD and MCI

ADNI cohort

Probable AD was diagnosed according to the NINCDS/ADRDA 1984 criteria (229) after a standardized clinical and cognitive examination and imaging and CSF-based biomarkers assessment.

KI cohort

Diagnosis of AD was made by consensus achieved during conferencing among specialists in neurology, geriatrics and psychiatry, nurses, neuropsychologists and speech therapist taking into account all available information, such as clinical examination, including neuropsychological, neurological and psychiatric examinations, CSF and blood tests and neuroimaging. Patients were diagnosed with AD dementia according to the International classification of Diseases – Tenth revision (ICD-10).

PRODE cohort

Patients were diagnosed with different types of dementia according to the International classification of Diseases – Tenth revision (ICD-10). MCI was diagnosed using Winblad’s criteria (230). Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE-16), which is a structured questionnaire of cognitive functioning during the last 10 years, was also administered.

3.4 NEUROPSYCHOLOGICAL ASSESSMENT

Participants of the studies I-III underwent the following neuropsychological assessment selected from the Wechsler Adult Intelligence Scale (WAIS) (231):

KI cohort

Mental State examination (MMSE) and Full-Scale IQ, Similarities, Information and Vocabulary, complex figure copying, Block design, Matrix, digit span, Rey auditory verbal learning test, Rey complex figure immediate retention, Trail making test A and B and Digit symbol.
**ADNI cohort**

MMSE, ADAS-cog (cognitive subscale of the Alzheimer's Disease Assessment Scale), American National Adult Reading Test (AMNART), Boston Naming tests, Digit span test, Ray auditory verbal learning test, Semantic fluency, Trail making test A and B and Digit symbol, Wechsler Adult Intelligence Scale (revised)

**PRODE cohort**

MMSE, Trail making test A and B, Controlled Oral Word Association Test (COWA) were measured at inclusion, at discharge and after the one-year follow-up period.

**Betula cohort**

Neurocognitive functions were extensively evaluated in the Betula cohort, using the following tests:

Episodic memory Face recognition, Name recognition, four-alternative forced choice, 16 items at study Sentence learning with encoding enactment, free recall, two cued recall tests, source recall, Sentence learning without encoding enactment, free recall, two cued recall, source recall, 16 items at study Word recall (T1–T4) with or without concurrent card sorting at study or test, free recall, 10 items Prospective memory, recall without and with reminding cue Memory for activities, free recall Semantic memory Knowledge recall, Word fluency, initial letter A Word fluency, initial letter M, five-letter words Word fluency, initial letter B, names of professions Word fluency, initial letter S, five-letter names of animals Word comprehension, 30 items Priming Name-stem completion, Word fragment completion. Other tasks Block design, Tower of Hanoi, MMSE, Letter digit substitution (228).

**3.5 LABORATORY DATA**

**3.5.1 CSF**

Participants of study I underwent CSF sampling.

For the ADNI cohort CSF was collected after an overnight fast using a 20- or 24-gauge spinal needle, frozen within 1 h of collection, and transported on dry ice to the ADNI Biomarker Core laboratory at the University of Pennsylvania Medical Center. The team involved in the analyses was unaware of the diagnoses and study hypotheses. The complete descriptions of the collection and transportation protocols are provided in the ADNI procedural manual at http://www.adni-info.org. CSF data were available for 79 subjects.

For the KI cohort CSF was collected in a subgroup (n=39) as a part of the memory clinic investigation. CSF was collected in the morning after an overnight fast. Participants underwent lumbar puncture in the L3–4 or L4–5 interspace. CSF was collected in 10 ml polypropylene tubes, and centrifuged within 2 h and then frozen until the analysis. CSF Ab42 was analyzed using a sandwich ELISA, constructed to specifically measure b-amyloid 42. CSF t-τ and p-τ were determined using a sandwich ELISA. All CSF samples were analyzed
at the department of clinical chemistry, Karolinska University Hospital, Stockholm, Sweden. The team involved in the analyses was unaware of the diagnoses and study hypotheses.

3.5.2 Cortisol

For the 49 participants of the study IV, cortisol measurements were available. Participants were instructed to sample saliva during 1 day at 07:00, 11:00, 16:00, and 23:00 h. Participants were instructed not to eat, brush their teeth, or smoke 1 h before sampling and to store their saliva in the refrigerator for up to 1 day before handing it over to the research staff. Saliva samples were frozen at −20°C in 4–6 years until cortisol levels were measured. Samples were analyzed in one batch with a chemiluminescence immunoassay (IBL International GmbH, Hamburg, Germany). The inter- and intra-assay coefficients of variation for the analysis were <8%. Outliers were defined by the cortisol level more than 3 SD around the mean and were excluded from further analysis. 16:00 and 23:00 cortisol measurements were missing for two subjects and replaced with the sample mean. The distributions of the cortisol measurements were skewed. In order to normalize them, logarithmic transformation was performed. Cortisol levels at different time-points were treated separately since morning and afternoon cortisol levels might have a different effect on depressive symptoms and cognitive functioning (232), (233), (234).

3.6 MRI METHODS

3.6.1 Acquisition parameters

Study I

ADNI: The data were collected at multiple ADNI sites using a standardized MRI protocol. For each subject T1-weighted MRI scans were collected using a sagittal volumetric magnetization-prepared rapid gradient echo (3D MP-RAGE) sequence with the following acquisition parameters: echo time of 4 ms, repetition time of 9 ms and resolution of 1.1×1.1×1.2 mm. The ADNI MRI quality control center at the Mayo Clinic (Rochester, MN) selected the MP-RAGE image with higher quality based on standardized criteria.

KI: High-resolution T1-weighted images were acquired with MPRAGE sequence on a 3-T TIM TRIO scanner (Siemens, Erlangen, Germany) at the Memory Clinic at Karolinska University Hospital in Huddinge (Sweden). These images were obtained using the following sequence parameters: 176 sagittal slices; time of repetition/time of echo=1900/ 2.57 ms; flip angle 9°; acquisition matrix=256×256, voxel size 1×1×1 mm3.

Study II and III

All MRI brain images were acquired using 1.5-Tesla scanners in study II, and 1.5 and 3-Tesla scanners in study III. The data were collected at several sites using a standardized MRI PRODE protocol. Sequence parameters are represented in the tables 2 and 3 below.
Tab. 2 MRI scanning parameters and number of participants Study II

<table>
<thead>
<tr>
<th>Scanner</th>
<th>Slice thickness</th>
<th>TR</th>
<th>TE</th>
<th>FA</th>
<th>AM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avanto</td>
<td>1.2mm</td>
<td>2400</td>
<td>3.79</td>
<td>8</td>
<td>192x192</td>
</tr>
<tr>
<td>Symphony</td>
<td>1.2mm</td>
<td>2400</td>
<td>3.71</td>
<td>8</td>
<td>192x192</td>
</tr>
<tr>
<td>Avanto</td>
<td>1.2mm</td>
<td>2400</td>
<td>3.61</td>
<td>8</td>
<td>192x192</td>
</tr>
<tr>
<td>Avanto</td>
<td>1.2mm</td>
<td>2400</td>
<td>3.61</td>
<td>8</td>
<td>192x192</td>
</tr>
<tr>
<td>Symphony</td>
<td>1.2mm</td>
<td>2400</td>
<td>2.88</td>
<td>8</td>
<td>192x192</td>
</tr>
</tbody>
</table>

Notes: TR-repetition time, TE-echo time, FA-flip angle, AM-acquisition matrix

Tab. 3 MRI scanning parameters and number of participants Study III

<table>
<thead>
<tr>
<th>center N</th>
<th>Scanner</th>
<th>Slice thickness (mm)</th>
<th>TR (ms)</th>
<th>TE (ms)</th>
<th>FA (°)</th>
<th>MFS (Tesla)</th>
<th>CS</th>
<th>MCI-DEM</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Symphony</td>
<td>1.2</td>
<td>2400</td>
<td>3.71</td>
<td>8</td>
<td>1.5</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Avanto</td>
<td>1.2</td>
<td>2400</td>
<td>3.79</td>
<td>8</td>
<td>1.5</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Avanto</td>
<td>1.2</td>
<td>2400</td>
<td>3.61</td>
<td>8</td>
<td>1.5</td>
<td>9</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>Symphony</td>
<td>1.2</td>
<td>2400</td>
<td>2.88</td>
<td>8</td>
<td>1.5</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Avanto</td>
<td>1.2</td>
<td>2400</td>
<td>3.61</td>
<td>8</td>
<td>1.5</td>
<td>14</td>
<td>12</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>Philips Intera</td>
<td>1.2</td>
<td>6.73</td>
<td>3.1</td>
<td>8</td>
<td>3</td>
<td>9</td>
<td>8</td>
<td>17</td>
</tr>
</tbody>
</table>

Notes: TR-repetition time, TE-echo time, FA-flip angle, MFS - Magnetic field strength
Study IV

Brain MRI data was collected at two time-points with four years interval for each participant from a 3 Tesla GE scanner equipped with a 32-channel head coil. High-resolution T1-weighted structural images were acquired using the following parameters: 180 slices; 1 mm thickness; TR 8.2 ms; TE 3.2 ms; flip angle 12°; FOV 25 × 25 cm.

3.6.2 Freesurfer analyses

Regional cortical thicknesses and volumetric measures were estimated in Freesurfer (v. 5.1 for studies I-III and v.5.3 for study IV), using The T1 3D brain MRI scans as input. The software is well documented and available for download online.

Freesurfer pipeline consist of two relatively independent streams:

The Surface-based Stream (estimates cortical thickness)

The surface-based pipeline consists of several stages (235), (236)

1) The volume is registered with the MNI305 atlas (237)
2) The variation in the white matter intensity is measured (bias field)
3) The intensity at each voxel is then divided by the estimated bias field at that location in order to remove its effect.
4) The skull is stripped
5) Voxels are classified as a white matter or not than white matter based on intensity and neighbor constraints.
6) Based on the expected MNI location of the corpus callosum and pons, the hemispheres are separated
7) The cerebellum and brain stem are removed
8) Initial surface is generated (a border between the gray and white matter)
9) Initial surface is refined according to the gradient (which should be between white and gray matter) – “white surface”
10) The pial surface (a border between the gray matter and CSF) is identified by following the intensity gradients between the gray matter and CSF
11) Cortical thickness is defined as the distance between the white and the pial surfaces at each location of the cortex (238). Locations of cortex obtained by separating the surface on vertices. The area of a vertex is defined by the average of the triangles surrounding the vertex.

The Volume-based (Subcortical) Stream (estimates volumes of subcortical structures)

The stream consists of five stages (239)

1) The volume is registered with the MNI305 atlas insensitive to pathology
2) Initial volumetric labeling
3) Correction of intensity variation
4) High dimensional nonlinear volumetric alignment to the MNI305 atlas
5) Labeling

Longitudinal Processing Stream
First, each scan is analyzed in the standard cross-sectional Freesurfer stream, after that, in the longitudinal processing stream unbiased within-subject template space and average image is created using robust, inverse consistent registration (240). Information from this subject template is used to initialize the longitudinal image processing in several locations to increase repeatability and statistical power.

3.7 STATISTICAL METHODS
Descriptive variables are presented as mean (SD) for normally distributed data and in the median and interquartile range (IQR) for non-normally distributed data. P<0.05 was set as a threshold of statistical significance if not specified otherwise. SPSS v 20 and 22, R programming language (R Core Team, 2016) v. 3.3.0, and Freesurfer (v.5.1 – Study I-III, and v.5.3 – study IV) qdec toolbox were used for statistical analysis in the project.

3.7.1 Univariate

3.7.1.1 Demographics (Studies I-IV)
For group comparison of the demographic and clinical data the following tests were used: Chi-square for categorical data, Student’s t-test (normally distributed) or Mann–Whitney U test (non-normally distributed) for two-group comparison of continuous data and 2-way ANOVA for three-group comparison of normally distributed continuous data with post-hoc pairwise comparisons.

3.7.1.2 Surface-based Statistical Analysis (Studies I, II and IV)
The surface-based statistical analysis was performed in Freesurfer by fitting a general linear model (GLM) at each vertex. GLM analysis on the surface enables to test models of how any surface-based measure (thickness) might change as a function of demographic and clinical variables, group membership (depressed vs nondepressed). FreeSurfer has a graphical toolbox for data analysis – Qdec, where the main statistical analysis for the studies was performed. This toolbox allows cortical thickness at each vertex to be used as a dependent variable in the linear model.
The smoothing kernel (full width at half maximum (FWHM)) for the reported results was 10 mm. Correction for multiple comparisons was carried out using the Monte Carlo simulation method. Data were tested against an empirical null distribution of maximum cluster size by running 10 000 synthesized Gaussian noise simulations with an initial cluster-forming threshold of p<0.05 with a vertex-level threshold of 200 vertices.

Study I
Cortical thickness was compared between AD patients with and without depressive symptoms, adjusting for age, gender, education and MMSE score. The interaction term was used to assess moderation effect of the depression on the relationship between CSF tau levels and cortical thickness.

**Study II**
We compared cortical thickness and HVs between LLD and controls and correlated these structural measurements with variables of interest (depression onset, MMSE, the number of depressive episodes). Interaction term was used to assess moderation effect of the depression on the relationship between age and cortical thickness, and to assess moderation effect of gender/depression onset on the relationship between depression and cortical thickness.

**Study IV**
We used two-stage model implemented in Freesurfer to compared atrophy rate among individuals with depressive symptoms at baseline vs controls and vs individuals who developed depressive symptoms during the follow-up period. On the first stage, temporal data within each subject was reduced to a single statistic. The rate of change of thickness was estimated as the difference per time unit, so rate = (thick2 - thick1) / (time2 - time1), here thickening in mm/year. Second, this measure was compared across groups and correlated with cortisol levels in a cross-sectional manner.

**3.7.1.3 Multivariable linear modeling of extracted ROI and variables of interest**
We correlated left and right hippocampal volumes with MMSE scores (study II, III) and cortisol levels (study IV). Other covariates in the models were: age, gender, education and intracranial volume (ICV) (if the volumes were not adjusted in advance). In study III we also correlated volumes of the right ventral diencephalon and mid-anterior corpus callosum with the number of depressive episodes, adjusting for age, gender, MMSE and the total CSF volume.

**3.7.2 Multivariate - Random Forest (study III)**
RF selected a bootstrapped subset of all observations – about 66% per tree and a random subset of all predictors/features (here: cortical thickness and subcortical volumes) at each node of the tree. The remaining 33% of the data, out-of-bag (OOB) data, was used to measure the RF performance.

RF was used to discriminate between the MCI-DEM/MCI and CS based on the CTH and SV measures separately and combined. In addition, demographic and clinical information was added to the models to test if performance could be further improved. Only the clinical information obtained at inclusion was used to assess if MCI-DEM statuses 1-year later could be predicted based on the earliest available clinical data and in the depressive state.

Five thousand decision trees were used in the RF classification models. RF models were trained to discriminate between MCI-DEM (n = 29) or MCI alone (n = 21) and corresponding CS patients. When discriminating between MCI vs CS, the number of CS was reduced to 30 (from 40) matched on age and gender and scanner field strength, in order to keep balance in
group class distribution. AUC, sensitivity/specificity, overall accuracy, and kappa, were used to assess the performance of the models. Confidence intervals (CI) were estimated using bootstrapping (n = 100). The most relevant structures for prediction of MCI-DEM or MCI were correlated with clinical variables related to depression in order to detect regions involved in both pathological processes. The previously established ADNI model (204) was implemented in the PRODE dataset to evaluate its ability to discriminate between MCI-DEM (n=29) and CS group (n=51).
4 SUMMARY OF THE FINDINGS

4.1 DEMOGRAPHICS

Baseline characteristics of the participants are provided in table 4. There were no significant differences in the main demographic and clinical characteristics between the participants with and without depressive symptoms, except in the KI cohort – there were significantly less women in the group of AD patients with depressive symptoms compared with AD without depressive symptoms.

<table>
<thead>
<tr>
<th>Study n</th>
<th>Cohort</th>
<th>Group</th>
<th>n</th>
<th>Age^</th>
<th>Gender (women)</th>
<th>Education (years)</th>
<th>MMSE</th>
<th>ADT¤</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KI</td>
<td>AD+Dp</td>
<td>16</td>
<td>66(6.25)</td>
<td>44%</td>
<td>12(2.75)</td>
<td>22(3.0)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>AD</td>
<td>AD</td>
<td>25</td>
<td>67.5(5.75)</td>
<td>72%</td>
<td>12(3.75)</td>
<td>25(2.5)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>AD+Dp</td>
<td>AD+Dp</td>
<td>84</td>
<td>76(4.85)</td>
<td>53%</td>
<td>16(2.0)</td>
<td>24(1.0)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>AD</td>
<td>AD</td>
<td>64</td>
<td>75(5.60)</td>
<td>52%</td>
<td>15(2.0)</td>
<td>23(1.5)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>LLD</td>
<td>LLD</td>
<td>49</td>
<td>75.6(6.46)</td>
<td>73%</td>
<td>9.7(2.80)</td>
<td>27.4(2.64)</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Controls</td>
<td>Controls</td>
<td>49</td>
<td>74.9(6.25)</td>
<td>71.4%</td>
<td>13.5(3.3)</td>
<td>28.8(1.44)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>PRODE</td>
<td>LLD-to-MCI/DEM</td>
<td>21+8</td>
<td>78.1(7.3)</td>
<td>75%</td>
<td>8.9(2.79)</td>
<td>24.6(3.0)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>LLD</td>
<td>LLD</td>
<td>40</td>
<td>76.4(5.8)</td>
<td>72%</td>
<td>10.1(2.61)</td>
<td>26.4(1.8)</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Betula</td>
<td>Dp symptoms*</td>
<td>23+2</td>
<td>58.15(15.04)</td>
<td>,</td>
<td>12.74(4.30),</td>
<td>28.3(1.41),</td>
<td>27.67(1.57)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>58.96(13.77)</td>
<td>73%, 56%</td>
<td>13.54(3.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls</td>
<td>35</td>
<td>59.63(12.8)</td>
<td>68%</td>
<td>13.69(3.69)</td>
<td>28.45(1.36)</td>
<td>3</td>
</tr>
</tbody>
</table>

^Mean (SD) if not specified;¤ADT – antidepressant treatment, significant differences are in bold; *21 MCI+8 dementia; **the first numbers in the row refers to the characteristics of the participants with depressive symptoms at baseline and the second describe participants with depressive symptoms at follow-up only, #Median (IQR)
4.2 STRUCTURAL BRAIN CHANGES ASSOCIATED WITH DEPRESSIVE SYMPTOMS IN FOUR INDEPENDENT COHORTS

We have found that depressive symptoms were associated with cortical thinning the same region – left supramarginal cortex in four independent cohorts of elderly people with and without AD (Figure 3, table 5). Depression-related thinning in ADNI, KI and the Betula cohorts was also observed in superior temporal cortex and temporal pole. In two non-demented cohorts (PRODE and Betula) angular cortex was also involved in depression. Longitudinal analysis revealed that thinning in these regions is secondary to depressive symptoms in the Betula cohort. In the ADNI and PRODE cohorts, fusiform cortex was involved in depression. In the Betula cohort, we also were able to assess thinning which developed in parallel with depressive symptoms. It covered medial superior frontal cortex and lingual cortex.

<table>
<thead>
<tr>
<th>Tab. 5 Structural brain changes associated with depressive symptoms in four independent cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
</tr>
<tr>
<td>Left Supramarginal</td>
</tr>
<tr>
<td>Left Superior Temporal</td>
</tr>
<tr>
<td>Left Temporal Pole</td>
</tr>
<tr>
<td>Left Angular</td>
</tr>
<tr>
<td>Left Fusiform</td>
</tr>
<tr>
<td>Left Parahippocampal</td>
</tr>
<tr>
<td>Left Medial Superior frontal</td>
</tr>
<tr>
<td>Right Superior temporal</td>
</tr>
<tr>
<td>Right Supramarginal</td>
</tr>
<tr>
<td>Right Angular</td>
</tr>
<tr>
<td>Right Angular</td>
</tr>
<tr>
<td>Right Lingual</td>
</tr>
</tbody>
</table>
4.3 CLINICAL CHARACTERISTICS OF DEPRESSION AND BRAIN STRUCTURE

The number of depressive episodes was associated with cortical thinning in the left temporal pole in women and reduced volume of the right ventral diencephalon in both – men and women.

We have found a moderating effect of gender on the relationship between cortical thickness and depression onset. Late-onset (LOD) (onset>65 years) women but not men had the widespread thinning in the prefrontal cortex compared with early onset depressed (EOD).

4.4 STRUCTURAL BRAIN CHANGES ASSOCIATED WITH COGNITIVE FUNCTIONS IN ELDERLY WITH DEPRESSION.

MMSE correlated positively with the volume of the right hippocampus (p=0.009) and superior frontal cortex (Cluster-wise p-value (CWP) =0.006) in the elderly with a depressive episode. This effect was more pronounced in the subgroup of late-onset depressed (SFC CWP=0.0001) (study II).

The volume of the right ventral diencephalon was associated with cognitive decline one year later in the elderly with a depressive episode (study III).
4.5 BIOLOGICAL CORRELATES OF STRUCTURAL BRAIN CHANGES (STUDY I AND IV)

CSF levels of total and phosphorylated tau were associated with cortical thinning in the cluster covering right posterior cingulate cortex and precuneus (CWP=0.008) and cluster covering right parahippocampal and fusiform gyri (CWP =0.001) in the AD patients with depressive symptoms (KI cohort, study I). No association was found in the non-depressed AD group.

Higher baseline saliva cortisol levels in non-demented individuals (study IV) were associated with widespread cortical atrophy in temporal, prefrontal and parietal cortex bilaterally and the right hippocampus, independently of age and MMSE.
5 DISCUSSION

5.1 INTERPRETATION OF THE RESULTS

5.1.1 Regions involved in depression

The regions involved in depression that were observed in this project were: superior temporal cortex (STC), Temporal pole (TP), supramarginal cortex (SMC), lingual cortex (LC), fusiform cortex (FF), parahippocampal cortex (PH) and superior frontal cortex (SFC). All these regions are involved in processing of information about other people, including faces, voices and touch (perceptual aspects of communication). Each is activated in response to faces (241,242). The fusiform gyrus is known to be a face-specific region activated in response to any face (243). The posterior STC (pSTC) and SMC are activated in response to personally familiar faces, and in response to emotionally significant faces (242,244), suggesting the role of these regions in emotional aspects of social perception. Studies using other modalities demonstrate similar results – recruitment of STC and SMC in perception of emotionally significant voices and gentle touch subjectively rated as positively valent (245,246). The relationship between pSTC and positive valence is supported by the fact that it gets suppressed when one is provoked to sadness (247). The fusiform gyrus is crucial for explicit face recognition. Consequently, when it is damaged patient do not recognize familiar people explicitly, but still develop autonomic response, with which recognition occurs implicitly. Similarly, STC could play a role in the context of positive valence attributing to social stimuli. Its alteration may play a role in social withdrawal observed in depression and AD. Multimodal EEG-fMRI studies should reveal if STC activation is a result or a cause of positively valent social perception. The temporoparietal junction (pSTC, SMC and angular cortex), integrates information from the temporal, parietal and occipital lobes, and is believed to play an important role in the theory of mind (what one believes in mental states of others).

Structural alteration in each of the regions investigated, have been found in different age groups of depressed (248–251) supporting the involvement of altered social processing regions in depression. However, these alterations are secondary to the depressive symptoms (181), (Lebedeva - manuscript), suggesting that impairment in social processing is a consequence of depression. Importantly, larger gray matter volume in the SMC correlates with faster recovered from depressive episode (182). Volume of the STC was important for prediction of depression severity (252) and discrimination between depressed and non-depressed (253). Functional alterations in STC, SMG, FC and LC during the sad face representation were important for discrimination between depressed and non-depressed (254). Interestingly, brain activation in response to slightly sad faces had higher discriminative ability compared with activation in response to very sad faces. Regions that develop atrophy first, in preclinical and mild AD, include the SFC, FF, SMC and LC). The most severe change after the AD onset is found among the STC, SMC, FF
regions (255). Consistent with topological mapping of such alterations, typical symptom of AD include prosopagnosia (LC, FF, TP), anosognosia (SFC, SMC - (256)), deficits in praxicon and prosody (257), (including affective prosody (258)), social-emotional agnosia in general (259), perception (SMC, pSTC), social withdrawal and apathy (SFC, pSTC). Together these findings suggest that mild (reversible?) alterations in these regions are associated with symptoms of depression (23) and severe, progressive damage in these regions is typical for AD. Depression might lead to functional deactivation of STC and SMC which may be responsible for perception of positive emotions. Decreased STC, SMC activity might lead to a subsequent atrophy of the region (probably through a hypoperfusion) and more persistent loss of its function (260). Future studies should uncover which pathways lead to inhibition of the STC.

Studies show that brain structure undergoes neuroplastic changes, for instance following the psychotherapy (168,261). But these studies were performed in younger adults – it is not known to what extent the elderly brain preserves neuroplastic abilities, especially those with high risk of AD. High levels of the neuronal loss markers in AD suggest low neuroplastic potential. We have found that tau levels were associated with cortical thinning in MTL in depressed AD. There was no association between a-beta and MTL in depressed. These findings are in consistency with previous studies showing that tau, but not a-beta is associated with hippocampal volume loss (262). Probably markers of neuronal loss should be helpful in assessment of therapeutic potential in individuals with at-high risk of AD, including depressed.

The medial superior frontal cortex is observed to atrophy in parallel with the development of depressive symptoms. Many studies have shown that the medial prefrontal cortex is involved in self-referential processes. A superior portion, in its turn, plays a role in awareness of own emotions, being activated during self-introspection promoted by some stimuli (263). Awareness of own emotions is known to be impaired in depressed, and in depressed elderly in particular (264). Interestingly, reduced metabolism in the temporoparietal junction (including pSTC, SMC and angular cortex), and the superior frontal cortex, is associated with inability of AD patients to perceive themselves from a third person perspective (256). SFC and LC could develop atrophy earlier because their dysfunction might be a contributor to depression, or because they might be less resistant to hypo-activity (more plastic/vulnerable).

5.1.2 Cortisol and brain atrophy

Depressive symptoms were associated with atrophy of temporal neocortex; however, we did not find any association between hippocampal volumes and depression. It has been shown that hippocampal volume loss can be observed in patients with clinical depression, whose duration of illness is longer than 2 years or who had more than 1 disease episode (265), which could be explained by elevated cortisol levels and cortisol toxic effect on the hippocampus. Indeed, our results demonstrate that cortisol is associated with subsequent atrophy of the right hippocampus independently of depressive symptoms. Cortisol was also associated with atrophy in most cortical regions, including medial and lateral temporal cortex, parietal and
prefrontal cortex, but not occipital cortex. Remarkably, this pattern of atrophy is typical for sporadic AD (201). The role of cortisol in AD development is supported by the results of a recent study where cortisol elevation predicted increased risk of Alzheimer’s disease on average of 6 years before the onset (266).

Mechanisms of aging and damaging effects of cortisol on the brain are not perfectly understood. Being a steroid, cortisol is able to alter gene expression in the brain cells in order to create an adaptive response. Cortisol stimulates activation of NMDA-receptors (N-methyl-D-aspartate), which enhance learning and memory (267). However, such a change in a catabolic direction (including blockage of protein synthesis, including nerve growth factors) have a short-term benefit, but, if occur too often the cell will not be able to restore (allostatic load) and becomes damaged (partially due to excessive Ca2+ influx - excitotoxicity). Accumulation of damage is a typical characteristic of aging. Moreover, hyperactivation of neurons leads to a hyperproduction of amyloid (90). Recent evidence suggests that cortisol may also promote aging directly upregulating the transcriptional activity in the “clock circadian regulator” genes (268). Another pathway might be that cortisol elevation during the depressive episode leads to a deficiency in adaptive immunity and subsequent chronic inflammation (with depression as its clinical representation).

Also, our findings of cortisol effect on the medial temporal atrophy are consistent with the “age-by-disease interaction hypothesis” and with the “glucocorticoid cascade hypothesis”.

Cortisol was associated with atrophy in STC and SMC, both of which are found to be altered in depressed. Interestingly, other steroid hormones like progesterone and testosterone are associated with atrophy in these regions in women (269,270), suggesting that plastic changes in the brain cortex may be ontogenetically adaptive and might aim to modify behavior according to actual need (puberty/pregnancy and maternal behavior/hostile environment). It has been shown that cortisol is a key regulators of neuronal activity under stress (271). For instance, social avoidance behavior is beneficial for individuals with a deficiency in adaptive immunity since it will decrease the risk of got sick. These might explain why such phenomenon as depression evolved. Naturally, stress reaction would be a result of a mismatch between the expectations and the actual outcome. In such a case plastic changes of the regions involved in "beliefs about others" may promote adaptive changes in beliefs and behavior. Another known phenomena in depressed —negative attention (272) and memory (273) bias may be beneficial in hostile conditions. Depression, it is hypothesized, evolved in order to accelerate disengagement from unreachable goals (274). This hypothesis nicely explains why brain cortex, especially regions involved in memory and social behavior are so sensitive to cortisol and other steroids. Probably, in a case of major changes in environment (loss/life-threat/childbirth) steroids promote brain "reload". It is possible that in mid-life such adaptive "reconstructions" are reversible; however, occurrence in late-life may be non-reversible due to decreased brain plasticity.
5.1.3 Depressive episodes and brain structure

The effect of cortisol on the hippocampus and brain cortex may explain why previous depressive episodes are strong predictors for future episodes (275). A depressive episode is associated with cortisol elevation, which damages the hippocampus, decreasing its ability to downregulate cortisol. Cortisol can be elevated even further, promoting immunological impairment, inflammation and instigating new depressive episodes. This circle leads to progressive HPA dysregulation and brain atrophy with each new episode. This may help to explains why cortisol elevation and atrophy in the hippocampus and temporal cortex are associated with the number of depressive episodes and disease duration (10,11). The progressive stress-related brain atrophy in the fronto-limbic area could explain the progression of negative symptoms (loss of cognitive, affective and social abilities) in depression, AD and even schizophrenia (276). Another important conclusion is that subcortical and cortical alterations, which develop in major psychiatric disorders via cortisol, may contribute to AD development. The number of depressive episodes may be a proxy measure for this process since it doubles the risk of dementia in depressed and bipolar patients (277). This is supported by our finding of the relationship between the number of depressive episodes and the right ventral diencephalon volume (RVD), given that the RVD volume was associated with a subsequent cognitive decline in LLD (278).

5.2 CONFLICTING RESULTS

There are numerous studies, including longitudinal and meta-analytical, showing a reduction of the brain structure in depression (153,154,174). However, some studies did not observe any brain changes associated with depression or depressive symptoms or observed positive associations.

Auning et al. did not find any relationship between preselected ROI (hippocampal volume, entorhinal, orbitofrontal, and anterior cingulate cortex), FTG PET metabolism, CSF biomarkers of AD and depressive symptoms (279). In contrast in all the analyses, they have found a trend towards less AD pathology in SCI and MCI with depressive symptoms compared with nondepressed. In our studies, we also have not found any relationship with the regions mentioned above and depressive symptoms. However trend towards less AD pathology in depressed in their results, could be explained by a confounding effect of the disease stage, since SCI tend to have more depressive symptoms but better glucose metabolism, less atrophy, and less abnormal CSF biomarkers, given that SCI and MCI were combined into one group. No significant results were observed when the SCI and MCI were analyzed separately, which could be explained with a low sample size (22 SCI in total).

Two studies have shown that subclinical depressive symptoms are associated with increased GM volumes and cortical thickness in various regions in healthy individuals (280,281). In contrast with many other studies, these two studies have used depression scale score as a continuous variable. In our understanding using scale scores as a binary measure is preferable since depression scale scores are usually quite skewed which might affect the results. There
was no significant association between depressive symptoms and brain structure when the total scores of depression scale were used. Depressive symptoms were associated with smaller hippocampal volumes in SCI but not in AD (282). One of the explanations might be that depressive symptoms were defined by a scale score or presence of antidepressant treatment. Evidence suggests that SSRI treatment may promote neurogenesis in the hippocampus (283), which could affect the result. However, one meta-analytical study which has found reduced GM in the hippocampus and parahippocampal region in LLD, also found increased lingual GM (176). The variability in the results may also reflect state-dependent manner of structural brain changes in depressed (183). Future research may benefit from integration between the results of structural and fMRI studies assessing the normal variability of structure and function with the imaging studies focused on clinical populations.

5.3 METHODOLOGICAL CONSIDERATIONS

General quality of the study can be assessed using internal validity (systematic error), precision (random error), causality and external validity (generalizability).

5.3.1 Internal validity

The lower the systematic error (bias): the higher the internal validity. Bias can be defined as a process of a systematic deviation from the true results. Systematic error consists of selection bias, information bias, and confounding effects. Selection bias occurs when the relationship observed in the study is not present out of the study. For instance in study I for the KI cohort, depressive symptoms were assessed for a subsample of patients, as a part of a routine medical examination (could be performed if clinician was concerned about depression), this constituted a potential source of selection bias, resulting in more depressive symptoms in those who underwent depression assessment compared with those who did not; however, even if present, such bias would lead to false-negative results. In the ADNI cohort, depressive symptoms were assessed in those, able to cooperate – understand and answer the questions. This could lead to a selection bias towards less cognitively impaired patients. However, this is, in general, known limitation of depression assessment – evaluation in severely demented patients lack validity. Some studies use imaging data that was collected as a part of a routine medical examination, which could also lead to a selection bias towards the presence of more severe or atypical clinical cases in the group who underwent brain imaging. However, for the KI, ADNI and PRODE cohort, neuroimaging data was collected specifically for research purposes. Every patient eligible for inclusion underwent brain MRI to minimize the probability of selection bias. Even though Freesurfer is a very robust imaging software (284), some segmentation errors might occur, especially in cases of severe brain atrophy. Two AD patients from the KI cohort we excluded due to inaccurate segmentation, creating some bias towards less severe cases being included in the analysis. Using a control group from another setting is another source of a selection bias. In the KI and ADNI cohorts, we did not have a control group in terms of cognitively
normal elderly (both depressed and non-depressed had AD). In the study III (PRODE cohort) control group was collected from the patients of the same hospitals from somatic units (mostly surgery). The presence of a control group from the same setting is an advantage (minimize the "healthy worker effect"). However, it could be suboptimal if we would measure anxiety or subsyndromal depression since patients undergoing a surgery usually have high levels of stress. PRODE patients had a clinical diagnosis of a depressive episode (many had moderate and severe depression). All controls were also assessed for depressive symptoms. The probability of selection bias due to the control group was minimized. In the study IV (Betula cohort) participants were randomly selected from the population registry in Umeå, Sweden. From the total sample, a subsample was randomly selected for the brain MRI acquisition. A possible source of selection bias here is that people who were more concerned regarding their health were more likely to participate in the study. However, demographic characteristics of the sample were similar to the characteristics of the Umeå population (228). The loss to follow-up is another potential source of selection bias. In the study III, 12 patients left the study before the one-year follow-up assessment. Even though there was no significant difference in main clinical and demographic characteristics between the baseline sample and the follow-up one, attrition in some cases happened due to death, suggesting some selection bias towards healthier patients. In the Betula cohort, there was no attrition during the four-year follow-up period.

Information bias suggests a presence of measurement errors. In the KI and ADNI cohorts, depressive scales, scored by a geriatrician or psychiatrist, were the basis for measurement, which suggests a proper quality of the data collection. The subjectivity of depression assessment is a general problem of the field; however, assessment becomes even harder when performed in demented patients. In the CDSS (KI cohort), caregivers’ impressions were taken into account. In contrast, in the GDS (ADNI cohort), the patients’ answers were used, which might make results less reliable in some cases (285). Another reason why results of the GDS might be less reliable in AD population is the inclusion of a valuation, “if the subject feels he/she has more memory problems than most”, which may be strongly interfered with by cognitive impairment. In the PRODE cohort, clinical data was collected using structured interviews with the patients and caregivers and case notes. A health-care professional involved in the data collection underwent standardized training prior to the study period and recurrently, twice a year, during the study period in order to secure reliable data, which strengthened the quality of data. Detailed clinical and neuropsychological assessments were performed for AD diagnostics in the KI and ADNI cohorts; however, no confirmation by postmortem examination of the AD diagnosis was available. Neuroimaging data was collected using standardized protocols in all the cohorts. In two multicenter cohorts – ADNI and PRODE MRI procedure was harmonized across centers based on American College of Radiology (ACR) phantom and healthy volunteers. Image quality was assessed by experienced radiologists. Some variation existed in the scanning parameters among the centers in the PRODE cohort, however, the intra-class correlation coefficient was acceptable and sensitivity analysis did not reveal any major variation in the results. In general quality of the Freesurfer segmentation was good. In the case of inaccurate segmentation or skull
stripping (the most common problem) the analysis was repeated several times until the result was sufficient, otherwise, the scans were excluded.

**Confounding** can be defined as a factor associated with an outcome and exposure, but not on a causal pathway. Unfortunately, sometimes there is not enough theoretical evidence in order to conclude if a factor is a confounder or mediator. A good example would be a relationship between cortisol, brain cortex and depressive symptoms. Numerous studies have shown the involvement of cortisol in depression. However, it is not clear if cortisol elevation is a result of depression or a causal factor for depression development. Strictly speaking, if one were to assume that cortisol is only the causal factor, then our results of the association between depression and cortical thickness would be confounded by an effect of cortisol on the cortical thickness. But, most probably, the relationship between depression and cortisol is an example of mutual causality.

Age is an important potential confounder, especially in the studies of elderly populations. Age- and gender- matching procedures were performed in the studies II, III and IV. Age was also taken into account in all the models, but one. In the RF model predicting MCI status in LLD, brain regions were not corrected for age. However, given that the volume of the right ventral diencephalon, which had the highest predictive scores (Gini index - GI) was not associated with age; we might suggest that absence of age correction did not affect the result significantly. Still it could affect the scores of the second most important variable – the right hippocampus, which volume correlated with age; however, hippocampus had much smaller GI compared with RVD (2.06 vs 8.26 respectively). Moreover, RVD alone discriminated between the MCI/DEM and non-converters with AUC=76% (unpublished data).

One of the main outcomes was cortical thickness (TH). Studies show variability in the structural measurements like TH and GM (VBM) related to the level time of the day (286). There was a significant decrease in TH and GM volumes in the evening compared with the morning. Interestingly, exactly, the most commonly associated with various psychopathologies prefrontal and temporal GM were the most sensitive. This might be another reason for high variability in results of different studies. The level of hydration is another physiological factor affecting cerebral gray matter (287). Hyperhydration was associated with larger volumes of the caudate nucleus and dehydration with smaller volumes of the temporal and parietal cortex.

Another potential confounder is a total intracranial volume (ICV). It is usually correlated with many subcortical volumetric measurements, like hippocampus. In the linear modeling, it was added as a covariate. In the multivariate analysis structures were ICV-corrected. However, the absolute size of the structures might also matter due to the variation in structural reserve (288).

Information regarding the antidepressant treatment was available for the KI, PRODE and Betula cohorts and was taken into account during the data analysis. This was not the case in the ADNI cohort, which should be considered as a limitation.
5.3.2 Precision

Precision is high when a random error is low. Random error is a variation in the data that cannot be explained by known factors. Precision was reported using only p-values, confidence intervals were not used. This could be misleading in the case of a large sample size, where even smallest effects will have low p-values. Large sample sizes were not a problem in the current project. In contrast – study III had a small sample size which could result in a low precision of the results. In study IV cortisol data was available for a subsample of 49 participants. Given the mild levels of depressive symptoms, it is possible that no direct effect between the depressive symptoms and cortisol was found due to the lack of power. Anyway, the conclusion about the mediating effect of cortisol can be made based only on the theoretical model. This should be considered as a limitation.

5.3.3 External validity

External validity examines the extent to which the results of a study provide a correct basis for applicability to other circumstances (289). For instance, the KI cohort was collected in the memory clinic, which means that the results can be generalized only on the memory clinic population, but not community elderly. In the context of the external validity, variability in tools for depression assessment, MRI vendors and fields strength, study populations, depression severity and cognitive status across the studies, may be considered as advantages, since similar structural changes related to depression were observed.

5.3.4 Causality

The most common way to establish a causal relationship in medical and social sciences is to use a temporality criterion (290). Studies I and II have cross-sectional design and no assumption regarding the causality can be made. However, studies III and IV have a longitudinal design. Study IV has longitudinal MRI data allowing us to suggest the temporal relationship between depressive symptoms and brain atrophy. However, in biological systems, mutual causality often exist, which is important to take into account on all the stages of a study.

5.4 GENERAL CONCLUSIONS (THE HEN AND THE EGG)

Claims of causality between depression and AD dementia would depend on the assumptions made regarding etiologies. Evidence suggests that dysregulation of the HPA system and inflammation are etiological factors for both depression and AD dementia. Depression, along with schizophrenia and bipolar disorder, is one of the major psychiatric disorders. All these disorders are associated with elevated cortisol levels (even on preclinical stages). Cortisol enhance cell metabolism. Excessive cortisol stimulation of the neurons prevents anabolic processes (restoration), accelerates brain aging (MTL in particular) and increases the risk of AD.
Chronic inflammation accompanying AD pathology may lead to depressive symptoms and depression-related brain changes. Such inflammation is more prominent in early stages of AD pathology, indeed, may start decades before the cognitive decline in AD (119). Thus, it is possible that even mid-life depression may be a result of inflammation related to AD pathology. The interpretation of studies that suggests that depression is associated with more severe AD pathology is counter to ours--more severe AD pathology is associated with depression. For instance, history of depression was associated with higher levels of amyloid plaque and tangle formation within the hippocampus and with progressive cognitive decline (291). The authors suggest that depression is a modifiable risk factor and its treatment may reduce risks of AD. But if AD-related inflammation is a cause of mid-life depression then etiological treatment for AD-related depression will be possible only when preventive/treatment measures for AD will be available. The relationship between APOE4 allele and depression (292) is evidence in favor of contributing role of AD-related neuroinflammation to depression. In case if inflammation is a contributing factor to both - depression and AD, it is important to identify its sources before starting any clinical trials. If inflammation has an infectious nature, anti-inflammatory drugs may be inappropriate.

Depression-related alterations of the brain (of the regions mainly involved in social- and self-perception abilities) may occur independently of AD or as a part of AD pathology (triggered by AD-related inflammation?) (Figure 4).

Fig.4
5.5 BROADER PERSPECTIVE

Major psychiatric disorders increase the risk of dementia (8,293,294), and so do common somatic diseases (cardiovascular, diabetes mellitus, infectious diseases) increase the risk of dementia (121,295,296), moreover deprived social groups (widowed, orphan, immigrants, low-income) are at higher risk of cognitive decline (297–300). This might be explained not only by higher occurrence of widely-established risk factors in these populations (low education, vitamin deficiency, alcohol consumption, unhealthy lifestyle, residence in polluted areas) but also by chronic stress (301,302) and decreased life satisfaction (303). Future studies should elucidate if there are shared biological pathways in these psychosocial factors.

Another conclusion is that an enormous amount of modifiable risk factors exist for dementia. Studies show that preventive measures like healthy diet, active lifestyle, education through life, stress reduction are effective in reducing risks of both – AD and depression. Considering the costs of dementia care, primary prevention is likely to be cheaper than management of dementia (304). Given the amount of knowledge about the AD, a complex evidence-based approach is already possible (305). Unfortunately, prevention of disease is not as profitable as disease treatment. This is an important ethical issue in medical sciences in general, making own health an individual responsibility.

Better management of depression and other psychiatric disorders in another way to decrease dementia risks. Depression in 15-50% it can be successfully treated with antidepressants (306) alone or, better, combined with psychotherapy. Psychedelic-assistant therapy is one of the most promising approaches against treatment-resistant depression (307,308).

5.6 FUTURE DIRECTIONS

In some cases, late-life depression may represent a predementia state (309). It is further supported by the presence of cortical thinning in the parahippocampal and fusiform regions in LLD, which was independent of the cognitive status (11). We have also shown that brain structure may be useful in prediction of cognitive decline in LLD (278). Importantly, adding of the MMSE scores improved the prediction a lot, which suggests that cognitive impairment accompanied LLD does not resolve with depressive symptoms but is persistent and progressive. It is also suggested that even in the depressive state; evaluation of the cognitive status is reasonable and may have a prognostic value. Future studies should assess whether more detailed neuropsychological tests (with a focus on hippocampal-dependent memory) alone, or combined with brain MRI would be able to predict cognitive decline during the depressive episode in elder patients.

It is also important to find out which factors affect brain plasticity (i.e. atrophy and hypertrophy of the brain structures), and which gray matter components contribute to it. Longitudinal brain MRI studies controlling for the effects of major psychosocial events, stress levels, biological factors (like steroid hormones and neurotrophins), age, SSRI treatment and markers of neuronal loss (like tau protein) may be useful in answering this
question. More studies assessing the relationship between brain structure and function are needed (168). This would give more insights about the plasticity (does increase in functional activity ultimately leads to the hypertrophy of the structure? and visa-versa), but would also improve our understanding and ability to interpret both – structural and functional findings. Studies of the brain plasticity in human (including ultra-high resolution brain MRI) would ultimately lead to better treatment strategies and treatment monitoring.

Many research questions could not be answered until the recent development and implementation of the ultra-high resolution brain MRI with the magnetic field strength of 7 Tesla and even higher, enabling structural evaluation of the key brain regions like the hypothalamus (psychiatric and endocrinologic diseases), substantia nigra (schizophrenia, Parkinson disease), hippocampal and amygdalae subdivisions (Alzheimer disease, depression, schizophrenia, PTSD) and many others.

Novel techniques allow assessing hippocampal neurogenesis in humans (310). It has been shown that in adult hippocampus around 700 neurons are added per day (311). Post-mortem studies show decreased number of mature neurons in the dentate gyrus in depressed (283). Future studies should assess a role of neurogenesis impairment for the development of depression in humans. It would be also important to study the relationship between hippocampal neurogenesis and ability of the hippocampus to suppress cortisol. Clinical studies of the factors promoting hippocampal neurogenesis (enriched environment, physical exercises, antidepressant treatment) may lead to a development of novel therapeutic strategies against both – depression and AD.

The negative impact of both – cortisol (312) and a-beta oligomers (313) on the long-term potentiation in the hippocampus suggests that these processes may be connected (314). Moreover, long-lasting NMDA receptor hyperactivation via cortisol possibly leads to amyloid hyperproduction (90). Future translational studies should help to assess the interaction between the two. Importantly, SSRI treatments improved LTP in animal models of depression (315). A large body of evidence in favor of the connection between excessive stress and cognitive decline, suggests, that clinical trials focused on stress reduction in the elderly population at risk for AD and other dementias may be positive. Controlling for the level of inflammation, cortisol, and antigens may add new mechanistic evidence for the mechanisms behind the relationship between affective states and cognitive decline.

Since AD pathology may start in mid-life, it is important to assess whether the mid-life MDD is connected with AD biomarkers, especially, considering that inflammation in AD may be an early event.
ACKNOWLEDGEMENT

I would like to acknowledge:

Professor Dag Aarsland Thank you for being so effective, enthusiastic and available to discuss any issues. It was great to work under the supervision of a person with a scientific mindset and interest in my research field.

My co-supervisors: Eric Westman for your support and experience in neuroimaging and also for being a good group leader, Bengt Winblad for your assistance, interest in your research field and, of course, for your role in Memantine studies! Lars-Olof Wahlund for being a very creative and open-minded researcher and a nice person.

My collaborators, for the great contribution to my project, in particular: Andy Simmons for your attention to details and English manners Tom Borza It was my pleasure to work with you! Thank you for all your assistance with the data, for being attentive and very responsible Asta Håberg you are great! Thank you for your contribution, time and high quality of your work Geir Selbaek thank you for the opportunity to work with you and your wonderful group Anna Sundström for your thoughtful comments and your interest in depression research Lenita Lindgren for your enthusiasm and assistance Lars Nyberg for our fruitful collaboration.

My group: Daniela Enache for our discussions, for being my PhD guide, and for your care about people in general and patients, in particular. Walid Tajeddinn for our scientific discussions and your explanations of the serotonin signaling Erica Berezckii for your support and care and for being detailed-oriented Michaela Karlstedt for being positive and for your research enthusiasm.

Olga Voevodskaya for being a good friend and a nice person Minutes of laugh we had and brainstorming discussions are priceless! (and for reviewing my kappa!)

Division of Clinical Geriatrics: Joana Braga Pereira It was great to work with you and to go out with you! Daniel Ferreira for nice conversations and your assistance, Gabriela Spulber for being dedicated and responsible and for being around when I just started my PhD Olof Lindberg for your curiosity, creativity and honesty Soheil Damangir for your assistance, our discussions, your critical mind and teaching talent Alejandra Machado Gustav Mårtensson, Una Smailovic it was nice to have you around Annette Eidehall for being a very nice, friendly person Farshad Falahati for our fun conversations Sara Garcia-Ptacek for interesting conversations and your ability to think out of the box Miia Kivipelto for your wonderful research Vesna Jelic for your care about patients.

Division of Neurogeriatrics: Alina Codita for being a nice person, your sympathy to cats and your help with our relocation, Maria Ankarcrona for your research and organizational talents and for being a very nice positive person Susanne Frykman Caroline Graff Lars Tjernberg Helena Karlström Dorota Religa for your important and interesting research Homira Behbahani for the PhD seminars Angel Cedazo Mínguez for your skepticism and ability to express directly your opinion Kevin Grimes for being open-minded and your assistance with my kappa! Ronny Handels for taking care of the environment Inga-Lill Haraldson for keeping us updated Eva Kallstenius and Gunilla Johansson Thanks a lot for all your help and support!!! Nuno Leal it was nice to have you around Jolanta Lundgren for your interesting research and high moral judgment Silvia Maioli for our nice baby-talks Torbjörn Persson for introducing me to the Swedish way of thinking Juraj Secnik for our conversations in the office on the 4th floor.

Division of Translational Alzheimer Neurobiology: Agneta Nordberg for being very consistent in your research, high research quality and hiring wonderful people Azadeh Karami for being dedicated, strong and beautiful, Elena Rodriguez-Vieitez for your wonderful studies, which I cite a lot, Konstantinos Chiotis for being a positive person, Erica Lana for our conversations, and inspiration to do some sports Ispit Srivastava for your sense of humor and nice pictures from the birthday party, Ruiqing Ni for being very friendly, Swetha Vijayaraghavan for interesting time in and out of the office and your ability to
think out of the box in your studies, Laetitia Lemoine for being an interesting person, Agneta Lindahl for being friendly, Taher Darreh-Shori for the PhD seminars and for your ability to ask striking questions even if it not your field of research, Maria Lindskog, for your important research, and interest in depression studies

Special thank you to Rita Almeida who’s competence, enthusiasm and teaching talent brought KI neuroimaging community to the new level.

A would like to acknowledge Robert Sapolsky for his outstanding research of the HPA axis impairment in primates.

To my friends in Stockholm Anais Louzolo for our wonderful times exploring the city and talking about everything Nataliya Tarasova for interesting discussions and nice time during courses Anna Mourskaia for fun times with our kids and nice lunches at KI Solna Alexandra Bernadotte for your enormous support and assistance during my pregnancy – Thank you!! Also for being such a bright interesting person

My colleagues, friends, and teachers from Russia (originally): Eugenia Uglova for our unforgettable work and research experience together and for our dreams and goals! Alexandra Yakovleva for our research, Ira Jarinova for our wonderful times at SS. Andrey Marchenko for introducing me to research in psychiatry, Arkadiy Korzenev for being the first researcher performing deep brain stimulation of the head of the caudate nucleus against treatment-resistant obsessive-compulsive disorder, Alexander Sofronov, Andrey Pashkovsky, Alexey Bocharov, Alla Dobrovolskaya, Andrey Saveliev, Ekaterina Shereshevskaya for unforgettable times as a research assistant and a resident doctor, Olga Savokina and Vladimir Samoylov for the exciting introduction to neurophysiology

I feel very thankful to my family: my mom, for everything, my dad, for inspiration, my grandma, for being an exceptional example of psychological and physical health at 86 years, Alex, for your competence and shared dreams and values, and, of course, Nicole, for great inspiration and understanding.

I want to thank all the patients and their families, and also say sorry for being so slow with solutions and with making the world a better place.
REFERENCES


23. Uekermann J, Abdel-Hamid M, Lehmkämper C, Vollmoeller W, Daum I. Perception of


36. Delgado P. Serotonin Function and the Mechanism of Antidepressant Action by. Arch


61. Liu Y, Ho RCM, Mak A. Interleukin (IL)-6, tumour necrosis factor alpha (TNF-α) and soluble interleukin-2 receptors (sIL-2R) are elevated in patients with major depressive disorder: A meta-analysis and meta-regression. J Affect Disord [Internet]. Elsevier B.V.; 2012;139(3):230–9. Available from: http://dx.doi.org/10.1016/j.jad.2011.08.003


78. World Health Organization (WHO). The ICD-10 Classification of Mental and Behavioural Disorders: 1–267.


158. Bidzinski J. PRELIMINARY ANALYSIS OF FUNCTIONS OF THE TEMPORAL LOBES IN


223. Greenberg S a. The geriatric depression scale (GDS). Best Pract Nurs Care to Older


235. Fischl B, Sereno M, Dale A. Cortical surface-based analysis: II: Inflation, flattening, and


248. Shad MU, Muddasani S, Rao U. Gray matter differences between healthy and


