Clinical, epidemiological, and functional neuroimaging perspectives on the association between depression and neurocognitive function

by

Eva Biringer

University of Bergen

Norway

2006

Clinical, epidemiological, and functional neuroimaging perspectives on the association between depression and neurocognitive function

by

Eva Biringer

Institute of Clinical Medicine, Department of Psychiatry

Faculty of Medicine, University of Bergen

Norway

2006

ISBN 82-3080235-1

Bergen, Norway 2006


EB 2001

Contents:

1. Acknowledgements

Supervisors and co-authors:

I want to thank my supervisor, Anders Lund, for his care and support during the work with this thesis. I also want to thank Arnstein Mykletun for his counselling on professional and methodological issues. Further, I want to express my sincere gratitude to each of the co-authors who are not mentioned elsewhere in "Acknowledgements". These are (in alphabetical order):

Ronald Bottlender, Alv A. Dahl, Jens Egeland, Rebecca Elliott, Knut Engedal, Lars Ersland, Åsa Hammar, Kenneth Hugdahl, Rune Kroken, Harald A. Nygaard, A. David Smith, Karsten Specht, Kjetil Sundet, Susanne Weis.

Colleagues:

I wish to thank my friend and colleague Kirsten I. Stordal for her support and advice during the first years of my work with this project. And for the qualified advice about neuropsychology and helping me organise data collection, I want to thank Astri J. Lundervold. Further, I would like to express my appreciation to Hugo Jørgensen, Jan Øystein Berle, Ketil Ødegaard, Marit Skare, Henriette Kvalheim, and my other colleagues at the Section of Psychiatry, University of Bergen, for their support during these years. I am also grateful to Jessica Marks for her help with language.

I especially want to thank the patients who participated in the study, and the co-workers who put a lot of effort into the collection of data.

Research groups and other institutions:

A sincere thanks goes to the leaders and co-workers of the following project groups for providing me with data and/or support and advice:

The Bergen-Oslo Project (by Bjørn R. Rund), the Bergen fMRI Group (by Kenneth Hugdahl), the Network of Psychiatric Epidemiology (NEPE) (by Tone T. Haug), and the Hordaland Health Study 1997-'99 (HUSK) (by Grethe S. Tell).

Funding:

I thank the Norwegian Research Counsil (NFR) for a three year career grant, and Helse Vest RHF for a three months additional grant. I am thankful to the following for providing me with further financial means: Dr. Einar Martens Fond til Forskning av Alvorlige Sinnslidelser (2001), Gerda Meyer Nyquist Gulbranson & Gerdt Meyer Nyquists legat (2001), Det Alminnelige Medisinske Forskningsfond (2002), Lundbeckstipendet (2002), Lege Bjarne Wilmanns Legat til Forskning av Hjernesykdommer (2005), and Section of Affective Disorders in Helse Bergen HF for allowing me to use 20% of the time spent at work for research purposes during Spring/Summer 2005. I also appreciate that the Bergen Mental Health Research Center (BMH-senteret) provided me with equipment and facilities during Autumn 2005, and I am grateful to H. Lundbeck AS and Wyeth Lederle for inviting me to participate in research conferences and seminars.

Finally, I wish to express my gratitude to my husband Frank, who has supported me 100% while I worked on this project, and the other members of my family. In particular, I am grateful to and for my father Harald Aaker and my stepfather Hjalmar Rønhovde, who both died in the Autumn of 2005, during the period I was working on this thesis.

2. Abbreviations

A	attention
ATC	Anatomical Therapeutic Chemical Classification System
BDI	Beck Depression Inventory
BDNF	brain-derived neurotrophic factor
BP	bipolar
BOLD	Blood-Oxygen-Level-Dependent
BOP	Bergen-Oslo Project
BRMS	Bech-Rafaelsen Melancholia Scale
C	healthy controls
CI	confidence interval
COWAT	Controlled Oral Word Association Test
Sem	semantic fluency sub-task, Controlled Oral Word Association Test
Phon	phonological fluency sub-task, Controlled Oral Word Association Test
DSM-IV	Dagnostic and Statistical Manual of Mental Disorders, 4th edition, 1994
ECT	electroconvulsive therapy
EF	executive function
fMRI	functional magnetic resonance imaging
GAD	General Anxiety Disorder
GAF	Global Assessment of Functioning Scale (DSM-IV)
HADS	Hospital Anxiety and Depression Scale
HADS-A	Hospital Anxiety and Depression Scale, Anxiety sub-scale
HADS-D	Hospital Anxiety and Depression Scale, Depression sub-scale
HUSK	Hordaland Health Study 1997-‘99

HAM-D	Hamilton Depression Rating Scale
ICD-10	International Classification of Diseases, 10th Revision
ICPC	International Classification System for Primary Care
IQ	Intelligence quotient
M	memory function
MADRS	Montgomery Åsberg Depression Rating Scale
MCI	mild cognitive deficit
MDD	major depression
M.I.N.I	MINI International Neuropsychiatric Interview
MMSE	Mini-Mental State Examination
NMDA	N-Methyl-d-Aspartate
PASAT2	Paced Auditory Serial Addition Test, 2 seconds sub-task
PASAT3	Paced Auditory Serial Addition Test, 3 seconds sub-task
PET	positron emission tomography
R	regression coefficient
r	Pearson's correlation coefficient
rCBF	regional blood flow
rMDD	remitted or recovered major depression
ROI	region of interest
S	psychomotor speed
SPECT	single photon emission computed tomography
SSRI	selective serotonin re-uptake inhibitor
Stroop C/W	Stroop Colour and Word Test, color-word sub-task
UP	unipolar
VeM	verbal memory function
ViM	visual memory function
VF	verbal fluency

WAIS-R	Wechsler's Adult Intelligence Scale -Revised
WAIS-R Dsb	Digit Symbol Test from WAIS-R, digit span backward sub-task
WCST	Wisconsin Card Sorting Test
Catc	categories completed variable, Wisconsin Card Sorting Test
Perr	perseverative errors variable, Wisconsin Card Sorting Test
NPerr	non-perseverative errors variable, Wisconsin Card Sorting Test
Ftms	failure to maintain set variable, Wisconsin Card Sorting Test

3. List of papers

Paper I: Biringer E, Lundervold A, Stordal KI, Mykletun A, Egeland J, Bottlender R, Lund A. Executive function improvement upon remission of unipolar major depression. Eur Arch Psychiatry Clin Neurosci 2005;255:373-80

Paper II: Biringer E, Mykletun A, Sundet K, Kroken R, Stordal KI, Lund A. Changes in neurocognitive function associated with remission of unipolar depression: a longitudinal study. Acta Psychiatr Scand (submitted)

Paper III: Hugdahl K, Specht K, Biringer E, Weis S, Elliott R, Hammar Å, Ersland L, Lund A. Increased parietal and frontal activation after remission from recurrent major depression: A repeated fMRI study. Cogn Ther Res (accepted for publication)

Paper IV: Biringer E, Mykletun A, Dahl AA, Smith AD, Engedal K, Nygaard HA, Lund A. The association between depression, anxiety, and cognitive function in the elderly general population -the Hordaland Health Study. Int J Geriatr Psychiatry 2005;20:989-97

4. Introduction

4.1 Depression

Depression is a highly prevalent psychiatric disorder (1-3). Life-time prevalence of Major Depressive Disorder (MDD) has been reported to be in the range of 7.9 to 17.8% (4, 5). Life-time cumulative probability of suffering a first episode of MDD has been found to be 27% in males and up to 45% in females (6).

Depressive symptomatology is generally associated with reduced quality of life (7, 8), lower level of functioning (7, 9), impaired work capacity (10, 11), and death (12, 13). The co-morbidity with other psychiatric conditions is high in depression (1, 7, 14-19). And the disorder is frequently co-existent with somatic conditions (20). Research during the past 30 years has made it clear that depression is also associated with lower neurocognitive function (21, 22). However, there are several unanswered questions with regard to the association between depression and neurocognitive function.

Depression is regarded as a spectrum disorder (23), with symptoms at all levels found in the population (24). The typical course of unipolar depression is depicted in Figure 1. The disorder often has a release-relapse course, with recurrent episodes of depression between non-symptomatic periods or periods of sub-threshold symptomatology (23, 25). The occurrence of one episode is associated with increased risk of further episodes (26, 27).

Diagnosis of depression is based on anamnestic information and observation of clinical characteristics, and not on etiology or evidence of underlying pathobiological changes (26). According to the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), diagnosis of a major depressive episode involves the presence of five or more of the following symptoms for a period of two weeks or more: 1) depressed mood most of the day, 2) diminished interest or pleasure in all, 3) weight loss, weight gain, or decrease or increase in appetite, 4) insomnia or hypersomnia, 5) psychomotor agitation or retardation, 6) fatigue or loss of energy, 7) feelings of worthlessness or guilt, 8) diminished ability to think or concentrate, or 9) recurrent thoughts of death or suicide (28). At least 1) or 2) should be present. When no anamnestic information about elevated mood (mania/hypomania) is present, MDD is referred to as ”unipolar” (as opposed to “bipolar”).

4.2 Neurocognitive function in depression

4.2.1 Depression is associated with lower neurocognitive function

Depressed patients' complaints concerning problems with memory or concentration are well known to the experienced physician. The association between depression and neurocognitive function has been subject to growing research interest over the past two to three decades. One of the main reasons for this interest, is the new examination techniques that have been introduced within the fields of neuroimaging, neurophysiology, and genetics. These techniques have opened possibilities of linking behavioural data from neuropsychiatry, neuropsychology, and cognitive neuroscience to evidence from these new fields. One of these new techniques, is functional magnetic resonance imaging (fMRI).

Previous research has made it clear that depressive symptomatology is associated with lower neurocognitive function (21, 29-31). Previous studies have reported deficits in memory function (30, 32-36), attention (22, 32, 37-40), executive function (22, 33, 37, 40-44), and psychomotor speed (32, 34, 37, 39-41, 45, 46) in depressed patients compared to healthy controls.

In clinical studies performed on younger samples of patients, effect sizes for the differences between depressed patients and healthy ones on tests of neurocognitive function have frequently been reported to be in the range of half a standard deviation (SD) to one SD of the sample mean in favour of the controls (31, 35, 37, 40, 47).

However, smaller clinical studies may be vulnerable for biases that emerge from the many differences that exist between severely ill patients and healthy controls, beyond level of depressive symptoms alone. It is obvious that depressed patients included in clinical studies are different from depressed subjects who are not included in such studies. Roness et al. (2005) recently showed that the majority of persons who suffer from depression do not seek professional help for their symptoms (48). So, then patients who are included in clinical studies are often those who have sought professional care or have been hospitalised (30, 32, 35, 49). These patients often have low levels of general functioning (50). They may have more co-morbid conditions, use more medication, have lower level of physical activity, higher rates of unemployment, and personality-related traits that may influence how they cope in a test situation (51-53) compared to patients not included in clinical trials. And the healthy controls that the patients are compared with may function better in several areas. These factors can all confound the association between depression and performance on tests of neurocognitive function, and, consequently, lead to over-estimation of the effect sizes for the differences in neurocognitive function between depressed patients and healthy controls.

However, two arguments favouring the correctness of estimates of effect from clinical studies exist: 1) diagnostic reliability is higher in clinical studies using diagnosis by specialist, and 2) in population studies, the most severely ill patients may be under-represented (54). Despite these two counter-arguments, the extent to which depression affects neurocognitive function is more likely smaller than the findings from clinical studies suggest.

Information about patients who have not sought help for their psychological problems can be found in population-based studies. In order to estimate more correctly the association between depression and cognitive function, the association should therefore be examined in such samples. This has been done in a few studies (19, 55-58). In these studies, neurocognitive tests which are normally used in clinical settings have been used. Several authors have reported that higher levels of depressive symptoms are associated with reduced performance on measures of neurocognitive function in general population samples (19, 55-58). However, in population-based studies the association found between depression and lower neurocognitive function has been relatively weak (19, 57).

When looked at along side one another, it seems reasonable to state that clinical and epidemiological study designs provide different kinds of information about the associations under investigation. Because they represent different methodology and samples, information from both types of designs could be useful to shed light on the associations studied.

4.2.2 Is neurocognitive function relevant for general functioning in depression?

As stated above, depressive symptoms are associated with problems on several dimensions of neurocognitive function. This said, it seems crucial to ask the following questions: What are the functional consequences of the reduction in neurocognitive function experienced by depressed patients? What dimensions of neurocognitive function are relevant for functioning? Is lower neurocognitive function associated with impairment of functioning also after recovery of the depressive symptoms? And can lower neurocognitive function within a depressive episode predict outcome later on?

It seems like the literature on the association between depression-related neurocognitive problems and functional disability is scarce. This is particularly true with literature that focuses on long-term effects and rehabilitation. However, previous studies indicate that patients' problems with coping in everyday situations and in their work-life may partly be due to lower neurocognitive function (59, 60). In the present study, attempts will be made to clarify some of the questions raised above concerning the association between neurocognitive function and general functioning. The predictive value of neurocognitive function within a depressive episode for symptomatic and functional recovery will also be assessed.

4.2.3 Does neurocognitive function improve upon remission of the depressive symptoms?

As mentioned earlier, it is now generally accepted that major unipolar depression is associated with lower performance on tests of neurocognitive function (21, 47). However, it is yet not clear, if remission of the depressive symptoms is followed by improvement of neurocognitive function in patients who recover from a depressive episode. This question is probably of major importance for patients, since they depend on being well-functioning cognitively in order to function at work and on other areas of their everyday lives after depression. If neurocognitive function normalises after a depressive episode, it can be regarded as a "state"-phenomenon. However, if the depression-associated reduction in neurocognitive function persists between episodes, it is considered to be a "trait"-factor. Consequently, evidence from previous studies investigating neurocognitive function during or after a depressive episode are frequently referred to as either supporting the "state"- or the "trait"-hypothesis (61, 62). Within this terminology, the term "scaring" is also used. The term "scaring" refers to a change that persists after a depressive episode and becomes progressively worse during future episodes (63).

Several authors have reported remission from a depressive episode to be associated with improvement in performance on tests of neurocognitive function (64-68). Improvement has mainly been reported on measures of verbal fluency (65, 66), attention (65, 67), and memory function (65). However, some authors have not found such associations between improvement in depression and improvement in tests of neurocognitive function (64, 69, 70). A few studies that have compared remitted patients with controls on tasks of verbal fluency and memory function have found that patients still perform more poorly than controls (67, 68, 71). Taking into consideration all these studies, it seems reasonable to conclude that the question if, and to what extent, neurocognitive function improves upon remission of depression, still is unanswered.

Tables 2-4 represent an overview of studies investigating the “state-trait”-question with regard to neurocognitive function in depression. These tables are based on literature searches performed in the databases MedLine, EmBase, and PsycInfo. Only studies targeting unipolar depression and using objective measures of neurocognitive function (neurocognitive tests) published after 1985 in English are included. Studies investigating the effect of electroconvulsive treatment (ECT) on neurocognitive function, studies which were primarily neuroimaging studies, and studies on elderly samples with development of MCI or dementia as the target of investigation were excluded from the tables. The studies in Tables 2 through 4 were scrutinised in terms of methodological approaches to the research questions investigated, sample quality and quantity, the aspects of neurocognitive function the tests targeted, and results, in addition to the conclusions of the respective authors. This was done in order to achieve highest possible consistency of the conclusions and interpretations of results in the tables.

A relevant objection to the results from the studies in the tables is the diversity within and between the studies concerning important design- and patient characteristics. The studies were heterogeneous with regard to type of patients included, duration of observation, and level of depressive rest-symptoms after remission. Yet, they included different diagnostic sub-groups, age-intervals, and psychotropic medications. Some studies supported the “state”-hypothesis (Table 3), other studies favoured the “trait”-hypothesis (Table 4), and some studies supported both the “state”- and “trait”-models (62, 64, 66, 70-73).

4.2.4 Does longer duration of depression lead to worsening of neurocognitive function?

Studies using computerised tomography (CT), or magnetic resonance imaging (MRI) have detected structural changes in the brain in patients with long-standing depression, as compared to healthy controls (74-77), or compared to first-episode patients (77). Further, some neurocognitive studies have found correlations between longer duration of depression and lower performance on tests of neurocognitive function (72, 78, 79). Three hypothesised mechanisms for how depression causes neuronal loss and, consequently, reduction in neurocognitive function, have been empirically supported by neuroimaging studies or neurophysiological studies (80):

1) Prolonged elevation of serum cortisol, either as part of a stress-response associated with depression (81), and/or due to a dysfunction in the feedback regulation in the

hypothalamic-pituitary-adrenal axis, may lead to neurotoxic damage to neurons (82-84). This hypothesised effect of elevated and dysfunctional regulation of secretion of corticosteroids has, in particular, been linked to loss of volume in the hippocampus, and because of this, to memory failure (76, 81, 82, 85). In line with this model, Egeland et al. (2005) recently reported an association between higher cortisol levels and lower memory function in a work using baseline data from the sample in Paper I and II (86).

2) Loss of glial cells, perhaps partly caused or mediated by glutamat neurotoxicity, has also been hypothesised as cause for a possible mechanism of progressive neuronal damage (74, 80). The fronto-temporal neuronal circuits may be vulnerable to such cell loss (74, 80).

Regarding possible localisation of dysfunctions in depression (irrespective of underlying cause), it should be mentioned that the frontal or prefrontal cortical areas and the frontal-striatal-thalamo-cortical loops associated with these seem to be of particular importance in depression. Dysfunctions within these loops have been linked to depressive mood and lower neurocognitive function (87-91). The psychomotor slowing frequently observed during a depressive episode could be caused by disturbances in the sub-cortical parts of these loops (92, 93). These parts are similar to those that are affected in basal ganglia disorders.

3) The neurotrophin brain-derived neurotrophic factor (BDNF), which is involved in growth and differentiation of cells, has recently been subject to interest within neurophysiological research. BDNF is produced by glial cells, and during exposure to stress, BDNF levels are reduced (83, 94). Based on animal studies using induced stress paradigms, it has been hypothesised that depression is associated with lower neurogenesis (83, 94). Impairment of neurogenesis leads to lower rate of cell repair after toxic damange to neurons. In depression, BDNF has particular significance because it has been linked to the increased neurogenesis that occurs during administration of antidepressant medication (94-96). In relation to antidepressants, the role of N-Methyl-d-Aspartate- (NMDA-) receptors should also be mentioned. Activation of these receptors seem to be involved in the long-term potentiation important for memory function, and antidepressants act as antagonists on them (96).

However, several neurocognitive studies have provided results that are contradictory to these above described hypothesised models that involve progressive alterations in neuronal functioning and worsening of neurocognitive function in depression (64, 66, 68, 70, 97, 98). These contradictory studies have not found correlations between estimates of duration of depression and results on tests of neurocognitive function. Thus, it still remains unclear, whether recurrent depressive episodes, or long-standing depression, lead to progressive worsening of neurocognitive functioning.

4.2.5 Abnormal patterns of regional brain activation in depression

Regardless of which cerebral dysfunctions it is that underlie the neurocognitive changes observed in depression, neurophysiological correlates of these changes must exist. Neuroimaging tools, such as functional magnetic resonance imaging (fMRI), Positron Emission Tomography (PET), and Single Photon Emission Computed Tomography (SPECT), represent unique possibilities of *in vivo* characterisation of the neurophysiologic mechanisms involved in depression. These techniques assess indicators of regional blood flow (rCBF) and metabolism in the brain.

In unipolar depression, functional neuroimaging studies have identified neurophysiological abnormalities in several areas of the brain. Studies using Positron Emission Tomography (PET) have shown abnormal patterns of regional blood flow and glucose metabolism in the prefrontal cortex, in the cingulate gyrus, amygdala, and related parts of striatum and thalamus (42, 75, 99-102). In patients with unipolar depression, a reduction of rCBF or metabolism in the left dorsolateral prefrontal cortex (Broca's areas (BAs) 9, 46) has been frequently reported in comparisons with healthy controls (93, 103-105). A reduction in the anterior cingulate gyrus (BA 24) has also been found (42, 93). Decreased activation have been reported in other areas (75, 93, 101).

However, it still remains unclear as to whether correlates of changes in regional brain activation during a depressive episode normalise when the depressive symptoms attenuate. Only few functional neuroimaging studies have been longitudinal in design, investigating levels of rCBF or metabolism in regions with pathological patterns of activation in symptomatic patients after remission of the depressive symptoms. Most of these studies (93, 102, 105-108), but not all (109) have reported increased metabolism upon remission in areas that have shown reduced activation in the symptomatic phase. Upon remission, increases have been detected in the left dorsolateral prefrontal cortex (BA 9, 46) (93, 106, 110), and in the left anterior cingulate gyrus (Broca 24) (93, 102). However, there is still a long way to go before it has been clarified how the patterns of cerebral activation vary when the level of depressive psychopathology changes.

4.3 Objectives of the study

The aim of Paper I was to investigate to what extent executive function changes upon remission of unipolar major depression. The aims of Paper II were firstly, to investigate if dimensions of neurocognitive function improve upon symptom remission in unipolar major depression, secondly, to examine if patients who recover completely from the depressive episode reach a level of function on these dimensions that is equal to that of healthy controls, and thirdly, to investigate if longer duration of depression is associated with lower neurocognitive function. The aim of Paper III was to investigate if patterns of regional cerebral activation, as measured by functional magnetic resonance imaging (fMRI), change during remission from major unipolar depression. The objective of Paper IV was to investigate how strong the association between depressive symptomatology and neurocognitive function was in an elderly population-based sample. Other aims of the present study not reported in the papers were: To investigate if neurocognitive function in major unipolar depression is associated with limitations in general functioning, to investigate if neurocognitive function predicts change in general functioning when the depressive symptoms attenuate, and to assess the predictive value of neurocognitive function within a depressive episode for symptomatic recovery later on.

5. Methods

5.1 Subjects

The present study presents data from two different samples. Papers I and II are based on a clinical sample consisting of thirty younger (mean age 36 years) patients with DSM-IV diagnosis of major unipolar depression of recurrent sub-type (111). These 30 were from an original baseline sample of 50 patients. They were re-examined with psychometric and neurocognitive measures two years after baseline examination in the Bergen-Oslo Project. The Bergen-Oslo Project was a collaboration study between several institutions in Bergen and Oslo, Norway, starting in 1998. All patients had suffered a minimum of two life-time episodes of depression at baseline (mean 3.8). At follow-up, 17 had recovered, and 13 were still symptomatic. At inclusion, 20 patients were hospitalised and ten were out-patients. At follow-up, all patients who had been hospitalised, had been discharged. Sixteen patients were employed or students at baseline. Fifteen were employed or students at follow-up. The other patients in the sample were either on sick-leave, received disability pension, or had no income. At inclusion, 26 patients were taking psychotropic medication (21 of these used selective serotonin re-uptake inhibitors (SSRIs)), at follow-up, 25 were on medication (20 on SSRIs). In the studies in Paper I and II, these 30 patients, who were examined twice, were compared to 50 healthy controls who were examined at baseline. The controls had been recruited through an advertisement in the local newspaper, or through personal network. They were comparable to the N=50 baseline patient sample with regard to age, gender, education, handedness, and intellectual abilities. For further information about background-data, please consult Paper II, Table 1, Stordal et al. (2004) (40), or Egeland et al. (2003) (112).

A sub-sample of nine patients underwent functional magnetic resonance neuroimaging (fMRI) scanning at baseline and at follow-up two years later (Paper III). These were compared to a sub-sample of healthy controls who were scanned at baseline. fMRI scannings and data-analyses were made by "the Bergen fMRI-group", located in Bergen, Norway.

Paper IV is an epidemiological study based on data from the elderly cohort in the Hordaland Health Study 1997-99 (HUSK). The HUSK study was one of the large-scale epidemiological studies performed in Norway during the late 1990's. The study was performed as a collaboration by the National Health Screening Service, the University of Bergen, and the local health services. In this study, all inhabitants aged 72-74 years old living within the city boundaries of Bergen, Norway, were invited to participate in general somatic examinations. Out of these, 2,203 subjects agreed to participate in an examination which involved tests of neurocognitive function. This amounts to 51% of the total age cohort. Twenty-five subjects who performed equal to or below a cut-off of nine points on a modified version of the Mini-Mental State Examination, which consisted of the 12 items most sensitive to dementia (113), were excluded due to the probable presence of mild cognitive impairment (MCI) or dementia. This cut-off corresponds to 23 points on the conventional Mini-Mental State Examination (MMSE) (114). The inferential analyses were performed on the 1,930 subjects who had provided valid answers to the Hospital Anxiety and Depression Scale (HADS) (115), and had completed all neurocognitive tests. This sub-sample amounted to 44% of the total age cohort. Prevalence numbers for mild and moderate levels of depression were found to be 9.3% and 2.2%, respectively, in this sample. These numbers are comparable to numbers from other studies recently performed in other Western societies (116-119).

5.2 Methods of measurement

5.2.1 Psychometric instruments

In Papers I, II, and III, evaluation of diagnosis and level of symptomatology was performed by trained psychiatrists. Ratings of symptoms of depression and anxiety were based on self-report in Paper IV.

Three commonly used structured measurement scales were used to measure symptoms of depression and/or anxiety. They all represent continuous approaches to levels of symptoms, but they can be transformed into dichotomous diagnostic tools by the introduction of cut-offs.

The 17-item Hamilton Depression Rating Scale (HAM-D) (120) and the 10-item Montgomery and Aasberg Depression Rating Scale (MADRS) (121) are both regarded as "gold standards" for assessment of depressive symptomatology. Patients included scored equal to or above 18 on both the HAM-D and MADRS in Papers I, II, and III. This corresponds to a moderate to severe level of depression.

The third rating instrument, the Hospital Anxiety and Depression Scale (HADS), is a structured self-report questionnaire that was developed by Zigmond and Snaith in 1983 to identify anxiety and depression among somatic in-patients (115). The instrument has shown good case-finding properties in various kinds of samples (122-124), including in general population samples (122, 123). It has good psychometric properties with regard to sensitivity, specificity, and factor structure in the normal population (122, 123). It should be suitable for detection of depression and anxiety in the elderly, since it does not focus on somatic symptoms or sleep-problems, which occur frequently in the elderly population. In the epidemiological study included in the present work, a cut-off of 8+ was used on the depression sub-scale (HADS-D). This corresponds to "mild" degree of depression.

5.2.1.1 Definition of remission and recovery

No general consensus exists in the literature with regard to definitions and nomenclature of remission of symptoms, treatment response, or recovery in depression (25, 125, 126). In the present study, any reduction of depressive symptomatology from baseline to follow-up is referred to as "remission", regardless of level and duration of reduction in symptoms. When the terms "response" or "responder" are used in studies cited, these are cited as "remission", or "in remission". However, in order to avoid misinterpretations, the accurate level of rest-symptomatology in the studies cited is most often reported (Tables 2-4).

Generally, the HAM-D is regarded as more sensitive to change than other scales (127). To define sub-groups as "recovered" or "non-recovered" according to symptom-status at follow-up, a cut-off of 8+ was chosen on the HAM-D in Papers I and II (126). These terms are used consistently through this work.

5.2.2 Neurocognitive function – constructs and measurement

The cognitive system may be seen as a complex functional system consisting of connected sub-systems that correspond to the major parts of what the mind performs (128). While cognitive psychology focuses on the theory of how the brain processes, stores, mentally manipulates, and expresses information with focus on normal functioning, the field of neuropsychology studies the brain-behaviour association in patients with various disorders with the purpose of identifying patterns, progression, and neuropsychological correlates of cognitive deficits (29, 129, 130).

Assessment of neurocognitive function is normally done by a battery of neurocognitive tests. The battery often includes both pen and paper tests as well as computerised tests. Tests are selected to represent different dimensions of neurocognitive functions. After testing the patient, the neuropsychologist elaborates performance profiles by comparing the patient's performance on the tests either to other tests he/she has completed, to healthy controls comparable with regard to age, gender, level of education, and intellectual abilities (IQ), or to norms generated from population samples (129, 130). In research using neurocognitive test batteries, significance testing is most often used to detect differences between groups of subjects (29).

Throughout the present work, the term "neurocognitive" is used when referring to cognitive function in general or performance on tests. The term has been chosen because it does not refer to any profession or underlying theoretic framework (as opposed to the terms "neuropsychological", "neuropsychiatric", or "cognitive"). When referring to group differences with regard to neurocognitive function, the words "lower" or "reduction" will be used (i.e. "neurocognitive function is lower in one group compared to another group"), as opposed to the terms "impaired" or "impairment", which require a defined cut-off for group differences (such as one group performing 1.0 or 1.5 SD below another group). Further, the word “dysfunction” has not been used when describing neurocognitive function, since this term does not specify directions for associations or group differences (this term is, however, used when describing pathobiological mechanisms).

5.2.2.1 Dimensions of neurocognitive function

In the literature, neurocognitive test measures are often grouped into the following domains of function: Attention, memory function, verbal skills, construction (performance), and concept formation and reasoning (29, 129, 130). However, it is essential to recognise that these constructs are theoretical, although supported by empirical evidence, and that one test measure does not necessarily represent one domain of function. Further, it is important to know that in order to perform a cognitive task, a composite of mental functions are necessary (29, 129, 131) and that considerable empirical overlap exists between construct dimensions.

Because of this, considerable effort was put into the operationalisation of the neurocognitive measures in the present study. Operationalisation was based on a priori theoretical asssumptions of the essential qualities of the test variables, in combination with evaluations of underlying factor structures. Test variables were added up to produce summary scales of neurocognitive function. This approach is empirically reasonable because it leads to an increase of construct reliability, thus representing a parallell to the approaches used in psychiatry, where latent constructs, such as depression, are measured by instruments with multiple construct indicators, rather than by asking the one question only: “Do you have reduced mood most of the day?”. In Paper II, the following summary scales of function were computed: Attention, verbal memory function, visual memory function, and psychomotor speed. In Paper I, a summary scale of executive function was made. Neurocognitive tests with literature references are presented in Table 1.

5.2.2.2 Attention

Attention refers to the processes by which subjects become receptive to and start processing incoming stimuli (29, 128, 129). It is a basic cognitive function that is the foundation for all test performance. It is closely related to activity rate (speed) and memory function, and it is regarded as a function with limited capacity (29, 129). The construct is sub-divided into 1) focused or selective attention (referred to as “concentration” in common language), i.e. the process of attending to the stimulus that is most important, while suppressing awareness of other distracting stimuli (128, 129), 2) sustained attention (vigilance), which refers to the capacity to maintain attentional activity for a period of time, 3) divided attention, which is the ability to respond to multiple tasks simultaneously, and 4) alternating attention, which refers to the ability to shift focus while performing a task (129). In Paper II, a summary scale of attention was created by adding up measures from two frequently used tests that are regarded as indicators of attention (132, 133).

5.2.2.3 Memory

The memory system can be explained by a hypothetical three-stage model that includes sensory memory, which very briefly (1-2 seconds) holds large amounts of incoming information while selecting and coding information; short time memory (includes immediate memory), which is a limited storage stage (7+/-2 bits of information, 30 seconds); and long-term memory, in which information has been organised and consolidated. This consolidation probably happens because of long lasting neurochemical changes at synaptic level. After successful encoding, stored information is retrieved by means of recognition or recall of learned material. Remembering thus implies both successful encoding and retrieval (129). Recently, Baddeley introduced the model of "working memory" to describe the dynamic part of short-time memory that is used for active manipulation of information during task performance (134). Working memory probably relies on neurophysiological activity in particular neuronal networks associated with the prefrontal- and parietal cortices (135-137). In Paper II, two scales of memory function were computed by adding up sub-tasks from two frequently used tests. One scale measured visual memory function (138). The other measured verbal memory function (139).

5.2.2.4 Executive function

Executive functions are thought of as higher-level cognitive functions that are involved in the control and regulation of lower cognitive operations (129, 140). They have been theoretically and empirically linked to functional neuronal circuits involving the prefrontal cortical areas (90, 129, 135, 140-142). No overall consensus exists with regard to the operationalisation of executive function. As with other dimensions of neurocognitive function, considerable conceptual overlap with other dimensions exists. Lezak has conceptualised it as having the following components: volition, planning, purposive action, and effective performance (129) (p. 650). In the present study, an operationalisation has been based on a theoretically and empirically funded model, which was introduced by Pennington and Ozonoff in 1996 (143).

This operationalisation includes indicators of set-shifting, planning, inhibition, working memory, and fluency (132, 133, 144-146).

5.2.2.5 Psychomotor speed

The rate at which information processing takes place is usually affected by disorders with brain dysfunction (129). This is peculiar to disorders involving the sub-cortical structures. Psychomotor speed is often assessed by simple reaction time tasks, but it can also be assessed by comparing tasks in which speed is essential (timed tasks) with non-speeded tasks (131). Slowing may occur at any place in the afferent or efferent systems during task performance (129). Considerable overlap with the dimension of attention is inevitable when assessing psychomotor speed. In Paper II, psychomotor speed was operationalised by adding up sub-tasks from two timed tasks (146, 147).

5.2.2.6 Neurocognitive changes in aging

Because one part of the present study is performed in an elderly sample (Paper IV), it should be mentioned that during aging, natural changes in neurocognitive function take place. In particular, increasing age is associated with a natural physiologic slowing (129, 148). However, memory function also seems to be affected by aging. Some evidence suggests that explicit memory is particularly affected, while other aspects of memory function remain more preserved (149). Notably, aging also entails problems with vision and hearing; and physical impairments such as these may impair neurocognitive test performance (129).

5.2.3 Functional magnetic resonance imaging (fMRI)

The functional magnetic resonance technique enables voxel-by-voxel mapping of patterns of cerebral activation by using magnetic resonance techniques that are sensitive to small local magnet-field variations (150). These magnet-field variations are caused by differences in the magnetic properties of oxygenated and de-oxygenated blood (referred to as the Blood-Oxygen-Level-Dependent- (BOLD-) effect). When neurons in particular areas become more active in response to sensory stimuli, this leads to increased local metabolism and blood flow. Thus, variations in BOLD-magnitude detected in the MR-scanner can be regarded as indirect indicators of level of neuronal activation in a particular area. Estimates of neuronal activation from the scanning process are subject to excessive statistical processing. In the final step of this processing, estimates from the transformations are projected onto a high resolution structural scan of a template brain, thereby creating a statistical "map" of areas (clusters) of levels of neuronal activation in the brain. fMRI requires a contrast between two conditions, which are typically a "resting state" and an "activated" state, the latter with cognitive or emotional stimulus. The "map" of levels of activation during activation is then contrasted with the subject's "map" of levels of activation in the resting state.

In the present study, scannings were performed by a 1.5 Tesla Siemens Vision MRI system. A "block"-design method for task presentation was used, in which a mental arithmetics task (i.e. stimulus) was presented to the test subject on special LCD-screen goggles in runs consisting of "ON-blocks" (stimulus presentation) interrupted by "OFF-blocks" (no stimulus presentation). The task used was a visual version of the Paced Auditory Serial Addition Test (PASAT). The PASAT is a mental arithmetics task (132). Performing the task also involves working memory. In previous studies, similar activation tasks have been associated with significant activations in the prefrontal and parietal cortices (135, 136, 151).

5.3 Designs

In Papers I and II, a longitudinal study design was applied. These papers were based on previously detected baseline differences between the depressed patient group and the control group (40, 45, 112). In the present study, a follow-up examination of the patient group was performed. At follow-up, patients were either partially or totally recovered. Between-group comparisons were made with regard to differences in change in neurocognitive function between sub-groups of recovered and non-recovered patients (groups defined according to symptom status at follow-up). Further, comparisons of recovered and non-recovered patients at follow-up with a healthy control group examined at baseline were made in order to assess if patients had reached the performance level of controls on neurocognitive tests. In addition to the categorical approach to symptoms (i.e. recovered vs. non-recovered sub-groups), analyses were performed with a continuous approach to level of depressive symptoms. Data from the healthy controls were collected at baseline only.

Paper III includes a similar design. A sub-sample of the depressed participants and healthy controls mentioned above were examined by fMRI. Within-group comparisons were made using estimates of levels of regional brain activation between baseline and follow-up.

Between-group comparisons were made between patients at baseline, patients at follow-up, and healthy controls (scannings of controls performed at baseline only).

Paper IV is a cross-sectional population-based investigation performed on the birth cohorts 1925 to 1927 living in Bergen, Norway. Comparisons between sub-groups scoring above or below cut-off for depression and/or anxiety were made. The associations between level of or caseness of depression and/or anxiety were explored.

5.4 Statistical procedures

In Papers I and II, Pearson's correlation coefficients *r* were calculated for the associations between the independent variables and the dependent variables. Independent-samples t tests or Mann-Whitney U tests were conducted to assess between-group differences. Indicators of neurocognitive function were added up to produce composite scores of neurocognitive dimensions. This approach was favoured because it increases construct reliability. It also reduces number of statistical comparisons, which is useful when statistical power is limited.

In Paper III, a three-group, one-way ANOVA model, containing the baseline and follow-up investigations of the patients, and the baseline investigation of the control subjects, was applied. At follow-up, linear regression analyses were performed in the depressed group to investigate if activation within particular regions of interest (ROIs) correlated with level of depressive symptomatology.

In Paper IV, linear regression analyses were performed to assess if depression and/or anxiety were associated with neurocognitive function. In analyses with categorical independent variables, dummy-variables were made and entered into the linear regression model. In a second step, adjustments for possible mediators or confounders were made.

To investigate to what extent neurocognitive function was associated with level of general functioning, linear regression analyses were performed at follow-up in the patient sample (N=30) described in Papers I and II. Level of general functioning was assessed using the Global Assessment of Functioning (GAF) Scale (28). Neurocognitive operationalisation was the same as in Paper II, and the summary scales of neurocognitive dimensions, which were computed cross-sectionally at follow-up, were used. These summary scales were entered as independent variables and the score on the GAF scale at follow-up was entered as dependent variable into the model. In order to adjust for the effect of depressive symptomatology on level of general functioning, HAM-D total score at follow-up was entered in a second step.

In the same sample, the degree to which neurocognitive function within the depressive episode predicted improvement in symptomatology and general functioning from baseline to follow-up two years later was investigated. Again, summary scales of neurocognitive function were computed by adding up test variables in line with the operationalisation used in Paper II. This time this was done with baseline scores. These summary scales of neurocognitive function were entered into a linear regression model with change in GAF or change in HAM-D from baseline to follow-up as dependent variables. Change-variables were made by subtracting scores at follow-up from scores at baseline. When change in GAF-score was entered into the model as dependent variable, the effect of symptomatic improvement on change in level of general functioning was adjusted for by entering change in HAM-D from baseline to follow-up into the model in a second step.

Statistical procedures were performed using the SPSS 11.5 (Papers I, II, IV) and the SPM99 software package (Paper III).

6. Summary of Papers I to IV

6.1 Paper I

Executive function has theoretically been linked to neuronal circuits associated with the frontal lobes. These systems may be affected in depression. Previous studies have reported that depressed patients perform poorer on tasks regarded as measures of executive function compared to healthy controls. To investigate to what extent executive function improved upon remission of depressive symptomatology, performance on executive function measures was examined on two separate occasions two years apart in patients with recurrent episodes of major unipolar depression. At baseline, the patients were moderately to severely depressed, at follow-up, they were partly or totally recovered. The main finding was that improvement in depression was followed by improvement in executive function. Improvement in depressive symptomatology explained 11% of the variance in improvement in executive function from baseline to follow-up. No significant difference between recovered patients and healthy controls was found. In conclusion, the study provided support for the “state”-hypothesis in depression.

6.2 Paper II

Conflicting previous literature has made it difficult to conclude whether remission of depression is associated with improvement on different dimensions of neurocognitive function. Yet, several hypotheses about long-lasting depression leading to progressive worsening of neurocognitive function have been proposed. The aims of this study were 1) to examine to what extent neurocognitive function improves upon remission of major unipolar depression of recurrent sub-type, 2) to investigate if neurocognitive function returns to normal level after recovery from depression (when normal is defined as the performance of healthy controls), and 3) to investigate if longer duration of depression is predictive of lower degree of improvement in neurocognitive function upon remission. The same sample and time-points of measurement as in Paper I were used. Operationalisation of measures of neurocognitive function was based on theoretical considerations and factor analysis, and test measures were grouped into four dimensions of neurocognitive function: Attention, verbal memory function, visual memory function, and psychomotor speed. A significant correlation between improvement in depressive symptomatology and change in verbal memory function over time was found, both when the association was investigated with categorical and dimensional approaches to level of depressive symptomatology. However, the possibility of persistent deficits in attention, visual memory function, and psychomotor speed could not entirely be ruled out by the study because mean performances in the recovered patients on these dimensions were still lower (although non-significantly) than the controls. Duration of depression was not predictive of improvement of neurocognitive function. Consequently, the study did not support a model in which longer duration of depression leads to progressive worsening in neurocognitive function.

6.3 Paper III

This paper provides a neurophysiological correlate to the findings of improvement in neurocognitive function associated with remission of depression in Paper I and II. A sub-group of patients from the sample used in those papers was examined with fMRI at baseline and at follow-up two years later while they were in remission. Scanning was done while the patients were performing a mental activation task that has previously been associated with increased activation in clusters in the frontal and parietal cortices in non-depressed subjects.

The most important finding was that the depressed patients showed significant increases in activation in areas related to task performance (in the left posterior cingulate gyrus (BA 31), right inferior frontal gyrus (BA 44), and bilaterally in the inferior parietal lobules (BA 40)) upon remission of the depressive symptoms. At follow-up, inverse correlations between level of depressive symptomatology and level of activation in these clusters were also found. These findings indicate that patterns of neuronal activation are altered in depression. The changes in activation seem to be related to change in depressive psychopathology. Because studies I and II showed improvement in neurocognitive test performance from baseline to follow-up in the sample from which the sub-group in Paper III was taken, it is reasonable to infer that the changes in level of activation seen in the present study represent a link to the pathobiological mechanisms that underlie both the depressive psychopathology and the reduction in neurocognitive function associated with it.

6.4 Paper IV

In this epidemiological study, in a cohort of elderly non-demented patients (aged 72-74 years), the previously established inverse association between depressive symptomatology and neurocognitive function was confirmed. An apparently inverse association between anxiety and reduced neurocognitive performance was explained by adjustment for co-morbid depression. Males were more cognitively affected by depressive symptoms than females. The inverse association between depressive symptoms and neurocognitive function was found to be close to linear, and also present in the sub-clinical symptom range. However, compared to effect sizes for the association between depression and neurocognitive function found in clinical studies, effect sizes for the association in this population sample were small at all levels of depressive symptom-load. In conclusion, the inverse association between depression and neurocognitive function was present, however weakly, in the elderly normal population. The association was also found at sub-clinical symptom levels. Thus, this inverse association between depressive symptoms and neurocognitive function can be regarded as a “normal”-phenomenon, that is, not only restricted to severely ill patients or to symptom-ranges above cut-off for caseness.

7. Results concerning neurocognitive function and general functioning

7.1 Neurocognitive function and general functioning

In the analyses performed to assess to what degree neurocognitive function was associated with level of general functioning, a medium-sized correlation was found between the psychomotor speed summary scale and the GAF-score at follow-up (R=0.35, R^2=0.12, beta=0.35, 95% CI= -0.18; 8.40, *p*=0.060). This marginally significant association between lower neurocognitive function and lower level of general functioning was found when patients were in remission (mean HAM-D 8.2 (SD 7.6)). After adjustment for the effect of depressive symptoms (as measured by the HAM-D) on level of general functioning, the association between lower neurocognitive function and lower level of general function was still present (R^2=0.55, beta=0.24, 95% CI=-0.37; 6.00, *p*=0.081) and still marginally significant.

The summary scales of verbal memory function and visual memory function did not correlate with GAF-scores at re-testing (*r*=0.08, and 0.01, respectively (n.s)). A small non-significant correlation between the summary scale of attention and GAF-score was found (R=0.20, R^2=0.04, beta=0.20, *p*=0.293). After adjustment for depressive symptomatology, this model still produced non-significant results (R^2=0.50, beta=0.06, *p*=0.681).

7.2 Neurocognitive function as predictor for outcome

A positive and significant association between psychomotor speed within the depressive episode and improvement in GAF-score from baseline to follow-up two years later was found (R=0.39, R^2=0.15, beta=0.39, 95% CI=0.167; 4.50, *p*=0.036). After the effect of improvement of HAM-D on improvement in GAF-score between baseline and re-test had been adjusted for, the association between baseline speed and GAF improvement was marginally significant (R2=0.29, beta=0.31, 95% CI= -0.22; 3.9, *p*=0.077). The other neurocognitive dimensions at baseline did not have any predictive value for improvement in general functioning, neither in the crude analyses, nor after the effect of change in HAM-D on change in general functioning had been adjusted for (crude R=0.16, 0.01, and 0.14 for attention, verbal memory, and visual memory, respectively (all *p*<0.05)).

Neurocognitive function within the depressive episode had no predictive value of improvement in depressive symptomatology (as represented by change in HAM-D from baseline to follow-up) (R in the range 0.01 to 0.23 , all (p<0.05)).

8. Discussion

8.1 Synopsis of results

The inverse association between depressive symptoms and neurocognitive function was found both in the clinical sample (Papers I and II) and in the population-based sample (Paper IV). In Papers I and II, empirical support for the "state"-hypothesis in major unipolar depression was found: In depressed patients, performance in several dimensions of neurocognitive function improved upon remission. After complete symptomatic recovery, patients' performance had improved to levels that were not significantly different from the performance of healthy controls. However, the presence of rest-deficits in neurocognitive function in the patients could not be completely excluded by these studies. The studies had limited statistical power, and mean test performance in the patient group that had recovered was still not equal to controls on several aspects of neurocognitive function. The improvement in depressive symptomatology from baseline to follow-up was probably pictured as increased levels of activation in certain cerebral regions in the fMRI-study (Paper III). These regions had shown reduced levels of activation at baseline when patients were severely depressed. No association of duration of depression with improvement of neurocognitive function was found (Paper II). In the population-based study, the inverse association between symptoms or caseness of depression and neurocognitive function found in an elderly sample was weak, compared to effect sizes from previous controlled clinical studies performed on severely depressed elderly patients. The inverse association was present at all levels of depressive symptoms, including in the lower sub-clinical symptom range typically seen in dysthymia.

A medium-sized correlation was found between higher psychomotor speed and higher levels of general functioning, as measured by the GAF-scale, at follow-up in the sample from Papers I and II.

A positive and significant association between psychomotor speed within the depressive episode and improvement in GAF-score from baseline to follow-up was found in the same sample. This association was only marginally significant after adjustment for the effect of improvement of depression on improvement on GAF-ratings.

8.2 How strong is the association between depression and lower neurocognitive function?

8.2.1 Strength of the association in clinical versus population-based samples

As mentioned earlier, the designs in clinical studies are vulnerable to biases which emerge from the many differences between severely ill patients and healthy controls. The population-based design used in Paper IV should theoretically avoid many of the effects of such biases on the association between depression and neurocognitive function. In Paper IV, the effect sizes for the group differences between depressed (HADS-D ≥8) were 0.2 SD for 'S'-task, 0.3 SD for m-DST, and 0.3 SD of the sample mean for KOLT in favour of the healthy subjects. These effect sizes for the group differences were considerably smaller than those found for the association in clinical studies (47, 152). Thus, the effect sizes for the inverse association between depression and neurocognitive test performance in the population-based study in Paper IV were smaller than findings from clinical studies suggest.

8.2.2 Possible explanations for the discrepancy in effect sizes

As stated above, the inverse association between depression and neurocognitive function found in the population-based study was weaker than findings from previous clinical studies have suggested. This discrepancy in effect sizes between clinical and epidemiological studies may be caused by different types of biases: If there is a dose-response relationship between depression and neurocognitive function, then the clinical studies represent the higher ranges of depressive symptoms, and epidemiological studies represent the lower ranges. Therefore, it seems reasonable to say that both study designs complement each other when the association between these factors is investigated.

Patients in clinical studies may be different from depressed patients who are not included in such studies with regard to a range of characteristics. Examples of factors that can potentially confound the associations between depression and neurocognitive function in clinical studies are: General level of functioning (59), work status (60), intellectual abilities (153), duration of illness, use of medication, sleep disturbances, level of physical activity (154), and personality and coping abilities (51-53). About half of the patients in Papers I and II were not working and almost all of them were using psychotropic medication. Their ratings of general level of functioning by the GAF-scale (28) suggested that they were severely impaired concerning function (see Paper I, Table 2). In addition, the control groups that the patient groups were compared to in such studies may be subject to other biases. These controls may be healthier and better functioning than the patients they are compared to. The presence of such factors mentioned above may lead to inflation of the effect sizes for the differences in neurocognitive function between depressed patients and healthy controls in clinical studies.

In epidemiological studies, however, the most severely ill patients could be under-represented (54). This may lead to weakening of effect sizes for the associations investigated.

8.3 Effects of antidepressant medication on neurocognitive function

Investigation of medication effects on neurocognitive function was beyond the scope of the present study. However, it should be mentioned, that in the population sample in Paper IV, 58 of the 1,930 subjects in the sample were taking antidepressant medication. Of these, 29 used SSRIs. When the linear regression analyses on the associations between depression and neurocognitive function were adjusted for use of antidepressant medication, no change in the magnitude of the estimates of effect were found (changes in standardised effect sizes betas $<2\%$). This indicates that antidepressant medication did not have impact on neurocognitive test results in this sample. The study designs used in Papers I, II, and III, however, did not allow for analyses with regard to medication effects. Almost all patients were on psychotropic medication. These were of different sub-types, although most patients were taking SSRIs. However, previous studies assessing neurocognitive function in medication free depressed patients, have also detected significant associations between depression and lower neurocognitive function (22, 71, 155). Studies comparing neurocognitive function in patients on antidepressants with patients not using antidepressants have not found differences in performance on tests of neurocognitive function (47, 70, 156). Because of the results of these studies it seems safe to say that the neurocognitive reduction in depression cannot be strongly associated with medication use. The use of tricyclic medication is, of course, an exception to this, since these agents have sedative effects due to their anticholinergic and antihistaminergic properties (83, 157).

8.4 The association between neurocognitive function and general functioning

8.4.1 Psychomotor slowing is related to lower general functioning

In the present study, medium-sized correlations were found between the summary scale of psychomotor speed and score on the GAF-scale (r=0.35), and between the summary scale of attention and the GAF-score (r=0.23) at follow-up in the patient sample used in Papers I and II. These correlations suggest that some dimensions of neurocognitive function are associated with lower level of general functioning in depression. Psychomotor slowing was most closely associated with lower level of general functioning. This finding may explain why depressed patients have problems with tempo-demanding work tasks.

8.4.2 The predictive value of neurocognitive function

A positive and significant association between psychomotor speed within the depressive episode and improvement in GAF was found. Psychomotor speed within the depressive episode explained 15% of the total variance in improvement in general functioning from baseline to follow-up in the patient sample included in Papers I and II. After the effect of improvement in depressive symptoms on GAF-score had been adjusted for, the association between baseline speed and GAF improvement was still marginally significant.

An alternative explanation for this association between slowing and lower tendency to functional recovery, could be confounding due to presence of personality traits that are associated with more hesitancy and insecurity of the patients in a test situation (51-53). Patients who have such personality traits may also have lower potential for functional improvement. However, the use of change-variables as dependent variables in the analyses referred to above probably made such a confounding effect on the association smaller, because the patients then served as his/hers own control in the analyses.

8.5 Discussion of improvement in neurocognitive function

8.5.1 Results in view of previous findings

In the present study, significant correlations between improvement in depressive symptomatology and improvement in verbal memory function and executive function over time were found. These findings were in line with several previous studies that found improvement on single tests regarded as indicators of these constructs: Trichard et al. (1995) and Beblo et al. (1999) both found remission of symptoms to be associated with improvement of semantic fluency during a short time interval (one month) (66, 158). Tarbuck and Paykel (1995) showed that remission was associated with improvement on measures of memory and semantic fluency (65). Further, Deuschle et al. (2004) found verbal memory function to be improved after recovery (159).

Contradictory to the findings of the present study, Neu et al. (2001) did not find any significant correlation between improvement in depressive symptomatology and improvement on tests of fluency and verbal memory during a three-month period of observation (64). Yet, also in that study was a significant within-group improvement on semantic fluency and verbal memory in the patient sub-group with unipolar depression.

Williams et al. (2000) reported a significant group difference between remitted patients and controls in favour of the controls on a task of short-term memory functioning (67). In this study, the follow-up interval was only ten days. Neu et al (2005) reported that remitted patients with recurrent episodes of depression still performed lower on verbal memory and semantic fluency after remission (9 months follow-up) (68), and Reischies and Neu (2000), and Nebes et al. (2000) found no significant group x time interactions between groups of remitted patients and healthy controls on measures of attention, verbal memory, visual memory, fluency, and psychomotor speed (69, 70).

8.5.2 Possible explanations for conflicting findings with regard to neurocognitive improvement

As shown in Table 3, several previous studies have also supported the "state"-model (65, 67, 72, 158, 160-162). However, a note of caution should be made with regard to the studies comparing remitted patients to healthy controls in Table 3: Small sample sizes make it difficult to detect significant differences between remitted patients and controls. As a result of this, some studies may have been wrongly classified as supporting the "state"-hypothesis instead of the "trait"-hypothesis (Type II error) (67, 72, 160).

However, as shown in Tables 2 and 4, there have also been several studies that have argued against the "state"-hypothesis (62, 64, 66, 68, 70, 71, 73, 155, 156, 163-165). Most of these studies were heterogeneous with regard to clinical characteristics of the patients included, such as diagnosis, duration of follow-up intervals, level of depressive symptomatology at re-test, co-morbidity, and use of medication (64, 66, 68, 70, 155, 156, 163). In several of the studies supporting the “trait”-hypothesis, short follow-up intervals (less than half a year) in combination with certain degrees of rest-symptomatology at follow-up may explain why neurocognitive function was not significantly improved (64, 68-70).

8.5.3 Does neurocognitive function return to “normal” after recovery?

All the studies shown in Table 2 found rest-deficits in neurocognitive function after remission of the depressive symptoms. Similar to this, mean performance in the patient group with complete recovery in Paper II was lower (although non-significantly) than the performance of the healthy controls on three of the four summary scales of neurocognitive function (Paper II, Figure 2). This may indicate that neurocognitive function does not return entirely to normal after a depressive episode, given that normal is defined as the performance level of healthy controls. Such rest-deficiencies in neurocognitive function have been reported in previous studies that compared remitted or recovered patients with healthy controls (62, 66, 71, 72, 163, 164). Thus, there exists some evidence supporting the “trait”-hypothesis, both in these previous studies and in the present study. However, it should be mentioned, that associations between depression and neurocognitive function at follow-up in these studies could be affected by biases of the kind mentioned in section 8.2.2.

8.5.4 Have patients reached their upper limit of their potential for improvement?

The premorbid level of functioning in the patient group was not known. That is, no estimates of the neurocognitive performance of the patient group relative to controls existed, neither prior to the actual episode, nor prior to the first depressive episode of depression experienced. Thus, even if patients performed lower than controls on several dimensions of neurocognitive function after recovery, their performance may actually have returned to their premorbid levels. An interesting parallell to this, would be Buist-Bouwman (2004)’s recent study, which showed that premorbid levels of different aspects of functioning were lower in subjects who later developed depression, and that post-episode functioning returned to these premorbid levels (166, 167). If patients’ starting point with regard to neurocognitive function is different from healthy controls’, this could either be due to a “trait” feature (either biological or psychological), or it could be caused by factors that confound the association between depression and neurocognitive function (section 8.2.2.).

8.5.5 No predictive value of duration of depression for improvement of neurocognitive function

In our study, duration of illness was not predictive of improvement of neurocognitive function during recovery (Paper II). This is in agreement with Neu et al. (2001) (64), and with findings from cross-sectional studies that have investigated the association between estimates of disease duration and neurocognitive performance in depressed or recovered patients (37, 70, 79, 98, 168). The findings that estimates of duration of depression did not correlate with neurocognitive improvement, suggest that longer duration of disease does not lead to progressive deterioration of neurocognitive function. This is contrary to the neurobiological models for progressive neuronal damage (“scaring”) presented in section 4.2.4. However, it should be kept in mind that the sample studied in Paper II was relatively young (mean age 35.8), and it is not known how these patients will perform as they age if they continue to suffer from recurrent episodes in future.

8.5.6 Conclusion about neurocognitive recovery after depression

In the present work, there is considerable evidence in favour of the model in which improvement of depression is associated with improvement of neurocognitive function. The following arguments support the ”state”-hypothesis: 1) Significant correlations between improvement in depression and improvement in neurocognitive function from baseline to follow-up were found in studies I and II. These associations were consistent both when linear and categorical approaches to depressive symptomatology were applied. Further, correlations most likely would have been stronger if measurement error in the estimates of depression and neurocognitive function had been completely absent. In these correlations, measurement error may have been present, both for the estimates of depression and for neurocognitive function, at baseline as well as at follow-up. Also, despite our presumption that construct reliability increased when the neurocognitive measures were added up to produce composite scales, correlations between change in depression and change in neurocognitive function would probably have been even larger if construct reliabilites for depression and for the neuropsychological constructs were perfect. 2) The patients’ mean performance after total recovery was less than half a SD of the sample mean lower than the mean performance of the healthy controls. Group differences between the controls and the recovered sub-groups at follow-up were statistically non-significant. Both findings indicate that neurocognitive rest-deficits were small. 3) Given that there exists an association between indicators of neuronal activation visualised by fMRI and depression (Paper III), the change in regional cerebral activation upon remission from the depressive episode is a further argument in favour of the view that changes in neurocognitive function are reversible.

However, the present study failed to completely reject the model with persistent neurocognitive changes in depression. In Paper II, patients' performance on several dimensions of neurocognitive function did not correlate with improvement of depression; and both in Paper I and II marginal (non-significant) rest-deficits in neurocognitive function were still present in the recovered patients, compared to controls, at follow-up.

Yet, based on the three findings initially mentioned in this section, the present study gives considerable empirical support for the model in which neurocognitive function normalises after depression ("state"-hypothesis), though the possibility of persistant reduction of some aspects of neurocognitive function ("trait"-effects) is not entirely ruled out.

8.5.7 Changes in patterns of regional brain activation upon remission

To the best of my knowledge, the study presented in Paper III is the first longitudinal fMRI-study that uses a cognitive activation paradigm on patients with unipolar depression. However, Davidson et al. (2003) used a paradigm in which the participants responded to emotional stimuli in a longitudinal study. In their study, significant increases in activation from baseline to follow-up eight weeks later were demonstrated in the left anterior cingulate gyrus (169).

In Paper III, significant within-group changes in activation were demonstrated in the depressed group upon remission of the depressive symptoms. The increases in activation were detected in the left posterior cingulate gyrus (BA 31), the right inferior frontal gyrus (BA 44), and bilaterally in the parietal lobes (BAs 40). The increase in the right inferior parietal lobule was seen in the same area that showed reduction in activation relative to healthy controls while patients were severely depressed.

Several PET resting-state studies performed on patients who were depressed at baseline and remitted at follow-up have been done (93, 102, 105, 107-110). Most of these studies have shown remission of depression to be associated with increase in metabolism in the left dorsolateral prefrontal cortex (BAs 9, 46) (93, 105, 110) or in the left anterior cingulate gyrus (BA 24) (93, 106). Similar to our study, Mayberg et al. (2002) and Mayberg et al. (1999) found increases in the parietal lobes (BAs 40) upon remission (102, 107). However, the recent study by Holthoff et al. (2004) did not find such an increase in activation in these regions upon remission after 12 weeks, but their finding of decrease in activation in cerebellum was in line with the present study (109).

The finding that depressive symptomatology at follow-up correlated with regional brain activation in the frontal and parietal lobes, is contradictory to Holthoff et al. (2004) and Rose et al. (2005), who did not find such correlations (109, 170). Yet, both those studies probably had low statistical power due to small sample sizes; and the analyses were performed with restrictions on the scales that measured depressive symptoms (only patients scoring within the severe symptom range were included). Therefore, Type II error cannot be excluded as cause for the lack of associations in these studies. However, in line with the findings in Paper III, several previous studies have found associations between levels of depressive symptomatology and activation (171-174).

Because the stimulation task given during the scanning sessions was a mental arithmetics task, which can also be regarded as a measure of working memory, this change in activation in the frontal and parietal lobes should theoretically be related to the improvement that was demonstrated on tasks such as the PASAT and the Backward Digit Span sub-task in Papers I and II.

In conclusion, the findings of significant changes in levels of activation in study III most likely provide a neurophysiological correlate to the remission of depressive symptomatology in Papers I and II. This change in activation may reflect change in the neurophysiological mechanisms involved in processing during performance of test tasks.

9. Methodological considerations

9.1 General methodological considerations

9.1.1 Measurement of depressive symptomatology

Patients included in Paper I and II were diagnosed by trained psychiatrists according to the Structured Clinical Interview for DSM-IV axis I disorders - patient edition (SCID I/P, version 2.0) (111) at inclusion. At follow-up, re-assessment of diagnosis was performed with the MINI International Neuropsychiatric Interview (M.I.N.I.) (175). Subjects who no longer fulfilled diagnostic criteria were then excluded (two patients had suffered from manic episodes and were excluded). In all papers, level of depressive symptomatology was measured by commonly used and well validated continuously scaled instruments (115, 120-123, 127). Two out of these are considered as international "Gold-standards" (120, 121). Inter-rater reliability for psychometric scales in these papers was assessed at baseline and found to be high (average intraclass correlations over 0.80) (112). However, the following considerations about assessment of psychiatric caseness and symptom levels should be discussed:

9.1.1.1 Dimensional versus categorical approach

In the present work, dimensional approaches to depressive psychopathology have been used extensively. Arguments in favour of the dimensional approach are: 1) Psychiatric syndromes are in their nature symptom continuums rather than categorical entities. Continuum models should be appropriate, both when looking upon one psychiatric condition separately (example: depression: low-medium-high levels of symptoms), or when overlap (co-morbidity) between syndromes is taken into account (example: depression-schizoaffective disorder-schizophrenia). 2) The dimensional approach also has the methodological advantage that it captures more of the variance in symptomatology than the categorical approach, and 3) it avoids errors arising from misclassification of individuals when diagnostic cut-offs are introduced on the measurement scales (information bias). Although the present diagnostic systems are based on categorical diagnoses, the dimensional approach to psychiatric symptomatology is increasingly used in research contexts, and it has been argued that it should also be introduced in future versions of the diagnostic systems used by clinicians (176).

In order to compensate for the low statistical power in Papers I and II, statistical analyses were performed on continuously scaled measures. But due to restriction of variances on the psychometric scales onto which customary cut-offs had been introduced, the effect estimates for the association between depression and lower neurocognitive function may have become under-estimated when categorical approaches to levels of depressive symptoms were used. This perhaps happened in the comparisons between recovered (N=17) and non-recovered (N=13) sub-groups in Papers I and II. However, it is important to note that findings from the categorical approaches were consistent with findings from the analyses using continuous approaches in all associations investigated.

9.1.2 Measurement and operationalisation of neurocognitive function

The construction of the neurocognitive test battery in Papers I and II was based on theory and tradition. The test battery was broad, and the tests included are frequently used and well validated (132, 133, 138, 139, 144-146, 177). Testing was performed by trained test-technicians under standardised conditions. Neurocognitive operationalisation in these papers was based on theoretic foundations, which were empirically supported by evaluations of underlying factor structures and estimates of internal reliability within dimensions. To produce composite scales of neurocognitive dimensions of function, single construct indicators (test variables) were added up. The composite scales of function were given names in line with the consensus that exists among clinicians and in the literature. This said, the following general considerations concerning operationalisation of neurocognitive test measures should be discussed:

9.1.2.1 Single tests versus composite scales

Traditionally, clinical neuropsychologists create neurocognitive performance profiles based on single test measures, and different test measures are regarded as indicators of different dimensions of neurocognitive function, regardless of the empirical overlap between test measures and neurocognitive dimensions (29, 129). This approach gives richness of detail, but lower reliability. In the present study, construct validity presumably increased when single test measures (construct indicators) were added up to produce summary scales of neurocognitive dimensions in Papers I and II. This approach also led to lower risk of making Type I errors due to multiple comparisons (discussed in 9.1.2.3). However, in Paper IV, only three test tasks were available, and lacking the advantage of more information, no summary scale of neurocognitive function was made.

Theoretically, operationalisation of neurocognitive construct indicators can be made on three levels of richness of detail: 1) Operationalisation based on single test variables. This approach is the conventional approach used in clinical neuropsychology and in research within this and related fields. This approach gives richness in detail, but low construct reliability. An advantage with analyses based on single tests is that profiles for patients' performance can be made. However, care should be taken when interpreting results based on single variables because confidence intervals often are overlapping. Another major concern about this approach, is the high risk of committing Type I error when multiple comparisons are performed. In the research literature, it seems to be a general problem that results from studies are reported based on findings from analyses on many single tasks, and that findings of statistical significance on one or few associations frequently are subject to over-interpretation. 2) An intermediate level of operationalisation was made in Papers I and II. In this approach, dimensions of neurocognitive functions were represented by summary scales consisting of several single test variables (construct indicators). This probably gave higher construct reliability, less error of measurement, and a tendency for increased correlations in the inferential analyses. 3) The third level of operationalisation would involve the computing of an "overall" composite scale including all available measures of neurocognitive function from the test battery. This approach, however, would lead to complete loss of nuances, and probably, no association between depression and the "global" scale of neurocognitive function would be detectable. In the present work, an attempt of demonstrating this mechanism was made in an exploratory analysis based on the 14 neurocognitive variables of change from baseline to follow-up used in the inferential analyses in Paper II. When this "global" 14-items composite scale was used as dependent variable, and the linear change in HAM-D score was used as independent variable, no significant correlation was found between these scales (r=0.23, p=0.122).

In conclusion, the increase in construct reliability achieved by computing summary scales should be regarded as favourable compared to using multiple indicators. However, if evaluations about neurocognitive test profiles were the aim of the investigation, analyses on single test measures would be useful, provided that statistical power were sufficient.

9.1.2.2 Intercorrelations and redundancy between dimensions

As mentioned in the introduction, intercorrelation (overlap) between constructs (dimensions) of neurocognitive function is considerable. An example of this intercorrelation is shown in Table 4 in Paper II, which reports Pearson's correlation coefficients *r* between summary scales of neurocognitive dimensions (variables of change from baseline to follow-up). Another example is given in Figure 2 in the present work, in which each neurocognitive dimension's relative effects on the association between improvement of depression and improvement in other neurocognitive dimensions are depicted. Because of the redundancy between neurocognitive dimensions, the number of neurocognitive measures used as construct indicators was reduced in Paper II. This reduction was achieved by omitting the single task measures with the weakest factor loadings, thereby increasing measurement reliabilities of dimensions.

Because confidence intervals for the neurocognitive single tests were overlapping in the analyses in Papers I and II, we were careful about making conclusions about neurocognitive improvement profiles based on analyses performed on single test measures. An example of this overlap, is shown in Figure 3 in the present work. Figure 3 shows overlapping of 95% confidence intervals for single task performances of the recovered and non-recovered sub-groups in Paper I. Conclusions about which aspects of neurocognitive function are more or less affected in depression should not be made based on analyses of single tests, because they are not truly different (overlap of confidence intervals). However, in Papers I and II, the results from the analyses performed on single test were presented and discussed with regard to neurocognitive profiles.

9.1.2.3 The pitfalls of significance testing

Most studies that have investigated the difference between patients with psychiatric disorders and healthy controls with regard to neurocognitive function have based their findings on significance testing (29). If the p-value is significant, most authors conclude that groups are different from each other with regard to the test variable(s). However, the p-value does not provide any of the following information: 1) What was the magnitude of the difference between the groups? 2) Does the difference in means between groups apply to all of the people in the group, or just to a sub-group within the sample? 3) Is the statistically significant group difference also clinically significant? (29, 178). In light of this, there are two points that will be discussed concerning the findings from papers included in the present study:

1) Significance depend on sample size (178). Significance is more frequent in larger samples, even when the magnitudes of associations are the same as in smaller samples. An implication of this, is that the possibility of positive findings is larger in larger samples (Type I error), and lower in smaller samples (Type II error) (178). This is easily seen in the present study, where Papers I and II, which had low numbers of participants, generated few statistically significant findings, and in Paper IV, which was based on a large sample, many p-values that were below the alpha-level used in Papers I and II were found. But in this paper, effect sizes for the associations were still small.

2) When multiple comparisons are performed, the risk of false positive findings increases (Type I error) (178). For instance, when performing four comparisons, the risk of getting one false positive becomes 20% by an alpha-level of 5%. Therefore, care should be taken to avoid such false positive findings when performing multiple comparisons. This can be done by a priori lowering of the alpha level (Paper IV), or by posthoc adjustment (Paper I).

In studies I and II multiple construct indicators were added up to produce summary scales of neurocognitive function, and the results were based on one (Paper I) or four (Paper II) comparison(s) only. Because of this no correction of alpha level was necessary. In addition, the use of continuous approaches to the measurements, thereby avoiding restriction of variances of the variables, most likely reduced the chances of Type II error due to the low statistical power of the studies. Still, in the present work, estimates of effect sizes for the associations demonstrated are reported, in addition to whether associations were statistically significant or not (*p*-values). Results were also frequently reported when they were non-significant. And, the general tendencies and consistencies between findings from different methodological approaches were emphasised.

9.1.3 Selection biases

In all the papers included in the present study, selection biases may have been present. In Papers I, II, and III these may have been present in several stages: Firstly, at baseline, because patients with complaints about cognitive function may have been more often referred to the study than patients without such complaints. And, the severely ill in-patients included in the studies may be, in general, different from other patient groups in level of functioning or other clinical characteristics. Secondly, selection biases may be present at follow-up because patients with disparate levels of depression or cognitive problems may not have been as likely to respond to the invitation to participate again. Fischer et al. (2001) previously reported that participants lost due to attrition in follow-up studies generally were more severely impaired during the baseline hospitalisation, and that males were more often lost than females (179). In Papers I and II, patients with longer duration of depression were over-represented at follow-up compared to the sub-group that was lost to follow-up. Most likely, this selection bias led to inflations of effect sizes for differences between depressed patients and healthy controls (see section 8.2.2). It should be noted, however, that there were no detectable differences with regard to important features such as age, gender, education, or intellectual abilities between those who were re-tested and those who were lost to follow-up.

Hansen et al. (2001) previously reported that patients with mental conditions and older age are more likely not to participate in health studies (54). In the epidemiological study presented in Paper IV, only 51% of the age cohort participated in the neurocognitive examination, and after those who had provided non-valid answers had been excluded from the data-file, only 44% of the cohort was left for the inferential analyses. This low participation rate is one of the major concerns in this paper. Subjects with female gender and lower level of education were under-represented in the study sample. As a consequence of these selection biases, prevalence of depression and anxiety may have been under-estimated, and variance of psychiatric symptomatology could have been restricted. This may again have led to under-estimation of effect sizes for the association between depression and test performance.

9.1.4 Confounding factors

In Papers I and II, no significant group differences were found between the recovered and non-recovered sub-groups with regard to age, gender, level of education, and level of general intellectual abilities when this was tested for. However, many other factors may confound the association between depression and neurocognitive function (see section 8.2.2). Examples of such factors are general level of functioning, absence from work due to sickness, motivation in the test situation, quality of sleep, and use of psychotropic medication or other substances (129, 153, 180). For instance, the non-recovered sub-group may, independent of the depressive symptomatology itself, in general, have lower level of functioning, less initiative, and perhaps also less offensive attitudes to the tasks given in the test situation. All of these potential confounding factors may exist independently of depression, or as parts of a vulnerability present (perhaps premorbid and independent of the depressive symptomatology), which also leads to poorer test performance.

In Paper IV, considerable effort was made to adjust the inferential analyses for possible confounders and mediators of the association between depression and test performance. Accurate information about a number of possible confounders was available, such as diagnosis (by ICPC- numbers), medication (by ATC-numbers), sleep disturbances, and physical activity. All of these factors were adjusted for in the inferential analyses. As expected, large changes in effect estimates for the association between depression and neurocognitive function occurred when level of education was entered into the statistical model. However, there may have been residual confounding caused by the effect of education or other factors adjusted for in the analyses. Further, factors that influence the association between depression and test performance not asked for in the test protocol may have been present. For instance, the protocol did not include tasks that made it possible to estimate the participants' general level of cognitive abilities (IQ). This is an important confounding factor in associations between psychiatric disorders and neurocognitive function (129, 153, 180).

9.2 Further strengths and limitations of papers I to IV

9.2.1 Papers I and II

To the best of our knowledge, only one previous study (by Abas et al. (1990) (156), see Table 4) has had a period of observation similarly long as the study that Papers I, II, and III were based on. By using a two year re-testing interval, our study had an advantage if neurocognitive recovery takes longer than symptom recovery after a depressive episode. In this timeframe, the possible delay in neurocognitive recovery after recovery from depression should be eliminated. The longitudinal design of the presemt study was in itself also a strength because intra-individual variables of change for depression and neurocognitive function could be computed, thus making each subject serve as his/her own "control" in the analyses. By using this approach, some of the effect of possible confouders on the association between depression and neurocognitive function was possibly avoided

Other strengths of this study include that the sample was well-characterised and homogeneous (only patients with unipolar major depression of recurrent sub-type were included). In addition, 30 out of the 50 patients in the original baseline sample were available at follow-up, which should be a satisfactory low rate of attrition.

Re-administration of a neurocognitive test task often leads to improved performance at re-testing as a result of learning (181-184). Also, ceiling effects may occur in the second test situation. This particularly applies to "one-shot"-tasks such as the Wisconsin Card Sorting Test (WCST) (144, 181, 185). To avoid such test re-test effects, alternate test-forms can be distributed at follow-up (64, 65, 70), or learning effects can be assessed with improvement in healthy controls as reference. When this was not done in the present study, it was because the test re-test interval was so long (two years) that the effect of learning on test performance was probably very minor. In addition, learning effects were most likely equally distributed in the recovered and non-recovered sub-groups, since they were equally depressed at baseline.

9.2.2 Paper III

Theoretically, findings of increase (or decrease) of activation in particular cerebral regions during or after depression improvement should be found in regions where 1) patients have shown a different level of activation at baseline compared to controls, and 2) healthy controls have shown changes in level of activation during stimulus processing (186). In the present study, changes of activation were detected in the parietal lobes. However, in the prefrontal region, where a significant increase in activation from baseline to follow-up was detected in the depressed group, no significant difference in activation was found at baseline in the depressed group compared to the controls. The lack of baseline reduction in the depressed patients compared to controls in this region, could, however, reflect a mechanism of over-compensation during task performance. Possibly, the patients "tried harder" to complete the neurocognitive task. This increased effort may have lead to more neuronal activation in prefrontal areas.

Another weakness of the present study is that the performance data at the second scanning was lost due to a technical error. Thus, the changes in activation at re-test compared to at baseline may have been confounded by better test-performance. Yet, it could be argued that unless the increase in activation from baseline to follow-up in the depressed group was caused by improvement in task performance (as opposed to improvement in depressive symptomatology), the loss of information about performance is not relevant for the main finding of increased activation over time. Previous studies have shown that level of estimates of cerebral activation is not correlated with task performance within depressed groups (151, 174). Thus, the findings of changes in patterns of activation in the present study can most likely be attributed to changes in neuronal activity associated with changes in level of depression, rather than test performance.

The comparison between the first and second measurement was based on the assumption of high test re-test reliability with regard to level of activation at both occasions. However, previous reports investigating pre- and post-test reliability have concluded that activation data are reliable with regard to this (186, 187).

Further limitations of the study were, firstly, that statistical power was low due to low number of participants. Secondly, patients were on different medications, which could influence neuronal transmission (83), and consequently, patterns of regional brain activation. Unfortunately, the design of the study did not allow for assessment of medication effects on changes in activation. Finally, it should be mentioned that the assumptions underlying One-Way ANOVAs with regard to normal distribution of the variances of the dependent variable, and independency of test variables (188), may not have been perfectly met.

9.2.3 Paper IV

In the present study, a self-report measure was used to assess caseness and level of depression. Therefore, diagnostic reliability in the analyses was probably lower compared to if diagnosis had been made by a specialist using a structured psychometric instrument. The HADS-D (115) focuses mostly on features of "anhedonia" in depression, but not on somatic symptoms that are related to depression. HADS-A covers anxiety symptoms corresponding to General Anxiety Disorder (GAD). This is different from many other "gold-standards". Therefore, some subjects may have been wrongly classified as not suffering from depression or anxiety. This may have led to weakening of the associations between depression and/or anxiety and lower neurocognitive function.

Despite this, and despite strong emphasis on psychometrics in measuring psychiatric psychopathology and neurocognitive function, error of measurement cannot be ruled out. For instance, in the correlations of depression or anxiety with neurocognitive test performances, such measurement errors could not be excluded. In these analyses, measurement errors could occur both for HADS as well as for the neurocognitive test measures. Measurement errors are likely to be random, and most likely resulting in under-estimation of the strength of the associations between depression or anxiety and neurocognitive test performance.

In this paper, the potential presence of un-detected cases with co-morbid dementia or mild cognitive deficit (MCI) may have confounded the association between depression and neurocognitive function. Depression can be seen as prodrome of or as an early clinical manifestation of dementia; and the degree of co-morbidity in these disorders is high (189). Although probable cases with dementia or MCI were excluded by an instrument which has shown high sensitivity when used as a screening instrument for these conditions (a modified version of the Mini-Mental State Examination (MMSE) (113, 114), there may still have been undetected cases included in the analyses on the association between depression and neurocognitive function. This may have led to over-estimation of the association. However, after further testing using Kendrick's Object Learning Test (KOLT), which is very sensitive to dementia (177), in addition to the modified version of the MMSE to exclude subjects with potential MCI or dementia, results regarding the association were not altered. It is therefore not likely that the associations found were confounded by the the presence of dementia or MCI.

10. Conclusions, implications, and directions for future studies

10.1 Conclusions

The present study confirmed the presence of an inverse association between depression and neurocognitive function. The inverse association between depressive symptoms and cognitive function was found to be close to linear, and also present in the sub-clinical symptom range. Thus, reduction of neurocognitive function in depression is not a phenomenon restricted only to severely depressed patients (in-patients). It can rather be regarded as a "normal"-phenomenon that also occurs frequently within the normal population. In the normal population, it is present in symptom-ranges below diagnostic threshold for depression, i.e. in the symptom-ranges often seen in dysthymia. Consequently, it may affect a considerable part of the population. Probably it is often seen in primary care settings, where patients with lower-range depressive symptoms frequently are seen.

The present study also generated empirical support for a model in which remission from depression is followed by improvement in neurocognitive function. However, longer duration of depression was not associated with poorer neurocognitive function. And it is therefore that the present study does not support a model in which recurrent episodes of depressed mood lead to reduction of neurocognitive function. Yet, the present study cannot completely exclude the persistence of rest-deficits in neurocognitive function after total recovery from the depressive symptoms, as the mean performance in the patient group that had become completely well was still marginally below the performance of healthy controls (although non-significantly).

Remission from depressive symptoms was also associated with changes in patterns of neuronal activation as measured by fMRI. These changes were seen in areas in which depressed patients previously have shown patterns of activation differently from healthy controls. Levels of activation in the patients after they had experienced improvement of depression were more similar to the activation of normal controls in the ROIs. This may indicate that a normalisation of underlying pathophysiological mechanisms had occured. One could speculate that this normalisation could be one of the explanations for the improved performance on test tasks after remission in the patients.

10.2 Clinical implications and generalisation of findings

The finding of an inverse association between depression and neurocognitive function is important because this may have implications for patients' level of functioning on several areas in life (59, 60). Clinicians who are responsible for patients who experience problems with concentration, memory, and tempo while they are depressed should try to keep the following in mind:

1) Because of the cognitive problems, it may be difficult for the patient to fully benefit from intensive psychotherapeutic intervention during the acute phases of episodes of major depression. Since these problems seem to improve upon remission of the depressive symptoms, such intervention should perhaps be made at a later stage, while the patient is in remission.

2) There exists no evidence suggesting that patients should not receive antidepressant medication when they suffer from lower neurocognitive function. The only exception to this rule, are tricyclic antidepressants. In the present study, it was shown that use of antidepressants did not have a negative impact on neurocognitive function (section 8.3).

3) Possibly, there exists a delay between improvement in depression and improvement in neurocognitive function. An implication of the existence of such a delay is that patients need longer sick-leaves, and/or particular arrangements at their work place during the first months of work after a depressive episode. It should be recommended that the patient has his/her own undisturbed work place, interruptions should be avoided, and work load should be tolerable. In particular, it is important to be aware that he/she may be sensitive to tempo-demanding tasks.

4) Clinicians in charge of depressed patients who experience memory and concentration problems, should reassure their patients that these problems are most likely going to improve when the patient becomes well. Such reassurance could avoid patients' speculations and fear about never "getting normal" cognitively again.

5) In some psychiatric or neurological disorders, rehabilitation programmes including neurocognitive "training" have been attempted (190-192). However, whether such programmes are relevant for depression is perhaps debatable because depressed patients are probably less affected by problems with cognitive tasks than other patient groups.

Knowledge generated by Papers I, II, and III can probably be generalised to other samples of patients who are severely affected by depression, or to other groups of depressed patients with low levels of functioning and need for professional care and medication. However, because of biases of participants at inclusion (only severely ill patients were selected to participate), findings may to a lesser extent be transferable to patients in ambulatory care (e.g. in primary health care), or to the general population. Findings from Paper IV are probably generalisable to the kind of elderly patients with lower levels of depressive symptomatology typically seen in primary health care settings. However, because there may exist confounders that affect the association between depression and neurocognitive function exclusively in older age groups, it may be debatable to what extent the findings can be transferred to other age groups.

The studies were not designed in a way that makes it possibly to make inferences about causal pathways. Consequently, no conclusions about causal relationships were made. However, if one should speculate, it seems appropriate that neurocognitive changes follow changes in depressive symptoms, and not vice versa. Also, it seems reasonable to think that both the depressive symptoms, the changes in neurocognitive function, and the changes in patterns of regional brain activation all are indicators of a common underlying pathophysiological dysfunction.

10.3 Directions for future studies

Studies within genetics, neurophysiology, and functional neuroimaging are going to be central in future research on the association between psychiatric disorders and neurocognitive function. Researchers using neurocognitive methodology should more often include estimates of biological correlates of depression in their studies in order to better clarify the associations between pathobiological changes, depressive psychopathology, and neurocognitive function.

The following should be taken into consideration in neurocognitive studies of the association between depression and neurocognitive function:

1) Studies should include sufficient numbers of participants in order to avoid problems related to low statistical power (multi-center studies). Sufficient power allows for testing of hypotheses concerning neurocognitive profiles (whether one dimension of neurocognitive function is significantly different from another) and differences between sub-groups.

2) Longitudinal study designs should be used, since these are more suitable for making inferences about causal relationships. In addition, longitudinal designs allow for within-subject comparisons (as opposed to between-subject comparisons). This reduces the effect of group biases on the associations under investigation, and makes it possible to create intra-individual profiles of change over time for the variables studied.

3) In order to avoid artificial inflation of effect sizes for the associations between depression and neurocognitive function, the associations should also be studied in samples from the normal population, and not only by comparing clinical groups with healthy controls. Further, an interesting question, which could be answered by using samples from longitudinal population studies, is whether premorbid levels of neurocognitive functioning are different in individuals who later develop depression compared to in people who do not become depressed.

4) Groups and sub-groups should be well-characterised with regard to sociodemographic factors, diagnosis, and clinical characteristics. This allows for adjustment for factors that may confound the association between depression and neurocognitive function. And it makes it possible to detect the aspects of depression that influences neurocognitive function the most. In addition, sub-groups that are impaired with regard to neurocognitive function can be characterised and compared to sub-groups of subjects who perform in the normal range.

5) Statistically, a dimensional approach to the variables studied is probably preferable to a categorical one, since a dimensional approach avoids misclassification of subjects scoring near to cut-off, and because it captures more of the variance of the variables. A consequence of this, is that subjects with all levels of depressive symptomatology would be included. Also,

conclusions about neurocognitive profiles based on statistically significant findings from one or two variables out of many variables tested should be avoided.

6) Finally, in order to treat and rehabilitate patients optimally, it should be clarified to what extent lower neurocognitive function is related to functional disability in depression.

11. References

1. Blazer D, Kessler R, McGonagle K, Swartz M. The Prevalence and Distribution of Major Depression in a National Community Sample: The National Comorbidity Survey. Am J Psychiatry 1994;151:979-86.
2. Judd L, Akistal H, Paulus M. The role and clinical significance of subsyndromal depressive symptoms (SSD) in unipolar major depressive disorder. J Affect Disord 1997;45:5-17.
3. Murphy J, Laird N, Monson R, Sobol A, Leighton A. A 40-Year Perspective on the Prevalence of Depression. Arch Gen Psychiatry 2000;57:209-15.
4. Kringlen E, Torgersen S, Cramer V. A Norwegian Psychiatric Epidemiological Study. Am J Psychiatry 2001;158:1091-98.
5. Angst J. Epidemiology of depression. Psychopharmacology 1992;106:71-4.
6. Rorsman B, Grasbeck A, Hagness O, et al. A prospective study of first-incidence depression.The Lundby study, 1957-72. Br J Psychiatry 1990;156:336-42.
7. Isacson D, Bingefors K, Knorring L. The impact of depression is unevenly distributed in the population. Eur Psychiatry 2005;20:205-12.
8. Group TCVAIPCS. How disabling is depression? Evidence from a primary care sample. British J Gen Pract 1999;49:95-98.
9. Judd L, Akistal H, Zeller P, et al. Psychosocial Disability During the Long-term Course of Unipolar Major Depressive Disorder. Arch Gen Psychiatry 2000;57:375-80.
10. Wang P, Simon G, Kessler R. The economic burden of depression and the cost-effectiveness of treatment. Int J Methods Psychiatr Res 2003;12:22-33.
11. Broadhead W, Blazer D, George L, Tse C. Depression, disability days, and days lost from work in a prospective epidemiologic survey. JAMA 1990;264:2549-50.
12. Cuijpers P, Smit F. Excess mortality in depression: a meta-analysis of community studies. J Affect Disord 2002;72:227-36.
13. Mykletun M, Bjerkeset O, Stewart R, et al. Anxiety, depression and mortality The HUNT Study. Am J Psychiatry (submitted) 2005.
14. Kessler R, McGonagle K, Shanyang Z, et al. Lifetime and 12-Month Prevalence of DSM-III-R Psychiatric Disorders in the United States. Arch Gen Psychiatry 1994;51:8-19.
15. Brown T, Campbell L, Lehman C, Grisham J, Mancill R. Current and Lifetime Comorbidity of the DSM-IV Anxiety and Mood Disorders in a Large Clinical Sample. J Abnorm Psychol 2001;110:585-99.
16. van Balkom A, Beekman A, de Beurs E, Deeg D, van Dyck R, van Tilburg W. Comorbidity of the anxiety disorders in a community-based older population in The Netherlands. Acta Psychiatr Scand 2000;101:37-45.
17. Beekman A, de Beurs E, Van Balkom A, Deeg D, van Dyck R, van Tilburg W. Anxiety and Depression in Later Life: Co-Occurrence and Communality of Risk Factors. Am J Psychiatry 2000;157:89-95.
18. Kessler R, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003;289:3095-3105.
19. Kizilbash A, Vanderploeg R, Curtiss G. The effects of depression and anxiety on memory performance. Arch Clin Neuropsychol 2002;17:57-67.

20. Stordal E, Bjelland I, Dahl A, Mykletun A. Anxiety and depression in individuals with somatic health problems. The Nord-Trøndelag Health Study (HUNT). Scand J Prim Health Care 2003;21:136-41.
21. Harrison J, Owen A. Cognitive Deficits in Brain Disorders. London: Martin Dunitz Ltd., 2002.
22. Porter R, Gallagher P, Thompson J, Young A. Neurocognitive impairment in drug-free patients with major depressive disorder. Br J Psychiatry 2003;182:214-220.
23. Angst J, Merikangas K. The depressive spectrum: diagnostic classification and course. J Affect Disord 1997;45:31-40.
24. Stordal E, Krüger M, Dahl N, Krüger Ø, Mykletun A, Dahl A. Depression in relation to age and gender in the general population: the Nord-Trøndelag Health Study (HUNT). Acta Psychiatr Scand 2001;104:210-216.
25. Keller M. Past, Present, and Future Directions for Defining Optimal Treatment Outcome in Depression: remission and beyond. JAMA 2003;289:3152-60.
26. Akiskal H. Mood disorders: Introduction and overview. In: Sadock B, Sadock V, eds. Kaplan & Sadock's Comprehensive Textbook of Psychiatry. Philadelphia: Lippincott Williams & Wilkins, 2000:1284-98.
27. Fava G, Ruini C, Sonino N. Treatment of recurrent depression: a sequential psychotherapeutic and psychopharmacological approach. CNS Drugs 2003;17:1119-22.
28. APA. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). Fourth Edition. Washington DC: American Psychiatric Association, 1994.
29. Zakzanis K, Leach L, Kaplan E. Neuropsychological differential diagnosis. Lisse: Swets & Zeitlinger, 1999.
30. Veiel H. A Preliminary Profile of Neuropsyhological Deficits Associated with Major Depression. J Clin Exp Neuropsychol, 1997:587-603.
31. Landrø N, Stiles T, Sletvold H. Neuropsychological Function in Nonpsychotic Unipolar Major Depression. Neuropsychiatry Neuropsychol Behav Neurol 2001;14:233-40.
32. Albus M, Hubmann W, Wahlheim C, Sobizack N, Franz U, Mohr F. Contrasts in neuropsychological test profile between patients with first-episode schizophrenia and first-episode affective disorders. Acta Psychiatr Scand 1996;94:87-93.
33. Austin M, Mitchell P, Goodwin G. Cognitive deficits in depression: Possible implications for functional neuropathology. Br J Psychiatry 2001;178:200-206.
34. Brand A, Jolles J, Gispen-de Wied C. Recall and recognition memory deficits in depression. J Affect Disord 1992;25:77-86.
35. Horan W, Pogge D, Borgaro S, Stokes J, Harvey P. Learning and Memory in Adolescent Psychiatric Inpatients with Major Depression:A Normative Study of the California Verbal Learning Test. Arch Clin Neuropsychol 1997;12:575-84.
36. Landrø N, Stiles T, Sletvold H. Memory functioning in patients with primary fibromyalgia and major depression and healthy controls. J Psychosom Res 1997;42:297-306.
37. Lampe I, Sitskoorn M, Heeren T. Effects of recurrent major depressive disorder on behavior and cognitive function in female depressed patients. Psychiatry Res 2004;125:73-9.

38. Mialet J, Pope H, Yurgelun-Todd D. Impaired attention in depressive states: a non-specific deficit? Psychol Med 1996;26:1009-20.
39. Purcell R, Maruff P, Kyrios M, Pantelis C. Neuropsychological function in young patients with unipolar major depression. Psychol Med 1997;27:1277-85.
40. Stordal K, Lundervold A, Egeland J, et al. Impairment across executive functions in recurrent major depression. Nord J Psychiatry 2004;58:41-7.
41. Degl'Innocenti A, Agren H, Backman L. Executive deficits in major depression. Acta Psychiatr Scand 1998;97:182-8.
42. Elliott R, Baker S, Rogers R, et al. Prefrontal dysfunction in depressed patients performing a complex planning task: a study using positron emission tomography. Psychol Med 1997;27:931-42.
43. Fossati P, Amar G, Raoux N, Ergis A, Allilaire J. Executive functioning and verbal memory in young patients with unipolar depression and schizophrenia. Psychiatry Res 1999;89:171-87.
44. Franke P, Maier W, Hardt J, Frieboes R, Lichtermann D, Hain C. Assessment of Frontal lobe Functioning in Schizophrenia and Unipolar Major Depression. Psychopathology 1993;26:76-84.
45. Egeland J, Rund B, Sundet K, et al. Attention profile in schizophrenia compared with depression: differential effects of processing speed, selective attention and vigilance. Acta Psychiatr Scand 2003;108:276-284.
46. Ilsley J, Moffoot A, O'Carroll R. An analysis of memory dysfunction in major depression. J Affect Disord 1994;35:1-9.
47. Kindermann S, Brown G. Depression and Memory in the Elderly: A Meta-analysis. J Clin Exp Neuropsychol 1997;19:625-42.
48. Roness A, Mykletun A, Dahl A. Help-seeking behaviour in patients with anxiety disorder and depression. Acta Psychiatr Scand 2005;111:51-8.
49. Massman P, Delis D, Butters N, Dupont R, Gillin J. The Subcortical Dysfunction Hypothesis of Memory Deficits in Depression: Neuropsycholgical Validation in a Subgroup of Patients. J Clin Exp Neuropsychol 1992;14:687-706.
50. Goethe J, Fischer E. Functional impairment in depressed patients. J Affect Disord 1995;33:23-9.
51. Beck A, Rush A, Shaw B, Emery G. Cognitive Therapy of Depression. New York: Guilford Press, 1979.
52. Elliott R, Sahakian B, Herrod J, Robbins T, Paykel E. Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment. J Neurol Neurosurg Psychiatry 1997;63:74-82.
53. Parker G, Malhi G, Mitchell P, et al. Progressing a spectrum model for defining non-melancholic depression. Acta Psychiatr Scand 2005;111:139-143.
54. Hansen V, Jacobsen B, Arnesen E. Prevalence of serious psychiatric morbidity in attenders and non-attenders. Am J Epidemiol 2001;154:891-4.
55. Paterniti S, Dufouil C, Bisserbe J, Alperovitch A. Anxiety, depression, psychotropic drug use and cognitive impairment. Psychol Med 1999;29:421-8.
56. Bryan J, Luszcz M. Measures of Fluency as Predictors of Incidental Memory Among Older Adults. Psychol Aging 2000;15:483-9.
57. Weiser M, Reichenberg A, Rabinowitz J, et al. Cognitive performance of male adolescents is lower than controls across psychiatric disorders: a population based study. Acta Psychiatr Scand 2004;110:471-4.

58. Palsson S, Johansson B, Berg S, Skoog I. A population study on the influence of depression on neuropsychological functioning in 85-year-olds. Acta Psychiatr Scand 2000;101:185-93.
59. McCall W, Dunn A. Cognitive deficits are associated with functional impairment in severely depressed patients. Psychiatry Res 2003;121:179-84.
60. McGurk S, Mueser K. Cognitive Functioning and Employment in Severe Mental Illness. J Nerv Ment Dis 2003;191:789-98.
61. Clark L, Goodwin G. State- and trait-related deficits in sustained attention in bipolar disorder. Eur Arch Psychiatry Clin Neurosci 2004;254:61-8.
62. Marcos T, Salamero M, Gutierrez F, Catalan R, Gasto C, Lazaro L. Cognitive dysfunctions in recovered melancholic patients. J Affect Disord 1994;32:133-7.
63. Ormel J, Oldehinkel A, Nolen W, Vollebergh W. Psychosocial Disability Before, During, and After a Major Depressive Episode A 3-Wave Population-Based Study of State, Scar, and Trait Effects. Arch Gen Psychiatry 2004;61:387-92.
64. Neu P, Kiesslinger U, Schlattmann P, Reischies F. Time-related cognitive deficiency in four different types of depression. Psychiatry Res 2001;103:237-47.
65. Tarbuck A, Paykel E. Effects of major depression on the cognitive function of younger and older subjects. Psychol Med 1995;25:285-96.
66. Trichard C, Martinot J, Alagille M, et al. Time course of prefrontal lobe dysfunction in severely depressed in-patients: a longitudinal neuropsychological study. Psychol Med 1995;25:79-85.
67. Williams R, Hagerty B, Cimpric B, Therrien B, Bay E, Hiroake O. Changes in directed attention and short-term memory in depression. J Psychiatr Res 2000;34:227-238.
68. Neu P, Bajbuij M, Schilling A, Godemann F, Berman R, Schlattmann P. Cognitive function over the treatment course of depression in middle-aged patients: correlation with brain MRI signal hyperintensities. J Psychiatr Res 2005;39:129-35.
69. Nebes R, Butters M, Mulsant B, et al. Decreased working memory and processing speed mediate cognitive impairment in geriatric depression. Psychol Med 2000;30:679-91.
70. Reischies F, Neu P. Comorbidity of mild cognitive disorder and depression-a neuropsychological analysis. Eur Arch Psychiatry Clin Neurosci 2000;4:186-93.
71. Weiland-Fiedler P, Erickson K, Waldeck T, et al. Evidence for continuing neuropsychological impairments in depression. J Affect Disord 2004;82:253-8.
72. Beats B, Sahakian B, Levy R. Cognitive performance in tests sensitive to frontal lobe dysfunction in the elderly depressed. Psychol Med 1996;26:591-603.
73. Jaracz J, Borkowska A, Chlopocka-Wozniak M, Rybakowski J. Cognitive functions in remitted unipolar female depressive patients during maintenance treatment with antidepressants. Arch Psychiatry Psychother 2002;4:15-23.
74. Shah PJ, Ebmeier KP, Glabus MF, Goodwin G. Cortical grey matter reductions associated with treatment-resistant chronic unipolar depression. Br J Psychiatry 1998;172:527-32.

75. Soares J, Mann J. The functional neuroanatomy of mood disorders. J Psychiatr Res 1997;31:393-432.
76. Sheline Y, Sanghavi M, Mintun M, Gado M. Depression duration but not age predicts hippocampal volume loss in medically healthy women with recurrent major depression. J Neurosci 1999;19:5034-43.
77. MacQueen G, Campbell S, McEwen B, et al. Course of illness, hippocampal function, and hippocampal volume in major depression. Proc Natl Acad Sci USA 2003;100:1387-92.
78. Harvey P, Le Bastard G, Pochon J, et al. Executive functions and updating of the contents of working memory in unipolar depression. J Psychiatr Res 2004;38:567-76.
79. Grant M, Thase M, Sweeney J. Cognitive Disturbance in Outpatient Depressed Younger Adults: Evidence of Modest Impairment. Biol Psychiatry 2001;50:35-43.
80. Sheline Y. 3D MRI Studies of Neuroanatomic Changes in Unipolar Major Depression: The Role of Stress and Medical Comorbidity. Biol Psychiatry 2000;48:791-800.
81. Sauro M, Jorgensen R, Pedlow C. Stress, Glucocorticoids, and Memory: A Meta-analytic Review. Stress 2003;6:235-45.
82. O'Brien J. The 'glucocorticoid cascade' hypothesis in man Prolonged stress may cause permanent brain damage. Br J Psychiatry 1997;170:199-201.
83. Kandel E, Schwartz J, Jessell T. Principles of Neural Science: McGraw-Hill, 2000.
84. Mitchell A. The Role of Corticotropin Releasing Factor in Depressive Illness: a Critical Review. Neurosci Biobehav Rev 1998;22:636-51.
85. Campbell S, MacQueen G. The role of the hippocampus in the pathophysiology of depression. J Psychiatry Neurosci 2004;29:417-26.
86. Egeland J, Lund A, Landrø N, et al. Cortisol levels predict executive and memory function in depression, symptom level predicts psychomotor speed. Acta Psychiatr Scand 2005;112:434-41.
87. Cummings J. Frontal-Subcortical Circuits and Human Behavior. Arch Neurol 1993;50:873-80.
88. Mega M, Cummings J. Frontal-subcortical circuits and neuropsychiatric disorders. J Neuropsychiatry Clin Neurosci 1994;7:271-2.
89. Royall D. Frontal Systems Impairment in Major Depression. Semin Clin Neuropsychiatry 1999;4:13-23.
90. Stuss D, Alexander M. Executive functions and the frontal lobes: a conceptual view. Psychol Res 2000;63:289-98.
91. Stuss D, Levine B. Adult Clinical Neuropsychology: Lessons from Studies of the Frontal Lobes. Annu. Rev. Psychol. 2002;53:401-33.
92. Rogers M, Bradshaw J, Phillips J, et al. Parkinsonian Motor Characteristics in Unipolar Major Depression. J Clin Exp Neuropsychol 2000;22:232-244.
93. Rogers MA, Bradshaw JL, Pantelis C, Phillips JG. Frontostriatal deficits in unipolar major depression. Brain Res Bull 1998;47:297-310.
94. Hashimoto K, Shimizu E, Iyo M. Critical role of brain-derived neurotrophic factor in mood disorders. Brain Res Reviews 2004;45:104-14.
95. Sapolsky R. Is Impaired Neurogenesis Relevant to the Affective Symptoms of Depression? Biol Psychiatry 2004;56:137-9.
96. Shelton R. Cellular mechanisms in the vulnerability to depression and response to antidepressants. Psychiatr Clin North Am 2000;23:713-29.

97. Burt D, Zembar M, Niedereche G. Depression and memory impairments: a meta-analysis of the association, its pattern, and specificity. Psychol Bull 1995;117:285-305.
98. Verdoux H, Liraud F. Neuropsychological function in subjects with psychotic and affective disorders. Relationship to diagnostic category and duration of illness. Eur Psychiatry 2000;15:236-43.
99. Drevets W. Neuroimaging Studies of Mood Disorders. Biol Psychiatry 2000;48:813-829.
100. Videbech P. PET measurements of brain glucose metabolism and blood flow in major depressive disorder: a critical review. Acta Psychiatr Scand 2000;101:11-20.
101. Kennedy S, Javanmard M, Vaccarion F. A Review of Functional Neuroimaging in Mood Disorders: Positron Emission Tomography and Depression. Can J Psychiatry 1997;42:467-75.
102. Mayberg H. Localization of Pathophysiology. Am J Psychiatry 2002;159:1979.
103. Dolan R, Bench C, Brown R, Scott L, Friston K, Frackowiak R. Regional cerebral blood flow abnormalities in depressed patients with cognitive impairment. J Neurol Neurosurg Psychiatry 1992;55:768-73.
104. Sackeim H. Functional Brain Circuits in Major Depression and Remission. Arch Gen Psychiatry 2001;58:649-50.
105. Drevets W, Bogers W, Raichle M. Functional anatomical correlates of antidepressant drug treatment assessed using PET measures of regional glucose metabolism. Eur Neuropsychopharmacol 2002;12:527-44.
106. Bench C, Frackowiak R, Dolan R. Changes in regional cerebral blood flow on recovery from depression. Psychol Med 1995;25:247-51.
107. Mayberg HS, Liotti M, Brannan SK, et al. Reciprocal Limbic-Cortical Function and Negative Mood: Converging PET Findings in Depression and Normal Sadness. Am J Psychiatry 1999;156:675-682.
108. Mayberg H, Brannan S, Tekell J, Silva JM, RK, McGinnis S, Jerabek P. Regional metabolic effects of fluoxetine in major depression: serial changes and relationship to clinical response. Biol Psychiatry 2000;48:830-43.
109. Holthoff V, Beuthien-Baumann B, Zundorf G, et al. Changes in brain metabolism associated with remission in unipolar major depression. Acta Psychiatr Scand 2004;110:184-94.
110. Brody A, Saxena S, Mandelkern M, Fairbanks L, Ho M, Baxter L. Brain Metabolic Changes Associated with Symptom Factor Improvement in Major Depressive Disorder. Biol Psychiatry 2001;50:171-8.
111. First M, Spitzer R, Gibbon M, Williams J. Structured clinical interview for DSM-IV axis I disorders- patient edition (SCID I/P, version 2.0). New York: Biometrics Research Department, New York State Psychiatric Institute, 1995.
112. Egeland J, Sundet K, Asbjørnsen A, et al. Sensitivity and spesificity of memory dysfunction in schizophrenia: A comparison with major depression. J Clin Exp Neuropsychol 2003;25:79-93.
113. Braekhus A, Laake K, Engedal K. The Mini-Mental State Examination: Identifying the Most Efficient Variables for Detecting Cognitive Impairment in the Elderly. J Am Geriatr Soc 1992;40:1139-45.
114. Folstein M, Folstein S, McHugh P. "Mini-Mental state" A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189-98.

115. Zigmond A, Snaith R. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983;67:361-70.
116. Copeland J, Beekman A, Dewey M, et al. Depression in Europe Geographical distribution among older people. Br J Psychiatry 1999;174:312-21.
117. Forsell Y, Winblad B. Major depression in a population of demented and nondemented older people: prevalence and correlates. J Am Geriatr Soc 1998;46:27-30.
118. Rosenvinge B, Rosenvinge J. Forekomst av depresjon hos eldre-systematisk oversikt over 55 prevalensstudier fra 1990-2001. Tidsskr Nor Lægeforen 2003;123:928-9.
119. Palsson S, Larsson L, Tengelin E, et al. The prevalence of depression in relation to cerebral atrophy and cognitive performance in 70- and 74- year-old women in Gothenburg. The Women's Health Study. 2001.
120. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56-62.
121. Montgomery S, Aasberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:382-9.
122. Bjelland I, Dahl A, Haug T, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale An updated literature review. J Psychosom Res 2002;52:69-77.
123. Mykletun A, Stordal E, Dahl A. Hospital Anxiety and Depression (HAD) scale: factor structure, item analyses and internal consistency in a large population. Br J Psychiatry 2001;179:540-4.
124. Olssøn I, Mykletun A, Dahl A. The hospital anxiety and depression rating scale: A cross-sectional study of psychometrics and case finding abilities in general practice. BMC Psychiatry 2005;5.
125. Fava M. Past, Present, and Future Directions for Defining Optimal Treatment Outcome in Depression Remission and Beyond. JAMA 2003;289:3152-60.
126. Frank E, Prien R, Jarrett R, et al. Conceptualization and Rationale for Consensus Definitions of Terms in Major Depressive Disorder. Arch Gen Psychiatry 1991;48:851-5.
127. Bagby R, Ryder A, Schuller D, Marshall M. The Hamilton Depression Rating Scale: Has the Gold Standard Become a Lead Weight? Am J Psychiatry 2003;161:2163-77.
128. Hampson P, Morris P. Understanding Cognition. Oxford: Blackwell Publishers Ltd, 1996.
129. Lezak MD. Neuropsychological Assessment. New York: Oxford University Press, 1995.
130. Howieson D, Lezak M. The Neuropsychological Evaluation. In: Yudofsky S, Hales R, eds. Textbook of Neuropsychiatry. Washington D.C.: The American Textbook Press, 1992:127-50.
131. Howieson D, Lezak M. Separating memory from other Cognitive Problems. In: Baddeley A, Wilson B, Watts F, eds. Handbook of Memory Disorders. Chichester: John Wiley & Sons Ltd., 1995:411-26.
132. Gronwall D. Paced auditory serial-addition task: a measure of recovery from concussion. Percept Mot Skills 1977;44:367-73.
133. Wechsler D. Wechsler adult intelligence scale-revised. New York: The Psychological Corporation, 1981.
134. Baddeley A. Working memory. Oxford: Oxford University Press, 1986.

135. Harvey P, Fossati P, Pochon J, et al. Cognitive control and brain resources in major depression: An fMRI study using the *n*-back task. Neuroimage 2005;26:860-9.
136. Landrø N, Rund B, Lund A, et al. Honig's model of working memory and brain activation: an fMRI study. Neuroreport 2001;12:4047-54.
137. Goldman-Rakic P, Friedman H. The Circuitry of Working Memory Revealed by Anatomy and Metabolic Imaging. In: Levin H, Eisenberg H, Benton A, eds. Frontal Lobe Function and Dysfunction. New York: Oxford University Press, 1991:72-91.
138. Meyers J, Meyers K. Rey Complex Figure Test and Recognition Trial Professional Manual. Psychological Assessment Resources, Inc. 1995.
139. Delis D, Kramer J, Kaplan E, Ober B. California Verbal Learning Test (CVLT) Manual. New York: The Psychological Corporation 1987.
140. Bryan J, Luszcz M. Measurement of executive function: considerations for detecting adult age differences. J Clin Exp Neuropsychol 2000;22:40-55.
141. Oscar-Berman M, McNamara P, Freedman M. Delayed-Response Tasks: Parallels Between Experimental Ablation Studies and Findings in Patients with Frontal Lesions. In: Levin H, Eisenberg H, Benton A, eds. Frontal Lobe Function and Dysfunction. Oxford: Oxford University Press, 1991:230-55.
142. Shallice T, Burgess P. Higher-Order Cognitive Impairments and Frontal Lobe Lesions in Man. In: Levin H, Eisenberg H, Benton A, eds. Frontal Lobe Function and Dysfunction. Oxford: Oxford University Press, 1991:125-38.
143. Pennington B, Ozonoff S. Executive Functions and Developmental Psychopathology. J Child Psychol Psychiatry 1996;37:51-87.
144. Heaton R, Chelune G, Talley J, Kay G, Curtiss G. Wisconsin Card Sorting Test. Manual: Psychological Assessment Resources Inc., 1993.
145. Benton A, Hamsher K. Multilingual aphasia examination. Iowa: AJA Associates, 1989.
146. Mitrushina M, Boone K, D'Elia L. Handbook of normative data for neuropsychological assessment. New York: Oxford University Press, 1999.
147. Miller E. California Computerized Assessment Package Manual. Los Angeles, 2001.
148. Fisher D, Duffy S, Katsikopoulos K. Cognitive slowing among older adults: what kind and how much? In: Perfect T, Maylor E, eds. Models of Cognitive Aging. New York: Oxford University Press, 2000:87-124.
149. Parkin A, Java R. Determinants of age-related memory loss. In: Perfect T, Maylor E, eds. Models of Cognitive Aging. New York: Oxford University Press, 2000:188-203.
150. Parry A, Matthews P. Functional magnetic resonance imaging (fMRI): A "window" into the brain. Oxford: University of Oxford, 2002:1-22 (www.fmrib.ox.ac.uk/fmri_intro/fmri_intro.htm).
151. Hugdahl, K, Rund B, Lund A, et al. Brain Activation Measured With fMRI During a Mental Arithmetic Task in Schizophrenia and Major Depression. Am J Psychiatry 2004;161:286-93.
152. Lyness S, Eaton E, Schneider L. Cognitive Performance in Older and Middle-Aged Depressed Outpatients and Controls. J Gerontol 1994;49:P129-36.
153. Mortensen E, Sorensen H, Jensen H, Reinisch J, Mednick S. IQ and mental disorder in young men. Br J Psychiatry 2005;187:407-15.

154. Malmstrom T, Wolinsky F, Andresen E, Miller J, Miller D. Cognitive Ability and Physical Performance in Middle-Aged African Americans. J Am Geriatr Soc 2005;53:997-1001.
155. Portella M, Marcos T, Rami L, Navarro V, Gasto C, Salamero M. Residual cognitive impairment in late-life depression after a 12-month period follow-up. Int J Geriatr Psychiatry 2003;18:571-6.
156. Abas M, Sahakian B, Levy R. Neuropsychological deficits and CT scan changes in elderly depressives. Psychol Med 1990;20:507-20.
157. Amado-Boccara I, Gougoulis N, Poirier Littre M, Galinowski A, Loo H. Effects of Antidepressants on Cognitive Functions: A Review. Neurosci Biobehav Reviews 1992;19:479-93.
158. Beblo T, Baumann B, Bogerts B, Wallesch C, Herrmann M. Neuropsychological Correlates of Major Depression: A Short-term Follow-up. Cogn Neuropsychiatry 1999;4:333-41.
159. Deuschle M, Kniest A, Niemann H, et al. Impaired Declarative Memory in Depressed Patients Is Slow To Recover: Clinical Experience. Pharmacopsychiatry 2004;37:147-51.
160. Bulbena A, Berrios G. Cognitive Function in the Affective Disorders: A prospective study. Psychopathology 1993;26:6-12.
161. Dozois J, Dobson K. A Longitudinal Investigation of Information Processing and Cognitive Organization in Clinical Depression: Stability of Schematic Interconnectedness. J Consult Clin Psychol 2001;69:914-25.
162. Ercoli L, Heaton R. Neuropsychological impairment in unipolar depression: Nature, severity and stability of deficits in outpatients. Diss Abstr Int-B 1996;57:4026.
163. Hammar A, Lund A, Hugdahl K. Long-lasting cognitive impairment in unipolar depression: A six months follow-up study. Psychiatry Res 2003;118:189-96.
164. Paradiso S, Lamberty G, Garvey M, Robinson R. Cognitive Impairment in the Euthymic Phase of Chronic Unipolar Depression. J Nerv Ment Dis 1997;185:748-54.
165. Kuny S, Stassen H. Cognitive performance in Patients recovering from Depression. Psychopathology 1995;28:190-207.
166. Biringer E. Functional impairment recovers after episodes of major depression. Evid Based Ment Health 2005;8:65.
167. Buist-Bouwman M, Ormel J, de Graaf R, Vollebergh W. Functioning after a major depressive episode: complete or incomplete recovery? J Affect Disord 2004;82:363-71.
168. Burt T, Prudic J, Peyser S, Clark J, Sackeim J. Learning and Memory in Bipolar and Unipolar Major Depression: Effects of Aging. Neuropsychiatry, Neuropsychol Behav Neurol 2000;13:246-53.
169. Davidson R, Irwin W, Anderle M, Kalin N. The Neural Substrates of Affective Processing in Depressed Patients Treated With Venlafaxine. Am J Psychiatry 2003;160:64-75.
170. Rose E, Simonotto E, Ebmeier K. Limbic over-activity in depression during preserved performance on the *n*-back task. Neuroimage 2005;29:203-15.
171. Baxter L, Schwartz J, Phelps M, et al. Reduction of prefrontal cortex glucose metabolism common to three types of depression. Arch Gen Psychiatry 1989;46:243-50.

172. Smith K, Morris J, Friston K, Cowen P, Dolan R. Brain mechanisms associated with depressive relapse and associated cognitive impairment following acute tryptophan depletion. Br J Psychiatry 1999;174:525-9.
173. Dolan R, Bench C, Brown R, Scott L, Frackowiak R. Neuropsychological dysfunction in depression: the relationship to regional cerebral blod flow. Psychol Med 1994:849-57.
174. Ravnkilde B, Videbech P, Clemmensen K, et al. The Danish PET/depression project: cognitive function and regional cerebral blood flow. Acta Psychiatr Scand 2003;108:32-40.
175. Sheehan D, Lecrubier Y, Harnett-Sheehan K, et al. Reliability and Validity of the MINI International Neuropsychiatric Interview (M.I.N.I.): According to the SCID-P and its reliability. Eur Psychiatry 1997;12:232-41.
176. Kupfer D, First M, Regier D. A research agenda for DSM-V. Washington D.C.: APA, 2002.
177. Kendrick D. Kendrick cognitive tests for the elderly: NFER-NELSON Publishing Company Ltd., 1985.
178. Altman D. Practical statistics for medical research. London: Chapman & Hall, 1991.
179. Fischer E, Dornelas E, Goethe J. Characteristics of People Lost to Attrition in Psychiatric Follow-up Studies. J Nerv Ment Dis 2001;189:49-55.
180. Collie A, Shafiq-Antonacci R, Maruff P, Tyler P, Currie J. Norms and the effects of demographic variables on a neuropsychological battery for use in healthy ageing Australian populations. Aust N Z J Psychiatry 1999;33:568-75.
181. Rossell S, David A. Improving performance on the WCST: variations on the original procedure. Schizophr Res 1997;28:63-76.
182. Seidman L, Peppleet J, Faraone S, et al. Wisconsin Card Sorting Test performance over time in schizophrenia Preliminary evidence from clinical follow-up and neuroleptic reduction studies. Schizophr Res 1991;5:233-42.
183. Dikmen S, Heaton R, Grant I, Temkin N. Test-retest reliability and practice effects of the Expanded Halstead-Reitan Neuropsychological Test Battery. J Int Neuropsychol Soc 1998;5:346-56.
184. Paolo A, Tröster A, Ryan J. Test-retest stability of the California verbal learning test in older persons. Neuropsychology 1997;11:613-16.
185. Basso M, Bornstein R, Lang J. Practice effects on Commonly Used Measures of Executive Function Across Twelve Months. Clin Neuropsychol 1999;13:283-92.
186. Thase M. Neuroimaging Profiles and the Differential Therapies of Depression. Arch Gen Psychiatry 2001;58:651-3.
187. Specht K, Willmes K, Shah N, Jancke L. Assessment of Reliability in Functional Imaging Studies. J Magn Reson Imaging 2003;17:463-71.
188. Green S, Salkind N, Akey T. Using SPSS for Windows. New Jersey: Prentice-Hall International, 2000.
189. Jorm A. History of depression as a risk factor for dementia: an updated review. Aust N Z J Psychiatry 2001;31:776-81.
190. Tesar N, Bandion K, Baumhackl U. Efficacy of a neuropsychological training programme for patients with multiple sclerosis - a randomised controlled trial. Klin Wochenschr 2005;117:21-22.
191. Lopez-Luengo B, Vazquez C. Effects of a neuropsychological rehabilitation programme on schizophrenic patients' subjective perception of improvement. Neuropsychol Rehabil 2005;15:605-18.

192. Kern R, Green M, Mintz J, Liberman R. Does 'errorless learning' compensate for neurocognitive impairments in the work rehabilitation of persons with schizophrenia? Psychol Med 2003;33:433-42.

12. Tables and figures

Table 1. Neurocognitive tests, neurocognitive dimensions measured, and literature references

Neurocognitive test	Neurocognitive dimension(s)	Reference
Paced Auditory Serial Addition Test (1)	Attention, psychomotor speed, working memory	Gronwall (1977)
WAIS-R Digit Span (2)	Attention, working memory	Wechsler (1981)
Wisconsin Card Sorting Test (3)	Executive function, concept formation, set-shifting	Heaton (1993)
California Verbal Learning Test (4)	Verbal memory function, conceptual organisation	Delis (1987)
Rey Complex Figure Test (5)	Visual memory function (recognition, recall)	Meyers and Meyers (1995)
Stroop Colour and Word Test (6)	Attention, psychomotor speed, inhibition	Mitrushina (1999)
California Computerized Assessment Package (7)	Attention, psychomotor speed, reaction time	Miller (2001)
Controlled Oral Word Association Test (8)	Verbal fluency, psychomotor speed, attention	Benton and Hamsher (1989)
WAIS-R Picture Completion (2)	General intellectual abilities	Wechsler (1981)
WAIS-R Similarities (2)	General intellectual abilities	Wechsler (1981)
WAIS-R Digit Symbol Test (2)	Attention, psychomotor speed, visuomotor coordination	Wechsler (1981)
Kendrick's Object Learning Test (9)	Visual memory function (object recall)	Kendrick (1985)

References Table 1.

1. Gronwall D. Paced auditory serial-addition task: a measure of recovery from concussion. Percept Mot Skills 1977;44:367-73.
2. Wechsler D. Wechsler adult intelligence scale-revised. New York: The Psychological Corporation, 1981.
3. Heaton R, Chelune G, Talley J, Kay G, Curtiss G. Wisconsin Card Sorting Test. Manual: Psychological Assessment Resources Inc., 1993.
4. Delis D, Kramer J, Kaplan E, Ober B. California Verbal Learning Test (CVLT) Manual. New York: The Psychological Corporation 1987.
5. Meyers J, Meyers K. Rey Complex Figure Test and Recognition Trial Professional Manual. Psychological Assessment Resources, Inc. 1995.
6. Mitrushina M, Boone K, D'Elia L. Handbook of normative data for neuropsychological assessment. New York: Oxford University Press, 1999.
7. Miller E. California Computerized Assessment Package Manual. Los Angeles, 2001.
8. Benton A, Hamsher K. Multilingual aphasia examination. Iowa: AJA Associates, 1989.
9. Kendrick D. Kendrick cognitive tests for the elderly: NFER-NELSON Publishing Company Ltd., 1985.

Table 2. Overview of findings with regard to neurocognitive recovery after unipolar depression. Cross-sectional studies comparing remitted or recovered depressed patients with healthy controls

Author year	Sample	Patients' age years mean (SD)	Duration of illness mean (SD)	Symptom status patient group mean (SD)	Medication	Main findings	Conclusion with regard to neurocognitive recovery/ comment
Weiland -Fiedler 2004 (1)	28 rMDD 23 C	37.8 (12.2)	3.7 (2.4) previous episodes	MADRS 2.1 (2.3)	No	rMDD<C: A, S, EF rMDD=C: VeM	Recovery not complete
Jaracz 2002 (2)	21 rMDD 17 C	46.3 (8.7)	10.8 (8.0) years	HAM-D≤7	Yes	rMDD<C: EF rMDD=C: EF, VF	Recovery not complete Females only
Tham 1997 (3)	10 rMDD	48.0 (19.0)	16.3 (12.1) years		8/10	30 to 56% of rMDD performed in impaired range	Patients' T-scores compared to T-scores from general population
Paradiso 1997 (4)	20 rMDD 19 C	55.9 (11.3)	7.5 (5.1) years	HAM-D 9.2 (4.8)	Yes	rMDD<C: A, VeM, S, EF	Recovery not complete Males only
Marcos 1994 (5)	28 rMDD 19 C	54.3 (9.9)	Long		22/28	rMDD<C: ViM, M, construction rMDD=C: A	Recovery not complete Melancholic sub-type

rMDD: remitted or recovered major depression; C: healthy controls

A: attention; S: psychomotor speed; M: memory function; VeM: verbal memory function; ViM: Visual memory function; VF: Verbal fluency; EF: executive function

HAM-D: Hamilton Depression Rating Scale

References Table 2.

1. Weiland-Fiedler P, Erickson K, Waldeck T, et al. Evidence for continuing neuropsychological impairments in depression. J Affect Disord 2004;82:253-8.
2. Jaracz J, Borkowska A, Chlopocka-Wozniak M, Rybakowski J. Cognitive functions in remitted unipolar female depressive patients during maintenance treatment with antidepressants. Arch Psychiatry Psychother 2002;4:15-23.
3. Tham A, Engelbrektson K, Mathe´ A, Johnson L, Olsson E, Aberg-Wistedt A. Impaired Neuropsychological Performance in Euthymic Patients With Recurring Mood Disorders. J Clin Psychiatry 1997;58:26-48.
4. Paradiso S, Lamberty G, Garvey M, Robinson R. Cognitive Impairment in the Euthymic Phase of Chronic Unipolar Depression. J Nerv Ment Dis 1997;185:748-54.
5. Marcos T, Salamero M, Gutierrez F, Catalan R, Gasto C, Lazaro L. Cognitive dysfunctions in recovered melancholic patients. J Affect Disord 1994;32:133-7.

Table 3. Overview of sample characteristics, duration of observation, and results in longitudinal studies reporting improved or normalised neurocognitive function upon remission of unipolar depression

Author year	Sample	Patients' age years mean (SD)	Duration of illness mean (SD)	Mean period of observation (SD)	Symptom status in remitted/recovered at re-test mean (SD)	Medication	Main findings	Comment
Deuschle 2004 (1)	24 MDD (3 sub-groups)	51.1 (14.2) 45.2 (18.4) 58.1 (11.8)	4.2 14.2 13.2 years	24.5 (7.2) months	HAM-D 1.3 (2.4)	Yes	Significant within group improvement on VeM	
Dozois 2001 (2)	45 MDD	39.8 (11.0)	rMDD 5.3 (4.0) nrMDD 6.0 (4.1) previous episodes	rMDD 5.6 (1.3) nrMDD 5.9 (1.3) months	BDI-II 9.1 (7.5)	35/45	rMDD improvement: A, M, S Time x depression-status interaction A, M, S	23 rMDD/ 22 nrMDD at re-test Emotional processing, females only
Williams 2000 (3)	25 MDD 27 C	31 (11.0)		10 weeks	BDI≈ 12	No	rMDD=C: A	Multiple assessments
Beblo 1999 (4)	27 MDD	Median 56 (range 32-65)	Median 2 episodes (range 1-30)	4.5 weeks (range 4-13)	Median BDI=14 (range 0-40) rMDD BDI ≤ 17	Yes	rMDD within-group improvement: M, VF, EF, construction	
Beats 1996 (5)	24 MDD 15 C	72 (5.9)	3.4 episodes (range 1-10)		HAM-D 4.7 (2.6) MADRS 6.5 (4.5)	21/24	rMDD=C: A, M rMDD<C: S	

Ercoli 1996 (6)	50 MDD 50 C			4 months		Not at inclusion	rMDD=C: A, M, S	
Tarbuck 1995 (7)	37 MDD	Younger 41.0 (11.1) Older 69.8 (7.0)		≈ 260 days	HAM-D Younger 2.7 (2.4) Older 4.4 (2.4)	Yes	Within-group improvement: A, M, S, VF	34 re-tested, Two sub-groups: younger (N=18) older (N=19)
Bulbena 1993 (8)	24 MDD (9 BP) 17 mania 30 C	50 (16)	5 (8) previous hospitalisations	3-12 months			Within-group improvement: A, M	17 previous ECT Clinical data refer to total patient sample

MDD: major depression; rMDD: remitted or recovered major depression; nrMDD: non-remitted major depression; C: healthy controls; BP: bipolar

A: attention; M: memory; VeM: verbal memory function; S: psychomotor speed; VF: Verbal fluency, EF=executive function

HAM-D: Hamilton Depression Rating Scale; MADRS: Montgomery Aasberg Depression Rating Scale;

BDI: Beck Depression Inventory

ECT=electroconvulsive treatment

References Table 3.

1. Deuschle M, Kniest A, Niemann H, et al. Impaired Declarative Memory in Depressed Patients Is Slow To Recover: Clinical Experience. Pharmacopsychiatry 2004;37:147-51.
2. Dozois J, Dobson K. A Longitudinal Investigation of Information Processing and Cognitive Organization in Clinical Depression: Stability of Schematic Interconnectedness. J Consult Clin Psychol 2001;69:914-25.
3. Williams R, Hagerty B, Cimpric B, Therrien B, Bay E, Hiroake O. Changes in directed attention and short-term memory in depression. J Psychiatr Res 2000;34:227-238.
4. Beblo T, Baumann B, Bogerts B, Wallesch C, Herrmann M. Neuropsychological Correlates of Major Depression: A Short-term Follow-up. Cogn Neuropsychiatry 1999;4:333-41.
5. Beats B, Sahakian B, Levy R. Cognitive performance in tests sensitive to frontal lobe dysfunction in the elderly depressed. Psychol Med 1996;26:591-603.
6. Ercoli L, Heaton R. Neuropsychological impairment in unipolar depression: Nature, severity and stability of deficits in outpatients. Diss Abstr Int-B 1996;57:4026.
7. Tarbuck A, Paykel E. Effects of major depression on the cognitive function of younger and older subjects. Psychol Med 1995;25:285-96.
8. Bulbena A, Berrios G. Cognitive Function in the Affective Disorders: A prospective study. Psychopathology 1993;26:6-12.

Table 4. Overview of sample characteristics, duration of observation, and results in longitudinal studies reporting rest-deficits in neurocognitive function upon remission of unipolar depression

Author year	Sample	Patients' age years mean (SD)	Duration of illness (SD)	Mean period of observation (SD)	Symptom status in remitted/recovered at re-test mean (SD)	Medication	Main findings	Comment
Neu 2005 (1)	27 MDD 34 C	53.4 (10.8)	Median 5 episodes	9.0 (1.9) months	HAM-D 5.1 (1.9)	Yes	rMDD<C: VeM, VF	MRI scan included
Hammar 2003 (2)	21 MDD 20 C	42 (10.0)		6 months	HAM-D 8.5 (5.5)	20/21	rMDD<C: A, S No group x time interaction: A, S	
Portella 2003 (3)	30 MDD 15 C	72.1 (5.9)	Late onset	12 months	rMDD HAM-D $\leq$ 8	Yes	No group x time interaction between rMDD and nrMDD: A, construction, EF	Two sub-groups at follow-up: rMDD (N=21) nrMDD (N=9)
Neu 2001 (4)	27 MDD 62 C	49.8 (8.8)	Median 3 episodes	3.2 (1.2) months	BRMS 5.0 (2.3)	2/3 of sample	No correlation change BRMS versus change in test performance Within-group improvement: VeM, VF	Also supporting "state"-hypothesis
Nebes 2000 (5)	20 MDD 19 C	70.8 (7.1)		12 weeks	HAM-D 6.2 (3.0)	Yes	No group x time interaction: A, S	

Reischies 2000 (6)	32 MDD 62 C	54.4 (11.7)		4.4 (1.9) months	BRMS 4.6	14/32	No within-group improvement: A, VeM, VF Improvement: ViM No group x time interactions	Several additional diagnostic groups included
Trichard 1995 (7)	23 MDD (UP/BP) 15 C	47 (14.0)		29 (10) days	MADRS 7.0 (4.0)	18/23	rMDD <C: A, S, EF; rMDD=C: VF Correlation change VF versus change MADRS Within-subject improvement: EF	3 assessments, results from second assessment (N=19) referred. Improvement also found
Kuny 1995 (8)	30 MDD 30 C	49.7 (16.0)	12.3 (10.3) years	≥ 30 days	Improvement HAM-D ≥ 50%	Yes	rMDD< C: A	Various depressive sub-types Eleven patients left at last of seven assessments
Abas 1990 (9)	20 MDD (16 UP/ 4 BP) 20 C	70.4 (6.1)	>1 previous episode	2 years	MADRS 6.2 (4.3)	6/20	rMDD<C: A, M, S	CT scan included

MDD: major depression; rMDD: remitted or recovered major depression; nrMDD: non-remitted major depression; C: healthy controls; UP: unipolar; BP: bipolar

A: attention; M: memory; VeM: verbal memory function; ViM: Visual memory function; S: psychomotor speed; VF: Verbal fluency; EF: executive function

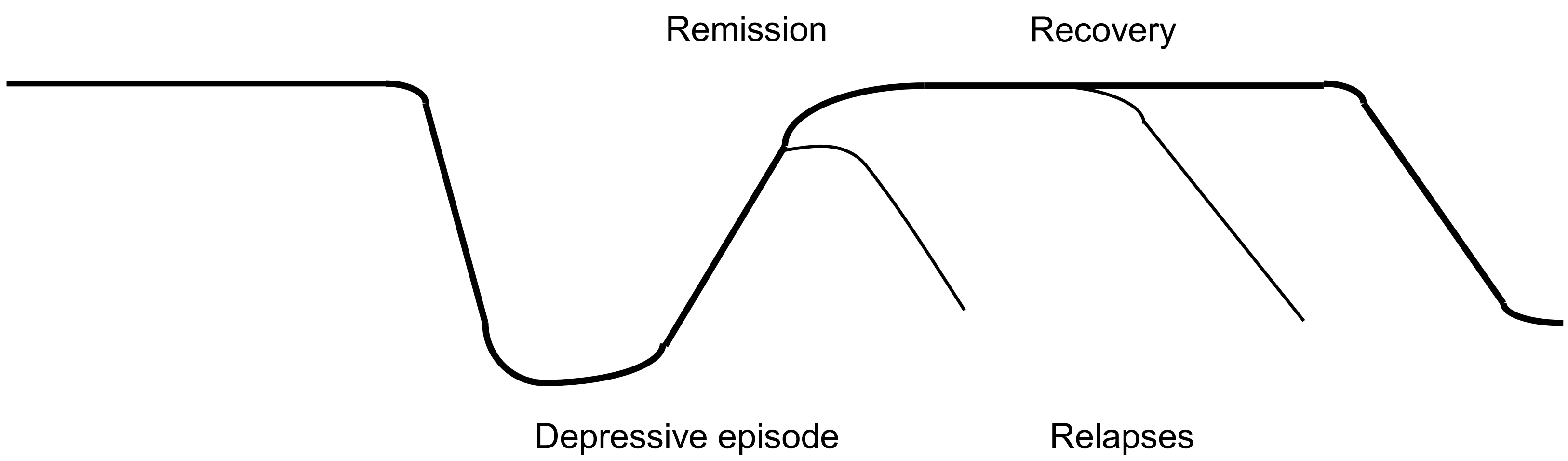
HAM-D: Hamilton Depression Rating Scale; MADRS; BDI: BRMS: Bech-Rafaelsen Melancholia Scale

MRI: magnetic resonance imaging; CT: computerised tomography

References Table 4.

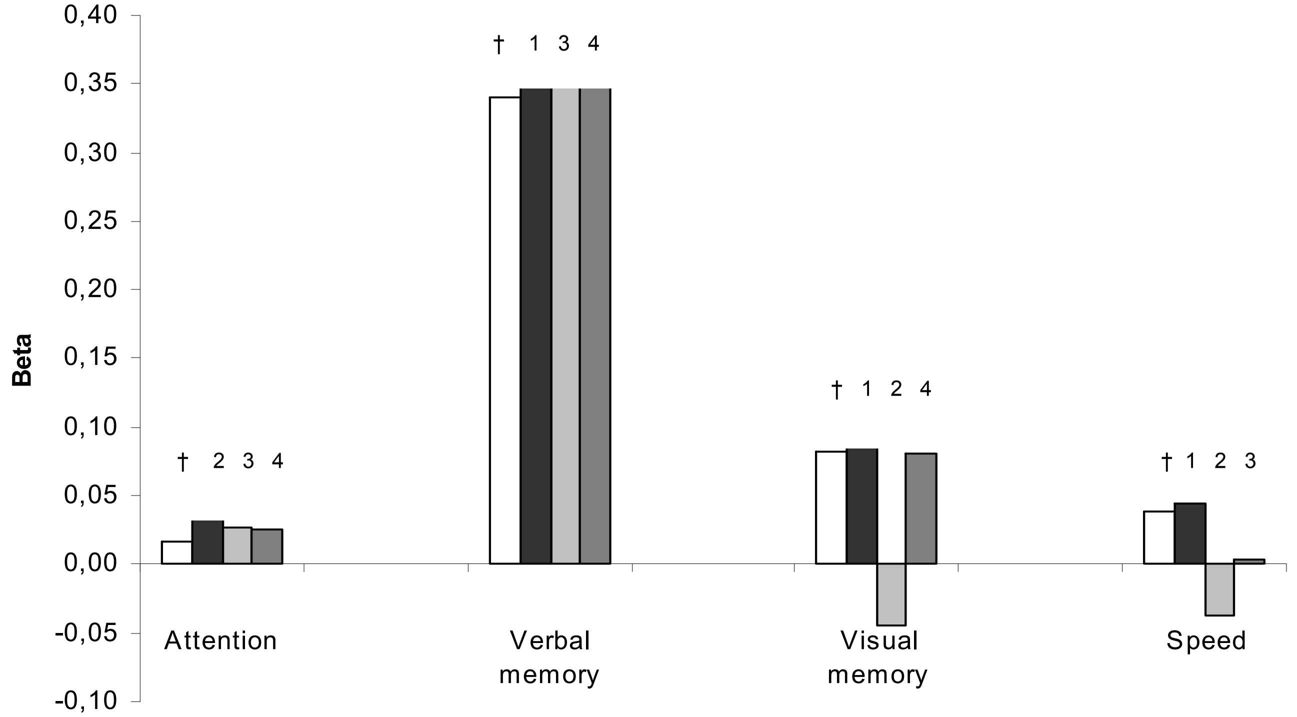
1. Neu P, Bajbuij M, Schilling A, Godemann F, Berman R, Schlattmann P. Cognitive function over the treatment course of depression in middle-aged patients: correlation with brain MRI signal hyperintensities. J Psychiatr Res 2005;39:129-35.
2. Hammar A, Lund A, Hugdahl K. Long-lasting cognitive impairment in unipolar depression: A six months follow-up study. Psychiatry Res 2003;118:189-96.
3. Portella M, Marcos T, Rami L, Navarro V, Gasto C, Salamero M. Residual cognitive impairment in late-life depression after a 12-month period follow-up. Int J Geriatr Psychiatry 2003;18:571-6.
4. Neu P, Kiesslinger U, Schlattmann P, Reischies F. Time-related cognitive deficiency in four different types of depression. Psychiatry Res 2001;103:237-47.
5. Nebes R, Butters M, Mulsant B, et al. Decreased working memory and processing speed mediate cognitive impairment in geriatric depression. Psychol Med 2000;30:679-91.
6. Reischies F, Neu P. Comorbidity of mild cognitive disorder and depression-a neuropsychological analysis. Eur Arch Psychiatry Clin Neurosci 2000;4:186-93.
7. Trichard C, Martinot J, Alagille M, et al. Time course of prefrontal lobe dysfunction in severely depressed in-patients: a longitudinal neuropsychological study. Psychol Med 1995;25:79-85.
8. Kuny S, Stassen H. Cognitive performance in Patients recovering from Depression. Psychopathology 1995;28:190-207.
9. Abas M, Sahakian B, Levy R. Neuropsychological deficits and CT scan changes in elderly depressives. Psychol Med 1990;20:507-20.

Figure 1. The course of major unipolar depression: Depressive episode, remission, recovery, and relapses



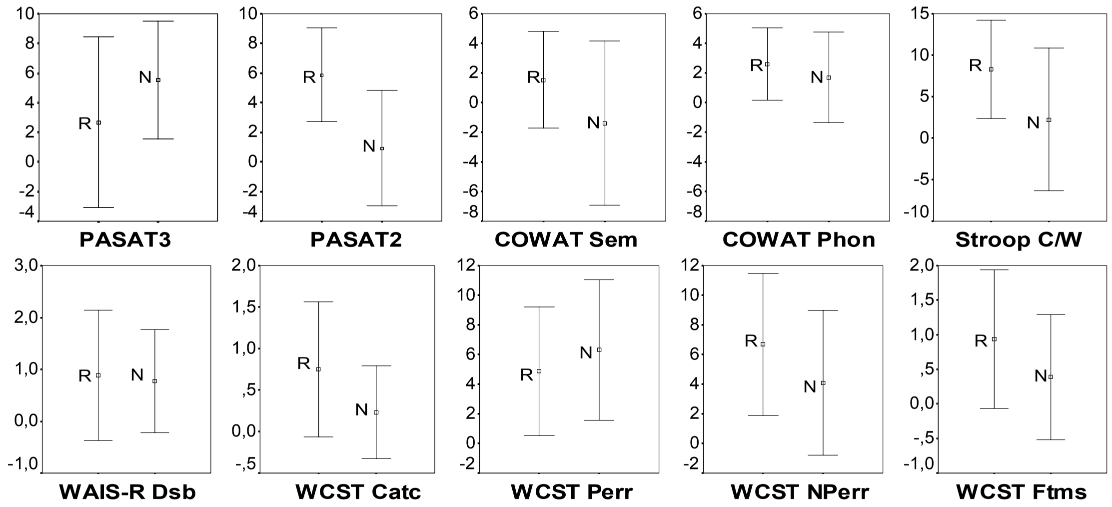
Modified from "Depresjonshåndboka"
(Gyldendal Akademisk 1999)

Figure 2. Relative contributions (partial effects) of the neuropsychological dimensions attention, verbal memory, visual memory, and psychomotor speed as shown by standardised effect sizes (betas) for the association between improvement in HAM-D and summary scales of change from baseline to follow-up



† Crude analysis
1 Adjusted for attention memory summary scale
2 Adjusted for verbal memory summary scale
3 Adjusted for visual memory summary scale
4 Adjusted for psychomotor speed summary scale

Figure 3. Means and 95% confidence intervals of raw scores of change between T1 and T2 for the ten EF measures for the Recovered (R) and the Nonrecovered (N) groups respectively



Abbreviations Figure 3

EF	Executive function
PASAT2:	Paced Auditory Serial Addition Test, 2 seconds sub-task
PASAT3:	Paced Auditory Serial Addition Test, 3 seconds sub-task
COWAT Sem:	Controlled Oral Word Association Test, semantic fluency sub-task
COWAT Phon	Controlled Oral Word Association Test, phonological fluency sub-task
Stroop C/W:	Stroop Colour and Word Test, color-word sub-task
WAIS-R Dsb:	Digit Symbol Test from Wechsler's Adult Intelligence Scale-Revised, digit span backward sub-task
WCST Catc:	Wisconsin Card Sorting Test, categories completed test variable
WCST Perr:	Wisconsin Card Sorting Test, perseverative errors test variable
WCST NPerr:	Wisconsin Card Sorting Test, non-perseverative errors test variable
WCST Ftms:	Wisconsin Card Sorting Test, failure to maintain set test variable

13. Errata

In section 6.4.

Page 27: ”In this epidemiological study, (a cohort of elderly non-demented patients

(aged 72-74 years))..” is replaced with “In this epidemiological study, in a cohort of elderly non-demented patients (aged 72-74 years)..”

In section 8.1.

Page 29: “A medium-sized correlation was found between higher degree of depressive symptoms and lower levels of general functioning…” is replaced with “A medium-sized correlation was found between higher psychomotor speed and higher levels of general functioning..”

In section 9.1.2.1

Page 39: “..the high risk of committing Type II error when multiple comparisons are performed” is replaced with “..the high risk of committing Type I error when multiple comparisons are performed.”

In section 9.1.4

Page 43: “In Papers I and II, significant group differences…” is replaced with “In Papers I and II, no significant group differences…”.

In section 9.2.3

Page 46: “..in the correlations of change in depression with change in neurocognitive test performances, such measurement errors could not be excluded. Here such errors could occur both at baseline, at re-testing, and for HAM-D as well as for the neurocognitive measures. Measurement errors are likely to be random, and most likely resulting in under-estimation of the strength of the associations between improvement in depression and improvement in neurocognitive performance.” is replaced with “..in the correlations of depression or anxiety with neurocognitive test performances, such measurement errors could not be excluded. In these analyses, measurement errors could occur both for HADS as well as for the neurocognitive test measures. Measurement errors are likely to be random, and most likely resulting in under-estimation of the strength of the associations between depression or anxiety and neurocognitive test performance.”