Dose-escalation in the picture: pharmacological and imaging studies in depression
Ruhé, H.G.

Citation for published version (APA):
Dose-escalation in the picture
Pharmacological and Imaging studies in depression
DOSE-ESCALATION IN THE PICTURE

PHARMACOLOGICAL AND IMAGING STUDIES IN DEPRESSION

Eric Ruhé
The work described in this thesis was mostly performed at the Academic Medical Center, Program for Mood Disorders in the Netherlands. The studies were conducted in close collaboration with the departments of Nuclear Medicine, General Practice, Pharmacology and Pharmacotherapy, Clinical Epidemiology, Biostatistics and Bioinformatics and Radiology.

Most of the studies in this thesis were financially supported by grants from the Academic Medical Center (SFA.07.012), the Netherlands Organisation for Health Research and Development (ZonMw), program Mental Health, education of investigators in mental health (OOG; #100-002-002), and from the Dutch Brain Foundation (14F06.45).

Additional support for the completion of this thesis was provided by Amicale Facel Holland.
Dose-Escalation in the Picture
Pharmacological and Imaging studies in depression

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingesteldecommissie,
in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 5 november 2008, te 10:00 uur

doort

Henricus Gerardus Ruhé
geboren te Amsterdam
Promotiecommissie

Promotor: Prof. dr. A.H. Schene

Co-promotor: Dr. J. Booij

Overige leden: Prof. dr. D.A.J.P. Denys
               Prof. dr. R. Dierckx
               Prof. dr. B.L.F. van Eck-Smit
               Prof. dr. C.G. Kruse
               Prof. dr. W.A. Nolen

Faculteit der Geneeskunde
‘The brain is not a chemical factory but an extremely complicated survival machine’

Arvid Carlsson, Nobel lecture 2000
Contents

Part I. General introduction 9
   1. General introduction and outline of the thesis 11
   2. Design and methods of the DELPHI-study 31

Part II. Studies to guide clinical treatment of Major Depressive Disorder 41
   3. Clinical use of the Hamilton Depression Rating Scale: Is increased efficiency possible? A post hoc comparison of HDRS, Maier and Bech subscales, CGI and SCL-90 scores. 43
      *Comprehensive Psychiatry* 2005; 46: 417-427
   4. Dose-escalation for insufficient response to a standard dose of a selective serotonin reuptake inhibitor in major depressive disorder: a systematic review. 59
      *British Journal of Psychiatry* 2006; 189: 309-316
   5. Switching antidepressants after a first selective serotonin reuptake inhibitor in major depressive disorder: a systematic review 73
      *Journal of Clinical Psychiatry* 2006; 67: 1836-1855

Part III. Serotonin in the etiology of Major Depressive Disorder 99
   6. Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: a meta-analysis of monoamine depletion studies 101
      *Molecular Psychiatry* 2007; 12: 331-359
   7. Serotonin transporter binding with $[^{123}I]β$-CIT SPECT in major depressive disorder versus controls: effect of season and gender. 139
      Submitted
Part IV. Neurobiological effects of paroxetine in the treatment of Major Depressive Disorder

8. Serotonin transporter gene promoter polymorphisms modify the association between paroxetine serotonin transporter occupancy and clinical response in major depressive disorder
   Pharmacogenetics and Genomics 2008; in press

9. Evidence why paroxetine dose-escalation is not effective in major depressive disorder: a randomized-controlled trial with assessment of serotonin transporter occupancy
   Neuropsychopharmacology 2008; in press

10. Successful pharmacological treatment of major depressive disorder attenuates amygdala activation to negative facial expressions. An fMRI study
    Submitted

11. Effect of the selective serotonin reuptake inhibitor paroxetine on platelet function is modified by a SLC6A4 serotonin transporter polymorphism
    Journal of Thrombosis and Haemostasis 2008; in press

Part V. Summary and general discussion

12. Summary, conclusions and general discussion

13. Samenvatting en conclusies

Part VI. Abbreviations, color figures, publications, acknowledgement and curriculum vitae

Abbreviations

Color figures

Publications

Dankwoord

Curriculum Vitae
GENERAL INTRODUCTION
General introduction

The present thesis concerns the pharmacological treatment of Major Depressive Disorder (MDD). It especially focuses on what to do when patients do not respond to a standard dose of an antidepressant.

In this thesis, the results of studies that were performed since 2001, while working at the Program for Mood Disorders of the department of Psychiatry of the Academic Medical Center will be reported. These studies first aimed at systematically reviewing the existing literature about treatment strategies for non-response to the first antidepressant (included in Part II). These reviews identified an important gap between equivocal evidence for dose-escalation and the firm recommendation of this strategy in clinical guidelines. This was the starting point for the experimental studies to further investigate the clinical efficacy and the mechanism behind dose-escalation, described in Part IV. Meanwhile, the findings, and further reading of conflicting results of others raised questions about the pathophysiological origin of MDD. Therefore, the generally known serotonin hypothesis for MDD, will be addressed in two additional studies (Part III). Part V of the thesis comprises the general discussion and will address directions for future research.

In this introduction chapter, the reader will be introduced to some backgrounds of MDD, its pharmacological treatment, the monoamine hypothesis about the etiology of MDD, and briefly on the mechanisms of action of antidepressant drugs. After a summary, the separate studies of the thesis will be introduced. In chapter 2, the methods of the Dose-Escalation Legitimate? Pharmacology and Imaging studies in depression (DELPHI-study) will be described. However, let’s start with some history.

MELANCHOLY, the subject of our present discourse, is either in disposition, or habit. In disposition is that transitory melancholy which comes and goes upon every small occasion of sorrow, need, sickness, trouble, fear, grief, passion or perturbation of the mind, any manner of care, discontent or thought, which causeth anguish, dullness, heaviness and vexation of spirit, any ways opposite to pleasure, mirth, joy, delight, causing forwardness in us, or a dislike. In which equivocal and improper sense, we call him melancholy, that is dull, sad, sowl, lumpish, ill disposed, solitary, any way moved or displeased. And from these melancholy dispositions, no man living is free. [...] Melancholy, in this sense, is the character of mortality. [...] It falleth out oftentimes that these dispositions become habits, and many affects contemned make a disease. [...] for that which is but a flea-biting to one, causeth unsufferable torment to another; and which one by his singular moderation and well composed carriage can happily overcome, a second is no whit able to sustain; but upon every small occasion of mis-conceived abuse, injury, grief, disgrace, loss, cross, rumour, etc. (if solitary, or idle) yields so far to passion, that his complexion is altered, his digestion is hindered, his sleep gone, his spirits obscured, and his heart heavy, his hypocondries mis-affected; wind, crudity, on a sudden overtake him, and he himself overcome with melancholy. [...] If any discontent seise upon a patient, in an instant, all other perturbations will set upon him; and then, like a lame dog or broken-winged goose, he droops, and pines away, and is brought at last to that ill habit or malady of melancholy it self. [...] But all these melancholy fits, [...] displeasing, violent and tyrannizing over those whom they seise on for the time—yet these fits, I say, or men affected, are but improperly so called, because they continue not, but come and go, as by some objects they are moved. This melancholy, of which we are to treat, is an habit, morbus sonomicus, or chronicus, a chronick or continue disease, a settled humour, not errant, but fixed; and as it was long increasing, so, now being (pleasant or painful) grown to an habit, it will hardly be removed. [...] INVETERATE melancholy, howsoever it may seem to be a continuante, inexorable disease, hard to be cured, accompanying them to their graves most part, yet many times it may be helped, even that which is most violent, or at least it may be mitigated and much eased. Nil desperandum. It may be hard to cure, but not impossible for him that is, most grieviously affected, if he be but willing to be helped.”
Burton’s work *The anatomy of melancholy* (first published in 1621) was a best-seller. It describes many aspects of the disease that we currently recognize as depression: the sometimes difficult distinction of the illness melancholy (depression) from ‘normal’ sadness, pain or sorrow after misery in life, the relation between stressful life-events and the development of the illness, the interaction of life-events and coping styles to develop the illness or remain well, the clinical symptoms, the enormous impact of the illness on a depressed person’s life, social dysfunction and the tendency of recurrence and chronicity.

Nowadays, MDD is one of the most prevalent and disabling illnesses in psychiatry, after ischemic heart disease the second most common cause of disability worldwide and expected to be the world’s second cause of disability by 2030, behind HIV/AIDS. MDD is diagnosed by the occurrence of a cluster of different symptoms affecting mood, pleasure, attention, activities, vital somatic functions (eating, sleep), and (ruminative) thoughts over oneself, guilt or suicide over a prolonged period of time. For clinical and scientific uniformity, MDD is currently defined by criteria specified in the Diagnostic and Statistical Manual (4th edition; DSM-IV), which is a non-theoretical classification-system.

**Epidemiology of MDD**

The 12-month prevalence of MDD is estimated to be 5.8% in the Netherlands, and 5.3% in the United States. Prevalence rates are two times higher in women, and higher in widowed, divorced, unemployed or disabled people, with somatic disease and first degree relatives of patients with MDD. In the general Dutch population, the median duration of MDD is 3 months, with 63% of new MDD episodes recovering within 6 months (regardless whether treatment is offered or the natural course of the disease is awaited), however, this prognosis is less favourable for people who seek help. Estimations of the lifetime prevalence are 15.4% in the Netherlands, >16% in Europe, and 13.2%-16.2% in the USA, with the same 2:1 female-male distribution. Most epidemiological studies indicate that nowadays 51-69% of depressed patients seek or receive treatment, which has increased over the last decade relative to the nineties of the previous century. However, only 50% of patients who are treated for MDD meet diagnostic criteria of MDD or meet indication criteria for treatment, and it is estimated that only 21.7% of MDD-patients in the USA receive adequate treatment. Therefore, depending on one’s view, MDD is frequently over-treated, under-treated or mis-treated.

MDD is a recurrent and potentially chronic disease. After the first episode, 50-80% of patients have a recurrence within 10 years. Of incident episodes of MDD 15-20% will become a chronic depression (defined as a duration of ≥2 years), although higher percentages have been reported. MDD is associated with high direct treatment costs as well as indirect costs of loss of productivity and quality of life. In the Netherlands, treatment costs of MDD (€250 million) in 1994 were estimated to comprise 1% of the total healthcare budget, and 13.6% of the total mental health budget. In 2003, MDD treatment costs in the Netherlands had increased to €660 million, still consuming 1.1% of the total healthcare budget. In 1999 10% of newly assigned disability-payments (WAO) in the Netherlands was due to MDD.

**Treatment of MDD**

In order to treat major depressive disorder effectively, many national clinical guidelines were developed. Besides various forms of psychotherapy, pharmacotherapy is often applied as monotherapy or in combination with psychotherapy. Antidepressants are effective drugs for MDD when compared to placebo, although effect sizes are moderate, and some critics claim that effects are exaggerated by pharmaceutical companies. Antidepressants are prescribed for MDD, but also for other psychiatric disorders, and are of considerable financial interest. In 2006, approximately 780,000 inhabitants (4.7%) of the Netherlands used

---

* Melancholy is derived from the Greek Μελαγχολια (melancholia), which refers to the black choler (bile) that was seen as the material cause of the illness those days.
antidepressants, mostly Selective Serotonin Reuptake Inhibitors (SSRIs). Annual prescriptions of antidepressants increased over the last decade by ~6% every year. The costs of prescribed antidepressants were €156 million in 2006, again, most costs represent the prescription of SSRIs.

A range of antidepressants have been developed, all aiming at the stimulation of serotonergic and/or noradrenergic neurotransmission. Four classes of antidepressants exist: SSRIs, Tricyclic antidepressants (TCAs), Monoamine-oxidase inhibitors (MAOIs) and a miscellaneous group. TCAs and SSRIs, but also ‘dual action’ antidepressants block the transporters at the pre-synaptic nerve ending, which causes inhibition of the reuptake of either serotonin (5-HT), noradrenaline (NA = norepinephrine), or both. The classic MAOIs irreversibly inhibit the enzymes mono-amine oxidase A and B, which degradeserotonin, noradrenaline and dopamine. The miscellaneous group harbours ‘dual action’ antidepressants (e.g. venlafaxine or duloxetine), and antidepressants with pre- and post-synaptic targets different from serotonin or noradrenaline transporters.

By nature of their largely similar, but also slightly distinct pharmacological affinities for transporters and receptors, antidepressants have (almost) comparable efficacy, but are different in their adverse effects. Because SSRIs are tolerated better than TCAs, SSRIs have become the antidepressants of first choice in most countries. Within their class, different SSRIs also have comparable efficacy, but different side effect profiles.

Achieving remission is the aim of treatment, because persistence of (residual) symptoms increases the risk of a future relapse or recurrence and is associated with ongoing psychosocial dysfunction. If remission is achieved, patients are advised to continue their antidepressants (at least 6 months) in order to prevent relapse, although guidelines’ recommendations are inconsistent. Nevertheless, only 50% of the patients respond (defined as a 50% decrease in depression severity) to the initial antidepressant, and remission rates (defined as a depression severity score below a certain cut-off) are lower (28-35%). Therefore, clinicians who treat MDD are often confronted with non-response, for which evidence-based approaches are necessary.

Insufficient response to antidepressants: A challenge for the clinician.

The clinician faces several dilemmas when a patient shows insufficient response to an antidepressant. First, how does one ascertain clinical improvement? A variety of measurement scales is available to measure depression severity and improvement of therapy. Psychopathology rating scales with clear cut-off scores provide a more direct representation of the effects of treatment than quality of life scales or a functional outcome. Thus, the use of a symptom-rating scale is highly recommended to assess treatment effects, either clinician-rated (e.g. the Hamilton Depression Rating Scale (HDRS)), Montgomery Åsberg Depression Rating Scale (MADRS) or self-rated (Inventory for Depressive Symptoms (IDS-SR)), or Beck Depression Inventory (BDI)). Despite criticism on internal validity, selection and rating of items, and sensitivity to change, most scales correlate with each other. Frank et al. proposed tentative definitions for response (≥50% decrease of pretreatment rating scale scores) and remission (e.g. HDRS ≤7, MADRS ≤7, IDS-SR ≤13, BDI ≤9). Therefore, measurements should be routinely performed before the start of treatment and at consecutive critical decision points. However, scales are not routinely implemented in daily practice. Apparently, the current challenge of implementation is to persuade clinicians to measure MDD severity repeatedly with appropriate scales. The next dilemma is when to measure the effects of the antidepressants. Although a recent meta-analysis indicated that the effects of antidepressants on depressive symptoms can be observed within the first weeks of treatment, patients may need 6-10 weeks to achieve substantial improvement. Therefore, the initiated treatment should not be changed too early, neither too late. Changing a regimen early may forestall a potentially late effect of the current treatment, but may also offer an increased chance of response. Contrary, unchanged treatment may increase despair due to the failure of response, but may prevent the (possible premature)
start of a more aggressive treatment with increased side effects. Over the last decades, the time
to declare that an antidepressant trial failed, increased from 3 weeks\textsuperscript{71} to 6-8 weeks\textsuperscript{70,72}
Interestingly, recommendations on this lag-time were never derived from studies primarily
designed to quantify the timing of the determination of response. Current guidelines for
depression differentially recommend a minimal duration of antidepressant therapy between 4
and 6 weeks\textsuperscript{13,21,23,25,26,28,29,67,73} At these time points (‘critical decision points’\textsuperscript{67}) a next treatment-
strategy must be considered.

A final dilemma is to adequately re-evaluate the diagnosis, without the (counter-) transference
of helplessness. Critical decision points are important for the reassessment of somatic or
psychiatric co-morbidity, re-evaluation of persistent psychosocial problems and treatment
adherence. However, when none of these issues apply, the patient may be accused of ‘a failure to
cooperate’, or blamed for the non-response by ‘having a personality disorder’ or ‘not being
motivated for treatment’. Although these reproaches occasionally may be justified, they could
also indicate the patient’s and/or clinician’s helplessness, which is often transferred from a
depressed person to the persons surrounding him or her. Most depressed patients want to
improve, but are indeed disabled in their coping-styles, their motivation and their problem solving
capacities. As such, routinely administered behavioural interventions (registration, planning and
gradual reactivation)\textsuperscript{74} are valuable in the pharmacological treatment of MDD, also deferring
increased helplessness.

**Insufficient response to antidepressants: What to do when the miracle doesn’t happen?**

Five pharmacological strategies for insufficient response to an antidepressant can be distin
guished: 1. prolongation, 2. dose-escalation, 3. switching, 4. augmentation and 5. combina
tion, all regarding the antidepressant initially given. Of course, a sixth strategy could be the addition
of psychotherapy to the monotherapy of pharmacotherapy\textsuperscript{75,76} but as this is not a topic of this
thesis, only the first 5 strategies are described.

1. **Prolongation of the trial for another 2-4 weeks.** This strategy is applicable only in case of doubt
   of the nature of the non-response (e.g. amendable psychosocial stressors, non-adherence),
   but is not favourable given the relative importance assigned to critical decision points.

2. **Dose-escalation.** With this strategy the dose of the prescribed antidepressant is increased to
   the maximal tolerable dose. Dose-escalation assumes a dose-response relationship, which is
equivocal for SSRIs. The advantage of dose-escalation is that it is easy and quick to apply.

3. **Switching of the antidepressant.** With this strategy the prescribed antidepressant is
   discontinued (eventually after tapering and a wash-out period) and the antidepressant is
   switched to another antidepressant either belonging to the same class, or to a different class
   compared to the initial antidepressant. The obvious advantage of switching is that it facilitates
   a new pharmacological approach, possibly targeting different neurotransmitter systems. Its
   disadvantage is that it requires some time, and that one may unwontedly discard achieved
   improvement when the chosen antidepressant appears to be a less effective one for this
   patient.

4. **Augmentation of the prescribed antidepressant.** With this strategy a drug is added that has no
   strong antidepressant effects by itself, but is known to increase the effects of antidepressants
   (e.g. lithium). The advantage of augmentation is that the effect of the antidepressant initially
   prescribed is maintained. Furthermore augmentation might result in faster responses than
   switching (as tapering is not required, nor a wash-out). Disadvantage is the risk of poly-
   pharmacy and the related interactions.

5. **Combination of antidepressants.** With this strategy another antidepressant drug, mostly with
   different pharmacological properties compared to the antidepressant already prescribed, is
   added to the initial antidepressant. This strategy has the same advantages and disadvantages
   as augmentation strategies, but the risks of severe interactions are higher.
Neurotransmission, the serotonergic system and neuronal networks

The brain passes information through electrical signals in neurons. Synapses form the functional contacts between neurons and mainly communicate through neurotransmitter secretion. Neurotransmitters are stored in synaptic vesicles. The vesicles are released into the synaptic cleft in response to presynaptic depolarisation. After the release of neurotransmitters in the synapse, they are either metabolized or transported back into the terminal to be used again.\textsuperscript{77}

Neurotransmitters bind to specific receptor proteins on the membrane of the pre- and postsynaptic neurons. Postsynaptically, receptors either induce a change in postsynaptic ion channels (ionotropic receptors), causing a depolarisation, or they activate G-protein mediated complexes (metabotropic receptors), which activate one or more metabolic steps (‘second messengers’). Activation of metabotropic receptors is generally responsible for forming enzymes that regulate gene expression, neurotransmitter synthesis, receptors and neuroplasticity. Which of the mechanisms is used depends on the particular postsynaptic receptor type.\textsuperscript{78} Three major ‘monoamine’ neurotransmitters are associated with psychiatric disorders: serotonin, noradrenaline and dopamine, of which the latter is traditionally related to psychotic disorders. SSRIs target the serotonergic system and increase serotonergic neurotransmission. From here, the serotonergic system and the neuronal networks that are believed to be involved in MDD will be described.\textsuperscript{79}

Serotonin

The serotonin neurons of the brain start in the raphe nuclei in the midbrain and project to the neocortex, basal ganglia, temporolimbic zones, hypothalamus, cerebellum and the brain stem (Figure 1.1).\textsuperscript{80,81} Serotonin is involved in several functions: sleep and wakefulness, appetite, nausea, migraine, headaches and regulation of mood.\textsuperscript{77} The serotonin receptors comprise of 5-HT\textsubscript{1A}, 1B, 1C, 1D, 1E, 1F, 5-HT\textsubscript{2A/2C}, 5-HT\textsubscript{4}, 5-HT\textsubscript{5A}, 5-HT\textsubscript{6} and 5-HT\textsubscript{7} (metabotropic) and the 5HT\textsubscript{3} (ionotropic) receptors.\textsuperscript{78} Serotonin plays an important role in brain development via regulation of neurite outgrowth, synaptogenesis and cell survival.\textsuperscript{82,83} Serotonin that is released into the synaptic cleft is either taken up back into the presynaptic nerve ending by the serotonin transporter (SERT), or degraded by MAO-A.\textsuperscript{77,78}

Figure 1.1. Serotonergic pathways through the human brain.

Serotonergic system. From the raphe nuclei in the midbrain, neurons project to the neocortex, basal ganglia, temporolimbic zones, hypothalamus, cerebellum and the brain stem.

The serotonin (and noradrenaline) deficiency hypothesis

In the late fifties of the previous century, TCAs appeared to be effective in treating MDD by increasing serotonergic and noradrenergic neurotransmission. This discovery led to the monoamine hypothesis: MDD might etiologically be explained by a deficiency in monoamine neurotransmitters: serotonin or noradrenaline. Since then, the working mechanism of AD is believed to be by (1) increased neurotransmission by increased synaptic levels of serotonin, noradrenaline and/or (2) specific agonistic effects on serotonin or noradrenaline (sub-)receptors.
Depletion of the available serotonin and noradrenaline is used as a model to test the involvement of monoaminergic systems in MDD. Serotonin depletion can be achieved by rapidly lowering the essential amino-acid tryptophan which cannot be synthesized by the body and must be ingested to enable formation of serotonin. To achieve depletion, a tryptophan free amino-acid mixture is administered (acute tryptophan depletion).\(^84\) Depletion of noradrenaline and dopamine occurs simultaneously, and uses the same concept (acute depletion of the essential amino-acids phenylalanine and tyrosine).\(^85\) As an alternative to induce a state of depletion, enzyme-blocking agents decrease the production of the monoamines. Para-chlorophenylalanine blocks serotonin synthesis,\(^86\) and Alpha-methyl-para-tyrosine blocks noradrenaline and dopamine synthesis.\(^87\)

Since 1975 an increasing number of depletion studies have been conducted, with different effects in different study-populations. In general in healthy controls no clear mood-effects were found, unless they had relatives with MDD. In remitted MDD patients who used antidepressants (or shortly after tapering) approximately 50% of the patients experienced a relapse after depletion. However this occurred only after depletion of the monoamine that their antidepressant targeted. In depressed patients no consistent deterioration of the mood effects were found.\(^88\)-\(^93\)

Thus, the monoamine-deficiency theory, in its purest form, states that depression can be cured by the increase of serotonergic and/or noradrenergic neurotransmission. However, the reverse train of thought, that depression is bio-etiologically caused by a deficiency of monoamines (e.g. serotonin and/or noradrenaline) has attractive face-validity, but probably is an untenable, superficial simplification.\(^75\);\(^94\) Therefore, the current, less pertinent view is that the monoamine hypothesis only partially explains MDD and the response to AD.\(^95\)-\(^98\)

The limbic-cortical dysregulation hypothesis

In a more multidimensional, systems-level model, MDD can be viewed as a disorder affecting discrete but functionally integrated pathways; neural networks, which can be identified by neuroimaging techniques.\(^1\) In such a network, dysfunction in one or more of the elements (e.g. after cognitive or somatic stress), will initially be tried to be influenced (or compensated) by other, remaining parts of the network, that try to maintain homeostatic emotional control. Therefore, results from neuroimaging studies investigating differences between healthy controls and MDD patients, must be considered as the identification of regions to be either etiologically abnormal or regions involved in (mal-)adaptive compensatory processes.\(^99\)

Because MDD is an affective disorder, the neurobiology of emotion processes is likely involved. For the processing of emotions two systems are important: a ventral system (consisting of amygdala, insula, ventral striatum, ventral anterior cingulate gyrus, and ventral prefrontal cortex) and a dorsal system (consisting of hippocampus, dorsal anterior cingulate gyrus, and dorsal prefrontal cortex). The ventral system serves to identify the emotional significance of a stimulus, the production of mood states, and automatic regulation of emotional responses, while the dorsal system serves to effortfully regulate mood states and subsequent behaviour.\(^100\)

Initial lesion-deficit studies, early Positron Emission Tomography (PET) studies (measuring regional resting state glucose metabolism or blood flow) and later functional Magnetic Resonance Imaging (fMRI)-studies identified several brain regions to be affected by MDD: the limbic structures (amygdala, hippocampus, hypothalamus and brainstem), the subcortical (basal ganglia and thalamus) and the cortical (dorsolateral prefrontal, ventrolateral prefrontal and orbitofrontal cortex (DLPFC, VLPFC, OFC respectively) structures. In these, the most consistent finding is a hypoactivity of the (dorsal) frontal lobe, while often a hyperactivity in the (ventral) VLPFCand OFC is found.\(^99\);\(^101\) Additionally, an increased activity of the (rostral, subgenual) anterior cingulated gyrus\(^99\);\(^101\) and the amygdala, anterior insula, and ventral striatum was found, although less consistently.\(^101\) Furthermore, fMRI studies point to a increased sensitivity of the amygdala for

\(^1\) Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) use a variety of radioligands with various half-lives to quantify different targets (transporters, receptors, or blood-flow or metabolism). Functional Magnetic Resonance Imaging (fMRI) is a technique to image brain activity (hemodynamic response) related to a specific task or stimulus.
emotional stimuli, and a bias to interpret these stimuli in a negative context. Interestingly, induction of sad mood in healthy volunteers also produces increased blood flow in the insula, subgenual anterior cingulate gyrus and decreased blood flow in the dorsomedial prefrontal cortex (DMPFC).

The above abnormalities were combined in the limbic-cortical dysregulation model, which includes a dorsal neocortical hypofunction, which results in ventral (para)limbic hyperactivity, with a reciprocity in this inverse relationship.

Which effects in the brain cause antidepressants to be antidepressants?
The various treatment forms available to relieve MDD and their moderate efficacy poses the question: How do antidepressants bring about their therapeutic effects? For this question, three levels of action must be distinguished: 1. direct neurotransmission effects, 2. second messenger effects, and 3. change in neuronal networks. For levels 1. and 2. see Figure 1.2. At all three levels, there appears to be a time dependent differentiation of the effects that occur, e.g. not all effects are the same in the consecutive weeks after the initiation of treatment. This may explain why it sometimes takes several weeks before the therapeutic effects become apparent. Because most of the recent research was done with SSRIs and SNRIs, the effects of these drugs are described hereafter, unless indicated differently.

Direct effects of antidepressants on neurotransmission
When an SSRI is ingested, the blockade of the target transporter (i.e. the serotonin transporter (SERT)) occurs within several minutes. Therefore, the antidepressant effects cannot only be based on increased neurotransmission by decreased serotonin reuptake, as these effects take much longer than this direct pharmacological effect. Further research revealed that after two days of treatment, the firing-rate of SERT-containing neurons in rats decreased, but that this firing-rate was restored within 2 weeks of continued treatment. This was attributed to somatodendritic 5-HT1A autoreceptors, which normally have a negative feedback on the neuron’s firing rate. These 5-HT1A autoreceptors appeared to desensitize. Microdialysis-experiments in rats showed that the restored firing-rate after 14 days was responsible for a 6-fold increase in intrasynaptic serotonin level, while after the acute blockade of the SERT this increase of the serotonin level was only small and transient. Furthermore, in humans, serotonin autoreceptors (5-HT1B/1D) (and also noradrenergic α2 autoreceptors) in the synapse normally inhibit serotonin release by feedback-mechanisms as well. Prolonged treatment with SSRIs again desensitize these receptors, resulting in increased serotonin release in the synaptic cleft. As such, desensitisation of HT1A autoreceptors in the raphe nuclei in the midbrain may have effects on serotonergic neurotransmission in critical brain areas where these serotonergic neurons project to. Finally, in rats, after prolonged administration of SSRIs the SERT itself is downregulated by 80-90%. This is probably caused by trafficking and internalization of the SERTs instead of altered SERT-gene-regulation, because mRNA expression in the cells studiedwas unaltered by the treatment.

Other antidepressants have rather different effects. TCAs (except clomipramine) do not change the pre-synaptic serotonin-containing neurons, but appear to sensitize the postsynaptic 5-HT receptors for serotonin. Along with serotonin, the responsiveness for noradrenaline was also found to be enhanced, likely due to enhanced α-adrenoreceptor-mediated transmission. MAOIs also increase serotonergic transmission by desensitisation of 5-HT1A autoreceptors, but do not desensitize other 5-HT autoreceptors, and desensitize noradrenergic α2 autoreceptors, which indirectly enhances serotonergic transmission.
Figure 1.2. Transporters, receptors and second messenger systems involved in the effects of antidepressants.

Figure adapted from Belmaker and Agam. The left half of the presynaptic neuron represents a serotonergic neuron, the right half a norepinephrinergic neuron. For color figure see page 277.

In the presynaptic neuron, serotonin is synthesized from tryptophan by tryptophan hydroxylase and stored in vesicles. Likewise, norepinephrine is synthesized from tyrosine by tyrosine hydroxylase. These vesicles merge with the cell membrane when the neuron is depolarized, thereby releasing their contents into the synaptic cleft.

After release, serotonin and norepinephrine are transported back into the presynaptic neuron by serotonin and norepinephrine transporters. Furthermore, serotonin and norepinephrine are catabolized by the monoamine-oxidase A (MAO-A) enzyme. In the synaptic cleft, serotonin and norepinephrine affect both the pre- and post-synaptic neuron. The presynaptic 5-HT1A and 5-HT1B auto-receptors decrease serotonin release by inhibitory feedback; the $\alpha_2$-adrenergic receptor does the same for the release of norepinephrine.

Post-synaptically, serotonin and norepinephrine bind to G-protein-coupled monoamine receptors (MARS): the cyclicAMP (cAMP)-coupled receptor, which activates protein kinase A (PKA), and the Phosphatidylinositol (PI)-coupled receptor, which activates phospholipase C (PLC) which thereafter form inositol triphosphate (IP$_3$) and diacylglycerol (DAG). IP$_3$ and DAG activate protein kinase C (PKC). Both PKA and PKC finally activate cAMP responsive element binding (CREB) protein, which stimulates DNA transcription. For example, this might result in the production of brain derived neurotrophic factor (BDNF).
Second messengers effects of antidepressants

The secondary effects of antidepressants can be divided in post-synaptic effects on the sensitivity and availability of receptors, and the effects caused by the activation of post-synaptic receptors. Effects on post-synaptic receptors include a decrease of 5-HT\textsubscript{1A} and 5-HT\textsubscript{2A} receptor density,\textsuperscript{113} and 5-HT\textsubscript{1D},\textsuperscript{114} and 5-HT\textsubscript{2C/B} responsivity.\textsuperscript{113,115} Because these effects are either investigated in animals or by using indirect measures (e.g., growth-hormone release), it is important that Meyer et al. only demonstrated 10% in-vivo 5-HT\textsubscript{2A} receptor downregulation in MDD patients between 20-30 years, but not in older patients.\textsuperscript{116} The desensitisation of post-synaptic receptors might explain the tolerance to adverse effects that occurs after some days of exposure to antidepressants.\textsuperscript{109} Interestingly, in rodents, the activation of post-synaptic 5-HT\textsubscript{1A} receptors in the hippocampus stimulates neurogenesis, which appears to be required for therapeutic effects of antidepressants.\textsuperscript{117}

After the activation of the G-protein metabotropic receptors (either by serotonin or noradrenaline), second messenger systems mediate signals to the neuron’s nucleus, where cAMP responsive element binding protein (CREB) regulates CREB-directed gene transcription.\textsuperscript{79,118} Most evidence indicates that CREB is upregulated by chronic antidepressant use, but not for all antidepressants, and results of studies appear to be biased by the cell type investigated, the brain region where these cells originate from and the timing after the first antidepressant exposure.\textsuperscript{118} One of the genes that is (positively) influenced by CREB is the brain derived neurotrophic factor (BDNF) gene.\textsuperscript{119} BDNF is a plethoric growth factor, which regulates neuronal survival, migration, differentiation, axonal and dendritic growth and synapse formation. The genomic structure of BDNF is complex, which facilitates differential activation by diverse and variable stimuli, which can be different in different brain regions and even in different parts of the cell.\textsuperscript{83}

Changes in neuronal networks

At the neuro-anatomical level, the novel neuroimaging techniques have ‘opened’ the brain to study the in-vivo effects of psychiatric treatment.\textsuperscript{104} In MDD, early PET studies found time dependent changes during the treatment with fluoxetine.\textsuperscript{120} After 1 week, glucose metabolism increased in the hippocampus and brainstem, and decreased in posterior cingulate gyrus, striatum and thalamus compared to the pre-treatment scan. After 6 weeks, patients who responded to the treatment had a decrease in metabolism in the subgenual cingulate gyrus, the hippocampus, the pallidum and insula and an increase in the anterior and posterior cingulate gyrus, prefrontal and parietal cortex. However, the non-responders showed a persistent, unchanged pattern of change as seen in week 1. The changes in prefrontal cortex and subgenual cingulate gyrus correlated best with symptomatic improvement.\textsuperscript{99,101,120} These findings were replicated in patients treated with paroxetine.\textsuperscript{121} Moreover, patients who responded while on placebo-treatment showed somehow similar but also distinct\textsuperscript{‡} changes in brain metabolism compared to fluoxetine responders.\textsuperscript{99,122} This indicates that in neuroimaging studies, response and treatment effects may coincide, but both may also have their specific, distinguishable effects.

In recent fMRI studies, the increased activation of the amygdala to negative (sad, fearful, angry) and happy faces have been investigated after treatment of SSRIs (fluoxetine,\textsuperscript{123,124} sertraline\textsuperscript{125}), venlafaxine\textsuperscript{126,127} and bupropion.\textsuperscript{118} These studies rather consistently found decreased activation of the amygdala, insula and increased cortical activity after treatment. However, most patients in these studies were treatment responders at the end of the study. In contrast with previous PET-studies, none of the fMRI studies used a placebo-comparison measured twice, so no distinction between specific drug and response effects in fMRI can be made yet.

\textsuperscript{‡} Similar changes included increased metabolism in frontal, parietal and posterior cingulate, and decreased metabolism in subgenual cingulate gyrus. Distinct changes included no changes in subcortical brainstem, hippocampal, and caudate metabolism in placebo-responders.
Chapter 1

(Pharmaco-)genetic effects relevant for antidepressants

Several genetic polymorphisms have been investigated in relation to treatment response to SSRIs. The polymorphism studied most is the SERT promoter gene (5-HTTLPR), for which a long (L) and a short (S) variant were identified, with a recently discovered functional tri-allelic variant (rs25531). The 5-HTTLPR is associated with the transcriptional activity of the SERT gene. Cells homozygous for the L-allele produce higher concentrations of SERT mRNA, and the rate of serotonin uptake by the transporter is >2-fold higher than in cells containing one or two copies of the S-allele. A meta-analysis of 15 studies showed a pooled association between the 5-HTTLPR-polymorphism and SSRI efficacy, with MDD patients with at least one L-allele having higher response rates to SSRIs. However, in a large sample of patients treated with citalopram (an SSRI), treatment response was not associated with the tri-allelic 5-HTTLPR-polymorphism. Furthermore, individuals carrying the S-allele experience increased adverse events after SSRI treatment, have elevated risk of depression in relation to life events, but also show increased amygdala reactivity to fearful stimuli. A large MRI-study in healthy controls showed associations of the 5-HTTLPR S/S polymorphism with unfavourable alterations in anatomy and function of the amygdala-cingulate feedback circuit. These findings strongly argue for an important role of the 5-HTTLPR-polymorphism in the development and functioning of emotional networks involved in MDD. Other pharmacogenetic associations with clinical response have been investigated, but will not be further addressed in this thesis.

Summary and questions addressed in this thesis

Major Depressive Disorder is a prevalent and disabling illness, which potentially recurs and may become a chronic disease. It is the second most common cause of disability worldwide, and has a 12-month prevalence of ±5.5%, with lifetime prevalences of 12-14% in males and 22-24% in females. MDD is associated with high direct treatment costs as well as indirect costs of loss of productivity and quality of life. Besides various forms of psychotherapy, pharmacotherapy with antidepressant drugs is often applied.

Antidepressants are grouped in four different classes: Selective serotonin reuptake inhibitors (SSRIs), tricyclic antidepressants (TCAs), monoamine oxidase inhibitors (MAOIs) and a miscellaneous group including so-called dual action antidepressants. SSRIs, dual-action antidepressants and TCAs are used most often. Most antidepressants increase serotonergic and/or noradrenergic neurotransmission. When considered more precisely, different classes of antidepressants appear to have distinct additional pre- and postsynaptic effects. Secondary to increased serotonergic and/or noradrenergic neurotransmission, complex second messenger pathways are activated, which are only partly understood. Macroscopically, presumably relevant changes in neuronal networks following antidepressant treatment have been identified. Nevertheless, which changes are required and specific for the improvement of symptoms remains an enigma. Furthermore, the effects of increased neurotransmission might be affected by the genetic make-up of patients (i.e. the 5-HTTLPR polymorphism). Genetic polymorphisms appear to be associated with treatment effects, but might also influence the development and connectivity of the neuronal networks.

Despite comparable efficacy of antidepressants, only 50% of the MDD-patients respond to the first antidepressant trial given, while fewer achieve full remission of symptoms. Therefore, insufficient response to the first antidepressant is a relevant clinical problem that challenges clinicians.

The clinician's first dilemma is how to adequately and efficiently measure the changes in symptom severity, for which the further development and implementation of short, easy to use and valid clinician rated questionnaires would improve clinical decision making. The second dilemma is to time the change of the treatment initially started: this should not be changed too early, neither too late. The third dilemma is to balance helplessness, impaired functioning by the disease and (counter-) transference by ‘blaming the victim’, i.e. reproaching the patient of the moderate efficacy of antidepressants as ‘being unmotivated’ or ‘having a personality disorder’.
When ‘the miracle doesn’t happen’, five strategies for non-response are possible options: prolongation of the initial trial, dose-escalation, switching, augmentation and combination of antidepressants. For all these options the current evidence is either equivocal or fragmentary. In order to develop evidence-based treatment algorithms for the 50% of MDD-patients who turn out to be non-responsive to the first antidepressant trial, systematic overviews of the literature are required and gaps in the evidence-base need to be addressed by new research-projects.

More fundamentally, the perceived delayed onset of symptom improvement with antidepressants, their moderated efficacy, and their extensive effects in the brain beyond an increase in serotonergic or noradrenergic neurotransmission, summons the question what is the etiopathogenetic model for MDD. Or, more specifically, what is the evidence that corroborates the corner-stone of most antidepressant’s action: the monoamine hypothesis and the involvement of the serotonergic system as a causal explanation for MDD.

Research projects underlying this thesis

This thesis incorporates three research projects that were initiated and conducted by the Program for Mood Disorders of the Academic Medical Center (AMC). One additional project was initiated by the department of vascular medicine of the AMC.

First, a guideline-project was initiated with a grant from the Academic Medical Center (No SFA.07.012). Aim of this guideline project was to develop an evidence based guideline for non-response after a first SSRI.

Second, as a result of this guideline project, two grants were obtained from the Netherlands Organization for Health Research and Development (ZonMw), program Mental Health, education of investigators in mental health (Geestkracht-OOG; projects #100-002-001 and #100-002-002). These grants were applied to initiate the DELPHI-study (Dose-Escalation Legitimate? Pharmacology and Imaging studies in depression). The DELPHI-study was set up as a methodologically sound trial to study clinical effectiveness of dose-escalation as a strategy for non-response. As a novel extension to existent dose-escalation trials, we aimed to investigate the molecular neurobiological target of dose-escalation: the occupancy of the serotonin transporter (SERT) by paroxetine (a SSRI). This was done by the acquisition of two or three single photon emission computed tomography (SPECT) scans in a subgroup of drug-free patients of the total DELPHI-cohort.

A third project was started as an extension to the DELPHI-SPECT study. This third project could be planned after a grant from the Dutch Brain Foundation (Hersenstichting) was obtained (project #14F06.45). The DELPHI-fMRI aimed to investigate the neurobiological changes in brain activation by treatment with paroxetine in a functional Magnetic Resonance Imaging (fMRI) study. This study was superimposed on the SPECT-imaging of the last 22 drug-free patients participating in DELPHI.

The final fourth research project in this thesis was initiated after an almost fatal bleeding complication, which occurred in a patient who underwent surgery, and lost 12 litres of blood. This complication was afterwards attributed to fluoxetine use. This project was funded internally by mutual contributions of both participating AMC-departments.

The questions addressed in this thesis

1 Is a short, easy to use clinician rated questionnaires as effective and precise as the routine Hamilton depression rating scale (HDRS)?

The HDRS is most frequently used as the golden-standard in clinical trials, but is probably too extensive to use in clinical practice. To facilitate clinical measurements of depression severity and objectivity for critical decision points, we reanalyzed the treatment-outcomes of two antidepressant-psychotherapy trials, which were performed by the Mentrum research group. We therefore investigated whether the effect-sizes of two 6-item subscales of the HDRS (17-items) – the Maier and Bech subscales – were comparable to the original HDRS in the measurement of depression severity and the sensitivity to measure changes. Furthermore, we investigated whether this comparability was stable across the full range of
response to treatment, and across different treatments and for different baseline severity of depression. We also determined cut-off points for remission for these subscales compared to conventional HDRS definitions.\textsuperscript{66,142,143} See chapter 3.

2 \textbf{What is the evidence for dose-escalation as a strategy for non-response to a first SSRI?}
For this question, we performed a systematic review of the evidence for the dose response relationship for SSRIs in MDD. See chapter 4.

3 \textbf{What is the evidence for switching antidepressants as a strategy for non-response to a first SSRI?}
For this question, we performed a systematic review of the evidence for switching after failure of a first SSRI in MDD. Part of this study was a meta-analysis of three switch-studies. See chapter 5.

4 \textbf{Does the depletion of monoamine (5-HT and NA/DA) systems lower mood in humans, and is this lowering of mood different across different populations?}
For this question, we performed a systematic review of monoamine depletion studies reporting mood effects of depletion. As an extension of previous systematic reviews of monoamine depletion studies,\textsuperscript{88-93} we aimed to pool the results of the small-sized depletion studies, because they might not have detected small differences by a lack of power, and pooling would quantify the balance of positive versus negative studies. Therefore, we applied a pooling technique (modified from conventional meta-analyses of randomized controlled trial data and including an adjustment for small sample bias) to handle the statistically paired cross-over designs of these studies in formal, stratified meta-analyses.\textsuperscript{144} See chapter 6.

5 \textbf{Do MDD-patients and healthy controls differ in the number of central serotonin transporters, and is the amount of available SERTs correlated with depression severity?}
Despite the fact that the working-mechanism of antidepressants supports the monoamine deficiency theory, the pathogenesis of MDD remains unclear.\textsuperscript{95} Therefore, differences in SERT availability in patients and healthy controls have been studied previously, with conflicting results. Additionally, significant effects on SERT availability have been reported for gender,\textsuperscript{145} smoking behavior,\textsuperscript{146,147} aging\textsuperscript{148} and season of scanning.\textsuperscript{148} Therefore, we analyzed the baseline SPECT-scans of the DELPHI-SPECT participants versus age and sex-matched healthy controls. Because our sample size was large, we were able to properly account for potential confounders and possible interactions, of which the multivariate effects are reported. See chapter 7.

6 \textbf{Does a common genetic polymorphism of the promoter region of the serotonin transporter gene (SLC6A4) modify the association between the SERT occupancy by paroxetine and the clinical response?}
We performed this study because SSRI-response is likely associated with 5-HTTLPR polymorphisms, 5-HTTLPR polymorphisms might influence SERT availability (the target for SSRIs), and it is unclear how occupancy of the available SERTs is related to clinical response. Thus, we aimed to investigate the paroxetine treatment by genotype interaction regarding clinical response on the molecular level of SERT occupancy. We quantified the relation between SERT occupancy and clinical response, and studied how the 5-HTTLPR -polymorphism affected this SERT occupancy-response relationship. We performed this study in the open phase of the DELPHI-SPECT study, when patients were treated with paroxetine 20 mg/day for 6 weeks. See chapter 8.

7 \textbf{Is dose-escalation of paroxetine an effective clinical strategy for non-response in MDD?}
The systematic review of dose-escalation (chapter 4), identified methodological flaws in previous dose-escalation trials. In this study we reevaluated the clinical efficacy of dose-escalation of paroxetine without these flaws, and, considering the molecular target of SSRIs, we also tested whether paroxetine dose-escalation increased SERT occupancy more than placebo dose-escalation. We therefore performed a 6 week, multicenter, randomized study in depressed patients not responding to 6 weeks of paroxetine at 20 mg/day. As a novel extension to previous clinical trials, and in order to elucidate the neurobiological basis for an
expected lack of benefit of dose-escalation, we included a SPECT imaging approach. Herewith, we quantified whether paroxetine dose-escalation increased SERT occupancy more than placebo dose-escalation. This enabled us to relate clinical findings to the neurobiological correlate of SERT occupancy. See chapter 9.

8 Does treatment with paroxetine normalize amygdala hyperactivation in MDD?
We initiated this study after the first reports of attenuated amygdala activation after treatment with sertraline. Thereafter several groups replicated a baseline hyperactivation of the amygdala but less consistently reported the attenuation of amygdala-hyperactivation after treatment. We therefore investigated whether activation of the amygdala by (negative) facial expressions differed from healthy controls, this activation of the amygdala changed after 6 and 12 weeks of treatment with paroxetine, the activation of the amygdala and other brain areas merely changed by paroxetine treatment or in relation with clinical response, and whether dose-escalation of paroxetine in week 6 non-responders affected activations, compared to placebo-dose-escalation. For this study, we performed an fMRI study in 22MDD patients who participated in the DELPHI-fMRI study. Patients were treated with paroxetine (20 mg/day followed by a randomized dose-escalation for non-responders) and were scanned at baseline, 6 weeks and 12 weeks of treatment. We obtained a baseline scan for 21 matched controls, to contrast baseline amygdala activation in MDD-patients. See chapter 10.

9 What are the changes in hemostasis and blood platelet parameters when patients are treated with paroxetine, and are these changes modified by dose-escalation or a genetic polymorphism of the promoter region of the serotonin transporter gene?
In this study, we evaluated the effects of standard and increasing dosages of paroxetine on the bleeding tendency and hemostatic functions of platelets in patients who were drug-free before the start of paroxetine. In addition, we assessed whether these effects are modified by the 5-HTTLPR polymorphism. See chapter 11.

References


47. Simon G. Choosing a first-line antidepressant: equal on average does not mean equal for everyone. JAMA. 2001; 286: 3003-3004.


142. Prien RF, Carpenter LL, Kupfer DJ. The definition and operational criteria for treatment outcome of major depressive disorder. A review of the current research literature. *Arch Gen Psychiatry*. 1991; **48**: 796-800.


CHAPTER 2

DESIGN AND METHODS OF THE DELPHI-STUDY
Chapter 2

The DELPHI-study

DELPHI is an acronym for Dose-Escalation Legitimate? Pharmacology and Imaging studies in depression. The DELPHI-study is the backbone of the studies described in Part IV of this thesis. Because the DELPHI study in fact comprised three research projects, with slightly different patient populations and different numbers of participants, this chapter will describe the design of the DELPHI-study and indicate how the nested sub-studies relate to the major DELPHI-study. Furthermore the patient disposition over these projects will be given.

Objectives, rationale and research questions

The major aims of the DELPHI-study were to investigate the clinical efficacy of and the neurobiological mechanism behind dose-escalation. In order to determine the neurobiological effects of dose-escalation, we intended to study different biological measures during treatment: paroxetine serum concentrations (PSC), serotonin transporter (SERT) occupancy, functional activations in the cortico-limbic-network, awakening cortisol changes and changes in ω-3/ω-6 poly unsaturated fatty acids (PUFAs). Furthermore, we planned to study the effects of genetic polymorphisms on outcomes (i.e. the serotonin transporter gene promoter region (5-HTTLPR)). Results related to some of these neurobiological measures are described in this thesis.

As a study drug, we chose paroxetine for three reasons. First, paroxetine is the selective serotonin reuptake inhibitor (SSRI) which is prescribed most frequently in the Netherlands. Second, paroxetine is a potent inhibitor of the cytochrome P450 2D6 sub-enzyme, which also is the enzyme that is responsible for its metabolism. Therefore paroxetine inhibits its own metabolism, which causes an exponential rise in blood serum concentrations after dose-escalation. Third, when we initiated this study, Gilmor et al. reported noradrenergic reuptake inhibition by paroxetine, suggesting that paroxetine, like venlafaxine and duloxetine, in fact was a ‘dual action’ antidepressant, especially at higher doses.

Our main research-questions were:

1. Is a 6-week true dose-escalation of paroxetine (up to 30-50 mg/day) in patients non-responsive to a 6-week trial with standard dose (20 mg/day) more effective than placebo dose-escalation?
2. Does a 6 week true dose-escalation of paroxetine increase SERT occupancy more than a placebo dose-escalation?
3. What is the relation between SERT occupancy and clinical response to paroxetine (either at a standard dose or after dose-escalation)? Is this relation modified by other neurobiological parameters (i.e. genetic polymorphisms)?
4. Which changes in the functional activation of the cortico-limbic brain regions correlate with clinical response to paroxetine? Is there a common and/or differential effect of paroxetine exposure and/or clinical response?
5. Does a 6 week true dose-escalation of paroxetine generate additional changes (e.g. in noradrenergic reuptake inhibition, cortisol awakening responses, or ω-3/ω-6 PUFAs) compared to a placebo dose-escalation, and are these changes related to clinical response?

Only questions 1-4 will be discussed in this thesis, and the methods used for these questions will briefly be discussed. However, for specific technical we refer to the method-sections in the relevant chapters.

Design and interventions

The DELPHI-study was set up as a randomized clinical trial (ISRCTN register nr. ISRCTN44111488; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=193). The study was approved by the Academic Medical Center (AMC) medical ethical committee (03/120, 03/287 and addendum to 03/120), and all participants provided written informed consent.
Between October 2003 and February 2007 patients were recruited from primary care, our AMC Program for Mood Disorders, and public psychiatric settings. Patients were treated by their referring physician or were referred to our outpatient department. In DELPHI two nested sub-studies were embedded, the DELPHI-SPECT and the DELPHI-fMRI (Figure 2.1). These sub-studies had slightly different in- and exclusion criteria (see below).

**Figure 2.1.** Design of the DELPHI-study.

![Diagram of the DELPHI-study](image)

DELPHI-SPECT and DELPHI-fMRI represent sub-studies nested within the total study. For these 2 sub-studies drug-free patients were recruited and treated in the Academic Medical Center. For color figure see page 280.

All eligible patients were treated open-label with paroxetine 20 mg/day for 6 weeks. When severe adverse effects occurred, dosages were reduced to 10 mg/day and again increased to 20 mg/day after one week. After 6 weeks of treatment, patients who responded (defined as ≥50% reduction of the pretreatment Hamilton depression rating scale (HDRS)) continued paroxetine 20 mg/day. Treatment non-responders were randomized to a true dose-escalation or a placebo dose-escalation, added to paroxetine 20 mg/day in a double blind design. Dose-escalation was provided in blue capsules containing 10 mg paroxetine or placebo. Randomization was stratified for treatment setting (SPECT/fMRI-group, outpatient department AMC, primary care, and public psychiatry), gender and age. Within strata, we applied a minimization method to achieve a balanced distribution. We concealed allocation by using an independently operated computer program.

Dose-escalation consisted of incremental steps of one capsule every 5 days towards a maximum of 50 mg/day (20 mg + 3 capsules). Patients were allowed to increase at a slower pace (e.g. by 7 days) or stop further escalation (e.g. 20 mg + 2 capsules) according to adverse effects. No dosage adjustments were allowed during the last 3 weeks of the study. We checked adherence by pill-counts and anamnesis.
Patients

Figure 2.2 summarizes patient disposition in the DELPHI-study and the DELPHI-SPECT/fMRI neuroimaging sub-study.

Figure 2.2. Patient enrolment and disposition over sub-studies.

The patients in DELPHI-SPECT and DELPHI-fMRI sub-studies are combined (left). For disposition of patients randomized see chapter 9.

General in- and exclusion criteria

Eligible patients met the following inclusion criteria: Age between 18 and 70 years, major depressive disorder (MDD) determined by the structured clinical interview for DSM-IV (SCID), and a HDRS (17 items) score above 18. All participants were either drug-free and/or had undergone no more than one antidepressant treatment (other than paroxetine) at an effective dose for ≥ 6 weeks for the present MDD-episode. By the latter criterion, we avoided treatment resistance as potential bias for inefficacy of dose-escalation. Exclusion criteria, apart from pregnancy (or wish to become pregnant), were bipolar disorder, psychotic features, neurological cognitive impairments (i.e. dementia), primary anxiety and/or substance abuse disorders and acute, severe suicidal ideation. Contrary, we allowed secondary co-morbid anxiety and/or substance abuse to increase applicability of our findings. In total, 278 patients were referred, of whom 107 patients started treatment with paroxetine 20 mg/day. Twenty patients withdrew from the study before week 6, and 27 patients were responder by then. Sixty non-responders were randomized.
DELPHI-SPECT and fMRI

Patients who were drug-free (for >4 weeks and ≥5 half-lives of a previous antidepressant) were asked to additionally participate in the neuroimaging sub-studies. Patients were initially asked to participate in DELPHI-SPECT only, but from August 2005 onwards, when DELPHI-fMRI was approved, we asked them for both DELPHI-SPECT and DELPHI fMRI. Participation in only one of the neuroimaging studies was also possible. We limited age to 25-55 years to reduce variability in SERT-measurements by age. For participation in the fMRIs, patients had to be free of metal objects in their body. None of the included patients reported past or present use of 3,4-methylenedioxymethamphetamine. We treated DELPHI-SPECT and DELPHI-fMRI patients at the outpatient department of the Program for Mood Disorders. We supplied medication in pillboxes.

We included 51 patients in the DELPHI-SPECT study, of whom 33 non-responders were randomized. For DELPHI-fMRI we included 22 patients, of whom 16 were randomized. Twenty patients participated both in the DELPHI-SPECT and DELPHI-fMRI. Unfortunately not all (repeated) scans were analyzable adequately, the reason why in chapters 8-10 different numbers of patients are described.

Healthy controls

We recruited 53 healthy controls as reference for the study-entry scans. We individually matched each patient in DELPHI-SPECT and DELPHI-fMRI by gender and age (±2.5 years). Healthy controls were in good physical health, and had never used psychotropic medication. Exclusion criteria were current or lifetime psychiatric disorder(s) according to the SCID (including abuse or addiction disorders), a Beck Depression Inventory (BDI) score >9, alcohol use >4 units per day (last month) or a 1st-degree relative with psychiatric disorder(s). We allowed healthy controls to have incidentally used illicit drugs, unless criteria for a DSM-IV disorder was met, but we prohibited illicit drug use the month prior to scanning. Twenty healthy controls also participated both in the DELPHI-SPECT and DELPHI-fMRI.

Measurements

Time points and questionnaires

We administered the HDRS$_{17}$, Inventory for depressive symptoms (IDS-SR$_{30}$), the occurrence of adverse effects and health-related quality of life (MOS-SF36) at study-entry, randomization (T0), and 6 weeks after randomization (T1). Adverse effects and depressive symptoms were also monitored in the weeks 1, 2 and 4 after the initial start of treatment, and after randomization, using the Maier and Bech subscales and IDS-SR$_{30}$ (Figure 2.1). Three trained investigators administered clinician-rated questionnaires. Agreement between raters was good (intraclass correlation coefficient = 0.98). Raters and patients were blinded for treatment.

Definition of primary and secondary outcomes

Primary clinical outcomes were HDRS$_{17}$-scores and the proportion of patients achieving response (≥50% decrease in HDRS$_{17}$) or remission (HDRS$_{17}$ ≤7). Secondary outcomes were total and specific (adverse effects / inefficacy) dropout rates, the Maier and Bech subscales and IDS-SR$_{30}$-scores, the occurrence of adverse effects and health-related quality of life.

Neurobiological measurements

Paroxetine serum concentrations: See chapters 8 and 9 for details.

SERT-gene promoter polymorphism: See chapters 8 and 11 for details.

Saliva and additional blood specimens

At study-entry, at T0, and T1, we collected two saliva specimens at awakening and 30 minutes thereafter, to determine salivary cortisol, dehydroepiandrosterone-sulphate, and α-amylase. Furthermore, we collected blood-specimens to quantify [H]Noradrenaline and [H]Serotonin uptake in ex-vivo models. Finally we collected blood-specimens to determine levels of platelet and plasma ω-3/ω-6 PUFAs. These neurobiological measurements remain to be analysed in different papers not in this thesis.
Neuroimaging

Measurement of SERT occupancy

We performed SPECT imaging at study-entry, T0 and T1 (Figure 2.1), between 2 to 10 pm according to previously described procedures. We made all scans 230 ±18 (SD) minutes after intravenous injection of approximately 100 MBq $[^{123}]$I-β-CIT, when the radioligand is at equilibrium for SERT binding in brain areas expressing high densities of SERTs. We performed SPECT imaging using a 12-detector single slice brain-dedicated scanner (Neurofocus 810, Strichmann Medical Equipment; Cleveland, OH). After attenuation correction and reconstruction in 3D mode (http://www.neurophysics.com), we defined regions of interest (RoIs) for midbrain, diencephalon (regions rich of SERT) and cerebellum (as a reference) by using validated templates (Figure 2.3). For further details of SPECT-procedures see chapters 7-10.

Figure 2.3. Regions of Interest (RoI) for Midbrain, Cerebellum and Diencephalon.

Magnetic resonance imaging of functional activation of the cortico-limbic-network

We acquired fMRI scans at study-entry, T0 and T1 (Figure 2.1), between 2 to 10 pm. FMRI-sessions lasted 50-60 minutes, each including a cognitive task (Tower of London), a structural scan and a facial expression task, reported in this thesis. We used a 3Tesla Intera MRI scanner (Philips, Eindhoven, NL), with a 6 channel head-coil for radiofrequency reception. Two magnet compatible response boxes were used to record subject’s performance and reaction times. For further details of fMRI-settings and the parameters of the faces paradigm: see chapter 10.

Power and interim analysis

For the randomization-phase of the total DELPHI-study, we performed a-priori power-calculations for two co-primary endpoints. We planned an interim analysis after SPECT data had been collected on at least 30 patients. Stopping criteria were predetermined using the O’Brien and Fleming approach, and were $p<0.0026$ in case of superiority and $p>0.50$ for futility. See chapter 9 for further details.
References

STUDIES TO GUIDE CLINICAL TREATMENT
OF MAJOR DEPRESSIVE DISORDER

PART II
CLINICAL USE OF THE HAMILTON DEPRESSION RATING SCALE: IS INCREASED EFFICIENCY POSSIBLE?
A POST HOC COMPARISON OF HDRS, MAIER AND BECH SUBSCALES, CGI AND SCL-90 SCORES.

Comprehensive Psychiatry 2005; 46: 417-427

Henricus G. Ruhé1,2
Jack J. Dekker1,3
Jaap Peen1
Rebecca Holman4
Frans de Jonghe1,2

1 Mentrum, Amsterdam, The Netherlands
2 Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
3 Department of Clinical Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
4 Department of Clinical Epidemiology and Biostatistics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
Abstract

Background
The 17-item Hamilton Depression Scale (HDRS) is used as a semi gold standard in research. In treatment guidelines, the HDRS-measurements serve to determine response, remission and guide clinical decision-making for non-responders. However, its use in clinical practice is limited, possibly because the HDRS is time-consuming. Additionally, the multidimensional HDRS is criticized for not measuring a unidimensional aspect as depression severity. The Maier and the Bech, two 6-item severity subscales extracted from the HDRS, are relatively unknown.

Aim
To investigate whether the measurements obtained with Maier and Bech subscales are comparable with the original HDRS-measurements.

Methods
Data from two randomized controlled trials in 482 male and female patients, diagnosed with a major depression (with or without dysthymia) according to DSM-III-R, of whom 219 participated in the trials, were reanalyzed. A standardized stepwise psychopharmacological treatment was compared with a combination of pharmacotherapy with Short Psychodynamic Supportive Psychotherapy in a psychiatric outpatient department. Outcome measures were internal consistency and concurrent validity of HDRS, Maier, Bech, Clinical Global Impression (CGI) scales and Symptom CheckList depression subscale. Effect sizes of HDRS, Maier and Bech were used to compare measured treatment effects for the randomized subjects participating in the trials. Item Response Theory was used to obtain conversion tables for the HDRS, Maier, Bech and SCL-90 depression subscale.

Results
We found moderate internal consistence (Cronbach $\alpha \approx 0.6-0.7$) and high correlations of the Maier and Bech subscales with overall HDRS-scores. Overall there were no clinically relevant differences in effect sizes between Maier, Bech and HDRS, although some differences were statistically significant. Receiver operating characteristic curves showed no difference between Maier and Bech to define remission, but showed the CGI ratings to be unreliable. A cut-off $\leq 4$ corresponded with a HDRS $\leq 7$ criterion in both subscales.

Conclusions
In clinical practice, both Maier and Bech scales can be used as equivalents of the HDRS, but will be more efficient.
Introduction

Major depressive disorder is a severe, disabling illness, expected to be the world’s second health-problem in 2020. Depression is associated with high costs, regarding direct treatment and indirect costs of loss of productivity and quality of life. Several clinical guidelines were developed to guide the treatment of this disorder, both psychotherapy and pharmacotherapy (or in combination) appear effective.

The use of self-report or clinician-rated symptom-scales is recommended to assess severity and response to treatment. Some experts claim clinician-rated symptom scales to have a larger validity and reliability than self-reporting scales, especially in patients with cognitive impairment, and more severe or psychotic depressions. Specific symptom-scales are more reliable than global rating scales. Especially, rating scales can be used to objectively determine specific cut-off points for response and remission.

In most clinical trials the Hamilton Depression Rating Scale (HDRS) – a clinician rated symptom-scale – is used as a standard to determine severity and response. Many versions of the HDRS exist, with the number of items usually varying between 17 and 24, however up to 36 items have been described. Longer versions were especially developed to cover reverse neurovegetative (atypical) symptoms. The Clinical Global Impression (CGI) – a clinician-rated global scale – is also frequently used. In clinical practice, although recommended, rating scales are not used routinely. Explanations for this discrepancy could be ignorance of existing scales, a strong belief in one’s clinical judgment, an unsystematic approach of depression, but also the amount of time needed for rating-scales (e.g. 15-20 minutes for the HDRS) and the necessity of training.

The HDRS is criticized as being sensitive to somatic symptoms (e.g. somatic illness or side-effects of drugs), for not rating all 9 DSM-IV domains, its unequal weightings of different symptoms and for the multidimensionality of the HDRS total score. Multidimensionality is important to cover the maximum range of clinical features of major depressive disorder, but does not necessarily measures depression severity. Multidimensional scales can be misleading when measurement of severity and treatment response is concerned, especially when the measured depressive symptoms do not change proportionally with depression severity. Finally, some reports emphasize that the HDRS systematically favors (sedative) Tricyclic Antidepressants (TCAs) above Selective Serotonin Reuptake Inhibitors (SSRIs). Sleep and somatic items may appear to be 'improved' by side-effects of TCAs, but worsened by side-effects (e.g. insomnia, gastrointestinal complaints, agitation) of SSRIs.

In order to overcome the problems of the multidimensional HDRS mentioned above, a more unidimensional subscale from the HDRS covering core-symptoms of severity is desired. Also, from a clinical point of view, fewer items will be less time consuming for application by busy clinicians. However, for the purpose of reference, subscale scores must remain anchored to the original HDRS. To identify shorter unidimensional subscales, Maier et al. used Rasch- and Mokken-analyses and Gibbons et al. used Factor-analysis. Bech and colleagues developed another 6-item subscale. This scale initially emerged from an analysis with experienced psychiatrists as a validity criterium, and was validated psychometrically thereafter using Rasch-analyses. This Bech subscale was combined with four items of the Cronholm-Ottosson Depression Scale to form the Bech-Rafaelsen Melancholia Scale. Santor and Coyne examined the score-performances of individual HDRS-items as a function of depression severity with a nonparametric Item Response Theory (IRT) approach, retaining 14 items. These 14 items included all 6 items of the Maier subscale and all 8 items of the Gibbons subscale. However one item from the Bech subscale (13, somatic symptoms) was not included.

In a meta-analysis of individual patient-data, Faries et al. evaluated the responsiveness of total HDRS and subscale scores in TCA and SSRI pharmacotherapy trials, finding a maximal sensitivity for the Maier subscale. In a similar reanalysis, Entsuah et al. found larger effect sizes for the Bech, Maier and Gibbons subscales compared to the HDRS in trials comparing SSRIs or venlafaxine. O’Sullivan et al. found comparable sensitivity to detect changes for the six-item
Bech subscale compared to the 17-item HDRS. Hooper et al. found equal sensitivity to change during treatment for the 6-item Bech subscale compared to the HDRS 17 item version. Möller and Bech et al. used the Bech subscale to reexamine treatment efficacy of SSRIs and mirtazapine (versus TCAs or placebo). The latter publications did not provide data for the Maier subscale.

In this paper we describe a secondary analysis of our trial data, in order to answer the following questions:

1) Are the Maier, Bech and HDRS comparable in the measurement of depression severity and the sensitivity to measure changes in severity? 2) Is this comparability stable across the full range of response to treatment (e.g. non-response, partial and full response), across different treatments and different baseline severity of depression? and 3) What are clinical cut-off points for the subscales to determine remission compared to conventional definitions.

We hypothesized that the differences between Maier, Bech and HDRS-scales would be small and that there would be no apparent effect modification across neither treatments nor baseline severity. In contrast, we hypothesized that for non-responders and partial responders the effect sizes would be smaller than for responders. This would additionally prove the hypothesis of sensitivity to change.

**Methods**

**Patient selection**

In the present analyses we use data from two published randomized controlled trials conducted between 1993 and 1998 which were published or accepted for publication. The first trial aimed at efficacy and effectiveness of pharmacotherapy versus the combination of pharmacotherapy with Short Psychodynamic Supportive Psychotherapy (SPSP) (16 sessions). The second trial investigated efficacy and effectiveness of a combination of pharmacotherapy with 8 versus 16 sessions of SPSP. Pharmacotherapy in both trials consisted of three successive steps in case of intolerance or inefficacy. Both trials started with fluoxetine (20 mg/day), when this was unsuccessful (CGI-I >2, only 'minimally improved' or worse) after 6 weeks amitriptyline (≥150 mg/day, dependent of plasma-levels) was initiated in trial 1 and nortriptyline (≥150 mg/day, dependent of plasma-levels) in trial 2. If again unsuccessful after 6 weeks, moclobemide (300-600 mg/day) was started in trial 1 and mirtazapine (30-45 mg/day) in trial 2.

Inclusion criteria for participation in the trials were age between 18 and 60 years, DSM-III-R defined Major Depression (with or without dysthymia) assessed in a structured clinical interview, a 17-item HDRS baseline score of at least 14 points and written informed consent. Patients were excluded in case of psycho-organic or psychotic or dissociative disorders, drug abuse, or when the patient was considered to be too unreliable to participate in a clinical trial. Other axis 1 comorbidity was not excluded. Further exclusion criteria were if there was a serious communicative or practical problem (e.g. language barrier or the patient will soon leave the country), if there was a contraindication for one of the antidepressants used, if the patient was adequately treated with antidepressants during the present depressive episode, if the patient used other psychotropic medication, or if the patient was or planned to become pregnant. Additional exclusion criteria were of the usual kind in drug research: “too ill” (e.g. antidepressants must be started immediately) and/or “too suicidal” (e.g. hospitalization is unavoidable) to participate in a clinical trial. The study was approved by the medical ethics committee. After complete description of the study to the subjects, written informed consent was obtained.

Of 3226 newly registered outpatients, 988 patients had a depressive disorder. By initial screening 503 of these 988 patients were excluded by the above exclusion-criteria leaving 485 subjects (including patients that later refused to participate or had a HDRS below 14; further referred to as the **diagnostic sample**). To enter the trials, a second exclusion check was performed by a psychiatrist (excluding 73 patients), and 142 subjects with a HDRS-17 <14 were excluded,
leaving 270 patients for randomization. After randomization 51 patients refused participation, leaving 219 patients who started the proposed therapy (further referred to as the per protocol sample).45,46

In this manuscript we used the diagnostic sample for most cross-sectional analyses, and the randomized patients in the per protocol sample for analyses of sensitivity of response-data. For non-completers, the last observation was carried forward (LOCF).

**Outcome measures**

Primary outcome-measures were the 17-item HDRS,18,19 the Maier subscale of the HDRS (containing items 1,2,7-10),28 the Bech subscale of the HDRS (items 1,2,7,8,10 and 13),37 the Clinical Global Impression Severity (CGI-S) and Improvement (CGI-I) scale,24 and the Symptoms check list, 90 items (SCL-90) depression subscale (SCL-90 dep).51,52 Thus, three levels of information were obtained: data from 1) an independent, trained, supervised and blinded research-assistant (HDRS-17 & Maier, Bech), 2) the treating clinician (CGI-S/I) and 3) the patient (SCL-90 dep). The HDRS was administered using a semi-structured interview.53 Before participating in the study, the reliability of the HDRS-assessments was established. During the study, in order to avoid slippage, audiotaped assessments were discussed monthly.

In the analyses of treatment-efficacy, response was defined as a ≥50% HDRS-score reduction, partial response as ≥20-50% reduction in HDRS-score and remission as a HDRS score of 7 points or less.16,54

**Statistics**

Cronbach-α coefficients and mean inter-item correlations were used to express internal consistency. To check whether the increased number of items in the HDRS accounted for a higher Cronbach-α coefficient than in the subscales (with only 6 items), we applied the Spearman-Brown formula.55 Next we calculated concurrent validity as Pearson correlation coefficients between total HDRS, Maier and Bech subscale scores and SCL-90 dep scores. Linear regression models calculated variance of HDRS-scores explained by the subscales.56 These analyses were performed in our diagnostic sample. Concurrent validity between CGI-S/I and HDRS subscale-ratings was determined also, however, to avoid low correlations due to limited dispersion, this was done for the last observation in the per protocol sample. The CGI improvement scale was compared with changes expressed as percentages of the baseline score.

In order to compare differences in sensitivity to measure treatment effects (also referred to as responsiveness), in data from the per protocol sample effect sizes (E-S) for HDRS, Maier and Bech subscales were calculated per subject as the within-subject changes in scale-scores divided by the pooled standard deviation of the mean change in scale score $(\frac{\text{TT}_{\text{end}} - \text{TT}_{\text{LOCF}}}{\text{pooledSD}})$.20 In this way, differences in effect sizes could be tested and 95% confidence intervals (95% CI) could be calculated. Differences in E-S between the scales were tested by paired T-tests. In order to determine significant effect-modification, the above analyses were repeated while data were stratified. For stratification we used initial HDRS scores of at least 19 for severe depression,19 criteria for response as described above and treatment-condition. Differences in E-S between strata were tested by Analysis of variance-models (ANOVAs).

The Partial Credit Item Response Theory (IRT) model57 was used to estimate the relationships between total scores on the HDRS and total scores on the Maier and Bech subscales of the HDRS. The scores were those obtained at exit (per protocol sample). The computer program OPLM58 was used to obtain a set of weights for each item in the HDRS using conditional maximum likelihood methods. The same software and the item weights were used to obtain estimates of the latent trait associated with each score on the HDRS, the Maier subscale and the Bech subscale. The total scores for the pairs of scales were equated by matching the total scores for which the latent trait scores were most similar.59 These methods are very similar to those used in a previous publication about the Quick Inventory for Depressive Symptomatology IDS.60 The range of SCL scores associated with each HDRS score was obtained directly from the original data.
Finally, Receiver operator characteristic (ROC) curves were constructed to summarize validity of cut-off points. Differences in areas under the curve (AUC) were tested with attention for interrelation (because we studied these tests within the same subjects) as described by Hanley.

For all data analysis except the IRT analysis, SPSS for Windows version 10.1 was used. For all tests two-tailed significance levels were applied.

Table 3.1. Studied populations.*

<table>
<thead>
<tr>
<th></th>
<th>Diagnostic sample (n= 485)</th>
<th>Per protocol sample (n= 219)</th>
<th>Combined I + II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial I</td>
<td>AD (n= 57)</td>
<td>AD+SPSP(16) (n= 72)</td>
</tr>
<tr>
<td></td>
<td>AD (n= 57)</td>
<td>10 (13.9)</td>
<td>8 (11.1)</td>
</tr>
<tr>
<td></td>
<td>AD+SPSP (n= 72)</td>
<td>11 (14.5)</td>
<td>12 (16.7)</td>
</tr>
<tr>
<td></td>
<td>AD+SPSP (n= 45)</td>
<td>11 (24.0)</td>
<td>10 (22.2)</td>
</tr>
<tr>
<td></td>
<td>AD+SPSP (n= 45)</td>
<td>11 (24.0)</td>
<td>10 (22.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital status (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>married</td>
<td>57 (20.2)</td>
<td>14 (24.6)</td>
<td>14 (24.6)</td>
</tr>
<tr>
<td>divorced</td>
<td>60 (12.5)</td>
<td>8 (11.1)</td>
<td>8 (11.1)</td>
</tr>
<tr>
<td>widowed</td>
<td>10 (12.5)</td>
<td>2 (2.8)</td>
<td>2 (2.8)</td>
</tr>
<tr>
<td>unmarried</td>
<td>28 (26.8)</td>
<td>5 (5.6)</td>
<td>5 (5.6)</td>
</tr>
<tr>
<td>other</td>
<td>29 (6.2)</td>
<td>2 (2.8)</td>
<td>2 (2.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>57 (11)</td>
<td>13 (14.3)</td>
<td>14 (14.3)</td>
</tr>
<tr>
<td>intermediate</td>
<td>57 (18.7)</td>
<td>12 (10.9)</td>
<td>12 (10.9)</td>
</tr>
<tr>
<td>high</td>
<td>57 (17.5)</td>
<td>11 (12.5)</td>
<td>11 (12.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of episode</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 1 year</td>
<td>57 (20.2)</td>
<td>12 (14.3)</td>
<td>12 (14.3)</td>
</tr>
<tr>
<td>1-2 years</td>
<td>57 (20.2)</td>
<td>13 (14.3)</td>
<td>14 (14.3)</td>
</tr>
<tr>
<td>&gt; 2 years</td>
<td>57 (17.5)</td>
<td>11 (12.5)</td>
<td>11 (12.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatric treatment during this episode (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antidepressants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤3 mths before (%)</td>
<td>77 (16.4)</td>
<td>10 (14.3)</td>
<td>9 (20.5)</td>
</tr>
<tr>
<td>adequate (%)</td>
<td>77 (16.4)</td>
<td>10 (14.3)</td>
<td>9 (20.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive episodes (prev. 5 years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>296 (63.1)</td>
<td>48 (72.3)</td>
<td>58 (72.3)</td>
</tr>
<tr>
<td>1-2</td>
<td>138 (29.4)</td>
<td>22 (31.4)</td>
<td>22 (31.4)</td>
</tr>
<tr>
<td>3 or more</td>
<td>35 (7.5)</td>
<td>8 (11.1)</td>
<td>8 (11.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-west Europe</td>
<td>414 (86.3)</td>
<td>63 (87.5)</td>
<td>63 (87.5)</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>18 (3.8)</td>
<td>3 (5.6)</td>
<td>3 (5.6)</td>
</tr>
<tr>
<td>Caribbean</td>
<td>22 (4.6)</td>
<td>1 (1.8)</td>
<td>1 (1.8)</td>
</tr>
<tr>
<td>Other</td>
<td>26 (5.4)</td>
<td>2 (3.5)</td>
<td>2 (3.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline scores of rating scales</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDRS-17</td>
<td>17.1 ±6.5</td>
<td>21.0 ±6.8</td>
<td>20 ±4.9</td>
</tr>
<tr>
<td>Maier</td>
<td>9.2 ±3.6</td>
<td>11.0 ±2.9</td>
<td>10.9 ±2.8</td>
</tr>
<tr>
<td>Bech</td>
<td>9.4 ±3.7</td>
<td>11.5 ±2.8</td>
<td>11.2 ±2.7</td>
</tr>
<tr>
<td>CGI-S</td>
<td>4.7 ±0.7</td>
<td>4.8 ±0.6</td>
<td>4.7 ±0.7</td>
</tr>
<tr>
<td>SCL-90 depression subscale</td>
<td>45.9 ±11.8</td>
<td>48.7 ±11.7</td>
<td>47.8 ±9.8</td>
</tr>
</tbody>
</table>

* Data represent means (±SD) unless indicated. Denominators of percentages vary due to missing values.
† Total diagnostic sample.
‡ No significant differences between treatment-groups (per protocol sample) (ANOVA or χ²).
§ Significant differences (p< 0.05) between in- and excluded patients (indep. T-Test).
¶ n= 241

References to studies: Trial I45 and Trial II46
Results

Patient characteristics

Table 3.1 shows demographics for the diagnostic and per protocol samples. There were no significant differences observed between the diagnostic and per protocol sample (tested as excluded versus included), except from a lower mean HDRS-score (and Maier, Bech and SCL90 depression scores) in the diagnostic sample. This difference was due to application of the entrance criterion (HDRS ≥14) for randomization. No significant differences existed between the different treatment-groups. The studied population existed of mainly unmarried, mid-thirty, moderately to highly educated, female, Caucasian adults, with moderate to severe depressive episodes of less than 1 year of duration. More than 75% of the subjects were not treated for the current depressive episode before, 16 percent received an inadequate trial of an antidepressant.

Table 3.2. Internal validity and concurrent validity of HDRS-17, Maier, Bech, SCL-90 depression subscale and CGI-S.

<table>
<thead>
<tr>
<th>Internal consistency</th>
<th>Concurrent validity: Pearson’s r (% explained variance)</th>
<th>Moderate depression</th>
<th>Severe depression*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cronbach-α Mean inter-item corr.</td>
<td>Maier</td>
<td>Bech</td>
</tr>
<tr>
<td>Diagnostic sample</td>
<td></td>
<td>Overall</td>
<td></td>
</tr>
<tr>
<td>Maier†</td>
<td>0.62</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Bech†</td>
<td>0.67</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>HDRS-17†</td>
<td>0.73</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.86 (75%)</td>
<td>0.86 (73%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.57 (57%)</td>
<td>0.76 (58%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.65 (42%)</td>
<td>0.60 (36%)</td>
<td></td>
</tr>
<tr>
<td>SCL-90 depression</td>
<td>0.88</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>subscale†</td>
<td>0.64 (41%)</td>
<td>0.64 (40%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.45 (21%)</td>
<td>0.45 (20%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.40 (16%)</td>
<td>0.42 (18%)</td>
<td></td>
</tr>
<tr>
<td>Per protocol sample</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CGI-S endpoint‡</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CGI-I endpoint§</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Severe depression defined as initial HDRS-17 ≥19 (n= 221).
† Maier, Bech, HDRS: n= 482; SCL-90 dep† n= 473.
‡ CGI-S: n= 229.
§ Compared with change expressed as percentage of baseline rating

Internal and concurrent validity

Data for internal and concurrent validity are presented in Table 3.2. Cronbach-alphas were slightly lower for the Maier and Bech subscales. If a 17-item scale is reduced to 6 items, the expected alpha is 0.49 (Spearman-Brown formula). Thus the observed values of 0.62 and 0.60 show increased internal validity for the subscales. The mean inter-item correlation was markedly higher for the Maier and Bech subscales. The correlation between Maier and Bech subscales was high. Both Maier and Bech subscales explained about 75% of the variance of the total HDRS-score. The self-rated SCL-90 dep was reasonably well correlated with the HDRS (r= 0.67) and the Maier and Bech subscales (r= 0.64). Concurrent validity of the scales was overall slightly less in the more depressed sub-group (HDRS ≥19; n= 194) compared to moderately depressed subjects, except for the correlation between HDRS and Maier subscale. The CGI-S at study-endpoint was moderately correlated with the HDRS (r= 0.57), as with the Maier and Bech subscales. The CGI-S showed higher correlation with the Bech subscale, especially in those severely depressed. The CGI-I at study-endpoint was less well correlated with the percentage change in HDRS (r= 0.42) and the subscales.
Sensitivity to change

In Table 3.3 and 3.4, overall and stratified E-S in the per protocol sample are presented. In these tables the 95% CI of the E-S indicates whether the E-S significantly deviates from 0 (no effect measured). Comparisons between E-S may produce significant differences between E-S, even when the 95% CIs between the two E-S overlap.

Of the 9 comparisons between the Maier and Bech subscales made in these tables, 5 were not significant. The Maier was significantly different from the HDRS in 4 out of 9 comparisons, while the Bech was significantly different from the HDRS in only 1 of the 9 comparisons. Differences between E-S were small.

### Table 3.3

Pre- and post-treatment Maier, Bech and HDRS-scores with corresponding effect sizes in per protocol sample. Stratification by depression-severity and final treatment response.

<table>
<thead>
<tr>
<th></th>
<th>Mean ±SD baseline</th>
<th>Mean ±SD endpoint (LOCF)</th>
<th>Mean decrease (95% CI)</th>
<th>SD of decrease</th>
<th>Effect size [E-S] (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All subjects (n= 219)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maier</td>
<td>10.9 ±2.75</td>
<td>6.2 ±4.46</td>
<td>4.7 (4.1 - 5.3)</td>
<td>4.54</td>
<td>1.03 (0.89 - 1.16)*†</td>
</tr>
<tr>
<td>Bech</td>
<td>11.1 ±2.69</td>
<td>6.2 ±4.50</td>
<td>4.9 (4.3 - 5.5)</td>
<td>4.54</td>
<td>1.08 (0.95 - 1.22)</td>
</tr>
<tr>
<td>HDRS-17</td>
<td>20.2 ±4.56</td>
<td>12.0 ±7.62</td>
<td>8.2 (7.2 - 9.2)</td>
<td>7.45</td>
<td>1.10 (0.96 - 1.23)</td>
</tr>
<tr>
<td><strong>Moderate depression (initial HDRS-17 &lt;19; n= 93)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maier</td>
<td>12.3 ±2.25</td>
<td>6.9 ±4.75</td>
<td>5.4 (4.6 - 6.2)</td>
<td>4.66</td>
<td>1.19 (1.01 - 1.37)*</td>
</tr>
<tr>
<td>Bech</td>
<td>12.5 ±2.16</td>
<td>7.0 ±4.75</td>
<td>5.5 (4.7 - 6.3)</td>
<td>4.65</td>
<td>1.21 (1.03 - 1.39)</td>
</tr>
<tr>
<td>HDRS-17</td>
<td>19.5 ±4.25</td>
<td>19.9 ±4.55</td>
<td>-0.4 (-1.2 - 0.3)</td>
<td>3.08</td>
<td>-0.06 (-0.16 - 0.05)*</td>
</tr>
<tr>
<td><strong>Severe depression (initial HDRS-17 ≥19; n= 126)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maier</td>
<td>11.5 ±2.86</td>
<td>7.4 ±3.07</td>
<td>4.1 (3.4 - 2.7)</td>
<td>2.58</td>
<td>0.90 (0.76 - 1.04)</td>
</tr>
<tr>
<td>Bech</td>
<td>11.4 ±2.59</td>
<td>7.4 ±3.19</td>
<td>4.0 (3.4 - 2.7)</td>
<td>2.71</td>
<td>0.90 (0.75 - 1.04)</td>
</tr>
<tr>
<td>HDRS-17</td>
<td>19.9 ±4.28</td>
<td>15.8 ±4.61</td>
<td>4.1 (3.7 - 5.5)</td>
<td>1.88</td>
<td>0.96 (0.90 - 1.03)</td>
</tr>
<tr>
<td><strong>Final Non-responders (n= 65)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maier</td>
<td>10.6 ±2.69</td>
<td>2.1 ±2.00</td>
<td>8.5 (7.8 - 9.1)</td>
<td>3.12</td>
<td>1.87 (1.72 - 2.01)**</td>
</tr>
<tr>
<td>Bech</td>
<td>10.9 ±2.71</td>
<td>2.1 ±1.89</td>
<td>8.8 (8.1 - 9.4)</td>
<td>3.14</td>
<td>1.93 (1.79 - 2.08)</td>
</tr>
<tr>
<td>HDRS-17</td>
<td>19.9 ±4.28</td>
<td>15.1 ±4.61</td>
<td>4.8 (4.1 - 5.5)</td>
<td>4.89</td>
<td>2.02 (1.89 - 2.16)**</td>
</tr>
<tr>
<td><strong>Final Partial-responders (n= 64)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maier</td>
<td>11.5 ±2.86</td>
<td>7.4 ±3.07</td>
<td>4.1 (3.4 - 2.7)</td>
<td>2.58</td>
<td>0.90 (0.76 - 1.04)</td>
</tr>
<tr>
<td>Bech</td>
<td>11.4 ±2.59</td>
<td>7.4 ±3.19</td>
<td>4.0 (3.4 - 2.7)</td>
<td>2.71</td>
<td>0.90 (0.75 - 1.04)</td>
</tr>
<tr>
<td>HDRS-17</td>
<td>19.9 ±4.28</td>
<td>15.8 ±4.61</td>
<td>4.1 (3.7 - 5.5)</td>
<td>1.88</td>
<td>0.96 (0.90 - 1.03)</td>
</tr>
<tr>
<td><strong>Final Responders (n= 90)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maier</td>
<td>10.6 ±2.69</td>
<td>2.1 ±2.00</td>
<td>8.5 (7.8 - 9.1)</td>
<td>3.12</td>
<td>1.87 (1.72 - 2.01)**</td>
</tr>
<tr>
<td>Bech</td>
<td>10.9 ±2.71</td>
<td>2.1 ±1.89</td>
<td>8.8 (8.1 - 9.4)</td>
<td>3.14</td>
<td>1.93 (1.79 - 2.08)</td>
</tr>
<tr>
<td>HDRS-17</td>
<td>19.9 ±4.28</td>
<td>15.1 ±4.61</td>
<td>4.8 (4.1 - 5.5)</td>
<td>4.89</td>
<td>2.02 (1.89 - 2.16)**</td>
</tr>
</tbody>
</table>

Note that the overlap of two 95%CI of E-S does not rule out a statistical significant difference between these E-S (see text).

* Significantly different from E-SHDRS (paired T-test; p <0.05)
† Significantly different from E-SBech (paired T-test; p <0.05)
‡ Significant differences of E-SMaier, E-SBech and E-SHDRS between moderate and severe depression (ANOVA; p< 0.05)
§ Response criteria: decrease in HDRS-scores: <20%= Non-response, 20-50%= Partial response and ≥50%= Response. Significant differences of E-SMaier, E-SBech and E-SHDRS between categories of response (ANOVA; p< 0.001)
¶ Significant differences between E-SMaier-E-SBech, E-SHDRS-E-SMaier and E-SHDRS-E-SBech (paired T-test; p <0.05)** Significant difference between E-SHDRS-E-SMaier (paired T-test; p <0.05)
In the total per protocol sample the Maier subscale was significantly less powerful to observe treatment effects: the E-S assessed by the Maier was significantly lower than the E-S of the Bech and HDRS. When stratified for depression severity, the E-S of Maier, Bech subscales and HDRS were larger in severe compared to moderate depression. A significant difference between these strata was observed for all E-S (ANOVA). Within the group of severely depressed subjects, the Maier was significantly less sensitive compared to the HDRS (paired T-test). Within the moderately depressed group, the Bech outperformed the Maier (paired T-test). Across different strata of final response significant differences in E-S were found (ANOVA). Within strata, the Bech subscale performed less in final non-responders, while the Maier performed significantly less than the total HDRS in final responders (paired T-tests).

In Table 3.4 it is shown that no overall differences in effect sizes were found between treatment modalities (ANOVA). Within the group of patients treated with anti-depressants only, the Maier subscale was significantly less sensitive to detect treatment differences than the HDRS, however the Maier did not differ significantly from the Bech subscale (paired T-test).

### Conversion of HDRS scores, criteria for remission and depression severity

Table 3.5 shows the conversion between HDRS scores and Maier, Bech and SCL90_{dep} scores. Maier and Bech cut-off scores to define remission (bold), mild, moderate and severe depression can be identified (italic). Figure 3.1 shows the ROC-curves for Maier, Bech CGI-S, CGI-I and SCL-90_{dep} cut-off scores, with HDRS ≤7 as the reference criterion. The difference in AUC for the Maier and Bech subscales was not significant (z= 1.25; p= 0.21). The difference in AUC between Maier and Bech subscales compared to SCL90_{dep} and both CGIs was highly significant (z> 3.8; p< 0.001). In the table below Figure 3.1 sensitivity and specificity for cut-off scores ≤3 and ≤4 for the Maier and Bech subscales are given.
**Figure 3.1.** Receiver operating characteristic curves for Maier and Bech subscales, SCL90 depression and CGI-S/I at endpoint compared to HDRS-17 'Remission' in per protocol sample.

<table>
<thead>
<tr>
<th>Sub-scale</th>
<th>Maier</th>
<th>Bech</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥3</td>
<td>88.9</td>
<td>93.2</td>
</tr>
<tr>
<td>≥4</td>
<td>97.2</td>
<td>85.7</td>
</tr>
</tbody>
</table>

Reference: HDRS-17 score ≥7 (Remission), Sens. = sensitivity, Spec. = specificity
AUC (SE): Maier 0.972 (.009), Bech 0.963 (.011), SCL-90p 0.862 (.028), CGI-S 0.743 (.036), CGI-I 0.738 (.035)

**Discussion**

**Major findings**

This study examined the relative effectiveness of the HDRS subscales as developed by Maier et al.\(^\text{18}\) and Bech et al.\(^\text{37}\) in monitoring severity and treatment effects in depression. We found that the Maier and Bech subscales gave results comparable to the original 17-item HDRS, with high concurrent validity and increased mean inter-item correlations and internal consistency. Maier and Bech subscales were highly comparable to each other in the measurement of treatment changes. Differences between E-S were rather small, and clinically irrelevant. For interpretation a conversion table linking HDRS-scores and Maier and Bech scores are provided. The Maier had a slightly (non-significant) higher sensitivity and specificity to predict the reference criterion for remission (HDRS ≥7). Both Maier and Bech subscales differentiated non-responders from partial and final responders.

A significant difference in sensitivity to change existed between the Bech and Maier within the group treated with antidepressants only. We were unable to find the reason for this difference compared to other treatment modalities, where the difference between Maier and Bech was not found or was not significant. The question arises whether there is a difference in sensitivity between the Maier and Bech subscales across different treatment-modalities or that other (post-randomization) differences between the groups or mere chance explain this observation. Because this difference was not found in the other groups (treated with both antidepressants and psychotherapy) we think it cannot be ascribed to a difference in detecting pharmacological (side-) effects. If a Bonferroni correction would be applied for the number of comparisons tested (p < 0.01), the observed difference would not remain its significance.
The relevance of the difference between the Maier and Bech subscales

The only difference between the Maier and Bech subscales is the inclusion of agitation (e.g. running thoughts or restlessness, 0-4 points) in the Maier versus the inclusion of general somatic symptoms (e.g. tiredness, 0-2 points) in the Bech. It could be argued that one scale is comparable with the other scale without the different item, e.g. the Maier subscale would then be comparable to the Bech subscale minus the ‘general somatic’ item. In our (diagnostic) sample the item agitation contributed 1.3 (SD= 0.9) points to the total Maier-score (9.2, SD= 3.6). The general somatic item contributed 1.5 (SD= 0.8) points to the total Bech-score (9.4, SD= 3.7). Thus, overall

Table 3.5. The conversion between the HDRS total scores and the Maier subscale, Bech subscale and the range of SCL scores using IRT analysis (per protocol sample).*

<table>
<thead>
<tr>
<th>HDRS</th>
<th>Maier</th>
<th>Bech</th>
<th>Range SCL90 dep</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>1 – 2</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>3 – 4</td>
<td>2</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>7 (remission†)</td>
<td>4</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>8 – 9</td>
<td>5</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>10 – 11</td>
<td>6</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>13 (mild†)</td>
<td>7</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>8</td>
<td>8</td>
<td>31 – 61</td>
</tr>
<tr>
<td>15</td>
<td>8</td>
<td>9</td>
<td>22 – 59</td>
</tr>
<tr>
<td>16</td>
<td>9</td>
<td>9</td>
<td>26 – 61</td>
</tr>
<tr>
<td>17</td>
<td>9</td>
<td>10</td>
<td>25 – 59</td>
</tr>
<tr>
<td>18 (moderate†)</td>
<td>10</td>
<td>10</td>
<td>30 – 62</td>
</tr>
<tr>
<td>19</td>
<td>10</td>
<td>10</td>
<td>28 – 60</td>
</tr>
<tr>
<td>20</td>
<td>11</td>
<td>11</td>
<td>38 – 67</td>
</tr>
<tr>
<td>21</td>
<td>11</td>
<td>11</td>
<td>30 – 72</td>
</tr>
<tr>
<td>22</td>
<td>12</td>
<td>12</td>
<td>39 – 61</td>
</tr>
<tr>
<td>23</td>
<td>12</td>
<td>12</td>
<td>36 – 61</td>
</tr>
<tr>
<td>24 (severe†)</td>
<td>13</td>
<td>13</td>
<td>38 – 67</td>
</tr>
<tr>
<td>25 (very severe†)</td>
<td>13</td>
<td>13</td>
<td>45 – 64</td>
</tr>
<tr>
<td>26</td>
<td>13</td>
<td>13</td>
<td>47 – 71</td>
</tr>
<tr>
<td>27</td>
<td>14</td>
<td>14</td>
<td>46 – 72</td>
</tr>
<tr>
<td>28</td>
<td>14</td>
<td>14</td>
<td>42 – 79</td>
</tr>
<tr>
<td>29</td>
<td>15</td>
<td>14</td>
<td>55 – 71</td>
</tr>
<tr>
<td>30</td>
<td>15</td>
<td>15</td>
<td>43 – 43</td>
</tr>
<tr>
<td>31</td>
<td>15</td>
<td>15</td>
<td>63 – 70</td>
</tr>
<tr>
<td>32</td>
<td>16</td>
<td>15</td>
<td>62 – 65</td>
</tr>
<tr>
<td>33</td>
<td>16</td>
<td>15</td>
<td>57 – 71</td>
</tr>
<tr>
<td>34 – 35</td>
<td>17</td>
<td>16</td>
<td>-</td>
</tr>
<tr>
<td>36</td>
<td>17</td>
<td>17</td>
<td>-</td>
</tr>
<tr>
<td>37 – 39</td>
<td>18</td>
<td>17</td>
<td>-</td>
</tr>
<tr>
<td>40 – 42</td>
<td>19</td>
<td>18</td>
<td>-</td>
</tr>
<tr>
<td>43 – 44</td>
<td>20</td>
<td>19</td>
<td>-</td>
</tr>
<tr>
<td>45</td>
<td>21</td>
<td>19</td>
<td>-</td>
</tr>
<tr>
<td>46</td>
<td>21</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>47 – 48</td>
<td>22</td>
<td>20</td>
<td>-</td>
</tr>
</tbody>
</table>

* The only valid conversions that can be made from this table are between 1) HDRS and Maier, 2) HDRS and Bech and 3) HDRS and SCL90 dep.
† Cut-offs as provided by Yonkers.11
tiredness was more present than agitation in this sample, and agitation was not rated near its maximum like tiredness. Both items occurred intraindividually at the same time, but were not interchangeable. This means that the Maier and Bech subscales show different perspectives on depressive symptomatology. In this respect, it is noteworthy to mention that the agitation item was dropped beforehand when the Bech subscale was developed and validated, because this item showed limited variance (i.e. was not found to be scored) in the two studied samples.\textsuperscript{36,37} Furthermore, in the Maier subscale the items psychomotor agitation and psychomotor retardation are included, which -at first sight- seem to represent two opposed polarities. However, these items also co-occurred within the same individuals. This can be explained by the broad definitions of agitation (both restlessness or running thoughts) and retardation (both retardation in activities or in thinking) in our semi-structured interview.

The original HDRS is often criticized to measure these somatic symptoms.\textsuperscript{11,28} Although the Bech subscale was designed as an unidimensional scale, the 'general somatic' item is still among the 6 items. However, in the Rasch-analysis this item was the least contributive and showed a ceiling effect for moderate and severe depression.\textsuperscript{37} Though both a DSM-IV and ICD-10 criterion for diagnosis, the aspecific 'tiredness' symptom may also be caused by physical illnesses. The Maier subscale does not include this item. Thus, the Maier subscale might especially be useful in patients suffering from somatic complaints or illnesses. Additional, methodological support comes from the exclusion of this somatic item by Santor et al.\textsuperscript{21} This hypothesis of better performance in patients with somatic complaints or illness needs further investigation, e.g. in comparison with the Hospital Anxiety and Depression Scale.\textsuperscript{63}

Previous comparative studies

Our findings are in line with findings of previous studies,\textsuperscript{11,20,26,40-44} and extends the evidence to support the Maier and Bech subscale as a valid alternative for the HDRS. This is relevant for the planning and conduction of clinical trials,\textsuperscript{40-42} but also for clinical practice.\textsuperscript{20,26} Hooper et al.\textsuperscript{26} found equal performance of the Bech subscale compared to the Montgomery Asberg Depression Rating Scale (MADRS).\textsuperscript{30} Because the MADRS was not used in our trials, we were unable to examine the performance of the Maier subscale compared to the MADRS. Hooper et al. questioned whether a possible ceiling effect in the Bech subscale would limit its usefulness in severely depressed patients.\textsuperscript{26} In our study more than 57\% of the per protocol patients had an initial HDRS greater than 18 (indicative for severe depression).\textsuperscript{11} We did not find a ceiling effect in our diagnostic sample (data not shown), and found consequently higher effect sizes for the Maier and Bech subscales in initially severely depressed patients, indicative for an adequate sensitivity to measure (larger) changes due to treatment. In addition to the observed ability to predict remission,\textsuperscript{41} we proposed cut-off scores for remission and various ranges for classification of depression severity.

In two publications Bech proposed the Bech subscale as an alternative measure to overcome the confounding influence of drug related side-effects in the comparison with placebo or active drugs.\textsuperscript{43,44} However this problem is not fully solved, as tiredness may be induced by histaminergic effects from antidepressants (e.g. tricyclics and mirtazapine).\textsuperscript{43} On the other hand, agitation (included in the Maier subscale) is known as a (mostly transient) SSRI-induced side-effect.

An extra dimension of our study is that it extends the data for use of the Maier and Bech subscales in populations treated with psychotherapy. Hooper et al. and O’Sullivan et al. already demonstrated the usefulness of the Bech subscale in pharmacological treatment of melancholia, dysthymia and typical and atypical depression.\textsuperscript{20,26} An alarming point of our study is the moderate correlation of the Maier, Bech and HDRS with the CGI-S and the CGI-I. Previous reports mentioned correlations between HDRS and CGI varying between 0.65 and 0.90.\textsuperscript{11,28,60} Our results underscore the need of a HDRS or subscale rating instead of the CGI. We consider the validity of the CGI to be questionable, as most CGI-raters (subjectively) evaluate their own treatment. Apparently, the clinician’s judgment does not coincide with scale scores. In this respect, the performance of the (self-rated) SCL-90\textsubscript{dep} is better. This was also illustrated in the ROC-curves regarding the criterion of remission. Further research is needed to investigate whether correlations with the HDRS of other self-rated scales (e.g. the Beck
Depression Inventory\textsuperscript{64) are higher than the SCL-90\textsubscript{dep}. In addition to this, a major limitation in our study and in any study investigating depression 'severity', is that there is no definite gold standard. We used HDRS data as the gold standard, which means that scales under investigation can never be judged to be better than the HDRS, however this would be reversed if the CGI was used as a gold standard.\textsuperscript{65}

**Conclusion**

We conclude that both Maier and Bech subscales of the HDRS are equivalent to the HDRS, and can easily be used to increase efficiency to measure treatment response in clinical practice. On theoretical grounds we have a slight preference for the Maier subscale. The use of subscales would improve the efficiency and objectivity of measuring response in clinical practice, where often no scale (instead of a Clinical Global Impression) is used at all. This would further bridge the gap between clinical practice and research based treatment-recommendations for non-response in depression. Maier and Bech subscales should be compared in patients suffering from co-morbid somatic illnesses, or patients treated with psychotherapy only. The impact of the difference of the one somatic item versus the agitation-item between the Maier and Bech subscales and the consequences for their applicability in clinical sub-groups needs further research.

**Acknowledgement**

The original randomized controlled trials were supported by an unrestricted educational grant from Eli Lilly Netherlands. All studies were carried out by the Mentrum Depression Research Group. We thank all psychotherapists, psychiatrists and residents for their excellent work.

**Conflicts of interest**

External funding did not support these post-hoc analyses.

**References**


17. Pien RF, Carpenter LL, Kupper DJ. The definition and operational criteria for treatment outcome of major depressive disorder. A review of the current research literature. Arch Gen Psychiatry. 1993; 48: 796-800.


DOSE-ESCALATION FOR INSUFFICIENT RESPONSE TO A STANDARD DOSE OF A SELECTIVE SEROTONIN REUPTAKE INHIBITOR IN MAJOR DEPRESSIVE DISORDER: A SYSTEMATIC REVIEW.

British Journal of Psychiatry 2006; 189: 309-316

Henricus G. Ruhé
Jochanan Huyser
Jan A. Swinkels
Aart H. Schene

1 Program for Mood Disorders, Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
Abstract

Background
Although SSRIs are frequently used for major depressive disorder, only 50-60% of patients respond to a standard dose. For non-responders dose-escalation is often applied.

Aim
To systematically review the evidence for dose-escalation of SSRIs.

Methods
A systematic literature search in MEDLINE, EMBASE, CINAHL and PsycInfo was performed. Randomised controlled trials and meta-analyses investigating dose-escalation of SSRIs were identified. Relevant articles were retrieved and critically appraised. Results were summarized in an evidence table. Pooling was not justified due to heterogeneity of the identified studies.

Results
Eight true dose-escalation studies and 3 meta-analyses were identified. The available data provided no unequivocal base for dose-escalation. Dose-escalation before 4 weeks of treatment at a standard dose appeared to be ineffective.

Conclusions
Dose-escalation of SSRIs is equivocally supported by evidence of RCTs, but methodological difficulties in the studies may account for this lack of evidence. Clinical implications and methodological considerations for future studies are discussed.
Introduction

Many countries developed national clinical guidelines for the treatment of Major Depressive Disorder (MDD). In these guidelines pharmacotherapy is among the most important treatments, while in many of the countries, Selective Serotonin Reuptake Inhibitors (SSRIs) have become the first line antidepressants. It is less clear what should be done in those 40-50% of the patients who do not respond to the first antidepressant given. Strategies in case of non-response have been published in several narrative reviews and in one systematic review. Three major strategies for non-response are recommended: 1) increase of the dose of the antidepressant (dose-escalation), 2) augmenting the antidepressant by adding a second drug, 3) switching to another antidepressant of the same or a different class.

Available dose-finding studies do not provide evidence to initiate pharmacotherapy for MDD with SSRIs in higher than standard doses. For non-responders, all guidelines recommend dose escalation as the appropriate strategy, instead of continuing an apparently insufficient regimen. Only the recent NICE guideline is less pronounced in this recommendation, and advises that if 'there are no significant side effects, a gradual increase in dose should be considered'. Moreover, surprisingly little systematic evidence is provided to underscore these recommendations.

Due to the above recommendations and its simplicity, dose-escalation is widely practiced and often the first strategy applied. The aim of our study was to systematically review the evidence for dose-escalation of SSRIs in MDD.

Methods

Design of studies to be included

Ideally the design of dose-escalation studies is randomisation of non-responders to higher doses of an antidepressant or placebo after some weeks of a standard dose. In this review we consider three other methodological requirements for those studies. First, dose-escalation should be deferred, e.g. 3-6 weeks after the initiation of the treatment because antidepressants need several weeks to have a clinical effect. The practice of dose-escalation and the possibility to demonstrate a dose-response relationship is based upon a selection of 'true' non-responders. As this might take 6-10 weeks, dose-escalation studies with early randomisation unintentionally diminish the possibility to prove the usefulness of dose-escalation. The inclusion of unidentified late responders in both arms of the study reduces the contrast between the intervention and control. Second, an outstanding study will have sufficient power to be able to demonstrate a clinically relevant difference (e.g. 20%) between treatment arms, and third will describe the method of dose-escalation and describe the early drop-out rates due to dose-escalation.

Identification and selection of articles

First, systematic literature searches (updated until February 10th 2005) were performed in four databases (MEDLINE, EMBASE, CINAHL, PsychInfo; all indexed years). As there are no specific keywords for dose-escalation studies, 'sensitive' searches were performed with the following terms: (((dose[textword(tw)] or dosage[tw]) and increase[tw]) or ((dose[tw] or dosage[tw]) and maxim*[tw]) or (upward[tw] and titrat*[tw])) OR dose-response relationship, drug[MeSH] in combination with the Cochrane Collaboration search-filter for RCTs and systematic reviews, the Cochrane Collaboration Depression Anxiety and Neurosis group (CCDAN) search-filter for MDD and MeSH-terms and text words for SSRIs. Primary selection (independently by the first and second author) was based on design and focus on dose-response relationships for SSRIs, by screening title and abstract of the article. Agreement on predominantly exclusion of irrelevant articles was 99.1%, with Cohen’s Kappa for interrater agreement of 0.62 (a 'substantial' agreement). Discrepancies between initial selection was resolved by discussion and concensus.
Second, all potentially relevant articles were judged according to specific inclusion and exclusion criteria (criteria available on request). In case of doubt an article was read fully and assigned afterwards. Additionally, relevant cross-references were retrieved. Double-publications were considered together to reveal the maximum of available information.

Critical Appraisal and summary

Selected articles were critically appraised and abstracted by the first author, using standardized forms derived from the Dutch Institute of Healthcare Improvement\(^\text{39}\) and the AHCPR.\(^\text{4}\) The items used for critical appraisal were the same as proposed by SIGN\(^\text{40}\) and Sackett et al.\(^\text{41}\) Each study was assigned a ‘level of evidence’ (LoE; Table 4.1).\(^\text{39}\) Levels of evidence are based upon the methodological robustness of studies. For the results, the highest LoE of the supporting scientific evidence (A1-D) is indicated.

**Table 4.1. Levels of Evidence: Therapeutic studies.**

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Systematic review including at least some studies of A2-level. Consistent results (homogeneity) across the included trials.</td>
</tr>
<tr>
<td>A2</td>
<td>Randomized controlled (double-blind) trial of good methodological quality, adequate size and consistence of results.</td>
</tr>
<tr>
<td>B</td>
<td>Randomized clinical trial of lower methodological quality or inadequate size. Other comparative research (e.g. non-randomized trial, comparative cohort-study, case-control study)</td>
</tr>
<tr>
<td>C</td>
<td>Uncontrolled, open study</td>
</tr>
<tr>
<td>D</td>
<td>Expert opinion, e.g. guideline-panel members</td>
</tr>
</tbody>
</table>

Source: Dutch Institute of Healthcare Improvement.\(^\text{39}\)

To assess judgement-bias by one person who performed the critical appraisal, interrater variation was determined in a slightly different set of 12 publications. Every other author critically appraised 4 publications, agreement for the appraisal-items was expressed by Cohen’s Kappa. Kappa values were 0.49 (for ‘validity of the study’), 0.86 (for ‘concealment of allocation’), while complete agreement existed for the appraisal-items ‘randomization of the study’, ‘level of evidence’ and ‘data extraction’ (Kappa = 1.0). This is in line with other reports of interrater agreement in appraisal of psychiatric research.\(^\text{42}\)

A qualitative summary with discussion of the results, restrictions, methodological flaws and external validity of the studies was described in an evidence table and a separate document, of which a summary is provided in this paper. Because of the apparent heterogeneity in timing of the dose-escalation between the studies, results were not pooled in a meta-analysis.

Results

Search results and selection of studies are presented in Figure 4.1. The 11 studies selected for this review are summarized in Table 4.2. A table of excluded studies is available on request.

![Figure 4.1. Selection of reported studies.](image-url)
Characteristics of the studies

Our searches identified 8 dose-escalation studies that increased dosages after at least 3 weeks of standard dosages.\(^{43-50}\) We further found 3 systematic reviews about dose response relationships, which included respectively three,\(^{91}\) three,\(^{52}\) and four\(^{53}\) of the eight identified dose-escalation studies.

Across the studies different outcome definitions for end-points were used. In 7 articles response was defined as a reduction of \(\geq 50\%\) in the Hamilton depression rating scale (HDRS).\(^{43-46,50-53}\) A Clinical global impression (CGI) improvement or severity score \(\leq 2\) was used for response in one study.\(^{48}\) Partial response was used in 3 studies and defined as \(25\%-50\%\) decrease in HDRS.\(^{45-49}\) In 7 studies remission-rates were reported. These were defined as a HDRS \(\leq 7\),\(^{46,49-50}\) or HDRS \(\leq 8\).\(^{48}\)

Different criteria were used to decide whether a patient should be randomized: non-response (by CGI,\(^{47}\) \(<50\%\) decrease in HDRS\(^{43-46,49}\)), or no remission (HDRS \(\leq 8\)).\(^{48}\) In the present studies no genetic information of the CYP P450 system, nor drug blood levels were reported.

The three previous reviews all had some methodological problems: Bollini et al. pooled studies with completely different designs and drug-classes, and applied a dose-equivalence strategy that lumped differential doses of SSRIs together.\(^{51}\) Baker et al. also pooled heterogeneous studies with different moments of dose-escalation, and used an unusual low reference dose of fluoxetine (5mg).\(^{53}\) Corruble et al. did not use an adequate search strategy and only described the dose-response relationships found in their identified studies as ‘flat’ ‘curvilinear’ or ‘linear’.\(^{52}\)

Outline of dose-escalation studies

We will briefly outline the dose-escalation studies. Dorseif et al. first investigated week 3 non-responders (n= 371 outpatients) to fluoxetine who were randomized to continuation with 20 mg or increase to 60 mg/day for 5 weeks. Response rates were 40.5\% and 44.7\% respectively and remission rates 33.3\% and 36.2\%. Drop-out rates due to side-effects were significantly different with 5.3\% and 11.6\% respectively.\(^{43}\) Schweizer et al. investigated 77 non-responsive outpatients after 3 weeks of fluoxetine 20 mg/day, with a randomization to placebo-increase or dose-escalation up to 60 mg/day for 5 weeks. Response rates were 51.2\% and 50\% respectively, with nonsignificant drop-out rates of 4.9\% versus 16.7\%.\(^{44}\) In a similar study Schweizer et al. studied dose-escalation of sertraline in outpatient nonremitters after 3 weeks of sertraline 50 mg/day (n= 75). Doses were randomly either kept at 50 mg/day or increased to 150 mg/day. Remission rates after 5 weeks were 32\% and 47\% respectively (non-significant). Specified drop out rates due to side effects were not reported.\(^{48}\)

Fava et al. first openly treated 15 outpatients (who were week 8 non-responders to fluoxetine at 20 mg/day) with increased doses of fluoxetine titrated up to 80 mg/day for 4 weeks. No response rates were given, but the mean HDRS\(_{17}\)-score decreased 6.2 points in week 8 non-responders and 10.1 points in partial responders.\(^{45}\) In a second study Fava et al. randomized week 8 non-responders to fluoxetine 20 mg/day (n= 41) to either fluoxetine 40-60 mg, desipramine addition or lithium-addition for 4 weeks. No placebo-increase was practised. Remission rates were 53\%, 25\% and 29\% respectively, but these differences were nonsignificant. Initial partial responders appeared to benefit most from fluoxetine dose-increases (data nonsignificant). Drop-out rates for side effects were 0\%, 17\%, and 7\% respectively.\(^{46}\) In a third study Fava et al. repeated the 3-arm randomized design from their 1994 study with a stratification for partial or non-response at week 8 (n= 101). After 4 weeks, the high-dose fluoxetine group showed increased but nonsignificant remission rates (42.4\%) compared to desipramine addition (29.4\%) and lithium-addition (23.5\%). Again initial partial responders appeared to benefit more from fluoxetine dose-increases compared to initial non-responders (differences nonsignificant). No specific data on dropout due to side effects were given.\(^{49}\)
### Table 4.2. Effectiveness of increasing the dose: Selected studies.

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Reference</th>
<th>LoE</th>
<th>N</th>
<th>Design (follow-up)</th>
<th>Intervention†</th>
<th>Comparison†</th>
<th>Outcome†</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2a. Dose-escalation studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benkert et al. (1997)</td>
<td>47</td>
<td>B</td>
<td>544</td>
<td>MDD, MinD OutP</td>
<td>RCT of week 3 nonresponders‡ (n = 86) (5 weeks)</td>
<td>PAR 40mg</td>
<td>PAR 20mg</td>
<td>Response (≥50% ↓ in HDRS) all: NNT PAR40mg = 100 (5.1-∞) For MDD only: NNT PAR40mg = 8 (2.4-∞) For baseline HDRS ≥24: NNT PAR40mg = 6 (1.7-∞)</td>
</tr>
<tr>
<td>Dornseif et al. (1989)</td>
<td>43</td>
<td>B</td>
<td>572</td>
<td>MDD OutP</td>
<td>RCT of week 3 nonresponders‡ (n = 371) (5 weeks)</td>
<td>FLX 60mg</td>
<td>FLX 20mg</td>
<td>Response (≥50% ↓ in HDRS) NNT = 25 (6.5-∞) Remission (HDRS≤7) NNT = 36 (7.3-∞) Response (CGI-I ≤2) NNT = 20 (6.5-∞) Drop-out SE NNH 16 (8.3-144)</td>
</tr>
<tr>
<td>Fava et al. (1992)</td>
<td>45</td>
<td>C</td>
<td>15</td>
<td>MDD OutP</td>
<td>Open trial of nonresponders³ to 8-12 weeks of FLX 20mg (4 weeks)</td>
<td>FLX 40-80mg (if tolerated)</td>
<td></td>
<td>Decrease in HDRS-scores 6 NR (6.2) and PR (≥10.1) (p&lt;0.05) Decrease in CGI-S NR (-0.9; n.s.) and PR (≥2.0)</td>
</tr>
<tr>
<td>Fava et al. (1994)</td>
<td>46</td>
<td>B</td>
<td>101</td>
<td>MDD Setting?</td>
<td>RCT of nonresponders‡ to 8 weeks of FLX 20mg (4 weeks)</td>
<td>FLX 20mg + DES 25-50mg FLX 20mg + Li 300-600mg</td>
<td></td>
<td>Remission (HDRS≤7) NNTall = 4 (1.6-∞) NNTNR = 6 (1.5-∞) (vs. Li) NNTPR = 2 (0.9-2.8) (vs. Li) Drop-out SE NNH 6 (2.6-∞) (vs. Li)</td>
</tr>
<tr>
<td>Licht et al. (2002)</td>
<td>50</td>
<td>A</td>
<td>1629</td>
<td>MDD OutP</td>
<td>RCT of week 6 nonresponders‡ (n = 295) (5 weeks)</td>
<td>SER 200 mg SER 100 mg + PLAC</td>
<td></td>
<td>Response (≥50% ↓ in HDRS) NNT = 10 (2.0-∞) Remission (HDRS≤7) NNH = 4 (2.4-∞) (vs. Li) NNTNR = 6 (1.5-∞) (vs. Li) NNTPR = 2 (0.9-2.8) (vs. Li) Drop-out SE NNH 6 (2.6-∞) (vs. Li)</td>
</tr>
<tr>
<td>Schweizer et al. (1990)</td>
<td>44</td>
<td>B</td>
<td>108</td>
<td>MDD OutP</td>
<td>RCT of week 3 non-remitters¶ (n = 77) (5 weeks)</td>
<td>FLX 60mg</td>
<td>FLX 20mg</td>
<td>Response (≥50% ↓ in HDRS) NNT = 10 (2.0-∞) Remission (HDRS≤7) NNT = 6 (3.4-16.4)</td>
</tr>
<tr>
<td>Schweizer et al. (2001)</td>
<td>48</td>
<td>B</td>
<td>91</td>
<td>MDD OutP</td>
<td>RCT of week 3 non-remitters¶ (n = 75) (5 weeks)</td>
<td>SER 150mg</td>
<td>SER 50mg</td>
<td>Remission (HDRS≤7) NNT 7 (2.7-∞) Response (CGI-I ≤2) NNT 5 (2.3-61.5) Dropout NNH = 34 (8.5-∞)</td>
</tr>
</tbody>
</table>
Table 4.2. Effectiveness of increasing the dose: Selected studies. (Continued)

<table>
<thead>
<tr>
<th>Study (year)reference</th>
<th>LoE</th>
<th>N</th>
<th>Design (follow-up)</th>
<th>Intervention</th>
<th>Comparison</th>
<th>Outcome</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baker et al. (2002)53</td>
<td>A2</td>
<td>573</td>
<td>MDD Setting</td>
<td>Meta-analysis of 1) 4 fixed dose RCTs (3-7 weeks) and 2) 4 dose-escalation RCTs (3-5 weeks) of NR in week 3-8. (SSRIs only). 1) Medium / High dose 2) Dose-escalation 1) Low dose 2) adding PLAC (or LI/IDES)</td>
<td>Increase in Response rate (≥50% ↓ in HDRS) across dose-range: 1) ITT: -9.5% (n.s.) DT: 7.8% (p&lt;0.01) 2) ITT: 6% (n.s.) DT: 9.3% (n.s.) Change of HDRS-decrease across dose-range: 1) ITT: 2.0 (p&lt;0.001). DT: unavailable. 2) ITT: 1.93 (p&lt;0.01). DT: unavailable.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bollini et al. (1999)51</td>
<td>A2</td>
<td>584</td>
<td>InP &amp; OutP</td>
<td>Meta-analysis of 33 RCTs (1975-1997) with various antidepressants. (3-156 weeks) Higher doses IMI-equivalent 201-250 (a) and &gt;250mg (b) Average daily dose (IMI-equivalent 100-200mg)</td>
<td>Efficacy in ITT analysis: 53.3% (ref), 46.3% (a), 48.3% (b) Completers analysis: 69% (ref), 67.3% (a), 76% (b) Dropout rates: 22% (ref), 28% (a), 35% (b). SE: 30% (ref), 36% (a), 48% (b)</td>
<td>Meta-analysis using regression models. Highly heterogenous studies (i.e. designs) pooled, no separation of various antidepressant classes, nonsystematic bias especially for SSRIs by conversion to IMI-equivalents.</td>
<td></td>
</tr>
</tbody>
</table>

*All dosages in mg/day. † Intention to treat (ITT)-results unless specified. ‡ CGI-efficacy index: minimal or no change in depression with no or non-interfering side effects. § <50% reduction in HDRS-score. ¶ HDRSxx >8

Abbreviations: CGI-I/S = Clinical global impression improvement/severity, DT = Dose tolerant sample, HDRSxx = Hamilton depression rating scale (xx denote number of items used), InP = Inpatients, ITT = Intention to treat, MDD = Major depressive disorder, MinD = Minor depression, N/A = not applicable, NNH/T = Number needed to harm/treat, NR = Nonresponders, OutP = Outpatients, PR = Partial responders, RCT = Randomized controlled trial, SE = side effects.

Drugs: CIT = citalopram, FLX = fluoxetine, IMI = imipramine, PAR = paroxetine, PLAC = placebo, SER = sertraline, TCA = Tricyclic antidepressants.
Benkert et al. investigated dose-escalation of paroxetine 20 mg/day in outpatients that were depressed or had minor depression. Those who did not respond after 3 weeks of treatment (n=86) were randomized to receive 40 mg paroxetine for 3 additional weeks or placebo-increase. Response rates were 75% in the placebo-increased group and 74% in the 40 mg group.47 Licht et al. investigated randomized dose-escalation of sertraline (up to 200 mg/day) versus sertraline 100 mg/day (placebo-increase) or mianserin addition in 295 outpatients non-responsive to sertraline 50 mg for 4 weeks and additionally increased to 100 mg for 2 more weeks. Response rates 5 weeks after randomization were significantly lower in the dose-increase group (56%) than in the sertraline 100 mg group (70%) and the mianserin-addition group (67%). Drop-outs due to side effects were not specified.50

Strengths, flaws and other details of all selected studies are provided in table 4.2. In summary we mention several methodological problems we encountered: absence of placebo controls,45-49 inclusion of minor depression,47 insufficient data-presentation,44-46 insufficient power,44-49 uncertainty about blinding,43-48 earlier dose-escalation before the randomisation,50 inadequate pooling of heterogeneous data50-53 and problems with conversion to dose-equivalents.51-53 None of the studies provided information about the method of dose-escalation nor described the early drop-out rates due to dose-escalation.

Evidence for dose-escalation?

From 4 of the 8 dose-escalation studies it appeared that dose-increments before 4 weeks were not effective (LoE: A2).43-45,47-49,51-53 However, in the meta-analysis of some of these studies by Baker et al., a potential dose-response relationship was found for dose-escalation if dropouts due to side effects were excluded from the analysis (a so called dose-tolerant sample).53 Baker et al. proposed that differential dropout due to side effects in the dose-escalation group (compared to placebo-increase) gave a substantial (negative) bias of the potential dose-response relationship. They argued that by applying a last-observation carried forward approach (often used in the original studies), more early drop-outs (due to side-effects) in the high-dose groups would unequally increase average severity scores and decrease response rates compared to the lower dose (or placebo) groups. This methodological problem could be overcome by analysing only dose-tolerant subjects (those not dropping out due to side effects).36

In the well performed study with sertraline by Licht et al. (not included in the three reviews), dose-escalation after 6 weeks was found to be less effective than continuation of the standard dose, or augmentation with mianserin (LoE: A2).50 After 8 weeks of treatment, increased dosages of fluoxetine were more effective than augmentation with lithium or desipramine,46-49 although in the latter study this was not significant (LoE: B). In these studies no placebo dose-escalation was performed. Both studies showed a non-significant trend of increased efficacy of dose-escalation compared to augmentation (lithium or desipramine) especially for partial responders (LoE: B).46-49

Across all studies, higher doses were related to increased dropout rates, which were associated with more side effects in some studies (LoE: A2).51 It appeared that the occurrence of side effects did not increase equally when dosages were gradually escalated in initial non-responders compared to fixed-dose trials. However, this could not be compared straightforwardly between the studies, and was not investigated specifically.

Additional concerns for clinicians

We identified no evidence to recommend on how dose-increase should be practiced. Also the maximum dosage to be achieved was not investigated well.
Discussion

Our systematic review provided 8 studies about dose-escalation in SSRIs. Only one of these approached our rather stringent criteria.\textsuperscript{50} We found no evidence for increased efficacy of dose-escalation within the first 4 weeks. Dose-escalation after 6 weeks appeared less effective than continuing the same dose. We found some, but limited evidence for efficacy of dose-escalation after 8 weeks, particularly in partial responders. This effect was seen within 4 weeks after dose-escalation. Irrespective of efficacy, dose-escalation unequivocally increased side-effects, but effects on drop-out rates due to side effects were less straightforward. Thus, in the absence of methodologically well designed studies we can neither unequivocally state that dose-escalation is useful, nor discard it as useless.

These findings may challenge the current beliefs and recommendations about dose-escalation, underscored by the fact that dose-escalation is generally practiced.\textsuperscript{33-34} Contrarily to this challenge, many patients that only partially responded are too often treated with long-term obviously insufficient treatments (e.g. standard doses of SSRIs). For these patients, one could argue that it is better to try dose-escalation than to continue inadequate treatment. Presumably, in the absence of clear guidance by trial-data, clinicians do not have many alternatives for non-responders or partial responders, and clinicians all have their case-histories of improvement after dose-escalation. A more sophisticated question must therefore also be asked: which subgroup of patients will benefit from dose-escalation?

So far, only the NICE guideline displayed some reserve in the general recommendation about dose-escalation.\textsuperscript{8} The British Medicines and Healthcare products Regulatory Agency’s Committee on Safety of Medicines examined the available evidence for dose-escalation as provided by pharmaceutical companies, and recommended the lowest efficacious dose (MHRA, 2004 Internet: www.mhra.gov.uk). From this report it was unclear which studies were taken as evidence. Three previous reviews concerning higher doses of antidepressants were published,\textsuperscript{51-53} of which the methodological shortcomings were already mentioned. The findings in these reviews previously challenged the belief of a dose-response relationship, but Baker et al. proposed a potential dose-response relationship, based on their dose-tolerant analysis. All reports summarized studies performed until 1997, thereafter the study by Licht et al. further challenged the efficacy of dose-escalation.\textsuperscript{50}

Limitations of the identified studies

Four major issues of concern in the 8 identified studies should be mentioned. First, the methodological quality of these studies varied between poor and good according to our classification. We summarized these methodological problems in the results section and table 4.2.

Second, and more in general, all dose-escalation studies (except the studies of Fava and colleagues, that lacked a placebo-control)\textsuperscript{45,46,49} suffered a methodological problem of the timing of dose-escalation.\textsuperscript{36} Even the most robust study by Licht et al. hampered its own design by a nonrandomized dose-increase 2 weeks prior to randomization.\textsuperscript{50} This problem might explain the high placebo response rates in some of the dose-escalation studies (up to 75%).

Third, in most studies no data were provided on the selective drop-out, nor the schedule of dose-increments.\textsuperscript{36} Because patients randomised to true dose-escalation might drop out more frequently and earlier after randomisation (with associated high severity scores) compared to those receiving placebo, this might have biased the intention to treat analyses in which last-observations are usually carried forward to study endpoints. This especially happens when dose-escalation is performed rapidly. The analysis of a dose-tolerant sample in such studies would indeed provide additional information, but these data were not provided.

Fourth, in the selected trials, mostly response was used as primary outcome, while currently remission of depression is the clinical aim of treatment.\textsuperscript{54} If dose-escalation would be effective, the question remains whether dose-escalation will also further improve initial responders that were nonremitters. So far only Schweizer et al. addressed this topic with equivocal effects of dose-escalation.\textsuperscript{48}
Possible explanations for a dose-response relationship

A possible explanation of the clinical observation that response might occur after dose-escalation, can be initial lower blood levels. This may be related to increased metabolism due to genetic polymorphisms of the cytochrome P450 (CYP P450) enzyme system. The incidence of increased metabolism by (multi-)duplicated genes of the CYP P450 2D6 genes varies between 1-2% in Swedish Caucasians, 3.6% in Germany and 7-10% in Spain and Sicily, and varies between ethnic groups (e.g. 29% in Black Ethiopians). A few studies showed equivocal evidence for the involvement of CYP P450 polymorphisms (responsible for rapid metabolism) as an explanation of non-response to a standard dose of SSRIs. However, a clear relationship between blood levels of SSRIs and response was never found. Perhaps, genetic variability of the central target of these drugs, the serotonin reuptake transporter, may be responsible more directly for the effects of SSRIs.

From in vitro and ex-vivo studies it appears that at higher doses selective antidepressants like SSRIs may become dual action agents that will also affect other monoamine systems like norepinephrine. From the current data of dose-escalation in SSRIs this theoretical hypothesis cannot be falsified nor proven. In addition, we are unaware of an acceptable method to test whether specific sites of action are responsible for the observed treatment-effects.

Limitations of the review

No meta-analysis was performed because the differences in timing of dose-escalation between the identified studies introduced substantial heterogeneity. An extension of the meta-regression approach as performed by Baker et al. was considered impossible to acknowledge this problem as the number of studies to explore this heterogeneity will lack power, moreover as age, sex, outcome definition and type of SSRI ideally should also be included in such a model.

The grading system for studies is not representing the appraised methodological dimensions of evidence. This improved the applicability of the results for busy clinicians, but reduced their strength.

Finally, patients studied in trials are generally selected populations, reducing external validity for the 'real world' clinical practice. All identified studies excluded psychotic depression, bipolar depression, depression in children or adolescents, and depressive disorder with severe psychiatric and somatic comorbidity.

Future dose-escalation studies

For future dose-escalation trials methodological issues should be considered. First, for optimal contrast in the study, an appropriate group of non-responders should be selected by postponing randomization and refraining from (additional) interventions before dose-escalation is applied. Six weeks is the minimum which can be reconciled with recommendations in current guidelines and is acceptable for clinical practice. Second, studies should have enough power to detect significant differences. This implies a large sample to start with, as approximately 50% of patients will show a response in the first 6 weeks. Third, the method of dose-escalation should be described and applied in a way that few patients drop-out. Fourth, adequate results should be presented: response and remission rates in intention to treat analyses and for the group that could be described as dose tolerant. Fifth, efficacy should be tested in predefined subgroups (e.g. partial responders at week 6). Sixth, genetic sampling (e.g. CYP P450 and SERT polymorphisms) and plasma SSRI level sampling would be interesting to further examine potential explanations for the clinically observed efficacy of dose-escalation and to identify potential prognostic variables.

Clinical implications and conclusion

For clinicians this review may challenge the belief in dose-escalation of SSRIs for patients who show insufficient response to a standard dose. This review points out that dose-escalation of SSRIs in a standard dose before the fourth week of treatment does not seem to be more
beneficial for patients. Thereafter dose-escalation is one of the easiest thing to do, but – to our opinion – should not be performed obligatory, as the evidence for this strategy is not compelling. Contrary dose-escalation appears more adequate to try than to continue an inadequate treatment, maybe partial responders will benefit most. Side-effects will most likely increase and might induce drop-out of treatment. For more clinical guidance well-performed new dose-escalation studies are needed. In order to fill this gap, for the introduction and registration of (new) antidepressants, clinical studies with an adequate dose-escalation design should be required by the U.S. Food and Drug Administration (FDA) and the European Agency for the Evaluation of Medicinal Products (EMEA).

Acknowledgements
This systematic review was realised by a grant for the development of a local evidence-based clinical practice guideline (No SFA.07.012) from the Academic Medical Center, Amsterdam, the Netherlands. This guideline considering strategies for non-response to a standard dose of a first SSRI is available on request.

Conflicts of interest
None

References


SWITCHING ANTIDEPRESSANTS AFTER A FIRST SELECTIVE SEROTONIN REUPTAKE INHIBITOR IN MAJOR DEPRESSIVE DISORDER: A SYSTEMATIC REVIEW

Henricus G. Ruhé
Jochanan Huyser
Jan A. Swinkels
Aart H. Schene

Program for Mood Disorders, Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
Abstract

Background
Selective Serotonin Reuptake Inhibitors (SSRIs) are frequently used as a first antidepressant for major depressive disorder, but have response-rates of 50-60% in daily practice. For patients with insufficient response to SSRIs, switching is often applied.

Aim
To systematically review the evidence for switching pharmacotherapy after a first SSRI.

Methods
A systematic literature search (updated until Feb. 10, 2005) in MEDLINE, EMBASE, CINAHL and PsycInfo (all indexed years) identified randomised controlled trials (RCTs) and open studies investigating switching strategies. In the absence of specific keywords for switching, we performed “sensitive” searches using free text words with wildcards ($) or (“alternat$” adj5 “treat$”) or (“alternat$” adj5 “therap$”) in combination with the Cochrane Collaboration search filter for RCTs, the Cochrane Collaboration Depression Anxiety and Neurosis Group search filter for major depressive disorder, and MeSH terms for antidepressants (in combination with additional text words for all antidepressive agents). Additionally, we included 4 recent STAR*D publications. We limited searches to adults and humans but did not apply language restrictions. Relevant articles were retrieved and critically appraised. The methodology of the studies, the results on efficacy and dropout due to side effects, and remarks were summarized in an evidence table. Three studies comparing a switch to venlafaxine or SSRIs were pooled.

Results
Eight RCTs and 23 open studies were identified, studying populations with different levels of treatment resistance. Definitions of response and remission rates varied between studies. Observed response rates after switching to any of the classes of antidepressants varied between 12% and 86%. Remission rates varied between 7% and 82%. The number of previous treatments with antidepressants was negatively correlated with treatment outcome. Rates of dropout due to side-effects varied considerably across agents (5-39%). Switching to venlafaxine showed a modest and clinically equivocal benefit over SSRIs (Number Needed to Treat = 13 (9.1-25.0))

Conclusions
After a first SSRI any switch within or between classes of antidepressants appear legitimate (second SSRI, novel dual acting antidepressants, selective noradrenergic or noradrenergic/dopaminergic agents, or TCA or mianserin). No unequivocal evidence is available to prove an advantage of a between class switch. More guidance by randomized empirical studies is needed. Clinical implications and methodological considerations for future studies are discussed.
**Introduction**

Major depressive disorder (MDD) is one of the most prevalent and disabling illnesses in psychiatry.¹ For the treatment of MDD, several national clinical guidelines were developed.²⁻⁹ In these guidelines, pharmacotherapy is among the most important treatments; mostly Selective Serotonin Reuptake Inhibitors (SSRIs) are the antidepressants of first choice. However, only 50 to 60% of patients respond to the first antidepressant given.¹⁰⁻¹¹ In a case of non-response, all treatment guidelines recommend three major strategies: 1) increasing the dose of the antidepressant (dose-escalation), 2) switching to another antidepressant of the same or different class, and 3) augmenting the antidepressant by adding a second drug that by itself is not an antidepressant. By various authors, a fourth strategy of combination of antidepressants is proposed.¹²⁻¹⁵

Surprisingly very little systematic evidence exists to date to underscore the recommendations for non-responders. One Cochrane review summarizes randomized, controlled trials (RCTs) of strategies in patients non-responsive to at least 4 weeks of an antidepressant at the recommended dose.¹⁶ With a thorough methodology, 16 RCTs were selected. Unintentionally, the studies included in this review represented more heterogeneous, difficult-to-treat populations, referred to as treatment-resistant depression (TRD). Although little information on previous treatments was found, the included studies especially considered tricyclic antidepressant (TCA) non-responders. The switch-options that were investigated in the included studies did not reflect clinical practice of switching to another antidepressant (one of the above recommendations), but used a variety of other drugs (oestrogen, benzodiazepines, ketoconazole, olanzapine). For the augmentation studies, meta-analyses were performed with 2 trials of lithium-augmentation and 3 pindolol-trials. A clinically significant benefit was found only for lithium augmentation. Thus, this review does not provide helpful information for clinicians in the case of non-response to a (first) SSRI.

Strategies for non-response have been summarized in several narrative reviews, focusing on all strategies together,¹⁷⁻²⁸ switching,¹²⁻¹³⁻²⁹⁻³² augmentation,¹²⁻¹³⁻³¹⁻³² or combination.¹²⁻¹⁵⁻³³ Dose escalation was summarized in 2 meta-analyses,³⁴⁻³⁵ 1 narrative,³⁶ and 1 recent systematic review.³⁷ The evidence for lithium augmentation was also summarized in meta-analyses by Bauer et al.³⁸⁻³⁹

After dose-escalation, switching antidepressants is widely practiced.⁴⁰⁻⁴² Switching to a different pharmacological class seems to be preferred by clinicians.⁴³ The above narrative reviews of switching strategies altogether provided a substantial overview. However, each review individually was limited in its presentation, predominantly by a lack of a well defined search strategy, and none of the reviews presented data on critical appraisal of the identified studies as proposed by the Cochrane collaboration.⁴⁴ The general conclusion today is that there is limited evidence available for switching antidepressants, and that there is no clear proven advantage of one switch option over the others. Additionally, recently the results of a large study designed to elucidate sequential treatment strategies after non-response became available (Sequenced Treatment Alternatives to Relieve Depression, STAR*D).⁴⁵ This study provided prospective data on response and remission rates after randomized treatment allocations in patients who were not in remission after one to four sequential steps of treatment (further referred to as STAR*D level I-IV).

Therefore, our primary objective was to systematically review and appraise the available research focusing on switching strategies for SSRI non-responders in MDD, including the recent STAR*D results. A secondary aim was to acknowledge and investigate the expected different levels of TRD as a source of variation between studies. Our principal question was whether the available evidence justifies distinct recommendations for next-step strategies after non-response to a first SSRI. We performed a systematic review following the Cochrane methodology and performed a meta-analysis of two switch options after a first SSRI: a second SSRI versus a serotonin-norepinephrin reuptake inhibitor (SNRI).
Methods

Studies included in the review

We expected very few randomised, controlled, switch-studies a-priori, despite the widespread availability of SSRIs during the last decade. As best-available evidence, we included open and randomized studies in which at least 50% of participants used an SSRI previously in the current depressive episode. Thus, we excluded studies describing switching from TCAs to SSRIs. Studies performed in populations with TRD were also included if previous use of an SSRI (in ≥50% of subjects) was unambiguously documented.

Identification and selection of articles

We performed systematic literature searches (updated until February 10th 2005) in four databases (MEDLINE, EMBASE, CINAHL, and PsychInfo; all indexed years). In the absence of specific keywords for switching, we performed ‘sensitive’ searches using free text words with wildcards ($)：“switch$” or (“alternat$” adj5 “treat$”) or (“alternat$” adj5 “therap$”) in combination with the Cochrane Collaboration search-filter for RCTs, the Cochrane Collaboration Depression Anxiety and Neurosis group search-filter for MDD and MeSH-terms for antidepressants (in combination with additional text words for all antidepressive agents). We limited searches to adults and humans, but did not apply language restrictions. Full queries are available on request. In addition, we included four identified studies released after these searches, including three studies from the STAR*D trial.46-49

The first and second authors (H.G.R., J.H.) independently screened titles and abstracts and selected articles on the basis of design and focus on switching antidepressants after SSRI-treatment. Agreement on exclusion of irrelevant articles was 99.1%, with a Cohen’s κ for interrater agreement of 0.62 (κ values between 0.45-0.75 indicate ‘substantial’ agreement; values above 0.75 indicate ‘almost perfect’ agreement).50 We resolved discrepancies between initial selection by discussion and consensus.

The first author (H.G.R.) judged all potentially relevant articles according to specific inclusion and exclusion criteria (full criteria available on request). In case of doubt the article was fully read and assigned thereafter. We retrieved additional cross-references, and checked reference lists of identified narrative reviews. We considered double-publications together to reveal the maximum of available information.

Critical Appraisal and summary

The first author (H.G.R.), a certified epidemiologist, critically appraised and abstracted the articles, using standardized forms derived from the Dutch Institute of Healthcare Improvement and the Agency of Healthcare Policy and Research (AHCPR).5 We used the same items for critical appraisal as proposed by the Scottish Intercollegiate Guidelines Network and Sackett et al.53 We assigned a ‘level of evidence’ (LoE; see table 4.1) to each study. Levels of evidence are based upon the methodological robustness of studies. In the results section, the LoE of the supporting scientific evidence (A1-D) is indicated. We extracted data on efficacy and tolerability from each study. As primary efficacy outcome we took the percentage response or remission on an intention to treat (ITT) basis. If several scales were used, we a priori preferred data for the Hamilton depression rating scale (HDRS; 17-item version or other versions); otherwise we used data from the Montgomery Åsberg depression rating scale (MADRS),55 clinical global impressions scale (CGI),56 or other applied scales (e.g. the 16-item quick inventory for depressive symptoms-self-rated (QIDS-SR16)).57 For tolerability, we took the dropout rate due to side effects as primary measure, followed by the overall dropout rate.

To assess judgement-bias by one person who performed the critical appraisal, we measured interrater variation in a slightly different set of 12 publications. Every other author (J.H., J.A.S., A.H.S.) critically appraised 4 publications. Cohen’s κ values for the appraisal-items were 0.49 (for ‘validity of the study’) and 0.86 (for ‘concealment of allocation’), while complete agreement
existed for the appraisal-items 'randomization of the study', 'level of evidence' and 'data extraction' ($\kappa = 1.0$). These results are in line with other reports of interrater agreement in appraisal of psychiatric research.\textsuperscript{34}

We first described a qualitative summary with discussion of the results, restrictions, methodological flaws and external validity of the studies in an evidence table and a separate document, of which a summary is provided in this article. For each study we indicated the level of treatment resistance as proposed by Thase et al.\textsuperscript{10,24} If possible, we calculated risk-differences and corresponding Numbers Needed to Treat (NNT) and Harm (NNH), with 95\% confidence intervals (95\% CIs). Because of the lack of homogenous randomised studies, we refrained from pooling in a meta-analysis, except for the three studies comparing the venlafaxine (SNRI) vs. second SSRI switch. We grouped antidepressants into six classes following the classification of the AHCPR.\textsuperscript{5}

## Results

We selected thirty-one studies for this review. Figure 5.1 shows the search results and selection of studies. Table 5.1 summarizes the included studies. A table of 8 excluded studies\textsuperscript{35,66} is available on request.

**Figure 5.1.** Selection of reported studies.

---

Abbreviations: BUP = bupropion, MAO-I= irreversible inhibitor of monoamine-oxidase, MIAN = mianserin, MIR = mirtazapine, NEF = nefazodone, REB = reboxetine, RIMA = reversible inhibitor of MAO, SSRI = selective serotonin reuptake inhibitors, TCA = tricyclic antidepressant, VLX = venlafaxine
### Table 5.1. Effectiveness of switching after ≥1 SSRI: Selected studies.

<table>
<thead>
<tr>
<th>Study (year) reference</th>
<th>LoE</th>
<th>N</th>
<th>Design (follow-up)</th>
<th>Intervention*</th>
<th>Comparison*</th>
<th>Outcome†</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baldomero et al. (2005)</td>
<td>C</td>
<td>112</td>
<td>MDD OutP TRD-I</td>
<td>SER 50-200mg</td>
<td>-</td>
<td>Response (CGI-I ≤ 2) 71.8% Dropout overall = 21.4%; Dropout SE = 9.8%</td>
<td>No placebo control. FLX-intolerance before switch only.</td>
</tr>
<tr>
<td>Brown et al. (1995)</td>
<td>C</td>
<td>55</td>
<td>MDD OutP TRD-I</td>
<td>CIT 20-40mg</td>
<td>-</td>
<td>Response (CGI-I ≤ 2) 65-5% Dropout overall = 5% Dropout SE = 0%</td>
<td>No placebo control. FLX-intolerance before switch only. No minimum level of HDRS required for study entrance. Start of CIT after placebo wash-out of 2-4 weeks.</td>
</tr>
<tr>
<td>Calabrese et al. (2003)</td>
<td>C</td>
<td>55</td>
<td>MDD OutP TRD-I</td>
<td>FLX (n= 12) 20-40mg</td>
<td>-</td>
<td>Response (CGI-I ≤ 2) 51% No significant differences between various combinations of switching. No SE reported.</td>
<td>No placebo control. Methodologically poor: retrospective study, unclear definition of initial non-response, small numbers, no characteristics of population. Unclear after how many weeks response was determined.</td>
</tr>
<tr>
<td>Joffe et al. (1996)</td>
<td>C</td>
<td>55</td>
<td>MDD OutP TRD-I</td>
<td>FLX (n= 12) 20-40mg</td>
<td>-</td>
<td>Response (CGI-I ≤ 2) 63% Dropout overall = 46%; Dropout SE = 5%</td>
<td>No placebo control. FLX-intolerance before switch only.</td>
</tr>
<tr>
<td>Poirier et al. (1999)</td>
<td>C</td>
<td>277</td>
<td>MDD OutP TRD-I</td>
<td>SER 50-200mg</td>
<td>-</td>
<td>Remission (HDRS&lt;sub&gt;24&lt;/sub&gt; ≤ 7) VLX = 24.8%, BUP = 21.3%, SER = 17.6%, NNT&lt;sub&gt;V LX, SER&lt;/sub&gt; = 14 (7.0-∞), NNT&lt;sub&gt;BUP, SER&lt;/sub&gt; = 28 (9.3-∞), NNT&lt;sub&gt;V LX, BUP&lt;/sub&gt; = 29 (9.2-∞) Response (≥50% ↓ in HDRS&lt;sub&gt;24&lt;/sub&gt;) VLX = 28.2%, BUP = 26.1%, SER = 26.7%, NNT&lt;sub&gt;V LX, SER&lt;/sub&gt; = 66 (10.6-∞), NNT&lt;sub&gt;BUP, SER&lt;/sub&gt; = 189 (13.4-∞) NNT&lt;sub&gt;V LX, BUP&lt;/sub&gt; = 49 (10.1-∞) Dropout overall = 18%; Dropout SE = 5%.</td>
<td>Unblinded study, blinded assessors, no placebo-group. Methodologically well performed effectiveness trial. Participants all received CIT 20-60mg as prior treatment (level II STAR*D). Due to high doses of CIT in 'level 1' (407/727 (56%) of subjects in this study were classified as CIT-intolerant. No washout applied.</td>
</tr>
<tr>
<td>Thase et al. (1997)</td>
<td>C</td>
<td>106</td>
<td>MDD OutP TRD-I</td>
<td>CIT 20mg (dosages could be increased to 60mg)</td>
<td>-</td>
<td>Response (≥50% ↓ in HDRS&lt;sub&gt;24&lt;/sub&gt;) Overall 62%; initial-intolerant subjects 71%, initial nonresponders 58%.</td>
<td>No placebo control. Nonresponse to SER was determined retrospectively (in most patients).</td>
</tr>
<tr>
<td>Thase et al. (2001)</td>
<td>C</td>
<td>57</td>
<td>MDD OutP TRD-I</td>
<td>FLX 20-60mg</td>
<td>-</td>
<td>Response (CGI-I ≤ 2) 63% Dropout overall = 18%; Dropout SE = 5%</td>
<td>Well performed open study. Unknown placebo response rate. Tolerance for CIT after FLX is good, despite direct switch. No increased rate of SE in first weeks of study.</td>
</tr>
</tbody>
</table>

*[designational abbreviations are explained in Section 5.2.C]
### Table 5.1: Effectiveness of switching after ≥1 SSRI: Selected studies. (Continued)

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>LoE</th>
<th>N</th>
<th>Design (follow-up)</th>
<th>Intervention*</th>
<th>Comparison*</th>
<th>Outcome†</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>5.1.B. to a TCA or mianserin</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fava et al. (2006)</td>
<td>A2</td>
<td>234</td>
<td>Multi center unblinded RCT (14 weeks)</td>
<td>NOR 50-200mg</td>
<td>MIR 15-60mg</td>
<td>Remission (HDRS≤7) NOR = 19.8%, MIR = 12.4% NNT_{MIR-NOR} = 14 (6.0-∞)  Not different in level II intolerant group. Response (≥50% in QIDS≤8) NOR = 16.5%, MIR = 15.5% NNT_{MIR-NOR} = 32 (8.1-∞)  Dropout SE NOR = 34.7%, MIR = 33.3%</td>
<td></td>
</tr>
<tr>
<td>Ferreri et al. (2001)</td>
<td>B</td>
<td>103</td>
<td>3 arm multicenter RCT of FLX 20mg-NR (6 weeks)</td>
<td>1. MIAN 60mg 2. FLX 20mg + MIAN 60mg 3. FLX 20mg</td>
<td></td>
<td>Response (≥50% in HDRS≤7) Mianserine = 48.5%, Fluoxetine = 37% NNT_{MIA-NF-FLX} = 9 (2.9-∞); NNT_{MIA+NF-FLX} = 12.1 (4.1-∞); Remission (HDRS≤8) Mianserine = 36%, Fluoxetine = 18% NNT_{MIA+NF-FLX} = 6 (2.6-∞); Dropout SE NHM_{MIA-F} = 5 (2.6-10.4); NHM_{MIA+NF-FLX} = 16 (6.8-∞)  Dropout overall NHM_{MIA-F} = 6 (2.6-∞); NHM_{MIA+NF-FLX} = 4 (4.9-∞)</td>
<td></td>
</tr>
<tr>
<td>Nierenberg et al. (2003)</td>
<td>C</td>
<td>92</td>
<td>Open phase of NOR treatment preceding a 2nd RCT (6 weeks)</td>
<td>NOR 100mg (adjusted to achieve 100 ng/ml)</td>
<td></td>
<td>Response (≥50% in HDRS≤7) = 42.4% Remission (HDRS≤5) = 11.9% Dropout overall = 34.7%</td>
<td></td>
</tr>
<tr>
<td>Nolen et al. (1988)</td>
<td>C</td>
<td>31</td>
<td>Blinded consecutive therapy after 4 weeks of FLV (4 weeks)</td>
<td>OXA100-300mg</td>
<td></td>
<td>Response (≥50% in HDRS≤7) = 38.7% Relapse = 19.4% within 6 months</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- **C**: Crossover study
- **MDD**: Major Depressive Disorder
- **OutP**: Outpatients
- **TRD-I**: Treatment-resistant depression, first episode
- **TRD-II**: Treatment-resistant depression, second episode
- **TRD-III**: Treatment-resistant depression, third or later episode
- **LoE**: Level of Evidence
- **Design (follow-up)**: Description of the study design and follow-up period
- **Intervention**: Description of the intervention(s) used
- **Comparison**: Description of the comparison(s) used
- **Outcome**: Outcome measures used in the study
- **Remarks**: Additional notes or remarks about the study
### Table 5.1. Effectiveness of switching after ≥1 SSRI: Selected studies. (Continued)

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>LoE</th>
<th>N</th>
<th>Design (follow-up)</th>
<th>Intervention*</th>
<th>Comparison*</th>
<th>Outcome†</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peselow et al. (1989)</td>
<td>C</td>
<td>15 + 10 MDD OutP TRD-I</td>
<td>Blinded cross-over design with original randomization (6 weeks)</td>
<td>IMI 65-275mg</td>
<td>-</td>
<td>Response (≥50% ↓ in HDRSunknown or CGI-I ≤2) PAR → IMI = 73%</td>
<td>Methodologically poor: unclear description of studied population, limited presentation of data. Data of initial nonresponders to IMI switched to PAR also provided.</td>
</tr>
<tr>
<td>Thase et al. (2002)</td>
<td>C</td>
<td>117 Chron MDD, MDD + Dys OutP TRD-I</td>
<td>Blinded, multicenter cross-over design with original randomization (12 weeks)</td>
<td>IMI 50-300mg</td>
<td>-</td>
<td>Response (CGI-I ≤2, ≥50% ↓ in HDRS15 and CGI-I ≤3) SER → IMI = 44.4% Remission (HDRS15 ≤7 and CGI-I ≤2) SER → IMI = 23% Dropoutoverall = 25%; Dropout SE = 9%</td>
<td>Well-performed study. Unknown placebo response rate. Data of initial nonresponders to 12 weeks IMI switched to SER also provided. Due to absence of second randomization only tentative comparisons with switch IMI → SER available.</td>
</tr>
</tbody>
</table>

#### 5.1.C. to mirtazapine, nefazodone, venlafaxine (dual action agents)

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>LoE</th>
<th>N</th>
<th>Design (follow-up)</th>
<th>Intervention*</th>
<th>Comparison*</th>
<th>Outcome†</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baldomero et al. (2005)</td>
<td>B</td>
<td>3097 MDD + MinD + Dys OutP TRD-I</td>
<td>Multicenter, open design; RCT of VLX vs CA in SSRI-NR or intolerant patients (24 weeks)</td>
<td>VLX 75-225mg</td>
<td>CA: FLX, PAR, CIT 20-60mg SER 50-200mg MIR 15-45mg</td>
<td>Response (≥50% ↓ in HDRS17): VLX = 77.3%, SSRIs = 71.1% NNTVLX-SSRI = 17 (10.5-35.0) Remission (HDRS17 ≤5): VLX = 5.9%, SSRIs = 52.1% NNTVLX-SSRI = 14 (9.1-29.3) HDRS17-scores differ significantly but clinically irrelevant at week 12 and 24 Dropoutoverall: VLX = 19.6%, CA = 23.3% NNHVLX-CA = 27 (15.1-120) Dropout SE: VLX = 2.3%, CA = 1.7% NNHVLX-CA 161 (62.1-∞)</td>
<td>Large randomized but unblinded study. Some methodological problems. 63.3% of included patients previously used a SSRI. Inclusion of 8.7% MinD. No differentiation of first-SSRI intolerant and unresponsive patients. Modified ITT analysis of ≥week 4 completers; in CA-treated switch group 22.7% received a non-SSRI. Baseline HDRS-scores sign. higher in VLX-group; differential loss to follow up 26.2% VLX vs 36.2% CA Only 3 time-points over 24 weeks. No separate dichotomous data for week 12 response/ remission</td>
</tr>
<tr>
<td>Fava et al. (2000)</td>
<td>81</td>
<td>94 MDD OutP TRD-I</td>
<td>Multicenter, open design (RCT of direct switch vs. washout) (8 weeks)</td>
<td>MIR 15-45mg</td>
<td>-</td>
<td>Response (≥50% ↓ in HDRS17) SSRI NR = 48%, SSRI Intol. = 53% Dropoutoverall = 43% Dropout SE = 26%</td>
<td>Methodologically sound. Unknown placebo response rate. Washout-phase offers no advantages.</td>
</tr>
<tr>
<td>Kaplan (2002)</td>
<td>C</td>
<td>73 MDD OutP TRD-I</td>
<td>Retrospectivenaturalistic study of SSRI-NR or non-sustaining SSRI responders (6-8 weeks + follow-up)</td>
<td>VLX 50-400mg</td>
<td>-</td>
<td>Response (HDRS25 ≤10 and CGI-21 ≥5) = 86% Remission (HDRS25 ≤8) = 82% Dropout SE = 5.5%</td>
<td>Methodologically very poor open, unblinded design, 1 researcher, retrospective data obtained. Mild depression included also. Unclear, but probable selection bias. Recruitment of SSRI responders who did not sustain their response (52%) might increase response rate. ITT-results not mentioned in study.</td>
</tr>
</tbody>
</table>
Table 5.1. Effectiveness of switching after ≥1 SSRI: Selected studies. (Continued)

<table>
<thead>
<tr>
<th>Study (year) reference</th>
<th>LoE</th>
<th>N</th>
<th>Design (follow-up)</th>
<th>Intervention*</th>
<th>Comparison*</th>
<th>Outcome†</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mischoulon et al. (2004)83</td>
<td>C</td>
<td>13</td>
<td>MDD OutP TRD–&gt;I</td>
<td>NEF 300-600mg</td>
<td>-</td>
<td>Response (&gt;50% ↓ HDRS6 and/or CGI-S ≤ 2) 31% Dropout SE = 39%</td>
<td>Small sample (pilot study). Unknown placebo response rate; 61.5% attrition; especially in previous FLX-users. Washout of 4-7 days applied. Heterogeneous group of TRD app. 46% ± stage II. No sign. differences in response rates and side effects compared to 13 patients treated with NEF as first applied antidepressant (but low power).</td>
</tr>
<tr>
<td>Mitchell et al. (2000)84</td>
<td>C</td>
<td>312</td>
<td>MDD Setting? TRD–&gt;I</td>
<td>VBX 75-300mg</td>
<td>-</td>
<td>Response (&gt;50% ↓ in MADRS) = 52.6% Remission (MADRS &lt;12) = 40.7% Dropout SE = 11%</td>
<td>Methodologically well. Unclear setting. Unknown placebo response rate. Unclear which proportion used ≥1 SSRI (41-68%). Probably chronically depressed subjects.</td>
</tr>
<tr>
<td>DE Montigny et al. (1999)85</td>
<td>C</td>
<td>152</td>
<td>MDD InP &amp; OutP TRD–I</td>
<td>VLX 75-375mg</td>
<td>-</td>
<td>Response (&gt;50% ↓ in HDRS21) = 58%; (≥50% ↓ in MADRS) = 62%; (CGI-I ≤ 3) = 66% Remission (≥75% ↓ in HDRS21) = 21% Dropout SE = 7.9%</td>
<td>TRD: at least 1 previous TCA or SSRI or MOC or TRAZ or NEF, majority of patients used a SSRI. Mean duration of episode 2 years (range 2 months - 12.5 years). Unknown placebo response rate.</td>
</tr>
<tr>
<td>Nierenberg et al. (1994)86</td>
<td>C</td>
<td>70</td>
<td>MDD InP &amp; OutP TRD–III</td>
<td>VLX 50-450mg</td>
<td>-</td>
<td>Response (&gt;50% ↓ in HDRS21) = 32.9%; (≥50% ↓ in MADRS) = 30%; (CG-I ≤ 3) = 40% Remission (≥75% ↓ in HDRS21) = 15.7%; (MADRS &lt;12) = 18.6%; (CGI-I = 1) = 22.9% Dropout SE = 9.6%</td>
<td>TRD: at least 3 drugs of ≥2 different classes, ≥1 TCA and ≥1 augmentation. Unclear what proportion of patients used ≥1 SSRI. Chronically depressed group (median duration 2.5 years). Unknown placebo response rate.</td>
</tr>
<tr>
<td>Poirier et al. (1999)74;87</td>
<td>A2</td>
<td>123</td>
<td>MDD InP &amp; OutP TRD–II</td>
<td>VLX 75-375mg</td>
<td>-</td>
<td>Response (&gt;50% ↓ in HDRS21): VLX = 45.0%, PAR = 29.0% NNTVLX-PAR = 7 (3.0-∞) Remission (HDRS21 &lt;10): VLX = 36.7%, PAR = 17.7% NNTVLX-PAR = 6 (2.9-8.9) Dropout SE: VLX = 8.2%, PAR = 4.8% NNNHLX-PAR = 30 (8.3-∞)</td>
<td>Methodologically well. Short follow-up during study. Dosing-schedules are different between VLX and PAR.</td>
</tr>
<tr>
<td>Reynaert-Dupuis et al. (2002)88</td>
<td>C</td>
<td>688</td>
<td>MDD InP &amp; OutP TRD–I</td>
<td>VLX 75-375mg</td>
<td>-</td>
<td>Response (&gt;50% ↓ in HDRS21) = 41% (in previous SSRI-treated patients)</td>
<td>86.3% of patients were switched to VLX due to inefficacy, of 41.7% who were switched from a SSRI separate response rates are given. Immediate switching applied (except from MAO-Is). Unclear presentation of data. Dropout rate not mentioned. Type of previous SSRI not significantly affected VLX- efficacy.</td>
</tr>
</tbody>
</table>

Rush et al. (2006)47 See under 5.1.A
### Table 5.1. Effectiveness of switching after ≥1 SSRI: Selected studies. (Continued)

<table>
<thead>
<tr>
<th>Study (year) reference</th>
<th>LoE</th>
<th>N</th>
<th>Design (follow-up)</th>
<th>Intervention*</th>
<th>Comparison*</th>
<th>Outcome†</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saiz-Ruiz et al. (1998)(^{89})</td>
<td>C</td>
<td>69</td>
<td>Multicenter, naturalistic, open design ((24 \text{ weeks}))</td>
<td>VLX 75–375mg</td>
<td>-</td>
<td>Response (≥50% \text{ in HDRS}<em>17) = 69.6%; (\text{CGI-I} \leq 2) = 63.8% Dropout (</em>\text{overall} = 30.4%); Dropout (_\text{SE} = 8.7%); (\text{SE occur in } 5.4%)</td>
<td></td>
</tr>
</tbody>
</table>

| Wan et al. (2003)\(^{90}\) | C | 24 | Retrospective chart review of consecutive subjects who failed response to ≥1 TCA and ≥1 SSRI \((2 \text{ weeks} - 3 \text{ years})\) | MIR 15–45mg | - | Response \((\text{CGI-I} \leq 2) = 20.8\%\); Dropout \(_\text{SE} = 20.8\%\) |

5.1.D. to reboxetine or buproprion (selective NRI or noradrenergic & dopaminergic agents)

| Fava et al. (2003)\(^{94}\) | C | 128 | Two center open design of prospectively determined FLX-NR \((8 \text{ weeks})\) | BUP-SR 150–400mg | - | Modified ITT \((n=26)\) Response \((\text{CGI-I} \leq 2) = 45.3\%\); Dropout \(_\text{SE} = 13.3\%\) |

| Fava et al. (2003)\(^{92}\) | C | 29 | Multicenter, open design of FLX-NR \((8 \text{ weeks})\) | REB 8–10mg | - | Response \((\text{CGI-I} \leq 2) = 44.5\%\); Dropout \(_\text{SE} = 13.3\%\) |

| Walker et al. (1993) \(^{93}\) | C | 39 | Open design of patients with sexual side-effects on FLX \((\text{partially FLX-NR}; n=16) \((8 \text{ weeks})\) | BUP 150–450mg | - | All \((n=36); \ \text{ HDRS}_{28}: 25.4 (\pm 5.8) \to 10.9 (\pm 10.8)\) |

Initial design of switching due to sexual side-effect; limited data provided for depressed subjects; e.g. no response rates; unknown placebo response rate. Improvement of orgasm function (84\%), satisfaction (78\%) and libido (78\%) after switch. 2 weeks washout applied; disappearance of sexual dysfunction linear with FLX-washout.

Methodologically very poor. In 17.2\% of eligible patients data were insufficient for inclusion. Highly treatment resistant population \((\text{average 7 previous drug-trials \[range 2-13\]}\}). Unclear what response indicated switch to MIR. CGI data determined by chart review. Chronic depression in 45.8\%. High level of comorbidity with anxiety disorders. Comedication with antidepressants and antipsychotics in 41.7\%.
Table 5.1. Effectiveness of switching after ≥1 SSRI: Selected studies. (Continued)

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>LoE</th>
<th>N</th>
<th>Design (follow-up)</th>
<th>Intervention*</th>
<th>Comparison†</th>
<th>Outcome‡</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGrath et al. (2006)⁴⁹</td>
<td>B</td>
<td>109 MDD</td>
<td>Multicenter unblinded RCT (12 weeks)</td>
<td>TCP 10-60mg</td>
<td>VLX-XR 75-300mg + MIR 15-45mg</td>
<td>Remission (HDRS₁₇ ≤ 7) TCP: 6.9%, VLX-MIR: 13.7% NNH = 15 (5.5-∞) Response (&lt;50% ↓ in HDRS) TCP: 12.1%, VLX-MIR: 23.5% NNH = 9 (3.9-∞) Dropout TCP: 4.1%, VLX-MIR: 21.6% NNH = 6 (2.7-3.5)</td>
<td>Unblinded study, blinded assessors. Participants received at least 1 SSRI, a 2nd SSRI or VLX or BUP or CIT augmentation (BUP or BUS) and some CBT and a 3rd treatment with NOR or MIR (level IV STAR*D).</td>
</tr>
<tr>
<td>Nolen et al. (1985)⁹⁶</td>
<td>B</td>
<td>26 MDD InP TRD-II</td>
<td>Randomized unblinded crossover design (4 weeks)</td>
<td>TCP 20-100mg</td>
<td>L₅HTP 20-200mg</td>
<td>Response (&lt;50% ↓ in HDRS₁₇) TCP: 42.9%, L₅HTP: 0% NNT = 3 (1.5-5.9) tranylcypramine side effects = 62% cardiovascular; 15% dropout</td>
<td>Small study, allocation of initial treatment not clearly described. Study-groups differ significantly in baseline HDRS-score. Stage II-III TRD-patients. In the paper data of second 4 weeks (cross-over phase) are also given. Limited presentation of data.</td>
</tr>
<tr>
<td>Nolen et al. (1988)⁹⁵</td>
<td>B</td>
<td>≥1 MDD InP TRD-II</td>
<td>RCT with secondary cross-over (4 weeks)</td>
<td>TCP 40-100mg</td>
<td>NOM 150-250mg</td>
<td>Response (&lt;50% ↓ in HDRS₁₇) TCP: 45.5%, NOM: 10.0% NNT = 3 (1.4-154.8) TCPSE = 58% cardiovascular, 33% insomnia</td>
<td>Stage II-III TRD-patients. Unclear method of randomization. In the paper data of second 4 weeks (cross-over phase) are also given. Limited presentation of data.</td>
</tr>
</tbody>
</table>

* All dosages in mg/day. † Intention to treat (ITT) results unless specified.

Abbreviations: AD = antidepressant, CA = ‘Conventional antidepressants’ (PAR, FLX, SER, CIT, MIR and other antidepressants not specified), CBT = Cognitive Behavior therapy, CGI-I = Clinical global impression –improvement scale, Dys = dysthymia, HDRS₁₇ = Hamilton depression rating scale ( rift indicates number of items), InP = In-patients, MADRS= Montgomery Åsberg Depression Rating Scale, MDD = major depressive disorder, MinD = Minor Depression, NH/T = number needed to harm/treat, NR = nonresponders, OCD = Obsessive compulsive disorder, OutP = Out-patients, PGI = Patient global improvement scale, QIDS-SR = Quick inventory of depressive symptomatology self-rated, SchA = Schizoaffective disorder, SE = Side-effects, TRD = treatment resistant depression, stages I: failure of 1 adequate trial of 1 major AD class; II: Stage I + failure of an adequate trial of 1 distinctly different AD class ; III: stage II + failure of an adequate trial of TCA; IV: Stage III + failure of an adequate trial of an MAOI; V: Stage IV +failure of a course of bilateral ECT; Drugs: BUP = bupropion, BUS = buspirone, CIT = citalopram, FLV= fluvoxamine, FLX = fluoxetine, IMI = imipramine, L₅HTP = L-5-hydroxytryptophan, MAO-I = irreversible inhibitor of monoamine-oxidase, MIR = mirtazapine, MOC = moclobemide, Nef = nefazodone, NOM = nomifensine, NOR = nortriptyline, NRI = noradreneline reuptake inhibitor, OXA = oxaprotiline, PAR = paroxetine, PLAC = placebo, REB = reboxetine, SER = sertraline, SSRI = selective serotonin reuptake inhibitor, TCA = tricyclic antidepressants, TCP = tranylcypromine, TRAZ = trazodone, VLX = venlafaxine. |
Second SSRI

We identified 7 open studies investigating a switch to a second SSRI.\textsuperscript{57,68,69-73} In one of these studies, the non-response to the initial SSRI was determined prospectively, and switching was applied immediately.\textsuperscript{69} In 4 studies, intolerance was determined retrospectively, with an (unclear) interval between the end of the previous SSRI and the next.\textsuperscript{67,68,70,72} In the remaining 2 (SSRI-intolerance) studies, patients either started a second SSRI soon after the first SSRI or had an SSRI-free interval.\textsuperscript{71,73}

Response rates of switching in SSRI non-responders varied between 46\% and 58\% in three uncontrolled studies of variable methodological quality.\textsuperscript{57,69} The response rate was lower (42\%) in a fourth study with a heterogeneous group of inpatients.\textsuperscript{70} However, response rates to a second SSRI varied between 56\% and 72\% when patients were intolerant to the first SSRI (4 studies).\textsuperscript{68,71,73} Dropout rates due to side effects were between 5\% and 21\%, in studies with initial non-responders\textsuperscript{69} and between 0\% and 10\% in SSRI-intolerant samples\textsuperscript{71,73} (LoE: C).

In the SSRI-arms of three RCTs, response rates varied between 26.7\% and 71.1\%, while remission rates were between 17.6-52.1\%.\textsuperscript{46,47,74} Dropout rates due to side effects varied between 4.8-21.0\%. For results on the comparisons with other arms see below (LoE: A2-B).

In summary, the data from the open studies and one of the RCTs\textsuperscript{46} suggest that, after 1 SSRI, non-responders and, notably, also SSRI-intolerant patients can benefit from a switch to a second SSRI with response rates of approximately 50\% and 70\%, respectively. However, the results in two RCTs\textsuperscript{47,74} indicated much less advantageous response and remission rates for a second SSRI (26.7-29.0\% and \textasciitilde17.6\% respectively).

Tricyclic Antidepressants and mianserin

We identified 2 RCTs with a switch to a TCA,\textsuperscript{48,75} with one having limited power due to a randomization into 3 arms.\textsuperscript{75} Four open studies investigated a switch from an SSRI to a TCA.\textsuperscript{76-80} The methodology of the open studies varied: one large cross-over study was methodologically sound,\textsuperscript{79,80} one small study was unequivocally poor,\textsuperscript{78} and two studies were of reasonable quality (investigating populations with TRD).\textsuperscript{76,77}

In the RCT of Ferreri et al. switching to mianserin (a noradrenergic tetracyclic) versus continuation of fluoxetine was investigated, with a third arm for their combination.\textsuperscript{75} No significant difference was found between switching to mianserin and continuation of fluoxetine (response 48.5\% and 36.8\% respectively; NNT= 9 (95\% CI 2.9-\infty)) in an ITT analysis. The combination of fluoxetine and mianserin performed better than continuation of fluoxetine (response 62.5\% in the combination group; NNT= 4 (95\% CI 2.1-34.1)). Dropout rates due to side effects were highest in the switch-group (24\%; NNH vs continuation= 5 (95\% CI 2.6-10.4)) (LoE: B).

The STAR*D level III study compared a switch to nortriptyline versus mirtazapine in a randomized unblinded design.\textsuperscript{48} All participants received citalopram plus either a switch to sertraline, venlafaxine or bupropion or citalopram augmentation with buspirone or bupropion. Response rates (\textasciitilde50\% decrease in QIDS-SR\textsubscript{16} score) were 16.5\% for nortriptyline and 13.5\% for mirtazapine (NNT= 32 (95\% CI 8.1-\infty)). Remission rates (HDRS\textsubscript{17} \leq 7) were 19.8\% vs 12.4\% for nortriptyline and mirtazapine, respectively (NNT= 14 (95\% CI 6.0-\infty)). There were no differences in remission rates for those intolerant to the level II treatments versus those who tolerated their second trial of antidepressants. Dropout rates due to side effects were high both for nortriptyline (34.7\%) and mirtazapine (33.3\%) (LoE: A2).

Thase et al. investigated a switch to imipramine in non-responders to sertraline in chronic depressive out-patients. They found a 44\% ITT response rate, with a dropout rate due to intolerable side effects of 9\%.\textsuperscript{79,80} The methodologically poor study by Peselow et al. (including also SSRI-intolerant patients) found a 73\% response rate after a switch to imipramine in outpatients.\textsuperscript{78} In the studies that recruited TRD populations, response rates after switching to nortriptyline\textsuperscript{76} and oxaprotiline\textsuperscript{77} decreased to 39\% in in-patients\textsuperscript{77} and 42\% in outpatients,\textsuperscript{76} with a 35\% overall dropout rate in the latter study (LoE: C).
In summary, for the switch to a TCA response rates of approximately 16.5 to 48.5% were found. Lower response rates were observed in studies that included more treatment resistant patients.

**Mirtazapine, nefazodon or venlafaxine (novel dual acting agents)**

We identified 13 switch studies to novel dual acting agents. The methodological quality varied. Four studies were randomized controlled trials: Poirier and Boyer compared a switch to paroxetine versus venlafaxine, Baldomero et al. compared a switch to venlafaxine extended release versus switching to any other antidepressant (77% of these switches used paroxetine, citalopram, sertraline, or fluoxetine) in an unblinded design, and the level II and III STAR*D switch studies, which were also unblinded studies, compared a switch after citalopram to venlafaxine extended release, bupropion, or sertraline and the switch thereafter to nortriptyline or mirtazapine. Other studies described open studies with mirtazapine, nefazodone, and venlafaxine. In seven of the studies, all patients received an SSRI before switching, in 2 studies, this was unclear. In contrast, one study included patients (52%) who initially responded to an SSRI but did not sustain their response.

In the RCT performed by Poirier and Boyer, switching to venlafaxine was more efficacious than paroxetine when remission (HDRS ≤ 10) was considered (remission rates: 36.7% and 17.7% respectively), with a NNT of 6 (95% CI 2.9-28.9). For a response criterion (≥ 50% reduction in HDRS), the difference was insignificant (response rates: venlafaxine= 45% and paroxetine= 29%; NNT= 7 (95% CI 3.0-∞)). Dropout rates due to side effects were comparable (8.2% for venlafaxine and 4.8% for paroxetine; NNH= 30 (95% CI 8.3-∞)) (LoE: A2).

In the randomized, unblinded study by Baldomero et al., venlafaxine showed a significantly increased remission (HDRS ≤ 57) rate (59.3%) compared with conventional antidepressants (51.5%) after 24 weeks of treatment, with a NNT of 13 (95% CI 8.9-23.7). In the conventional antidepressants group, 77.3% of the patients used a second SSRI; for SSRIs, the remission rate was 52.1% (NNT= 14 (95% CI 9.1-29.3)). Response (≥ 50% reduction in HDRS) rates also showed a modest but significant advantage: 77.3% for venlafaxine versus 71.1% for SSRIs (NNT= 17 (95% CI 10.5-35.0)). Overall dropout was slightly lower in the venlafaxine group when compared with all conventional antidepressants (28.3% vs 32.8%; NNH= 27 (95% CI 15.1-120). Dropout rates due to side effects were not significantly different between venlafaxine and conventional antidepressants (12% vs 7.3% respectively; NNH= 161 (95% CI 62.1-∞)) (LoE: B).

The level II STAR*D trial did not find significant differences between the switches to venlafaxine, bupropion, and sertraline. Before the switch, all participants received citalopram (20-60 mg for a maximum of 14 weeks). Patients were randomized over different randomization possibilities for which they were at equipoise. The assessors of the primary outcome (HDRS ≤ 7) were blind to the treatment. After 14 weeks of treatment, response rates (≥ 50% decrease in QIDS-SR16) were 28.2% for venlafaxine, 26.1% for bupropion and 26.7% for sertraline (not significant). Remission rates (QIDS ≤ 7) were not significantly different for venlafaxine, bupropion, and sertraline (24.8%, 21.3% and 17.6% respectively). For corresponding NNTs see table 5.1. The dropout rate due to side effects was not statistically different for venlafaxine (21.2%), bupropion (27.2%), and sertraline (21.0%) (LoE: A2).

The level III switch study was described earlier. Mirtazapine response, remission and side-effects related dropout rates were 13.5%, 12.4% and 33.3%, respectively (LoE: A2).

Dropout rates due to adverse effects varied between 5.5% and 11% for venlafaxine, between 20.8% and 26% for mirtazapine and was 39% in one study with nefazodone (LoE: A2, C).
We performed a meta-analysis of the three RCTs that compared switching to venlafaxine versus SSRIs,\textsuperscript{46,47,74} although the differences in duration of follow-up introduced some heterogeneity (ranging from 4 weeks by Poirier and Boyer\textsuperscript{74} to 24 weeks by Baldomero et al.\textsuperscript{46}). As shown in figure 5.2, the weighted difference in remission-rates (fixed effects model) was 8\% (4 – 11\%) in favour of venlafaxine (NNT= 13 (95\% CI 9.1 – 25.0), and for response 6\% (1 – 10\%), (NNT= 17 (95\% CI 10.0 – 100.0)). Omission of the methodologically poorer study of Baldomero et al. increased the difference in remission rates (10\% (95\% CI 3-16\%) fixed effects model; NNT= 10 (95\% CI 6.3-33.3)), but decreased the difference in response rates (4\% (-3-12\%) fixed effects model; NNT= 25 (95\% CI 8.3-\infty)). The dropout rate due to side effects was only reported in two studies\textsuperscript{47,74}; the weighted difference was 1\% (95\% CI -5-7\%) (fixed effects model) with more dropouts for venlafaxine.

In summary, heterogeneous studies considering switching to mirtazapine, nefazodone and venlafaxine showed response rates of approximately 28-50\% in subjects without obvious TRD, while in subjects with increased levels of TRD response percentages dropped (investigated for venlafaxine and mirtazapine). Pooling of results showed a modest and clinically equivocally advantageous increased remission rate for venlafaxine over SSRIs (NNT= 13 (95\% CI 9.1-25.0)).

**Bupropion and Reboxetine (agents specifically affecting dopaminergic and/or noradrenergic neurotransmission)**

We identified one RCT and two small open studies of switching to bupropion.\textsuperscript{47,92,93} The STAR*D level II switch study including bupropion was described earlier.\textsuperscript{47} There were no significant differences in remission or response rates for bupropion compared to venlafaxine or sertraline. In this study, bupropion had the (statistically insignificant) highest dropout rate (27.2\%) due to side effects (LoE : A2).
In two open studies with bupropion, Fava et al. prospectively determined fluoxetine nonresponse in a small but well performed study, and Walker et al. recruited patients that were primarily suffering sexual side-effects of fluoxetine, and only reported a decrease in 28-item HDRS-scores. One larger, well-performed, open study investigated the switch to reboxetine in fluoxetine non-responders.

Thus, switching from fluoxetine was investigated, with reported response rates of 34.6% for bupropion and 45.3% for reboxetine. For bupropion, specified dropout rates were not reported in one study. The side effect-related dropout rate was 10.3% in subjects with sexual dysfunction while taking fluoxetine. For reboxetine, the dropout rate due to side effects was 13.3% (LoE: C).

In summary, switching to bupropion or reboxetine was scarcely studied but was a possible option with response-rates of 26.1%-34.6% and 45.3% respectively. The remission rate of switching to bupropion was not different compared to venlafaxine or sertraline.

**Reversible Inhibitor of Monoamine-oxidase A**

We identified no studies that investigated switching from a SSRI to a reversible inhibitor of monoamine-oxidase A.

**Monoamine-oxidase A inhibitor**

We identified one RCT from STAR*D (level IV) and two small, interrelated randomized studies after 4 weeks of treatment with at least one SSRI (fluvoxamine) and oxaprotiline. We identified no studies of SSRI non-responders in atypical depression. Two studies were RCTs and one an unblended, randomized, cross-over study. The STAR*D study investigated outpatients; the studies by Nolen et al. were performed in treatment resistant in-patients.

Nolen et al. found tranylcypromine to be more efficacious than nomifensine, in both studies the response rate for tranylcypromine was 42.9 and 45.5%. All patients previously received at least fluvoxamine and oxaprotiline. Fifty-eight to 62% had side effects affecting their blood pressure levels (LoE: B).

The STAR*D level IV study included patients that had not been in remission after citalopram (level I), either venlafaxine, bupropion, sertraline or citalopram augmentation with buspirone or bupropion (level II), and additionally received nortriptyline or mirtazapine (level III). These patients were randomized between tranylcypromine and a combination of venlafaxine with mirtazapine. Of the included patients 32.1% were intolerant for the level III medication. Remission rates (HDRS$_{17} \leq 7$) were low for tranylcypromine (6.9%) and the combination treatment (13.7%; NNH= 15 (95% CI 5.5-∞)). Response rates ($\geq 50\%$ decrease in QIDS-SR$_{16}$) were also not significantly different: 12.1% vs 23.5% for tranylcypromine and venlafaxine with mirtazapine, respectively (NNH= 9 (95% CI 3.9-∞)). Dropout rates due to side effects were higher for tranylcypromine: 41.4% versus 21.6% for venlafaxine with mirtazapine (NNH= 6 (95% CI 2.7-35.2)).

**Additional concerns for clinicians regarding switching**

Little evidence is available about the optimal way to switch. Abrupt reduction or discontinuation of SSRIs may produce somatic and psychological withdrawal symptoms, of which occurrence is inversely related to the plasma half-life of the initial SSRI. Overlap of antidepressants during switching is generally avoided.

Direct switching (without a washout phase) from an initial SSRI (fluoxetine at the standard dose or citalopram at high dosages) to another SSRI (paroxetine, citalopram, sertraline), nortriptyline, mirtazapine, bupropion, reboxetine, or venlafaxine was well tolerated. Also, direct switching reduced the emergence of side effects compared with placebo in a 1-week washout phase (which might have been discontinuation symptoms). In case of higher than standard doses of SSRIs, some data for tolerance of direct switching were generated by STAR*D. However, the results published so far do not specify dropout rates in the first 2 weeks after switching. Also, because tapering of high doses of previous antidepressants was not applied in STAR*D, this trial was not designed to examine the optimal...
switch strategy if higher than standard doses were used before the switch. Thus, if necessary, direct switching of high-dose antidepressant therapy appears possible after citalopram as a first SSRI.47

In a case of switching from an SSRI to a TCA, other reviewers did not recommend a washout period.20;29 In one included study a direct switch to mianserin was less well tolerated.25 For switching to nefazodone, in one study a 4-day to 7-day washout was applied but not investigated.81 A 1-week washout period is suggested for switching to a reversible inhibitor of monoamine-oxidase A, and a 1-week to 2-week washout period is recommended for switching to a MAO-I.20;29 For fluoxetine these washout-periods should be prolonged to 5 weeks because of the long half-life of fluoxetine and norfluoxetine.20 The inhibition of cytochrome P450 subenzymes by SSRIs may increase the levels of some TCAs during the first to fifth (fluoxetine) week.20

Discussion

This report systematically reviewed and appraised the available research focusing on switching strategies for SSRI non-responders in MDD, including the recent STAR*D results. We found that the available evidence does not justify distinct recommendations for next-step strategies after non-response to a first SSRI. The pooled difference in remission rates of switching to venlafaxine (an SNRI) versus a second SSRI showed a modest and clinically equivocal advantage of venlafaxine (NNT= 13 (95% CI 9.1-25.0)), this difference increased when the largest and methodologically poorest study was omitted (NNT= 10 (95% CI 7-34)).

In summary, after a first SSRI, switching to any of the current classes of antidepressants has approximately a 50% chance of response. Still, a direct comparison of the rates across the predominantly open studies is methodologically not justified. In STAR*D response and remission rates were lower (respectively 26.8% and 21.3% at level II,47 15% and 16.2% at level III,48 and 17.4% and 10.1 at level IV49). Rush et al. attributed these lower remission rates to the inclusion of patients who were more chronically depressed, had lower socio-economic status and suffered from more co-morbid somatic and psychiatric diseases.47 In general, the level of TRD10 of included studies was inversely correlated with treatment outcome. Although this finding carries the risk of an ecological fallacy, it is worrisome, which further appears from the STAR*D results. After the second antidepressant the chances of response or remission by switching again are becoming rather low, challenging us to find new approaches.100-102

Dropout rates due to side-effects varied between 5-21% for a second SSRI and venlafaxine; 10-35% for TCAs, bupropion, and reboxetine; 20-33% for mirtazapine; 39% for nefazodone and 41.4% for tranylcypromine. It should be noted that these percentages cannot simply be compared with each other because of heterogeneous populations and open-study designs. In randomized comparisons, no significant differences in side-effect related dropout were found, except for tranylcypromine versus a combination of venlafaxine with mirtazapine.49

With eight RCTs,46-49;74;75;95;96 switching-options after a first SSRI were generally investigated with open studies. In these open studies, switching to a second SSRI (7 studies) and venlafaxine (7 studies) were studied most frequently. Furthermore, the studies were of variable methodological quality. In our opinion, the available evidence for switching strategies allows general recommendations only. Switching is open to all studied antidepressant classes (second SSRI, novel dual-acting antidepressants, selective noradrenergic and noradrenergic/dopaminergic agents, or TCA or mianserin) without clear recommendations other than those that apply for the selection of initial treatment. In the choice of an initial antidepressant, some reports promoted TCAs for treatment of inpatients;103-106 however, it is unclear what special feature is associated with inpatients (e.g. severity), and studies investigating switching strategies after an SSRI in inpatients were not identified. From the available studies it must be emphasized that side-effects to a first SSRI did not reduce the chance of response or increase the chance of intolerance for a
second SSRI. Because of side effects, we think that MAO-I should not be prescribed as a second antidepressant after a first SSRI. A possible exception – but not investigated after a first SSRI – is for atypical depression.

Switching from a failed TCA treatment was reviewed earlier. The response rates for within-class switching with SSRIs appear more favourable than a TCA-TCA switch: in two small trials, response-rates of a within-class TCA switch were 9% and 30%. The SSRI results challenge the belief that any within class switch should be considered illogical. The between classes switching strategies from a TCA to an SSRI (investigated in 10 trials; response rates varying between 4% (inpatients) and 75% (out-patients)), to a heterocyclic antidepressant (e.g. bupropion, trazodone, nomifensine, oxaprotiline; 6 studies; response rates between 10%-56%) and to a MAO-I (6 trials; response rates between 29-83%) showed similar broad ranges of response rates. These ranges reflect differences in heterogeneous study-populations as well. Again, it is inappropriate to simply compare these rates determined in different studies.

On theoretical grounds, it is logical (and often recommended) to switch to an antidepressant with different or combined sites of action (e.g. norepinephrine uptake inhibition after unsuccessful serotonergic uptake inhibition). Others pointed out the complex interaction of monoamine systems alone, proposed other possible etiologic mechanisms, and considered the monoamine hypothesis only partially explanatory for depression and the response to antidepressants. Six RCTs so far compared different pharmacologic approaches in non-responders (venlafaxine versus paroxetine, venlafaxine versus an SSRI, venlafaxine versus sertraline or bupropion, nortriptyline versus mirtazapine, fluoxetine versus mianserin or a mianserin-fluoxetine combination, and tranylcypromine versus a venlafaxin-mirtazapine combination). These RCTs found equivocal superiority of dual-action pharmacotherapy. However, in STAR*D the empirical proof of this theoretical strategy was not found.

Apart from switching, augmentation or combination, and addition of (or switching to) psychotherapy are possible options. Only Ferreri et al. and McGrath et al. compared switching versus combination (the latter at a higher level of TRD). In STAR*D, a switch to or augmentation with cognitive behaviour therapy was possible after citalopram (unpublished yet), and augmentation of citalopram with buspirone or bupropion was also studied. A direct comparison between switching and augmentation after citalopram was not feasible. Therefore, clear recommendations about choosing one of these strategies relative to each other are not possible. In most countries, SSRIs are generally prescribed as first line treatment, often provided in primary care. We think that switching-strategies after a first SSRI will be preferred, especially in primary care, in which augmentation and combination strategies may be unfamiliar to physicians. This hypothesis is supported by audits, even among psychiatrists.

Limitations of the identified studies

Well-designed switch-studies are difficult to carry out, and therefore, it does not surprise that the evidence to date is limited in several ways. We found predominantly open, uncontrolled studies, with a risk of more positive results than in blinded studies and without a possibility to actively compare strategies. There were few studies that clearly described the inclusion of prospectively determined SSRI non-responders. This finding is of importance, as in retrospectively determined non-responders, current depression may cause recall-bias. Furthermore, in some studies non-responders were not treated directly after cessation of the unsuccessful drug, which might have biased results; for example depression worsened after cessation, or –the other way around– depression may have improved because of the natural course of depression. Several other problems were encountered: unclear criteria for initial non-response, inclusion of mild or minor depression, possible selection-bias, limited presentation of results, absence of ITT-data, small sample sizes (n< 40), and low statistical power. In general, less robust studies found more positive results for the drug of interest. Table 5.2 presents a summary of these problems.
The STAR*D trials were randomized but unblinded effectiveness trials. The primary outcome (remission by HDRS) was determined by blind assessors; the secondary outcomes by the QIDS-SR were self-rated by the unblinded patients. The a priori definition of nonremission for missing data will have decreased remission rates because of attrition, but this a priori definition was considered noninfluential after sensitivity analyses. The aggressive dose increases in STAR*D trials prevented undertreatment, but might have increased attrition, and definitely increased the percentages of treatment-intolerant patients at all levels. Especially in the level IV trials the treating physicians might have been unfamiliar with the prescribed medication (tranylcypromine, venlafaxine-mirtazapine combination), reducing the vigour of the applied pharmacologic intervention.

### Future switching studies

After the STAR*D trials, the question arises as to whether many randomized direct comparisons between switches among drug classes are fruitful to develop fully evidence-based recommendations for switching. Results of some studies might still be published. Also, the predictors of poor response and nonremission need to be further clarified. In order to structure directions of research, the recommended approaches in guidelines should be evaluated for each treatment step. The Texas Medication Algorithm project proved that algorithms are beneficial for patient care however, our next challenge is to investigate which steps within these algorithms are better compared with each other.

Ideally, three or more armed studies should be designed. Switching within the same class or to different classes of drugs should be compared with an augmentation or new approach, while also an arm for continuation of the initial therapy should be included. The latter arm would then represent a form of placebo control. Naturally, these studies are hard to carry out, may have to overcome resistance and doubts concerning the ethics of the continuation arm, or may suffer from selective patient withdrawal from this continuation arm. The STAR*D-project has been a major step in this direction, especially by proving the feasibility of such large multicenter trials and the methodology of (equipoise) randomization. At the same time the effectiveness approach with many centers, high levels of co-morbidity, chronicity and many arms of treatment might have reduced the ability to find differences.

We found that the response rates in switch-studies decreased with increased levels of TRD. Therefore, future studies must consider the level of TRD as an important effect-modifying variable. Ideally, in future research, clear populations of prospectively determined treatment resistance should be selected, or analyzed in a priori defined subgroups to increase our knowledge about confounding or effect-modifying variables. Finally, to improve the acceptance of switching in daily clinical practice, more studies of patients’ perspectives of switching of antidepressants are needed.

### Limitations of the review

Several limitations of this review should be mentioned. First, a review like this cannot overcome the paucity of high-quality evidence to date. The Cochrane Collaboration primarily rejects open studies as high-quality evidence. If this criterion had been applied, only 8 studies would have qualified for the review, obviously limiting its applicability. The majority of included studies had methodological flaws, two studies were excluded for clear invalidity. We decided a priori to include open studies, and–even more–to include studies in which 50-100% of patients initially used an SSRI, introducing different levels of TRD. Of course, the latter decision is debateable from a methodological point of view.

Second, in the selected trials, mostly response was used as the primary outcome, while currently remission of depression is the clinical aim of treatment. Only 13 of 31 studies (42%) included remission as an outcome criterion. Only STAR*D primarily investigated the practice of switching in order to achieve remission.
Third, patients studied in the included trials represented selected populations, reducing the generalizability of the findings to the 'real world' clinical practice; as an effectiveness trial, the STAR*D results overcame this problem. Fourth, critical appraisal was performed by one reviewer (H.G.R.), while ideally this should have been performed by two raters. However, we found our interobserver agreement to be moderate to good and no worse than in previous interrater attempts in psychiatry. Fifth, the grading system for studies does not represent the appraised methodological dimensions of evidence. This improved the applicability of the results for busy clinicians, but reduced their strength.

**Strengths of the review**

This is the first review that applied the thorough methodology to search for, identify, and appraise articles as used in Cochrane reviews. The applied methodology and transparent presentation of data allow clinicians to make their own judgements and, if necessary, to retrieve the source of data. Apart from the relevant up-to-date information for clinicians, this review could well serve national guideline committees as a building stone for the development of treatment guidelines for MDD.

**Conclusion**

This systematic review about switching identified 8 RCTs and mostly open switch studies of variable methodological quality in heterogeneous populations. The STAR*D results largely increased the amount and quality of the available evidence, but did not show differential class effects to guide switching. After a first SSRI switching is open to all studied antidepressant classes (except irreversible MAO-inhibitors), without clear recommendations other than those that apply for the selection of initial treatment. For recommendations about when to choose between switching, augmentation, combination, or psychotherapeutic strategies as a next step, hardly any evidence of comparisons of these strategies relative to each other exists. Future algorithm-based switch studies and studies of patient perspectives regarding switching will have to improve our knowledge to guide treatment for SSRI non-responders.

**Acknowledgements**

This systematic review was realised by a grant for the development of a local evidence-based clinical practice guideline (No SFA.07.012) from the Academic Medical Center, Amsterdam, the Netherlands. This guideline considering strategies for non-response to a standard dose of a first SSRI is available on request from the first author.

**Conflicts of interest**

None
References


Chapter 5


Chapter 5

Review of switching in MDD
SEROTONIN IN THE ETIOLOGY OF MAJOR DEPRESSIVE DISORDER
Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: A meta-analysis of monoamine depletion studies

Molecular Psychiatry 2007; 12: 331-359

Henricus G. Ruhé¹
Nada S. Mason¹
Aart H. Schene¹

¹Program for Mood Disorders, Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
Abstract

Background
Dysfunction in the monoamine systems of serotonin (5-HT), norepinephrine (NE) and dopamine (DA) may causally be related to Major Depressive Disorder (MDD). Monoamine depletion studies investigate the direct effects of monoamines on mood. Acute tryptophan depletion (ATD) or para-chlorophenylalanine (PCPA) deplete 5-HT, acute phenylalanine/tyrosine depletion (APTD) or alpha-methyl-para-tyrosine (AMPT) deplete NE/DA. Available depletion studies found conflicting results in heterogeneous populations: healthy controls, patients with previous MDD in remission and patients suffering from MDD.

Aim
To review and quantify the decrease in mood after 5-HT and NE/DA depletion in humans.

Methods
Systematic search of MEDLINE and EMBASE (1966-October 2006) and cross-references. Randomized studies applying ATD, PCPA, APTD or AMPT versus control depletion were included. Pooling of results by meta-analyses was stratified for studied population and design of the study (within or between subjects).

Results
Seventy-three ATD, 2 PCPA, 10 APTD and 8 AMPT studies were identified of which 45 ATD and 8 APTD studies could be meta-analyzed. 5-HT or NE/DA depletion did not decrease mood in healthy controls. 5-HT or NE/DA depletion slightly lowered mood in healthy controls with a family history of MDD. In drug-free patients with MDD in remission, a moderate mood decrease was found for ATD, without an effect of APTD. ATD induced relapse in patients with MDD in remission who used serotonergic antidepressants.

Conclusions
Monoamine depletion studies demonstrate decreased mood in subjects with a family history of MDD and in drug-free patients with MDD in remission, but do not decrease mood in healthy humans. Although depletion studies usefully investigate the etiological link of 5-HT and NE with MDD, they fail to demonstrate a causal relation. They presumably clarify a vulnerability trait to become depressed. Directions for further investigation of this vulnerability trait are proposed.
Introduction

Major depressive disorder (MDD) is characterized by a lowered mood. MDD is a disabling disease which affects 20% of the world’s population.\(^1\) MDD is often treated with antidepressants (AD), mostly Selective Serotonin Reuptake Inhibitors (SSRIs), Serotonin Norepinephrin Reuptake inhibitors (SNRIs), Norepineprine Reuptake Inhibitors (NERIs) or Tricyclic Antidepressants (TCAs).\(^2\, ^5\)

The working mechanism of AD is believed to be either by (1) increased neurotransmission by increased synaptic levels of serotonin, norepinephrine and dopamine (monoamines) or (2) specific agonistic effects on serotonin or norepinephrine (sub-)receptors. The increased levels of monoamines were discovered in the late fifties, when the TCAs and Monoamine-oxidase A inhibitors (MAO-I) appeared to effectively treat MDD. This discovery led to the monoamine hypothesis: MDD might etiologically be explained by a deficiency in monoamine neurotransmitters: serotonin (5-HT), norepinephrine (NE) or (to a lesser degree) dopamine (DA). The monoamine systems in the brain have complex interactions. Therefore, the current, less pertinent view is that the monoamine hypothesis only partially explains MDD and the response to AD.\(^6\, ^10\)

Depletion of the available 5-HT, NE and/or DA is used as a model to test the involvement of monoaminergic systems in MDD. Two reviews recently described and reviewed the techniques of monoamine precursor depletion and enzyme-blocking methods.\(^11\, ^12\) In brief, 5-HT depletion can be achieved by rapidly lowering the levels of the essential amino-acid tryptophan which cannot be synthesized by the body and must be ingested to enable formation of 5-HT. To achieve depletion, a tryptophan free amino-acid mixture is administered (acute tryptophan depletion (ATD)).\(^13\) Depletion of NE and DA uses the same concept (acute depletion of the essential amino-acids phenylalanine/tyrosine (APTD)).\(^14\) As an alternative to induce a state of depletion, enzyme-blocking agents decrease the production of the monoamines. Para-chlorophenylalanine (PCPA) blocks 5-HT synthesis,\(^15\, ^16\) and alpha-methyl-para-tyrosine (AMPT) NE and DA synthesis.\(^17\) Since 1975 an increasing number of depletion studies have been conducted. Monoamine depletion showed differential effects in different study populations: healthy controls, healthy controls with a family-history of MDD, patients with MDD in remission using AD or after cessation of AD, and patients who had a current episode of MDD.

Previous reviews summarized the methodology of depletion tests.\(^18\) Others reviewed the findings of specific depletion tests across different psychiatric illnesses,\(^11\, ^18\, ^20\) or the prediction of response to depletion.\(^12\, ^13\) Booij et al. presented a study that pooled results across studies individual subject data of six ATD studies (‘mega-analysis’) in order to investigate the mediating role of clinical, demographic and biochemical characteristics in the mood response to ATD in remitted subjects who previously had MDD.\(^21\) However, a clear summary of the mood-effects of monoamine depletion across different populations is lacking. Except from the study of Booij et al., we are unaware of any attempt of pooling studies. Pooling is important because the small-sized depletion studies might not have detected small differences by a lack of power. Finally–as in drug research–pooling will quantify the weight of positive versus negative studies.

Therefore, we aimed to review the evidence for mood lowering properties of monoamine depletion studies as a model for MDD. Our main question was: does the depletion of monoamine (5-HT and NE/DA) systems lower mood in humans? A secondary question was: is the lowering of mood different across different populations? This paper reports our systematic review with a stratified meta-analysis of the mood effects of monoamine depletion studies.
Methods

Design of the study

We included all available randomized prospective monoamine depletion studies (tryptophan and phenylalanine/tyrosine) and enzyme blocking studies (PCPA or AMPT) in humans. Included studies measured a change in mood after depletion. Two study designs were included. First, randomized within-subjects studies, where each subject was exposed to a true depletion versus a sham intervention at a second occasion. This way, each subject served as his/her own control. Second, randomized between-subjects studies, where two groups of subjects were compared, with the intervention applied to one group and a control intervention to the other group. We excluded animal studies and studies that selected patients with apparent co-morbidity (e.g. substance abuse or dependence, studies with only smokers, psychotic disorders, anxiety disorders). Additionally, we excluded studies in depression subtypes with a supposed different etiology (e.g. seasonal affective disorder, bipolar disorder). We also excluded studies that combined depletion with other interventions (e.g. sleep deprivation) or studies that did not report mood effects. Finally, we excluded studies recruiting a patient group with both unipolar and bipolar depression when >25% of subjects had bipolar depression and when no separate data for the unipolar group were provided.

Literature searches and selection

We searched PubMed from 1966 to October 1st 2006 and EMBASE from 1980 to October 1st 2006 using a comprehensive search strategy (Table 6.1). We retrieved additional references identified in previous reviews and cross-references from identified studies. Two reviewers (H.G.R. and N.S.M.) independently selected articles based on the a priori inclusion and exclusion criteria as stated above. In case of doubt an article was retrieved and the full content was considered. In 85.4% the reviewers agreed on selection instantly. The initial kappa for agreement was 0.61 (95% confidence interval (95% CI): 0.52 - 0.71). Agreement was mainly influenced by a differential inclusion of research letters, studies in depression subtypes with a supposed different etiology and studies without apparent mood-scores in the abstract. Discrepancies in included studies were discussed between both reviewers until full consensus was reached.

Table 6.1. Search terms.

<table>
<thead>
<tr>
<th>#</th>
<th>Search terms used</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Depressive disorder[MeSH] OR depression[MeSH] OR (depressive[TW] AND symptoms[TW]) OR mood[TW]</td>
</tr>
<tr>
<td>2</td>
<td>Indoleamine$ OR serotonin[MeSH] OR 5-hydroxytryptamine[TW] OR 5-HT[TW]</td>
</tr>
<tr>
<td>3</td>
<td>((Tryptophan[MeSH] OR Tryptophan[TW]) AND depletion[All Fields]) OR (fenclonine[MeSH] OR PCPA[TW] OR parachlorophenylalanine[TW])</td>
</tr>
<tr>
<td>4</td>
<td>#1 AND #2 AND #3</td>
</tr>
<tr>
<td>6</td>
<td>((Tyrosine[MeSH] OR Tyrosine[TW]) AND depletion[All Fields]) OR (alpha-methyltyrosine[MeSH] OR AMPT[TW] OR alpha-methyl-para-tyrosine[TW])</td>
</tr>
<tr>
<td>7</td>
<td>#1 AND #5 AND #6</td>
</tr>
<tr>
<td>8</td>
<td>(#4 OR #7) NOT (Animal[MeSH] NOT Human[MeSH])</td>
</tr>
</tbody>
</table>

MeSH = Medical Subject Heading, TW = textword Note: PubMed search terms are provided; these terms were slightly modified for EMBASE-searches.

Data extraction and Assessment of the studies

Two authors (H.G.R. and N.S.M.) abstracted the retrieved studies as follows: population studied, study design (within-subjects or between subjects), applied mood scale(s), number of subjects (and–if applicable–number of subjects for each sequence of intervention and control depletion), male/female ratio, the intervention and its control condition (including dosage and duration). Furthermore, we extracted plasma levels of tryptophan, tyrosine or relevant levels of monoamine-metabolites before and after intervention and control condition and other relevant
data that could influence the outcome of the study (e.g., different consistency of control drink, unclearly reported data, combined interventions). Primary outcome variables were changes in mood-scale scores before and after intervention and control, and the number of relapses in the intervention and control group. If more than one mood-scale was used, we primarily used the Profile of Mood States (POMS).\textsuperscript{26} Otherwise, we used the Multiple Affect Adjective Checklist (MAACL)\textsuperscript{,27} Visual Analogue [Mood] Scales (VA[M]S)\textsuperscript{,28} Hamilton-depression rating-scale (HDRS)\textsuperscript{,29} and Montgomery-Åsberg Depression Rating Scale (MADRS).\textsuperscript{30} However, in studies in patients with previous MDD in remission we primarily chose the HDRS or MADRS followed by the POMS, because most of these studies used depression rating scales to measure effects of depletion. If subscales were used, we took the subscale representing depressed mood.

We validated the abstracted studies (criteria available on request). We based criteria on previous reviews and the Cochrane handbook\textsuperscript{11;12;18;31}. We assessed studies as poor, moderate, or good in the context of this meta-analysis. We particularly assessed quality as good when studies applied a randomized double-blind design, when the achieved depletion was judged to be sufficient, and when mood-scale ratings were provided.\textsuperscript{11;18} If more than one of these items was rated inadequate, we assessed the study as moderate. If studies had one of these items missing and at least one other aspect of the study was not well reported, we also assessed the study as moderate. If crucial data (e.g., scores and SDs) were missing the study was assessed as poor.

### Data synthesis

We qualitatively summarized all included studies in Tables 6.2-6.5 (ATD, PCPA, APTD and AMPT respectively), irrespective whether the studies were pooled in the meta-analysis. We acknowledged three clinically heterogeneous study populations\textsuperscript{a priori}: A. healthy controls (A1 negative/ A2 positive family history of MDD), B. patients with a previous MDD currently in remission (B1 currently not using AD / B2 currently using AD) and C. patients with a current episode of MDD (with or without AD). These populations were considered and analyzed separately.

### Statistical pooling

Mostly, continuous scores for changes in mood rating scales were reported. Some studies in patients with MDD in remission presented relapse rates. In order to pool continuous effect estimates from different scales, we applied a standardization (providing Hedges’ g). The effect-estimates (1 per study) with the corresponding standard error (SE) were entered in the inverse-variance statistics for pooling in Review Manager 4.2.\textsuperscript{32;33} The supplementary appendix gives the formulas used to determine Hedge’s g, the difference in relapse rates and corresponding SEs for different study designs.

Different study designs were not combined in pooling. Especially the within subjects design needs attention in meta-analysis. In this design, differences between the experimental and control condition are statistically paired. As paired data improve power, calculations with data from a within subjects (or cross-over) design in a weighted mean differences model would not be justified.\textsuperscript{33}

### Heterogeneity, effect modification, sensitivity-analysis and assessment of publication bias

We first performed the meta-analyses with fixed effects models. We assumed more homogeneous results after our \textit{a priori} attempt to reduce clinical heterogeneity by stratification. However, if effect-estimates and 95% CI for the individual studies showed graphical poor overlap or consistency, we interpreted this an indication of statistical heterogeneity. We used the chi-squared test and \textit{I} in addition. \textit{I} represents a chi-squared statistic relative to its degree of freedom. An \textit{I} value >50% is indicative of heterogeneity.\textsuperscript{34} We applied a conservative random effects model\textsuperscript{35} when we suspected statistical heterogeneity.
We investigated effect modification by gender, and the influence of a positive family history of MDD in healthy controls. Furthermore, we investigated whether different mood scales caused differences in outcomes. Therefore, we stratified analyses for these variables, and presented stratified Hedges’ g, with 95% CI. Differences between strata were tested by subtracting the chi-squared heterogeneity statistic per stratum from the total chi-squared heterogeneity statistic. This residual (Q_res) has a chi-squared distribution, with the total number of strata-1 degrees of freedom.

We imputed R= 0.5 to calculate missing pooled SDs within each intervention/control and for the changes in mood-scores between interventions (supplementary appendix). We checked the impact of this imputation on the calculated effect estimates. Therefore, we increased R to 0.8 (less conservative) and decreased R to 0.2 (more conservative) to compare the new effect estimates with the original findings for R = 0.5.

We assessed publication bias graphically with a funnel-plot, plotting Hedges’ g versus the precision of the study (the inverse of the standard error of Hedges’ g). Additionally, we tested publication bias with Galbraith’s radial plot, which regresses the standard normal deviate (Hedges’ g divided by its standard error) against the precision. For a set of studies not distorted by selection bias the intercept of the regression model will be close to zero.36

Results

Identified studies

Our systematic searches identified 392 articles. In total we selected 90 studies. Three studies applied both ATD and APTD and a control.37-40 Therefore, 73 studies with ATD, 2 with PCPA, 10 with APTD and 8 with AMPT as monoamine depletion method were identified (summarized in Tables 6.2-6.5 respectively). Several studies investigated contrasts in different populations.41-51 A list of excluded studies is available on request. Most studies had a within-subjects design (n= 74). The majority of studies investigated healthy controls (n=64).

Qualitative summary

The majority of the 90 studies also investigated other effects of monoamine depletion: 24 studies measured effects on cognitive functions,53;54;55;56;67;68;72 11 measured effects on other behavioral measures (pain, impulsivity, panic attacks, appetite, noise-stress, aggression),73-83 8 measured effects on neuroendocrine parameters,37;60;84-89 11 measured electric encephalogram (EEG) alterations and/or sleep effects,72;90-99 14 measured changes in brain function with Positron Emission Tomography (PET),97;98;100-104 Single Photon Emission Computed Tomography (SPECT),105 or Magnetic Resonance Imaging (MRI).66;68;70;106-108 Two studies stratified mood response to ATD by genetic polymorphism of the 5-HT transporter promoter region.46;51

We judged the overall methodological quality of the identified studies as good, with appropriate application of the ATD, APTD or AMPT depletion. Two studies with PCPA reported case series only and were rated ‘poor’.15;109 In 8 studies no data on the adequacy of the achieved depletion was reported.73;75;78;88;89;93;99 In 2 studies insufficient depletion was attained.63;111 In 10 studies no clear SDs or SEs were given for the observed mood-effects.13;76;85;88;92;103;110-112 In total, this were 14 patients. The reason that several apparently highquality studies were rated as moderate in this review was due to the lack of presentation of the mood effects, which in those studies was not the primary outcome.

In 34 of 90 studies the POMS was used, in 2studies the MAACL, in 34 studies a visual analogue scale, in 29 studies a version of the HDRS, and in 5 of the 90 studies the MADRS. In 5 studies only other or undefined mood-scales were used.28;80;97;98;113 Especially for the POMS and the VAS the direction of a decrease of mood in subjects was not always clearly reported.53;57;77;113 In our meta-analysis no differential effect of the applied mood-scale was observed (Q_res = 3.89, df= 2, p>0.05; data not shown).
Table 6.2. Included studies Acute Tryptophan Depletion (ATD).

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td>**Abbott (1992)**73</td>
<td>Rand., DB, Betw. SS</td>
<td>2*30 M</td>
<td>Not clearly reported</td>
<td>POMS-depression: ATD: BL: 2.50 (±0.58 SE), post-test: 1.73 (±0.71). CONT: BL: 2.50 (±0.58), post-test: 1.98 (±0.58).</td>
<td>Study primarily investigates effect of ATD on analgesia.</td>
<td>Moderate</td>
</tr>
<tr>
<td>**Allen (2006)**66</td>
<td>Rand., DB, Within SS, FH-</td>
<td>10 M+F</td>
<td>Good</td>
<td>VAS: No exact scores reported, no sign. difference in mood found</td>
<td>Study investigates effects of ATD on cognition (verbal fluency and working memory) and prefrontal activity by fMRI.</td>
<td>Moderate</td>
</tr>
<tr>
<td>**Amin (2006)**67</td>
<td>Rand., DB, Within SS</td>
<td>19 F</td>
<td>Good**</td>
<td>POMS and HDRS: No exact scores reported, no sign. difference in mood found on total POMS or HDRS</td>
<td>Study investigates effect of ATD on cognition (memory, visuospatial learning) and the protective effects of estrogen therapy (ET) in peri/post-menopausal women. Only pre ET-data used.</td>
<td>Moderate</td>
</tr>
<tr>
<td>**Barr (1997)**75</td>
<td>Rand., DB, Within SS, FH-</td>
<td>6 M+F</td>
<td>Good</td>
<td>VAS-happy: ATD: BL: 51.2 (±9.1), post-test: 42.7 (±13.8). CONT: BL: 51.7 (±25.9), post-test: 62.7 (±12.2); POMS not reported</td>
<td>Study investigates effect of fluoxetine treatment on ATD in healthy controls, only pre-fluoxetine data used.</td>
<td>Good</td>
</tr>
<tr>
<td>**Benkelfat (1994)**67</td>
<td>Rand., DB, Within SS, FH-</td>
<td>19 M</td>
<td>Good</td>
<td>POMS-depressed: ATD: BL: 57.2 (±2.0 SE), post-test: 58.9 (±2.0), change: 1.7 (±1.4). CONT: BL: 56.3 (±2.6), post-test: 60.0 (±2.1), change: 4.3 (±2.3). No one reached 10 point decline in POMS-D.</td>
<td>Study focuses on psychomotor and neuroendocrine effects of ATD.</td>
<td>Good</td>
</tr>
<tr>
<td>**Bhatti (1998)**80</td>
<td>Rand., DB, Within SS, FH-</td>
<td>12 M</td>
<td>Good</td>
<td>POMS-depressed: ATD: BL:0.9 (±1.3), post-test: 0.7 (±3.3), CONT: BL: 0.9 (±1.3), post-test: 0.4 (±1.0).</td>
<td>Control-drink is 25% intervention drink, ATD and CONT differ &gt;48hrs. Study investigates effects of ATD on sleep EEG.</td>
<td>Moderate</td>
</tr>
<tr>
<td>**Craen (2002)**74</td>
<td>Rand., DB?, Within SS, FH-</td>
<td>20 M</td>
<td>Good</td>
<td>POMS-depressed: No exact scores reported, no sign. difference in mood found</td>
<td>Study investigates difference in impulsivity after ATD between controls with FH+ and FH- for alcoholism.</td>
<td>Poor</td>
</tr>
<tr>
<td>**Danjou (1990)**84</td>
<td>Rand., DB?, Within SS, FH-</td>
<td>18 M</td>
<td>Not tested</td>
<td>Mood 'questionnaires' did not reveal differences in affect for ATD</td>
<td>Study investigates the effect of ATD on auditory evoked potentials with EEG.</td>
<td>Poor</td>
</tr>
<tr>
<td>**Debener (2002)**97</td>
<td>Rand., DB, Within SS, FH-</td>
<td>21 F</td>
<td>Good</td>
<td>POMS-depressed**: ATD: change -4 (±1.6 SE), CONT: change: 0.8 (±1.55); VAMS:</td>
<td>Study also examines stability of ATD response which appears to be poor</td>
<td>Good</td>
</tr>
<tr>
<td>**Ellenbogen (1996)**134</td>
<td>Rand., DB, Within SS, FH-</td>
<td>12 M</td>
<td>Good</td>
<td>VAS sadness: ATD: BL: 1.8 (±1.6), post-test: 2.0 (±1.6). CONT: BL: 1.0 (±0.7), post-test: 1.8 (±1.6).</td>
<td>Study investigates effect of ATD on amygdale reactivity to fearful faces with fMRI.</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
Table 6.2. Included studies Acute Tryptophan Depletion (ATD). (Continued)

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion</th>
<th>Results mood scores (scale, scores (±SD))</th>
<th>Remarks</th>
<th>Judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evers (2006a)</td>
<td>Rand., DB, Within SS</td>
<td>13 M</td>
<td>Good</td>
<td>POMS: No exact scores reported, no sign. difference in mood found</td>
<td>Study investigates effects of ATD on cognition (performance monitoring and response inhibition and memory) and prefrontal / hippocampal activity by fMRI</td>
<td>Moderate</td>
</tr>
<tr>
<td>Evers (2006b)</td>
<td>Rand., DB, Within SS, FH-</td>
<td>15 F</td>
<td>Good</td>
<td>POMS: No exact scores reported, no sign. difference in mood found</td>
<td>Study investigates effects of ATD on a combined cognitive and emotional task by fMRI</td>
<td>Moderate</td>
</tr>
<tr>
<td>Harrison (2002)</td>
<td>Rand., DB, Within SS, FH-</td>
<td>13 F</td>
<td>Good</td>
<td>POMS-depressed: ATD: BL: 2.07 (±0.63 SE), post-test: 0.76 (±0.63) CONT: BL: 1.61 (±0.52), post-test: 1.15 (±0.56).</td>
<td>Study also reports absence of effects of ATD and APTD on interleukin-6 activation</td>
<td>Good</td>
</tr>
<tr>
<td>Harrison (2004)</td>
<td>Rand., DB, Within SS, FH-</td>
<td>13 F</td>
<td>Good</td>
<td>VAMS: exact scores not reported, no sign. differences in mood found</td>
<td>Study also investigates memory and cognitive effects of ATD and APTD</td>
<td>Moderate</td>
</tr>
<tr>
<td>Hayward (2005)</td>
<td>Rand., DB, Betw. SS, 20 FH-</td>
<td>2*12 M+F</td>
<td>Good</td>
<td>HDRS: ATD: BL: 0.17 (±0.11 SE), post-test: 0.67 (±0.19), CONT: BL: 0.33 (±0.19), post-test: 0.41 (±0.19). BDI: ATD: BL: 1.5 (±0.56 SE), post-test: 0.58 (±0.33), CONT: BL: 2.13 (±0.57), post-test: 2.25 (±0.80). POMS and VAS: exact scores not reported. No significant effect on any scale found.</td>
<td>Study also investigates cognitive processing in formerly depressed versus healthy controls during ATD</td>
<td>Moderate</td>
</tr>
<tr>
<td>Hughes (2000)</td>
<td>Rand., DB, Within SS, FH-</td>
<td>20 M</td>
<td>Good</td>
<td>Undefined scale: no significant changes in mood found</td>
<td>Study investigates effects of ATD on auditory evoked potentials on EEG</td>
<td>Poor</td>
</tr>
<tr>
<td>Hughes (2003)</td>
<td>Rand., DB, Within SS, FH-</td>
<td>20 M</td>
<td>Good</td>
<td>VAS-sadness: ATD: BL: 90.7 (±3.5 SE), post-test: 66.7 (±4.7), CONT: BL: 94.8 (±2.6), post-test: 1.6 (±1.9). No significant main effects on any VAMS scale.</td>
<td>Study investigates effects of ATD on verbal and visuospatial learning and memory, attention and executive function</td>
<td>Good</td>
</tr>
<tr>
<td>Klaassen (1999a)</td>
<td>Rand., DB, Within SS</td>
<td>13 M+F</td>
<td>Good</td>
<td>POMS-depressed: ATD: BL: 0.6 (±1.3), post-test: 1.5 (±3.3), CONT: BL: 0.5 (±1.0), post-test: 0.2 (±0.6). VAS-depressed: ATD: BL: 2.7 (±6.6), post-test: 8.2 (±18.2), CONT: BL: 3.2 (±6.3), post-test: 4.5 (±3.0).</td>
<td>Study investigates specificity of ATD vs LYS-depletion on mood-effects and memory.</td>
<td>Good</td>
</tr>
<tr>
<td>Klaassen (1999b), Klaassen (2002)</td>
<td>Rand., DB, Within SS, FH+ &amp; FH-</td>
<td>11 FH- M+F</td>
<td>Good</td>
<td>POMS-depressed: scores not completely reported. ATD-CONT differences: FH- 0.0 (±0.45)</td>
<td>No baseline POMS-scores provided (‘non-significant different ces’). Study compared FH+ group with FH-. The FH+ ves responded to tryptophan depletion with a significant lowering of POMS scores, but FH- ves did not. Also memory effects measured.</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
Table 6.2. Included studies Acute Tryptophan Depletion (ATD). (Continued)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Koszycki (1996)*</td>
<td>Rand., DB, Betw. SS, FH-</td>
<td>2*20 M</td>
<td>Good</td>
<td>VAMS-depressed: ATD: BL: 1.5 (±1.0 SE), post-test: 0.8 (±0.6). CONT: BL: 3.4 (±1.8), post-test: 4.5 (±1.4).</td>
<td>Study investigates effect of administration of cholecystine-tetrapeptide after ATD on occurrence of panic-attacks</td>
<td>Good</td>
</tr>
<tr>
<td>Leyton (1999)*</td>
<td>Rand., DB, Betw. SS, FH-</td>
<td>15 ATD 14 CONT</td>
<td>Good</td>
<td>POMS**: ATD: change: -3.28 (±1.4 SE). CONT: change: -1.4 (±1.2 SE). VAMS-depressed: ATD: BL: 1.0 (±0.9), post-test: 0.8 (±0.8), change: -0.3 (±0.8). CONT: BL: 0.4 (±0.5), post-test: 0.0 (±0.4), change: 0.1 (±0.4).</td>
<td>POMS; no exact scores reported. Third arm with APTD included</td>
<td>Moderate</td>
</tr>
<tr>
<td>McAllister-Williams (2002)</td>
<td>Rand., DB, Within SS, FH-</td>
<td>14 M</td>
<td>Good</td>
<td>POMS: No exact scores reported, no sign. difference in mood found</td>
<td>Study investigates effects of ATD on event related brain potentials on EEG during an episodic memory task</td>
<td>Moderate</td>
</tr>
<tr>
<td>Leyton (2000)*</td>
<td>Rand., DB, Within SS, FH-</td>
<td>14 ATD 14 CONT</td>
<td>Good</td>
<td>POMS**: ATD: BL: 1.8, post-test: 1.2. CONT: BL: 2.9, post-test: 1.4. VAMS-depressed: ATD: BL: 1.0, post-test: 0.8, change: -0.2 (±0.8). CONT: BL: 1.0, post-test: 0.8, change: -0.2 (±0.8).</td>
<td>No clear presentation of mood-scores; study compares ATD + CO2-challenge in healthy controls vs. patients with panic disorder</td>
<td>Poor</td>
</tr>
<tr>
<td>Moreno (1999), Moreno (2000)</td>
<td>Rand., DB, Within SS, FH-</td>
<td>12 M+F</td>
<td>Good</td>
<td>HDRS-25, POMS and IDS scores not clearly reported. ATD did not cause significant increases in depressive symptoms in healthy control subjects.</td>
<td>Study compared ATD effects in healthy controls vs. patients with MDD in remission. Instead of placebo a 25% strength depletion drink was used as a control.</td>
<td>Poor</td>
</tr>
<tr>
<td>Murphy (2002)</td>
<td>Rand., DB, Within SS</td>
<td>11 F</td>
<td>Good</td>
<td>Scores reported separately for groups with ATD first and CONT first. VAMS-sadness: ATD first: ATD: change: -1.4 (±1.5). CONT: change: -1.0 (±1.5). CONT first: ATD: change: -1.0 (±1.5). CONT: change: -1.0 (±1.5).</td>
<td>Study investigates cognitive and emotional processing after ATD</td>
<td>Good</td>
</tr>
<tr>
<td>Neumeister (2002)*</td>
<td>Rand., DB, Within SS, FH-</td>
<td>24 F</td>
<td>Good</td>
<td>HDRS-21: s/s subtype (n= 4): ATD: change 8.5 (±2.9). CONT: change -1.0 (±1.5). CONT: change -1.0 (±1.5). I/l subtype (n=10): ATD: change -0.1 (±1.3). CONT: change -0.1 (±1.3).</td>
<td>Study investigates effect of bi-allelic serotonin transporter promoter (5-HTPR) genotype and effect of ATD</td>
<td>Good</td>
</tr>
<tr>
<td>Neumeister (2004)*</td>
<td>Rand., DB, Within SS, FH-</td>
<td>19 M+F</td>
<td>Good</td>
<td>HDRS-24**: s/s subtype (n= 7): ATD: BL: 0.5 (±0.1), post-test: 3.4 (±0.6). CONT: BL: 0.5 (±0.1), post-test: 2.6 (±0.5).</td>
<td>Study mainly investigates regional cerebral blood-flow changes during ATD with PET-scans</td>
<td>Good</td>
</tr>
<tr>
<td>Neumeister (2006)*</td>
<td>Rand., DB, Within SS, FH-</td>
<td>26 M+F</td>
<td>Good</td>
<td>HDRS**: s/s subtype (n= 12): ATD: BL: 0.4 (±0.4), post-test: 3.6 (±1.4). CONT: BL: 0.8 (±0.4), post-test: 2.4 (±0.8). I/l subtype (n=12): ATD: BL: 0.4 (±0.4), post-test: 5.2 (±1.1). CONT: BL: 0.4 (±0.4), post-test: 2.5 (±0.6). L/l subtype (n= 7): ATD: BL: 0.4 (±0.4), post-test: 2.4 (±1.3). CONT: BL: 0.2 (±0.4), post-test: 1.2 (±0.8).</td>
<td>Study investigates effect of tri-allelic serotonin transporter promoter (5-HTPR) genotype and effect of ATD on mood and regional cerebral metabolism with PET</td>
<td>Good</td>
</tr>
</tbody>
</table>
Table 6.2. Included studies Acute Tryptophan Depletion (ATD). (Continued)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oldman (1994)77</td>
<td>Rand., DB, Within SS</td>
<td>12</td>
<td>Good</td>
<td>POMS-depressed: ATD: BL: 11.1 (±0.52 SE), post-test: 0.9 (±0.58). CONT: BL: 1.9 (±0.92), post-test: 3.3 (±0.89). VASS-sad: ATD: BL: 15.0 (±4.99 SE), post-test: 9.2 (±4.5). CONT: BL: 11.7 (±4.73), post-test: 8.3 (±3.87).</td>
<td>In this study a second placebo-arm (plain water) is used; comparisons here are for balanced drink as CONT. Effects of ATD on appetite investigated also.</td>
<td>Good</td>
</tr>
<tr>
<td>Park (1994)75</td>
<td>Rand., DB, Within SS</td>
<td>12</td>
<td>Good</td>
<td>VAMS-sad: ATD: BL: 6.7 (±5.7 SE), post-test: 6.6 (±8.1). CONT: BL: 3.2 (±3.0), post-test: 5.6 (±6.8).</td>
<td>Study investigates cognitive performance after ATD</td>
<td>Good</td>
</tr>
<tr>
<td>Praschak-Rieder (2005)100</td>
<td>Rand., DB, Within SS, FH-</td>
<td>14</td>
<td>Good</td>
<td>No exact VAS-ratings provided in study; “none of the 14 subjects experienced a transient deterioration in mood [...] levels.</td>
<td>Study primarily investigates effects of ATD on serotonin transporter density/affinity during ATD</td>
<td>Moderate</td>
</tr>
<tr>
<td>Ravindran (1999), Knott (1999)85</td>
<td>Rand., DB, Betw. SS, FH-</td>
<td>2*13</td>
<td>Good</td>
<td>POMS-change:ATD -1.29 (±0.94) vs CONT +1.41(±0.55). Significant changes in mood, but this occurred in 5 of the 13 subjects only, while in others in the ATD-group no decrease of mood was observed.</td>
<td>No clear presentation of data on mood-scores; probably SDs reported. Study also investigates effect of fenfluramine administration after ATD and effect of ATD on immune measures and EEG</td>
<td>Moderate</td>
</tr>
<tr>
<td>Richell (2005)58</td>
<td>Rand., DB, Betw. SS</td>
<td>15 + 13</td>
<td>M+F</td>
<td>Good</td>
<td>No baseline data on POMS or VAS reported, only measures during the noise-stress paradigm reported after ATD, no sign. differences in mood found.</td>
<td>Study especially investigates difference in response to noise-stress by ATD vs CONT.</td>
</tr>
<tr>
<td>Rubinsztein (2001)56</td>
<td>Rand., DB, Betw. SS, 80% FH-</td>
<td>2*15</td>
<td>M+F</td>
<td>Good</td>
<td>POMS-depressed: no exact scores reported. Scores were similar between test and placebo groups, also on VAMS, no sign. differences in mood found.</td>
<td>Study investigates effects of ATD on cognitive test concerning attention</td>
</tr>
<tr>
<td>Schmeck (2002)79</td>
<td>Rand., DB, Within SS, FH-</td>
<td>12</td>
<td>Good</td>
<td>Eigenschaftswörterliste: no exact scores reported. High aggressive women were sign. more depressed after ATD than other groups. High aggressive men did not show an effect of ATD on mood. The mood of non-aggressive men and women did not change after ATD.</td>
<td>Study looked at different reaction to ATD in high and low-aggressive males and females.</td>
<td>Poor</td>
</tr>
<tr>
<td>Schmitt (2000)57</td>
<td>Rand., DB, Within SS, FH-</td>
<td>17</td>
<td>M+F</td>
<td>Good</td>
<td>POMS-depression: ATD: BL: 0.92 (±0.11 SE),post-test: 0.93 (±0.12). CONT: BL: 0.95 (±0.11), post-test: 0.96 (±0.13)</td>
<td>Study investigates cognitive performance after ATD</td>
</tr>
<tr>
<td>Shansis (2000)58</td>
<td>Rand., DB, Within SS, 7/12 FH-</td>
<td>12</td>
<td>Good</td>
<td>POMS-depressed &amp; VAS: no exact scores reported. Mood changes on control day not sign. different than on ATD day for any of the six dimensions in the POMS and VAS.</td>
<td>Previous mood-disorder in one subject (not excluded). Study inves tigates effects of ATD on attention and memory</td>
<td>Poor</td>
</tr>
<tr>
<td>Smith (1987)112</td>
<td>Rand.,DB, Within SS</td>
<td>80</td>
<td>Good</td>
<td>MAACL **:ATD: BL: 10.9, post-test: 18.2. CONT: BL: 9.5, post-test: 8.2. No SDs reported. Significant lowering of mood scores by ATD.</td>
<td>Study investigates ATD in negative and positive environment. Both conditions show a significant change on mood scores.</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
Table 6.2. Included studies Acute Tryptophan Depletion (ATD). (Continued)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith (1997)b</td>
<td>Rand., DB, Within SS, FH-</td>
<td>11</td>
<td>M+F</td>
<td>Good</td>
<td>VAS-sadness: no exact scores reported. No significant main effects were seen. VAS-happy for females (n=6) reported separately**: ATD: BL: 69.7 (±5.6 SE), post-test: 61.6 (±6.4). CONT: BL: 71.4 (±5.3), post-test: 66.1 (±6.1). For males (n=5)**: ATD: BL: 72.2, post-test: 67.8. CONT: BL: 77.2, post-test: 70.6. No SEs reported</td>
<td>Study also investigates effect of ATD when sad-mood is induced. Data for females VAS-happy used in meta-analysis</td>
</tr>
<tr>
<td>Stewart (2002)⁵⁰</td>
<td>Rand., DB, Within SS, FH-</td>
<td>27</td>
<td>M+F</td>
<td>Good</td>
<td>POMS depression: ATD: BL: 1.7 (±2.6), post-test: 1.1 (±2.3), change: -0.6 (±2.1). CONT: BL: 3.4 (±6.0), post-test: 2.4 (±4.8), change: -1.0 (±4.6).</td>
<td>Study also investigates effect modification by high vs low neuroticism scores during ATD (on mood and psychometric performance).</td>
</tr>
<tr>
<td>Talbot (2006a)⁷⁷</td>
<td>Rand., DB, Within SS, FH+</td>
<td>17 + 15</td>
<td>M+F</td>
<td>Good</td>
<td>VAMS: No exact scores reported, no sign. difference in mood found</td>
<td>Study investigates effects of ATD on decision making and learning</td>
</tr>
<tr>
<td>Talbot (2006b)⁷⁵</td>
<td>Rand., DB, Within SS, FH+</td>
<td>16</td>
<td>M</td>
<td>Good</td>
<td>VAMS: ATD: change -0.5 (±6.1). CONT: change -5.7 (±8.3). BDI: ATD: change -0.0 (±1.4). CONT: change 0.4 (±1.5)</td>
<td>Study investigates effect of ATD on Regional Cerebral Blood Flow by SPECT</td>
</tr>
<tr>
<td>Voderholzer (1998)⁹³</td>
<td>Rand., DB, Within SS, FH+</td>
<td>12</td>
<td>M+F</td>
<td>Good</td>
<td>HDRS-6: ATD: BL: 0.83 (±1.8), post-test: 0.42 (±1.2). CONT: BL: 0.17 (±0.4), post-test: 0.42 (±0.7).</td>
<td>Study investigates effects of ATD on sleep EEG</td>
</tr>
<tr>
<td>Weltzin (1994)⁹⁹</td>
<td>Rand., DB, Within SS</td>
<td>9</td>
<td>F</td>
<td>Good</td>
<td>Only peak-changes for ATD and CONT are provided for an unclearly defined mood-scale: ATD: 0.78. CONT 0.0. SDs not reported. SD of difference estimated conservatively from p&lt;0.05: ±1.01.</td>
<td>Study investigates effect of ATD on mood in bulimic patients (n=13) and healthy controls (n= 9).</td>
</tr>
<tr>
<td>Weltzin (1995)⁹⁹</td>
<td>Rand., DB, Within SS</td>
<td>10</td>
<td>F</td>
<td>Good</td>
<td>Only peak-changes between ATD and CONT are provided for an unclearly defined mood-scale: 0.0 (±1.3), this value used in meta-analysis</td>
<td>Study investigates effect of ATD on mood and short-term eating beha viour in bulimic patients (n=10) and healthy controls (n= 10).</td>
</tr>
<tr>
<td>Yatham (2001)⁹⁴</td>
<td>Rand., DB, Within SS, FH+</td>
<td>10</td>
<td>F</td>
<td>Good</td>
<td>HDRS-20, POMS: No exact scores reported, no sign. difference in mood found</td>
<td>Study investigates effects of ATD on 5-HT₂ receptors with PET</td>
</tr>
<tr>
<td>Young (1985)⁷³</td>
<td>Rand., DB, Betw. SS</td>
<td>3*12</td>
<td>M</td>
<td>Good</td>
<td>No clear MAACL-scores or SDs reported per group for ATD vs CONT. Sign. decrease of mood after ATD.</td>
<td>Study investigates effects of ATD on proofreading task with or without distraction</td>
</tr>
</tbody>
</table>

A2. Healthy controls with +ve family history for MDD

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benkelfat (1994)⁶⁹</td>
<td>Rand., DB, Within SS, FH+</td>
<td>20</td>
<td>M</td>
<td>Good</td>
<td>POMS-depressed: ATD: BL: 55.2 (±1.7 SE), post-test: 51.2 (±1.6), change: -4.0 (±1.6). CONT: BL: 54.0 (±1.6), post-test: 57.3 (±1.1), change: 3.3 (±1.6). Six (30%) of people showed a decline of 10 points or greater on POMS-D.</td>
<td>POMS changed in FH+ sample, not in FH- sample in same study.</td>
</tr>
<tr>
<td>Author (year)*</td>
<td>Design</td>
<td>N</td>
<td>Adeq. of depletion</td>
<td>Results mood scores (scale, scores (±SD))</td>
<td>Remarks</td>
<td>Judgement†</td>
</tr>
<tr>
<td>---------------</td>
<td>--------</td>
<td>---</td>
<td>------------------</td>
<td>------------------------------------------</td>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>Ellenbogen (1999)</td>
<td>Rand., DB, Within SS, FH+</td>
<td>13</td>
<td>F</td>
<td>Good</td>
<td>POMS-depressed**: 1st ATD: change: -0.3 (±1.7 SE). CONT: change: -1.1 (±1.5). VAMS no exact scores reported. No significant change in mood on any of the POMS or VAMS items.</td>
<td>Possible selection bias by exclusion of many former patients mentioned: extremely FH+ sample, but never depressed. Study found poor temporal stability of ATD in repeated testing.</td>
</tr>
<tr>
<td>Neumeister (2002)</td>
<td>Rand., DB, Within SS, FH+</td>
<td>21</td>
<td>F</td>
<td>Good</td>
<td>HDRS-21: s/s subtype (n=5): ATD: change 9.2 (±2.1). CONT: 0.8 (±0.5) 1/s subtype (n=7): ATD: change 10.0 (±4.5). CONT: 0.1 (±0.4) 1/l subtype “*” (n=9): ATD: change 0.4 (±1.2). CONT: -0.2 (±1.3)</td>
<td>Study investigates effect of bi-allelic serotonin transporter promoter (5-HTTPr) genotype and effect of ATD</td>
</tr>
<tr>
<td>Stewart (2002)</td>
<td>Rand., DB, Within SS, FH-</td>
<td>5</td>
<td>M+F</td>
<td>Good</td>
<td>POMS-depression:ATD: BL: 5.2 (±10.0), post-test: 5.6 (±10.4), change: 0.4 (±10.5). CONT: BL: 6.4 (±12.7), post-test: 4.2 (±7.4), change: -2.2 (±5.5).</td>
<td>See above in section “healthy controls”</td>
</tr>
</tbody>
</table>

B1. Patients with MDD previously, but currently in remission (without medication)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion</th>
<th>Results mood scores (scale, scores (±SD))</th>
<th>Remarks</th>
<th>Judgement†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cassidy (1997)</td>
<td>Rand., DB, Within SS 1-4 days after ECT</td>
<td>5</td>
<td>M+F Rx-free &lt;3 mths</td>
<td>Good</td>
<td>MADRS: ATD: BL: 4.2 (±3.5), post-test: 3.6 (±3.2). CONT: BL: 4.8 (±2.9), post-test: 4.0 (±3.5).</td>
<td>Study investigates recurrence of MDD after ATD in patients successfully treated with ECT</td>
</tr>
<tr>
<td>Delgado (1991)</td>
<td>Rand., DB, Within SS ≥2 weeks after ‘stability’</td>
<td>69</td>
<td>M+F Rx-free ≥3 mths</td>
<td>Good</td>
<td>HDRS-25: exact scores not reported; no sign. difference by ATD. 30% experienced an improvement of mood the day after ATD. Depressive relapse: No exact data provided.</td>
<td>In original studies Bipolar patients included; unclear whether these were excluded in this pooled analysis</td>
</tr>
<tr>
<td>Hayward (2005)</td>
<td>Rand., DB, Within SS ≥2 months remission (average 29 months)</td>
<td>27</td>
<td>M+F Rx-free ≥3 mths</td>
<td>Good</td>
<td>HDRS-25: ATD: BL: 1.0 (±0.41 SE), post-test: 1.38 (±0.36), CONT: BL: 0.42 (±0.23), post-test: 0.50 (±0.19). BDI: ATD: BL: 2.25 (±0.76 SE), post-test: 1.75 (±0.68), CONT: BL: 2.17 (±0.63), post-test: 1.63 (±0.53). POMS and VAS: exact scores not reported. No significant effect on any scale found.</td>
<td>Study also investigates cognitive processing in formerly depressed versus healthy controls during ATD</td>
</tr>
<tr>
<td>Moreno (1999), Moreno (2000)</td>
<td>Rand., DB, Within SS ≥3 months remission</td>
<td>12</td>
<td>M+F Rx-free ≥3 mths</td>
<td>Good</td>
<td>HDRS-25, POMS and IDS scores not clearly reported. ATD caused non-significant increases in depressive symptoms in history-positive subjectsPost-test: mean peak HDRS scores ATD: 14 (±6), ‘CONT’: 10 (±7).</td>
<td>Study compared healthy controls with patients with MDD in remission. Instead of placebo a 25% strength depletion drink was used as a control.</td>
</tr>
</tbody>
</table>
### Table 6.2. Included studies Acute Tryptophan Depletion (ATD), (Continued)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neumeister (2004)**</td>
<td>Rand., DB, Within SS ≥3 months remission</td>
<td>27 M+F Rx-free ≥3 mths</td>
<td>Good</td>
<td>HDRS-24**: ATD: BL: 1.2 (±0.2), post-test: 11.5 (±0.3). CONT: BL: 0.7 (±0.2), post-test: 2.1 (±0.3). 16/27 (59%) patients experienced transient return of MDD during ATD vs 0/27 during CONT</td>
<td>See above in section “healthy controls”</td>
<td>Good</td>
</tr>
<tr>
<td>Neumeister (2006)**</td>
<td>Rand., DB, Within SS ≥3 months remission</td>
<td>27 M+F Rx-free ≥3 mths</td>
<td>Good</td>
<td>HDRS**: s/s subtype (n= 7): ATD: BL: 0.4 (±0.4 SE), post-test: 7.0 (±1.4). CONT: BL: 0.6 (±0.4), post-test: 2.6 (±0.8). s/l subtype (n= 12): ATD: BL: 0.8 (±0.4), post-test: 11.4 (±1.0). CONT: BL: 0.8 (±0.4), post-test: 2.4 (±0.6). l/l subtype (n= 8): ATD: BL: 2.0 (±0.3), post-test: 14.6 (±1.2). CONT: BL: 0.9 (±0.4), post-test: 1.7 (±0.6).</td>
<td>Study investigates effect of tri-allelic serotonin transporter promoter (5-HTT) genotype and effect of ATD on mood and regional cerebral metabolism with PET</td>
<td>Good</td>
</tr>
<tr>
<td>O'Reardon (2004)**</td>
<td>Rand., DB, Within SS 5 months remission</td>
<td>10 M+F Post-CBT</td>
<td>Good</td>
<td>HDRS-24**: ATD: change: 4.97 (±1.0 SE). CONT: change: 3.22 (±1.0). 0/10 patients relapsed in either condition</td>
<td>Currently not using antidepressant medication, responders to CBT treatment.</td>
<td>Good</td>
</tr>
<tr>
<td>Smith (1997)**</td>
<td>Rand., DB, Within SS ≥6 months remission</td>
<td>15 F Rx-free ≥6 mths</td>
<td>Good</td>
<td>HDRS-12: ATD: change (7hr): 7.3. CONT: change: 0.1 SD of difference ATD-CONT calculated from 95% CI: 7.2 (±4.9)</td>
<td>5/15 patients experienced relapse during ATD, 0/15 during CONT</td>
<td>Moderate</td>
</tr>
<tr>
<td>Aberg-Wistedt (1998)**</td>
<td>Rand., DB, Betw. SS After response</td>
<td>12+8 M+F CIT</td>
<td>Insufficient</td>
<td>MADRS: exact scores not reported. Five of the 12 patients with ATD showed a Worsening of depressive symptoms in ATD: 5/12 (MADRS 12-71%). CONT: 0/8.</td>
<td>Study also investigates relation with relapse due to ATD and other variables (plasma cortisol, platelet-serotonin)</td>
<td>Moderate</td>
</tr>
<tr>
<td>Booij (2005a)**</td>
<td>Rand., DB, Within SS After ≥ partial remission</td>
<td>20 M+F SSRI SNRI</td>
<td>Good</td>
<td>MADRS: ATD: BL: 4.6 (±0.9 SE), post-test: 7.9 (±1.8). CONT: BL: 3.7 (±0.9), post-test: 3.7 (±0.9). Relapse of MDD: ATD: 7/20; CONT: 0/20 assumed.</td>
<td>Include of subjects with HDRS&lt;15 (partial remission). A 25% strength tryptophan deficient mixture was used as CONT.</td>
<td>Good</td>
</tr>
<tr>
<td>Booij (2006)**</td>
<td>Rand., DB, Within SS After ≥ partial remission</td>
<td>19 M+F SSRI</td>
<td>Good</td>
<td>HDRS-17 scores sign. lower after ATD: SI+: (n= 8): ATD: BL: 2.5 (±0.8 SE), post-test: 7.6 (±2.0). CONT: Not given SI- (n= 11): ATD: BL: 2.9 (±0.7), post-test: 3.4 (±0.6). CONT: Not given MADRS scores sign. lower after ATD: SI+: ATD: BL: 5.0 (±1.2 SE), post-test 12.00 (±2.4). CONT: Not given SI-: ATD: BL: 5.4 (±1.3 SE), post-test: 9.3 (±1.7). CONT: Not given</td>
<td>Inclusion of subjects with HDRS&lt;15 (partial remission). A 25% strength tryptophan deficient mixture was used as CONT. No data for ATD mood-ratings after CONT provided. Study also investigates effect of ATD on Heart Rate Variability (HRV) and impulsivity and the interaction by +/− history of suicidal ideation (SI+/SI−)</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

### B2. Patients with MDD previously, but currently in remission (currently using medication)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Booij (2005a)**</td>
<td>Rand., DB, Within SS After ≥ partial remission</td>
<td>19 M+F SSRI</td>
<td>Good</td>
<td>HDRS-17 scores sign. lower after ATD: SI+: (n= 8): ATD: BL: 2.5 (±0.8 SE), post-test: 7.6 (±2.0). CONT: Not given SI- (n= 11): ATD: BL: 2.9 (±0.7), post-test: 3.4 (±0.6). CONT: Not given MADRS scores sign. lower after ATD: SI+: ATD: BL: 5.0 (±1.2 SE), post-test 12.00 (±2.4). CONT: Not given SI-: ATD: BL: 5.4 (±1.3 SE), post-test: 9.3 (±1.7). CONT: Not given</td>
<td>Inclusion of subjects with HDRS&lt;15 (partial remission). A 25% strength tryptophan deficient mixture was used as CONT. No data for ATD mood-ratings after CONT provided. Study also investigates effect of ATD on Heart Rate Variability (HRV) and impulsivity and the interaction by +/− history of suicidal ideation (SI+/SI−)</td>
<td>Moderate</td>
</tr>
</tbody>
</table>
Table 6.2. Included studies Acute Tryptophan Depletion (ATD). (Continued)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delgado (1991)**</td>
<td>Rand., DB, Within SS ≥2 weeks after ‘stability’</td>
<td>46</td>
<td>M+F various</td>
<td>Good</td>
<td>HDRS-25: exact scores not reported. Depressive relapse: ATD: 24/46, CONT: 0/46.</td>
<td>Heterogeneous population currently using various AD. BUP or DES-treatment lower chance of relapse. In original studies Bipolar patients included; unclear whether these were excluded in this pooled analysis</td>
</tr>
<tr>
<td>Landolt (2003)**</td>
<td>Rand., DB, Within SS,PHZ After remission</td>
<td>5</td>
<td>M+F PHZ</td>
<td>Good</td>
<td>POMS-depression: no BL measures were given separately for ATD and CONT: BL: 0.4 (± 0.4 SE). ATD: post-test: 2.0 (±1.6). CONT: post-test: 1.0 (±0.4). Post-test measures taken next morning after ATD or CONT at 16:00h.</td>
<td>Study primarily investigates effect of ATD on recurrence of REM-sleep while using PHZ</td>
</tr>
<tr>
<td>Moore (1998)**</td>
<td>Rand., DB, Within SS ≥2 months remission</td>
<td>10</td>
<td>M FLX SER PAR</td>
<td>Good</td>
<td>POMS-depressed: ATD: BL: 3.0 (±1.3 SE), post-test: 2.7 (±1.4). CONT: BL: 3.0 (±1.3), post-test: 3.3 (±1.6). HDRS-25 ratings not exactly reported but did not change significantly</td>
<td>Study also investigates effects of ATD on REM-sleep. A 25% strength tryptophan deficient mixture was used as CONT. Subjects in remission &gt;2 months</td>
</tr>
<tr>
<td>Morris (1999)**</td>
<td>Rand., DB, Within SS Full remission</td>
<td>8</td>
<td>M SSRI AMI TRAZ PHZ TCP Li</td>
<td>Good</td>
<td>Modified HDRS-12, BDI. No exact scores given. HDRS increase &gt;5 points (relapse): ATD: 7/8, CONT: 0/8</td>
<td>Study investigates brain metabolism with PET during ATD and associations with MDD relapse. Six of 8 participants received ≥6 months antidepressants. No indication which drugs were currently used.</td>
</tr>
</tbody>
</table>
### Table 6.2. Included studies Acute Tryptophan Depletion (ATD). (Continued)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>O’Reardon (2004)</strong></td>
<td>Rand., DB, Within SS 5 months remission</td>
<td>9</td>
<td>M+F FLX PAR</td>
<td>Good</td>
<td>HDRS-24: ATD: change: 8.22 (±1.2), CONT: change: 2.56 (±1.2). 3/9 patients relapsed during ATD vs 0/9 during control</td>
<td>Currently using antidepressant medication, responders to SSRI treatment.</td>
</tr>
<tr>
<td><strong>Praschak-Rieder (2004)</strong></td>
<td>Rand., DB, Within SS 2 months remission</td>
<td>8</td>
<td>M CIT</td>
<td>Good</td>
<td>HDRS-21: ATD: BL: 3.88 (±2.6), post-test: 12.3 (±5.4). CONT: no scores reported. Relapse: ATD: 6/8 patients, CONT: no numbers reported 0/8 assumed for analysis</td>
<td>Study primarily investigates effects of ATD on regional 5-HT₁A binding potential with PET.</td>
</tr>
<tr>
<td><strong>Smith (1999)</strong></td>
<td>Rand., DB, Within SS ≥3 months remission</td>
<td>8</td>
<td>M various</td>
<td>Good</td>
<td>HDRS: no exact scores reported. Relapse: ATD: 6/8 patients, CONT: 1/8</td>
<td>Study primarily investigates effects of ATD on regional brain activity with PET. Subjects in remission &gt;6 months. 2 patients did not use antidepressants, 6 used various types (SSRI, TCA, MAO-I)</td>
</tr>
<tr>
<td><strong>Spillmann (2001)</strong></td>
<td>Rand., DB, Within SS ≥3 months remission</td>
<td>10</td>
<td>M+F FLX SER VLX PAR</td>
<td>Good</td>
<td>HDRS-6: ATD: BL: 1.1 (±1.2) post-test (7h): 4.7 (±3.4), change: 3.6 (±3.1); CONT: BL: 1.4 (±1.7) post-test: 1.0 (±1.4), change: -0.4 (±1.2). Increase in scores on the HDRS-6 was significantly higher for ATD than for CONT.</td>
<td>In remission after SSRI, currently using antidepressant medication. 5 initial participants dropped out.</td>
</tr>
</tbody>
</table>

C. Patients currently with MDD (with or without antidepressants)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adeq. of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delgado (1994)</td>
<td>Rand., DB, Within SS</td>
<td>43</td>
<td>M+F ≥ 2 wks Rx-free</td>
<td>HDRS-25: No exact mood-scores provided; One day after ATD 16/43 patients experienced improvement of mood, 10/46 decrease in mood relative to CONT</td>
<td>Improvement of mood ratings 24 hr after full ATD associated with treatment response after antidepressant treatment</td>
<td>Poor</td>
</tr>
<tr>
<td>Price (1997)</td>
<td>Rand., DB, Within SS</td>
<td>22</td>
<td>M+F Rx-free</td>
<td>VAS, HDRS-25: No exact scores provided; scores unaffected by ATD.</td>
<td>Study investigates the (biochemical) effects of m-Chloro phenyl piperazine (mCCP, serotonin-agonist) on HDRS, cortisol, prolactin and growth-hormone during ATD</td>
<td>Poor</td>
</tr>
<tr>
<td>Price (1998)</td>
<td>Rand., DB, Within SS</td>
<td>36</td>
<td>M+F ≥ 3 wks Rx-free</td>
<td>HDRS-25: ATD: BL: 28 (+7), post-test 28 (+7). CONT: 27 (+9). Study included 4 patients with BD. Study investigates the (biochemical) effects of rapid tryptophan infusion on HDRS, cortisol, prolactin and growth-hormone after ATD</td>
<td>Study included 4 patients with BD. Study investigates the (biochemical) effects of rapid tryptophan infusion on HDRS, cortisol, prolactin and growth-hormone after ATD</td>
<td>Good</td>
</tr>
</tbody>
</table>

AD = antidepressants; AMI = Amitriptyline; AMPT = Alpha-methyl-para-tyrosine; ATD = Acute Tryptophan Depletion; BD = Bipolar Disorder; BDI = Beck Depression Inventory; Betw.SS = Between-subjects parallel groups design; BL = Baseline; BUP = bupropion; CBT = Cognitive Behaviour therapy; CIT = citalopram; CONT = Control; DB = double-blind; DES = desipramine; ECT = Electro-convulsive therapy; EEG = electro-encephalography; FLX = fluoxetine; HAMA = Hamilton Anxiety Scale; HDRS = Hamilton Depression Scale; IMI = imipramine; LYS = lysine; Li = lithium; MAACL = Multiple Affect Adjective CheckList; MADRS = Montgomery Åsberg Depression Rating Scale; MAO-I = Monoamine oxidase inhibitor; MDD = Major Depressive Disorder; MEG = magneto-encephalography; MIR = mirtazapine; NOR = nortriptyline; PAR = paroxetine; PHZ = phenelzine; POMS = Profile of Mood States; Rand. = randomized; SD = Standard deviation; SE = Standard Error; TCA = tricyclic antidepressants; TCP = Tranylcypromine; TRAZ = Trazodone; VA(M)S = Visual Analogue (Mood) Scale; VLX = venlafaxine; Within SS = Within-subjects cross over design.

* studies included in the meta-analysis in bold
† Plasma tryptophan declined by at least 50%
‡ SD given unless indicated otherwise (e.g. SE)
§ Judgement in the context of applicability for this meta-analysis
** Estimated from figure
†† Unclear whether partially the same population is reported (Weltzin (1994 and 1995))
Table 6.3. Included studies with Para-chlorophenylalanine (PCPA).

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Design</th>
<th>N</th>
<th>Adequacy of depletion</th>
<th>Results mood scores (scale, scores (±SD))</th>
<th>Remarks</th>
<th>Judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shopsin (1975)</td>
<td>Within SS, MDD in remission by IMI</td>
<td>1</td>
<td>Not measurable; poor</td>
<td>PCPA given in small doses for relatively short periods caused a reversal of the antidepressant effect of IMI. HDRS, no exact mood scores reported.</td>
<td>Note: n = 1, 1 patient with BD and 3 patients treated with AMPT not included here.</td>
<td>Poor</td>
</tr>
<tr>
<td>Shopsin (1976)</td>
<td>Within SS, MDD in remission by TCP</td>
<td>3</td>
<td>Not reported; poor</td>
<td>All patients experienced a relapse. HDRS, no exact mood scores reported.</td>
<td>2 patients with BD not included here.</td>
<td>Poor</td>
</tr>
</tbody>
</table>

PCPA = Para-chlorophenylalanine. For other symbols and abbreviations see footnote in Table 6.2.
Table 6.4. Included studies Acute Phenylalanine/Tyrosine Depletion (APTD).

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adequacy of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judge-ment§</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A1. Healthy controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coupland (2001)^81</td>
<td>Rand., DB, Within SS</td>
<td>5 M</td>
<td>Not reported</td>
<td>POMS-depressed: APTD: change: -0.2 (±2.7) CONT: change: 0.2 (±2.6).</td>
<td>Study investigates effect of pentagastrine to induce anxiety after APTD</td>
<td>Moderate</td>
</tr>
<tr>
<td>Harrison (2002)^37</td>
<td>Rand., DB, Within SS, FH-</td>
<td>13 F</td>
<td>Good</td>
<td>POMS-depressed: APTD: BL: 0.92 (±0.56 SE), post-test: 1.15 (±0.55) CONT: BL: 1.61 (±0.52), post-test: 1.15 (±0.56).</td>
<td>Study also reports absence of effects of ATD and APTD on interleukin-6 activation</td>
<td>Good</td>
</tr>
<tr>
<td>Harrison (2004)^38</td>
<td>Rand., DB, Within SS, FH-</td>
<td>13 F</td>
<td>Good</td>
<td>VAMS, no significant changes observed; exact scores not reported</td>
<td>Study also investigates memory and cognitive effects of ATD and APTD</td>
<td>Moderate</td>
</tr>
<tr>
<td>Leyton (1999), Leyton (2000)^39,40</td>
<td>Rand., DB, Betw. SS, FH-</td>
<td>12 APTD 14 CONT F</td>
<td>Good</td>
<td>POMS**: APTD: change: -2.39 (±2.1 SE). CONT: change -1.04 (±1.2). VAMS-depressed: APTD: BL: 1.3 (±1.6 SD), post-test: 0.4 (±0.5), change: 0.4 (±0.4).</td>
<td>POMS; no exact scores reported. Third arm with ATD included</td>
<td>Moderate</td>
</tr>
<tr>
<td>Lythe (2005)^61</td>
<td>Rand., DB, Within SS</td>
<td>12 M+F</td>
<td>Good</td>
<td>VAMS, no significant changes observed; exact scores not reported</td>
<td>Study investigates effect of APTD on neuro psychological and subjective measures</td>
<td>Moderate</td>
</tr>
<tr>
<td>McLean (2004)^62</td>
<td>Rand., DB, Betw. SS</td>
<td>19+20M+F</td>
<td>Good**</td>
<td>VAMS-contentedness: APTD: BL: 10.2 (±0.7 SE), post-test: 11.5 (±0.5), CONT: BL: 11.8 (±0.5), post-test: 11.8 (±0.5).</td>
<td>Study investigates effect of APTD on neuro psychological and subjective measures</td>
<td>Good</td>
</tr>
</tbody>
</table>

**A2. Healthy controls with +ve family history for MDD**

Table 6.4. Included studies Acute Phenylalanine/Tyrosine Depletion (APTD). (Continued)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adequacy of depletion†</th>
<th>Results mood scores (scale, scores (±SD))</th>
<th>Remarks</th>
<th>Judge-ment§</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B1. Patients with MDD previously, but currently in remission (without medication)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McTavish (2005)*64</td>
<td>Rand., DB, Within SS ≥ 6 months re mission, aver age 41 months</td>
<td>15 F Rx-free ≥ 6 mths</td>
<td>Good Only (TYR+PHE)/NAA ratio provided</td>
<td>HDRS-6**: APTD: BL: 1.7 (±0.7 SE), post-test: 2.4 (±0.9). CONT: BL: 1.0 (±0.4), post-test: 2.7 (±1.1).VAS-depressed not significantly changed by APTD vs CONT; no exact scores provided.</td>
<td>Study investigates effects of APTD on subjective and a spatial recognition memory-task</td>
<td>Good</td>
</tr>
<tr>
<td>Roiser (2005)*65</td>
<td>Rand., DB, Within SS ≥ 6 months re mission</td>
<td>17 M+F Rx-free ≥ 6 mths</td>
<td>Insufficient TYR depletion; (TYR+PHE)/NAA ratio decreases 80%</td>
<td>HDRS-19: APTD: BL: 2.1 (±2.0), post-test: 2.2 (±2.0). CONT: BL: 1.4 (±1.3), post-test: 1.1 (±1.1).VAS-contentedness: APTD: BL: 12.4 (±1.5), post-test: 12.8 (±1.5). CONT: BL: 12.3 (±1.8), post-test: 12.8 (±1.8).</td>
<td>BD-II patients included (n=?). Study investigates effects of APTD on subjective and neurocognitive measures</td>
<td>Good</td>
</tr>
</tbody>
</table>

**B2. Patients with MDD previously, but currently in remission (currently using medication)**

-  

**C. Patients currently with MDD (with or without antidepressants)**

-  

APTD = Acute Phenylalanine/Tyrosine Depletion; NAA = neutral amino acids; PHE = phenylalanine; TYR = tyrosine† tyrosine levels declined by at least 60%for other symbols and abbreviations see footnote in Table 6.2.
<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adequacy of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A1. Healthy controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krahn (1999)88</td>
<td>Rand., DB, Within SS, FH-</td>
<td>10 M+F</td>
<td>Not reported</td>
<td>HDRS scores increased significant (change: 3 SD not reported) in AMPT compared to CONT (no more data provided) POMS no significant changes (no more data provided)</td>
<td>Study investigates effects of AMPT on melatonin secretion. Promethazine 25mg as CONT</td>
<td>Poor</td>
</tr>
<tr>
<td>McCann (1995)82</td>
<td>Rand., DB, Betw. SS</td>
<td>41 M</td>
<td>Not determined; prolactin rise of 188% in AMPT group</td>
<td>POMS- sadness, VAS-sadness; no exact scores provided; figure for calmness shown. AMPT caused maximal ratings of tension, which was reversed by addition of L-dopa/ carbi dopa. Non-significant increase in sadness reported in AMPT-group (regardless L-dopa/ carbi dopa).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zimmermann (1996)89</td>
<td>Rand., DB, Within SS, FH-</td>
<td>10 M+F</td>
<td>Sufficient MHPG decrease in M+F</td>
<td>HDRS, POMS: No significant differences. No exact scores reported</td>
<td>Study investigates effects of AMPT on prolactin and melatonin secretion, including gender-effects. Promethazine 50 mg as CONT</td>
<td>Moderate</td>
</tr>
<tr>
<td><strong>B1. Patients with MDD previously, but currently in remission (without medication)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 6.5. Included studies Alpha-methyl-para-tyrosine (AMPT). (Continued)

<table>
<thead>
<tr>
<th>Author (year)*</th>
<th>Design</th>
<th>N</th>
<th>Adequacy of depletion†</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B2. Patients with MDD previously, but currently in remission (currently using medication)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delgado (1993)(^{17}) Rand., DB, Within SS 'Various' length of remission</td>
<td>14 M+F DES MAZ FLX SER</td>
<td>Good**</td>
<td>HDRS-25; No exact scores reported. Significant effects of AMPT vs CONT in DES and MAZ treated patients, not in FLX or SER treated patients.</td>
<td></td>
<td>Moderate</td>
<td></td>
</tr>
</tbody>
</table>

** HDRS-25; No exact scores reported. Significant effects of AMPT vs CONT in DES and MAZ treated patients, not in FLX or SER treated patients.

### C. Patients currently with MDD (with or without antidepressants)

<table>
<thead>
<tr>
<th>Author (year)(^{115})</th>
<th>Design</th>
<th>N</th>
<th>Adequacy of depletion</th>
<th>Results mood scores (scale, scores (±SD‡))</th>
<th>Remarks</th>
<th>Judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miller (1996)(^{115})  Rand., DB, Within SS</td>
<td>17 M+F</td>
<td>Good</td>
<td>HDRS-25**; AMPT: BL: 29, post-test: 28.5. CONT: BL: 30.5, post-test: 29.5. SDs not reported. VAS: exact data not provided; sign decrease in subscales ‘tired’ and ‘energetic’ for AMPT vs CONT</td>
<td>Patients were not using antidepressant medication. 2/17 patients had bipolar depression, 6/17 had co-morbid anxiety, dysthymic or eating disorders</td>
<td>Moderate</td>
<td></td>
</tr>
</tbody>
</table>

AMPT = Alpha-methyl-para-tyrosine; IDS = Inventory of Depressive Symptoms; MAZ = mazindol; PLAC = additional placebo, SleepDepr. = Sleep deprivation† Decline in catecholamine metabolites homovanillic acid (HVA) by at least 60% or 3-methoxy-4-hydroxyphenylglycol (MHPG) by at least 40% For other symbols and abbreviations see footnote in Table 6.2.
Quantitative summary (pooling)

Fifty-eight percent (52/90) of the studies supplied enough data to be included in the meta-analysis. Meta-analysis was possible for ATD and APTD studies only. Table 6.6 gives an overview of the number of identified studies for each pre-defined population and eligibility for meta-analysis. Due to remaining heterogeneity we used random effects models for all meta-analyses.

ATD

In figure 6.1 the pooled results of ATD in healthy controls in studies with a within subjects design are presented. Overall Hedges' g (95% CI) was -0.27 (-0.45 to -0.09). We stratified results by family history for MDD (negative, positive or not reported in the studies). Pooled Hedges' g for healthy controls with a negative family history (-0.19 (-0.43 to 0.05)) was significantly different ($Q_{res}=6.59$, df= 1, $p=0.01$) compared to controls with a positive family history (-0.56 (-1.00 to -0.13)). The pooled result in studies that did not report the family history status resembled the studies with a negative family history (-0.28 (-0.57 to 0.00) $Q_{res}=1.36$, df= 1, $p=0.24$). In between subjects studies similar results were found. Pooled Hedges' g was -0.63 (-1.95 to 0.70) for controls with a negative family history and -0.06 (-0.57 to 0.45) in one study that did not report family history status ($Q_{res}=0.14$, df= 1, $p=0.71$). The large Hedges' g in controls with a negative family history was largely determined by one study. Leaving this study out reduced Hedges' g to 0.16 (-0.43 to 0.76) (data not shown).

### Table 6.6. Identified studies for different types of depletion and design and eligibility for meta-analysis.

<table>
<thead>
<tr>
<th>Depletion Type</th>
<th>Identified (in meta-analysis)</th>
<th>Within Subjects</th>
<th>Between Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Healthy controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FH-</td>
<td>41 (22)</td>
<td>10 (5)</td>
<td></td>
</tr>
<tr>
<td>FH+</td>
<td>5 (5)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>B. Patients with MDD in remission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without AD</td>
<td>8 (6)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>With AD</td>
<td>13 (12)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>C. Patients with Current MDD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without AD</td>
<td>3 (1)</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>With AD</td>
<td>1 (1)</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Para-chlorophenylalanine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Patients with MDD in remission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With AD</td>
<td>2</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Acute Phenylalanine/Tyrosine Depletion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Healthy controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FH-</td>
<td>5 (3)</td>
<td>2 (2)</td>
<td></td>
</tr>
<tr>
<td>FH+</td>
<td>1 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Patients with MDD in remission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without AD</td>
<td>2 (2)</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>C. Patients with Current MDD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without AD</td>
<td>1</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

AD = antidepressant; FH = family history; MDD = Major Depressive Disorder
Acute Tryptophan Depletion (ATD) in healthy controls studied in a with subjects design, stratified by status of family history for depression (negative or positive for Major Depressive Disorder). References to studies as indicated, different subgroups per study handled as separate studies with appropriate pooling weights. FH+ = family history negative, FH++ = family history positive, N/A = not reported.

Figures 6.2A and 6.2B show the modification of Hedges’ g by gender for healthy controls (within subjects design) with a negative or positive family history. The difference in Hedges’ g between males and females was most prominent in healthy controls with a negative family history (0.23 (-0.10 to 0.57) vs. -0.44 (-0.81 to -0.06) respectively; Q_{res} = 11.92, df = 1, p < 0.001). In controls with a positive family history, males experienced a larger decrease in mood after ATD in only one study (Hedges’ g -0.98 (-1.53 to -0.42) Q_{res} = 11.92, df = 1, p < 0.001). In contrast, females with a positive family history only had a slightly larger Hedges’ g (-0.56 (-1.43 to 0.31)) than females with a negative family history (Q_{res} = 0.62, df = 1, p = 0.43).

Figures 6.3A and 6.3B present the pooled Hedges’ g for patients with MDD in remission without or with current AD (within subjects design). In the remitted patients without current AD we stratified results by length of time without AD. Only 2 studies with 3-6 months without AD\(^{47}\) largely differed in Hedges’ g (-3.85 (-0.39 to -1.31) compared to 1 study directly after successful electroconvulsive therapy (Hedges’ g 0.04 (-0.85 to 0.94);\(^{47}\) and 3 studies with at least 6 months without AD (pooled Hedges’ g -0.60 (-1.38 to 0.18) (Q_{res} = 29.34, df = 2, p < 0.0001),\(^{48}\)\(^{33}\)\(^{18}\). Leaving one possible outlier\(^{47}\) out diminished the overall pooled Hedges’ g for remitted patients without AD from -1.90 (-3.02 to -0.78) to -1.66 (-1.83 to -0.29), but the observed effect modification by length of time without AD remained highly significant (p < 0.001; data not shown).
Acute Tryptophan Depletion (ATD) in healthy controls studied in a within subjects design, stratified by status of family history for depression and gender included in the studies. Family history negative for MDD (A) and family history positive for MDD (B). References to studies as indicated, different subgroups per study handled as separate studies with appropriate pooling weights. FH- = family history negative, FH+ = family history positive, MDD = major depressive disorder.
Figure 6.3A/B.

**A**

<table>
<thead>
<tr>
<th>Study or sub-category</th>
<th>Hedges' g (random)</th>
<th>Weight %</th>
<th>Hedges' g (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 3months</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cassidy 1997</td>
<td>14.37</td>
<td>0.04</td>
<td>[0.85, 0.94]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>14.37</td>
<td>0.04</td>
<td>[0.85, 0.94]</td>
</tr>
<tr>
<td>Test for heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $Z = 0.10$ ($P = 0.92$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02 3-6 months</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neumester 2004</td>
<td>7.79</td>
<td>-9.40</td>
<td>[-12.66, -6.95]</td>
</tr>
<tr>
<td>Neumester 2006 f/s</td>
<td>6.07</td>
<td>-3.87</td>
<td>[-6.62, -1.12]</td>
</tr>
<tr>
<td>Neumester 2006 s/f</td>
<td>11.81</td>
<td>-3.15</td>
<td>[-4.79, -1.51]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>13.02</td>
<td>-1.35</td>
<td>[-2.66, -0.04]</td>
</tr>
<tr>
<td>Test for heterogeneity: $\chi^2 = 28.36$, df = 3 ($P = 0.0001$), $I^2 = 89.4%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $Z = 2.81$ ($P = 0.005$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03 &gt;=6 months</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haynes 2004</td>
<td>15.26</td>
<td>0.00</td>
<td>[-0.55, 0.55]</td>
</tr>
<tr>
<td>O’Reardon 2004</td>
<td>14.95</td>
<td>-0.51</td>
<td>[-1.19, 0.17]</td>
</tr>
<tr>
<td>Smith 1997a</td>
<td>14.73</td>
<td>-1.39</td>
<td>[-2.16, -0.62]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>44.94</td>
<td>-0.60</td>
<td>[-1.38, 0.18]</td>
</tr>
<tr>
<td>Test for heterogeneity: $\chi^2 = 8.35$, df = 2 ($P = 0.02$), $I^2 = 76.0%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $Z = 1.50$ ($P = 0.13$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>100.00</td>
<td>-1.90</td>
<td>[-3.02, -0.78]</td>
</tr>
<tr>
<td>Test for heterogeneity: $\chi^2 = 66.05$, df = 7 ($P = 0.0001$), $I^2 = 89.4%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $Z = 3.33$ ($P = 0.0009$)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B**

<table>
<thead>
<tr>
<th>Study or sub-category</th>
<th>Hedges' g (random)</th>
<th>Weight %</th>
<th>Hedges' g (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 With current SSRI medication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bremner 1997 No Rel.</td>
<td>15.41</td>
<td>0.27</td>
<td>[-0.26, 0.81]</td>
</tr>
<tr>
<td>Bremner 1997 Relapse</td>
<td>10.44</td>
<td>-1.10</td>
<td>[-2.05, -0.16]</td>
</tr>
<tr>
<td>Moore 1998</td>
<td>14.73</td>
<td>-1.39</td>
<td>[-2.16, -0.62]</td>
</tr>
<tr>
<td>O’Reardon 2004</td>
<td>10.08</td>
<td>-1.33</td>
<td>[-2.31, -0.36]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>60.86</td>
<td>-0.60</td>
<td>[-1.28, 0.08]</td>
</tr>
<tr>
<td>Test for heterogeneity: $\chi^2 = 14.09$, df = 4 ($P = 0.007$), $I^2 = 71.6%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $Z = 1.72$ ($P = 0.08$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02 With current SSRNRI medication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Booj 2005a</td>
<td>18.74</td>
<td>-0.52</td>
<td>[-1.00, -0.05]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>18.74</td>
<td>-0.52</td>
<td>[-1.00, -0.05]</td>
</tr>
<tr>
<td>Test for heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $Z = 2.16$ ($P = 0.03$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03 With current BUP medication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evans 2002</td>
<td>14.01</td>
<td>-0.25</td>
<td>[-0.96, 0.47]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>14.01</td>
<td>-0.25</td>
<td>[-0.96, 0.47]</td>
</tr>
<tr>
<td>Test for heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $Z = 0.68$ ($P = 0.50$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04 With current PHZ medication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landolt 2003</td>
<td>6.39</td>
<td>-0.76</td>
<td>[-2.12, 0.60]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>6.39</td>
<td>-0.76</td>
<td>[-2.12, 0.60]</td>
</tr>
<tr>
<td>Test for heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $Z = 1.10$ ($P = 0.27$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>100.00</td>
<td>-0.49</td>
<td>[-0.89, -0.10]</td>
</tr>
<tr>
<td>Test for heterogeneity: $\chi^2 = 15.01$, df = 7 ($P = 0.04$), $I^2 = 53.3%$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: $Z = 2.43$ ($P = 0.01$)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Acute Tryptophan Depletion (ATD) in former depressed patients in remission, studied in a within subjects design without or with current medication. Figure 6.3A is stratified by length of time without medication; 6.3B is stratified by type of medication used by patients. References to studies as indicated, different subgroups per study handled as separate studies with appropriate pooling weights. BUP = bupropion; SNRI = Serotonin Norepinephrine Reuptake Inhibitor; SSRI = Selective Serotonin Reuptake Inhibitor; PHZ = phenelzine.
In remitted patients previously depressed and currently using AD, ATD caused a decrease in mood (pooled Hedges’ g -0.49 (-0.89 to -0.10). Hedges’ g varied slightly depending on type of AD ($Q_{res}$ = 0.92, df = 3, p = 0.82). Surprisingly, the pooled Hedges’ g for SSRIs showed a moderate point estimate, which did not reach significance (-0.60 (-1.28 to 0.08). Especially for bupropion treatment Hedges’ g was small and not significant (-0.25 (-0.96 to 0.47). No ATD studies with other ADs without a 5-HT mechanism of action were available for this comparison.

We stratified the results of ATD studies by length of remission, which revealed significant effect modification. In remitted patients using AD decreased mood after ATD was especially seen in the first 5 months after the achievement of remission (pooled Hedges’ g -0.55 (-0.90 to -0.21) ($Q_{res}$ = 47.18, df = 2, p < 0.0001). In contrast, remitted patients without AD showed more decrease in mood after ≥2 months of remission (Hedges’ g -1.65 (-2.60 to -0.69) ($Q_{res}$ = 13.81, df = 2, p = 0.001)) (figure available on request).

Relapse rates in remitted patients with AD were increased after ATD compared to control depletion (pooled difference in relapse rate 47% (28% to 66%); Figure 6.4). This increase in relapse rate was especially seen in patients using SSRIs (47% (27% to 67%)) or a SNRI (35% (14% to 56%)). The NE acting drug desipramine showed no significant difference in relapse rate (7% (6% to 9%)) after ATD. This effect modification by drug was statistically significant ($Q_{res}$ = 18.02, df = 2, p < 0.001).

In patients who were depressed at the time of ATD we found two studies for meta-analysis. These studies included patients who used,\textsuperscript{120} or did not use\textsuperscript{86} AD (Figure 6.5). The effects of ATD were opposed: Hedges’ g was 0.32 (-0.22 to 0.86) for patients using different types of ADs, and -0.12 (-0.45 to 0.21) for patients without AD. Two studies in depressed patients without AD were not suitable for meta-analysis. ATD did not decrease mood during depletion in these studies.\textsuperscript{86,121} Contra-intuitively, in one study mood improved the day after depletion in 16/43 patients,\textsuperscript{121} a result which was also found by one study in the meta-analysis.\textsuperscript{120}

**Figure 6.4.**

<table>
<thead>
<tr>
<th>Study or sub-category</th>
<th>Diff. relapse-rate (random)</th>
<th>Weight %</th>
<th>Diff. relapse-rate (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 SSRI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bremer 1997 (all)</td>
<td>11.49</td>
<td>0.29</td>
<td>[0.05, 0.52]</td>
</tr>
<tr>
<td>Delgado 1999</td>
<td>11.18</td>
<td>0.53</td>
<td>[0.28, 0.79]</td>
</tr>
<tr>
<td>O’Reardon 2004</td>
<td>10.21</td>
<td>0.33</td>
<td>[0.03, 0.64]</td>
</tr>
<tr>
<td>Praschak-Reeder 2004</td>
<td>10.36</td>
<td>0.75</td>
<td>[0.45, 1.05]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>10.36</td>
<td>0.75</td>
<td>[0.45, 1.05]</td>
</tr>
<tr>
<td>Test for heterogeneity: $I^2$ = 66.9, df = 3 (P = 0.08)</td>
<td>0.47</td>
<td>[0.27, 0.67]</td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 4.53 (P &lt; 0.00001)</td>
<td>0.47</td>
<td>[0.27, 0.67]</td>
<td></td>
</tr>
<tr>
<td>02 SSRI+SNRI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boog 2005a</td>
<td>11.90</td>
<td>0.35</td>
<td>[0.14, 0.56]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>11.90</td>
<td>0.35</td>
<td>[0.14, 0.56]</td>
</tr>
<tr>
<td>Test for heterogeneity: not applicable</td>
<td>0.35</td>
<td>[0.14, 0.56]</td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 3.28 (P = 0.001)</td>
<td>0.35</td>
<td>[0.14, 0.56]</td>
<td></td>
</tr>
<tr>
<td>03 DESIPRAMINE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delgado 1999</td>
<td>13.08</td>
<td>0.07</td>
<td>[-0.06, 0.19]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>13.08</td>
<td>0.07</td>
<td>[-0.06, 0.19]</td>
</tr>
<tr>
<td>Test for heterogeneity: not applicable</td>
<td>0.07</td>
<td>[-0.06, 0.19]</td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 1.04 (P = 0.30)</td>
<td>0.07</td>
<td>[-0.06, 0.19]</td>
<td></td>
</tr>
<tr>
<td>04 Various antidepressants mixed</td>
<td>12.85</td>
<td>0.52</td>
<td>[0.38, 0.67]</td>
</tr>
<tr>
<td>Delgado 1991</td>
<td>11.57</td>
<td>0.88</td>
<td>[0.65, 1.10]</td>
</tr>
<tr>
<td>Morris 1999</td>
<td>7.38</td>
<td>0.63</td>
<td>[0.14, 1.11]</td>
</tr>
<tr>
<td>Smith 1999</td>
<td>31.80</td>
<td>0.67</td>
<td>[0.41, 0.93]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>31.80</td>
<td>0.67</td>
<td>[0.41, 0.93]</td>
</tr>
<tr>
<td>Test for heterogeneity: $I^2$ = 64.5, df = 2 (P = 0.04)</td>
<td>0.67</td>
<td>[0.41, 0.93]</td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 5.01 (P &lt; 0.00001)</td>
<td>0.67</td>
<td>[0.41, 0.93]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for heterogeneity: $I^2$ = 54.31, df = 8 (P &lt; 0.00001)</td>
<td>0.47</td>
<td>[0.28, 0.66]</td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 4.80 (P &lt; 0.00001)</td>
<td>0.47</td>
<td>[0.28, 0.66]</td>
<td></td>
</tr>
</tbody>
</table>

Acute Tryptophan Depletion (ATD) in former depressed patients in remission, differences of relapse rates versus control depletion studied in a within subjects design in remitted patients stratified by current medication use. References to studies as indicated. SNRI = Serotonin Norepinephrine Reuptake Inhibitor; SSRI = Selective Serotonin Reuptake Inhibitor.
APTD

Figure 6.6A and 6.6B show results from APTD studies in healthy controls (within and between subjects). APTD did not decrease mood: pooled Hedges’ g were 0.10 (-0.23 to 0.43) in within subjects, and 0.12 (-0.43 to 0.68) in between subjects studies. However, in one study with healthy controls with a positive family history for MDD, a moderate but non-significant effect on mood was found (Hedges’ g -0.49 (-1.17 to 0.19)).63 In patients with MDD in remission without AD (Figure 6.7), no effect of APTD was observed in 2 studies (pooled Hedges’ g -0.02 (-0.50 to 0.47)).64,65 No studies inpatients with current MDD were identified.

Figure 6.5.

Acute Tryptophan Depletion (ATD) in depressed patients stratified by use of concurrent medication (within subjects design). References to studies as indicated.

Sensitivity analysis and publication bias

We examined our assumptions for intercorrelation of before-after mood ratings per condition in ATD studies in healthy controls without a family history of MDD. At the same examination, we also examined the assumed intercorrelation between two test conditions (ATD vs. CONT). As expected, increasing the assumed R to 0.8 (less conservative) increased the pooled Hedges’ g from -0.19 (-0.43 to 0.05) to -0.31 (-0.64 to 0.02). Reducing R to 0.2 (more conservative) decreased the pooled Hedges’ g to -0.14 (-0.34 to 0.06). The calculated R’s were higher than 0.5 in 80% of the studies that reported all relevant data. Therefore, we judged the imputed value of 0.5 for R as acceptable.

Inspection of the funnel-plot of the within subjects ATD studies in healthy controls without a family history of MDD revealed asymmetry (figure available on request). More studies with a decrease in mood after ATD vs. control depletion were published. In a Galbraith-plot of studies, the intercept in the regression equation was -2.08 (SE 0.90; p= 0.050). Therefore, we concluded that publication bias probably distorted our findings. We did not inspect funnel-plots in other populations of our review due to the limited number of studies.

Discussion

In this systematic review we performed the first meta-analysis of the mood-effects in ATD and APTD studies. The depletion of monoamine systems (both 5-HT and NE/DA) does not decrease mood in healthy controls. However, in healthy controls with a family history of MDD the results suggest that mood is slightly decreased, both by ATD and APTD. Additionally, healthy females are more affected by ATD than healthy males, especially in controls without a family history of MDD. In patients who were previously depressed but in remission without AD, ATD moderately decreases mood, while APTD does not significantly decrease mood. ATD has comparable mood lowering effects in patients with MDD in remission who are still using ADs. The site of action of
these ADs (5-HT or NE/DA) predicts the occurrence of a lowering of mood or a short relapse in MDD after depletion of the corresponding monoamine. Our findings are in line with the summaries in previous reviews.1728-2024

The most consistent finding from this review is the decrease of mood and relapse into a depressed state after ATD and APTD in remitted MDD patients who use AD. In remitted patients without AD relapses are less prominent. Because after remission medication is often continued, the difference in mood responses after depletion might be related to the duration of the achieved remission. Previous reviews discussed the relationship between the duration of the remission and the effect of monoamine depletion, with opposite conclusions. Bell et al. concluded that effects of ATD were more pronounced early in recovery.23 In contrast, Booij et al. concluded that duration of remission was not associated with mood response to ATD.21

Figure 6.6A/B.

Acute phenylalanine/tyrosine depletion (APTD) in healthy controls (6.6A. within subjects design; 6.6B. between subjects design), stratified by status of family history for depression. References to studies as indicated. FH+ = family history negative, FH+= family history positive, N/A = not reported.

Booij et al. investigated predictors of relapse in a pooled analysis of the individual patient data of some of the studies included in this review.21 This is often referred to as a ‘mega-analysis’. They found that recurrent depressive episodes (≥2), female gender, and a history of suicidal thoughts/ attempts predicted relapse. Duration of remission did not contribute to this prediction when confounding was considered. We found a modest relationship between relapse and the duration of the remission after ATD. We defined the duration of remission as the reported average duration, or – if not stated – the minimum duration of remission used as inclusion criterion for the studies. We found that especially in the first 5 months after remission, ATD caused lowering of
mood in remitted patients still using AD. However, the problem with our and Bell’s comparison is the intraindividual spread in duration within the studies. This spread is not considered at the same level of detail as in a ‘mega-analysis’. In addition, confounding can only be considered in ‘mega-analysis’. Therefore – although their study did not include all available studies – we think that Booij et al. provide the best available indication of riskfactors for mood lowering effects by ATD.

**Figure 6.7.**

Acute phenylalanine/tyrosine depletion (APTD) in former depressed patients in remission without medication, studied in a within subjects design. References to studies as indicated.

**Do depletion studies elucidate the pathogenesis of MDD?**

The absence of robust mood effects in healthy controls indicates that mood is not a direct correlate of 5-HT or NE levels in the brain. The only healthy controls who are modestly affected by monoamine depletion studies are healthy controls with a positive family history for MDD. This might be indicative of a biological vulnerability which is revealed by depletion studies. Of interest are the findings of recent studies that combined ATD with neuroimaging or genetic sampling, reviewed by Fusar-Poli et al.¹²¹ The intelligent approach of combining depletion with imaging or genotyping appears very promising. A summary of the complex results of these studies goes beyond the scope of our review. However, these neuroimaging and genotyping studies also suggest that monoamine depletion discloses rather a ‘trait’ vulnerability than a pure ‘state’ dependent change due to depletion.

Additionally, a depressive relapse after monoamine depletion in remitted patients who use AD, occurs only if the target of the depletion (5-HT, NE) coincides with the working mechanism of the antidepressant used. This emphasizes that AD indeed specifically affect their supposed target systems. However, we may only conclude that an undepleted 5-HT system is required for serotonergic AD. The same holds for the NE system and norepinephrinergic AD. Delgado proposed an alternative explanation for the decrease in mood after monoamine depletion in patients: depletion of e.g. 5-HT may give the same effect as abrupt discontinuation of SSRIs.¹²⁴ Rapid discontinuation is also associated with mood-effects, which are considered to be different from a depressive relapse.¹²⁵

What certainly cannot be concluded is that MDD is caused by low levels of 5-HT and/or NE/DA. This simplification, which is often used to promote the use of AD specifically affecting 5-HT or NE or both systems, represents a Catch-22 argument, and ignores the notion that serotonergic and norepinephrinergic AD presumably act by a final common pathway. In this pathway post-synaptic changes at the cellular level are supposed to be responsible for the remission of MDD.¹²⁶ Cellular changes include up or down regulated receptors, increased neuronal interconnections and sprouting and changes in levels of neuropeptides (e.g. corticotrophin releasing hormone). The clinical question of how to distinguish a patient who will respond to an AD with e.g. NE effects has not been solved. Descriptive variables at the symptom level have not yet sufficiently predicted the response to any selective agent. Therefore, a pragmatic approach to affect this final common pathway might be to prescribe ADs which target both 5-HT and NE. However, the effectiveness of this approach is still equivocal.¹²⁷-¹²⁹

In patients who are depressed at the time of monoamine depletion, no further decrease in mood is observed. A ceiling effect could be responsible for this finding. But, a more straightforward conclusion is that there is no simple relation between 5-HT or NE deficiency and mood or MDD. Nevertheless, this finding is complicated by the finding of some authors¹³⁰-¹³¹ that mood is lowered or elevated the day after ATD. A delayed decrease of mood was indicative for treatment refractoriness,¹³² a finding that was not yet replicated by others.¹³³-¹³⁵ The number of
comparable studies to date is limited. Therefore, we think no clear conclusions or explanations can be made, except that even in MDD patients mood is not a correlate of 5-HT or NE levels in the brain.

We agree with the conclusion of Booij et al. and Bell et al. that a relapse of depressive symptoms in remitted patients after depletion probably reflects a biological vulnerability of the 5-HT system in remitted patients.\textsuperscript{112,123} This vulnerability increases their risk to become depressed. This increased risk was further demonstrated in two prospective studies that used the response to ATD in remitted patients to predict later relapse/recurrence.\textsuperscript{130,131} However, also these results need replication. Moreover, the cause of an increased vulnerability remains uncertain. Probably the cause is a combination of genetic, environmental and other determinants (e.g. ‘scarring’ the brain after multiple depressive episodes).\textsuperscript{123}

In conclusion, monoamine depletion by ATD and APTD does not elucidate a causal factor in the pathogenesis of MDD. However, ATD and APTD remain useful models to safely and directly manipulate 5-HT, NE and DA function in living humans, and to study the behavioral consequences of this manipulation, especially in subgroups of humans with an apparent vulnerability.\textsuperscript{11,12,24}

Still, several methodological issues need to be addressed. First, because of the competition of aminoacids to pass the blood-brain barrier, ATD might unwillingly result in an intracerebral increase of tyrosine/phenylalanine.\textsuperscript{24,132} Vice-versa, APTD might increase intracerebral tryptophan availability. Levels of other amino-acids than those depleted are not provided in the studies. Second, ATD provides a specific net lowering of 5-HT. Contrary, depletion of tyrosine and phenylalanine lowers both NE and DA, which are synthesized in the same cascade (with DA thereafter transformed into NE by dopamine beta-hydroxylase). Also AMPT interferes early in this cascade. Although evidence from animal studies points to more DA depletion by APTD and more NE depletion by AMPT,\textsuperscript{13} it is impossible to truly distinguish between net NE and DA depletion. Third, test re-test reliability for monoamine depletion paradigms was only tested for ATD and was rather limited, which limits the robustness of the method.\textsuperscript{134,135} Fourth, also in healthy controls subtle cognitive effects of monoamine depletion in the brain occur: deficits in learning and memory consolidation and improvement in focused attention and executive function.\textsuperscript{136} These effects show similarity with symptoms of MDD. The question remains whether these effects might represent mild first symptoms of MDD or the starting symptoms of a cascade of altered brain functions leading to MDD. Fifth, in line with the fourth issue, MDD does not develop within one day. Therefore, the changes by experimental monoamine depletion by ATD and APTD may be too short to really induce a complex biological and psychological deregulation which is recognized as MDD. Patients suffering from gastrointestinal carcinoid tumors – 5-HT producing tumors with expected prolonged states of secondary tryptophan depletion – are generally not depressed but do show improved focused attention.\textsuperscript{137} However, carcinoid findings were not yet related to tryptophan levels over time. Sixth, the single depletion of one monoamine system by ATD or APTD/AMPT may be too simplistic, especially because of the complex interaction of monoamine systems. Five dual depletion studies were not included in this review.\textsuperscript{38,138-141} In healthy controls contrasting results were found. Hughes et al. found some decrease of mood on 3 VAMS-subscales,\textsuperscript{139} which was also found in another open study.\textsuperscript{140} However, no effects were found in 2 other studies.\textsuperscript{18,141} In unmedicated patients with MDD no significant increase of MDD was found after dual depletion.\textsuperscript{138} It would be interesting to investigate the effects of simultaneous depletion of 5-HT and NE/DA in other populations.

Limitations of the studies and the review

Some limitations should be mentioned. First, female gender is a risk-factor for MDD, and was also found to be a predictor of relapse after ATD in remitted patients. Based on several studies, gender differences in 5-HT metabolism are probable, and hormonal factors may play a role in 5-HT function.\textsuperscript{21} In our results we examined the effect of gender in healthy controls. The difference between males and females was most prominent in healthy controls without a family history of MDD. In the included studies hormonal status (pre or postmenopausal state) was not distinguished in the results nor used as inclusion criterion. Second, many small differences...
between the studies existed: different composition of depletion and control drinks, different presentation of tryptophan/tyrosine/phenylalanine or their ratios to other neutral amino acids, different presentation of free vs. total tryptophan/tyrosine/phenylalanine values, different measurements, different scales. Differences between studies undoubtedly introduced heterogeneity between the studies, which may bias the results of this review. Therefore, we recommend an international consensus protocol. Third, limited presentation of the data forced us to make assumptions, that may have influenced our findings. However, the assumptions appeared to have little influence on the results in a sensitivity analysis. Future studies need to address clear presentation of their data, preferably including a description of the directions of the effects, SDs, in cross-over designs the numbers of subjects having the control or sham first, and preferably include a SD of the pooled difference between the depletion and the control condition. For example, only 10 studies adequately reported the number of patients treated in each sequence in the within subjects design. Fourth, 6 studies included Bipolar Patients. Three of these studies were also included in the meta-analysis. This involved 6 patients of the total of 265 patients (2.2%) in the concerning meta analytical comparisons. Therefore, we consider the possibility of bias by the inclusion of this etiologically different disorder unlikely. Fifth, the rate of agreement in the selection of studies still may rise questions about the clarity of our selection criteria. However, discrepancies in agreement could easily be solved between the two reviewers. Therefore, we do not think our sensitive searches missed relevant studies. Finally, an indication of publication bias was found. If publication bias truly exists, the studies which found no effect or an increase in mood after ATD would not have been published. Indeed, the mood effects of studies that could not be included in the meta-analysis were mostly very small and non-significant. Furthermore, the exact information in studies required for meta-analysis forced us to exclude many studies. It seems natural that a non-significant result will not be given this level of detailed attention (change-scores with SDs), especially when mood-effects are not the primary outcome in these studies. Without this publication bias the pooled effect of ATD in healthy controls would probably have been lower. Because our conclusion already is that there is no apparent mood effect of ATD in healthy controls, we conclude that the observed publication bias will not severely distort the general conclusion. Therefore, despite these limitations, we consider the results of our review after our strict methods as valid.

Conclusion and future studies

We conclude that although ATD and APTD are important in the investigation of the monoamine systems, monoamine depletion does not directly decrease mood. Although previously the monoamine systems were considered to be responsible for the development of MDD, the available evidence to date does not support a direct causal relationship with MDD. There is no simple direct correlation of 5-HT or NE levels in the brain and mood. The depletion of 5-HT by ATD and NE/DA by APTD most clearly decreases mood in vulnerable patients who are in remission from their MDD, while still using AD. Furthermore, depletion affects mood in unmedicated patients in remission or healthy controls with a family history positive for MDD. Therefore, the monoamine systems are probably important systems in the vulnerability to become depressed. The changes in brain metabolism in remitted patients who relapse after ATD or AMPT suggest that 5-HT and NE systems give input to a final common pathway which needs further research to be clarified.

We suggest some lines for future research. First: three or four-armed depletion studies comparing ATD, APTD (and – if possible – their combination) versus sham depletion to investigate the differential effects of the 5-HT and NE systems on mood in remitted patients or controls with positive family history for MDD. Second: a further exploration of the relation between known genetic polymorphisms of the 5-HT, NE and DA systems (e.g. 5-HTPR) and biological cerebral responses to depletion paradigms, as measured by PET/fMRI (e.g. like Neumeister et al.146). Third: the relations between monoamine depletion indicating biological vulnerability and psychological vulnerability for MDD (see Booij et al.11) Fourth: Replication, further validation and standardization of the properties of ATD and other depletion paradigms as a diagnostic test for
recurrences in remitted patients (see Moreno et al.\textsuperscript{130} and Neumeister et al.\textsuperscript{131}). New studies will increase our knowledge of the 5-HT and NE systems, which are important targets in the current treatments of MDD. This knowledge will finally improve the treatment for MDD.

**Acknowledgement**

The authors wish to thank Mrs Natasha Wiebe, MMath PStat, Research Associate at the University of Alberta, Canada and especially Dr. Rob J.P.M. Scholten, M.D. PhD, epidemiologist and director of the Dutch Cochrane Centre at the Academic Medical Centre, Amsterdam for their indispensable help with the statistic methods used in this review. This study was partially funded by the program Opleiding Onderzoekers GGZ (OOG) (project number 100-002-002) by ZonMw, the Hague, the Netherlands.

**Conflict of interests**

None

**References**


105. Talbott PS, Cooper SJ. Anterior cingulate and subgenual prefrontal blood flow changes following tryptophan depletion in healthy males. Neuropsychopharmacology. 2006; 31: 09.


Differences between intervention and control measurements

In depletion studies changes in mood scores typically represented mean mood-scores both before (pre) and after (post) the depletion/challenge (experimental intervention) and the placebo/sham/control-intervention. Because the mood scores were not necessarily identical at the start of the experiment and the control, we first calculated the mean change in mood-score (pooled difference) for the experimental and control condition separately per study. Some studies also provided the standard deviation (SD) of the pooled differences. When the standard deviation was not reported, we calculated the standard deviation of the pooled difference for paired data:

$$SD_{\text{change}} = \sqrt{SD_{\text{pre}}^2 + SD_{\text{post}}^2 - (2 \times R \times SD_{\text{pre}} \times SD_{\text{post}})}$$  \hfill (1)

In this formula, the correlation coefficient $R$ was calculable in four studies only and ranged between 0.42 and 1.00 for the experimental and between 0.34 and 0.95 for the control condition. To be able to calculate the standard deviation of the change between pre and post test mood scores for the rest of the studies we imputed a correlation coefficient $R$ of 0.5. This value was considered to be a conservative assumption.

Statistics for studies with a within subjects design

The difference in the changes of mood scores between intervention and control were expressed as difference of change scores:

$$\text{Difference}_\text{changes} = \text{Change}_\text{CONT} - \text{Change}_\text{INT}$$  \hfill (2)

For this difference the SD of the difference was calculated by again applying formula (1), with an assumed $R$ of 0.5.

To acknowledge the different mood scales to measure change in mood, the difference in changes between experimental and control condition were standardized by calculating Hedges' adjusted $g$, which is similar to Cohen's $d$, but includes an adjustment for small sample bias:

$$g = \frac{\text{Change}_\text{INT} - \text{Change}_\text{CONT}}{SD_{\text{Difference\ changes}}} \times (1 - \frac{3}{4(n_{\text{AB}} + n_{\text{BA}}) - 9})$$  \hfill (3)

In this formula $n_{\text{AB}}$ and $n_{\text{BA}}$ represent the number of subjects randomized to intervention or control as first test in the study. If the numbers for $n_{\text{AB}}$ and $n_{\text{BA}}$ were not reported, we assumed that the sample was split half for the two sequences. For Hedges' $g$ SE was calculated:

$$\text{SE} = \sqrt{\frac{n_{\text{AB}} + n_{\text{BA}}}{4n_{\text{AB}} \times n_{\text{BA}}} + \frac{Hedges'g^2}{2(n_{\text{AB}} + n_{\text{BA}} - 3.94)}}$$  \hfill (4)
Statistics for studies with a between subjects design

For between subject studies comparable statistics were used to calculate the mean change in mood-scores. Because the between-subjects design is a parallel group design, Hedges’ $g$ was calculated with formula (3) in which for $n_{AB}$ and $n_{BA}$, $n_{INT}$ and $n_{CONT}$ were substituted. The formula for the SE was slightly different to acknowledge the absence of paired data:

$$SE = \sqrt{\frac{n_{INT} + n_{CONT}}{n_{INT} \times n_{CONT}} + \frac{Hedges' \ g^2}{2(n_{INT} + n_{CONT} - 3.94)}}$$  \hspace{1cm} (5).$$

Statistics for Relapse rates

For relapse rates of MDD after depletion provided in a within subjects design the difference in relapse rates was calculated as:

$$\text{Difference}_{\text{Relapse-rate}} = \frac{n_{\text{Relapse}_\text{INT}}}{N} - \frac{n_{\text{Relapse}_\text{CONT}}}{N}$$  \hspace{1cm} (6),$$

in which N is the total number of patients included. The standard error then is:

$$SE_{\text{Difference}_{\text{Relapse-rate}}} = \frac{1}{N}\sqrt{b+c - \frac{(b-c)^2}{N}}$$  \hspace{1cm} (7),$$

in which $b$ represents the number of patients with a relapse after the intervention but not the control condition and $c$ the number of patients with a relapse after the control but not the intervention. If the numbers of ‘pairs’ were not extractable from the paper, a conservative approach was used assuming the minimal number of patients relapsing both after the intervention and the control condition (maximal $c$), resulting in the largest SE.
SEROTONIN TRANSPORTER BINDING WITH $^{[123]}$I-$\beta$-CIT SPECT IN MAJOR DEPRESSIVE DISORDER VERSUS CONTROLS: EFFECT OF SEASON AND GENDER.

Submitted

Henricus G. Ruhé
Jan Booij
Johannes B. Reitsma
Aart H. Schene

1 Program for Mood Disorders, Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
2 Department of Nuclear Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
3 Department of Clinical Epidemiology, Biostatistics & Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
Abstract

Background
The serotonin system is undoubtedly involved in the pathogenesis of major depressive disorder (MDD). More specifically, the serotonin transporter (SERT) serves as a major target for antidepressant drugs. There are conflicting results about SERT availability in depressed patients versus healthy controls (HC).

Aim
To measure SERT availability and study effects of age, gender and season of scanning in MDD patients versus HC.

Methods
We included 49 depressed outpatients (42.3 ± 8.3 [SD] years) with a Hamilton Depression Rating Scale above 18, who were drug-naïve or drug-free for ≥4 weeks, and 49 age- (±2 years) and sex-matched HC. Subjects were scanned with single photon emission computed tomography (SPECT) using $[^{123}]$β-CIT. SERT availability was expressed as specific to non-specific binding ratios (BP_{ND}) in midbrain (MID) and diencephalon (DIENC) with cerebellar binding as a reference.

Results
In crude comparisons between patients and HC, we found no significant differences in MID or DIENC SERT availability. In subgroup analyses, depressed males had numerically lower MID SERT availability than HC, whereas in women SERT availability was not different (significant diagnosis*gender interaction; $p=0.048$). In DIENC we found a comparable diagnosis*gender ($p=0.002$) and an additional smoking*gender ($p=0.036$) interaction. In MID the season of scanning showed a significant main-effect ($p=0.018$) with higher SERT availability in winter.

Conclusions
Differences in MID and DIENC SERT availability in MDD patients compared with HC are modified by gender. The season of scanning is a covariate in MID. The diagnosis*gender and gender*smoking interactions on SERT availability should be considered in future studies of the pathogenesis of MDD.
Introduction

Major depressive disorder (MDD) is a highly prevalent and disabling disease, often treated by selective serotonin reuptake inhibitors (SSRIs). SSRIs block the serotonin transporter (SERT), which lowers the reuptake of serotonin (5-HT) from the synaptic cleft and increases neurotransmission. Despite the fact that the working-mechanism of antidepressants supports the monoamine deficiency theory, the pathogenesis of MDD remains unclear. Therefore, differences in SERT availability in patients and healthy controls (HC) have been studied previously.

Postmortem studies have shown reduced or unchanged concentrations of SERTs in MDD-patients compared with HC, but these studies may be biased by retrospective data collection, previous antidepressant use, suicidal behavior apart from MDD or non-selective ligands (reviewed by Stockmeier). Cerebral SERTs in humans can be quantified in-vivo with single photon emission computed tomography (SPECT) and positron emission tomography (PET). The first SERT radioligand, iodine-123-labeled 2β-carbomethoxy-3β-(4-iodophenyl)-tropane ([123I]β-CIT) binds both to SERTs and dopamine transporters (DATs). Diencephalon (DIENC) and brainstem (MID) [123I]β-CIT binding predominantly reflect SERT, while striatal [123I]β-CIT uptake reflects DAT.

Studies comparing depressed patients with HC have reported either decreased, unchanged, or increased SERT availability in MDD patients (Table S7.1). A negative correlation between SERT availability and severity of depression (measured by Hamilton Depression Rating Scale (HDRS) scores) was reported in patients with primary MDD or Wilson’s disease. Discrepant results among studies may be explained by differences in scanning techniques, analytic methods, and subject sampling, although the effects of additional variables and their interaction might also explain conflicting results. Staley et al. reported lower SERT availability in the DIENC of MDD female patients versus HC, and suggested that this interaction accounted for the contradictory results between studies. Furthermore, in HC significant effects on SERT availability have been reported for gender, smoking behavior, aging and season of scanning.

Our objectives were to quantify SERT availability in MDD patients versus HC while accounting for these potential covariates and possible interactions, and to correlate SERT availability with depression severity. Therefore, we compared [123I]β-CIT SPECT scans of drug-free MDD patients with age- and sex-matched HC.

Methods

Subjects

After approval by the institutional ethical committee and written informed consent, we recruited depressed patients from primary care, and our outpatient department (October 2003-August 2006). Patients were eligible if they were 25-55 years old, had a diagnosis of MDD (diagnosed by structured clinical interview for DSM-IV (SCID Patient Version)), had a HDRS score >18, and were drug-free or used no more than one antidepressant (stopped for >4 weeks and ≥5 half-lives of this antidepressant before scanning) for the present MDD-episode. Exclusion criteria were pregnancy (or desire to become pregnant), bipolar disorder, psychotic features, primary anxiety and/or substance abuse disorders and acute, severe suicidal ideation. We allowed secondary co-morbid anxiety and/or substance abuse.

We individually matched each patient by gender and age (±2 years). HC were in good physical health and had never used psychotropic medication. Exclusion criteria were current or lifetime psychiatric disorder(s) according to the SCID (including abuse or addiction disorders), a Beck Depression Inventory (BDI) score >9, alcohol use >4 units per day (last month) or a 1st-degree relative with psychiatric disorder(s). We allowed HC to have incidentally used illicit drugs unless
criteria for a DSM-IV disorder were met, but prohibited illicit drug use the month prior to scanning. Patients and HC received €50 and €40, respectively. No restrictions were made with respect to smoking behavior.

**Procedure & SPECT-imaging**

We performed all scans 230 ±18 (SD) minutes after intravenous injection of approximately 100 MBq \([^{123}\text{I}]\beta\text{-CIT}\), when the radioligand was at equilibrium for SERT binding in brain areas expressing high densities of SERTs.\(^{22}\) Radiosynthesis of \([^{123}\text{I}]\beta\text{-CIT}\) and image acquisition were described earlier.\(^{25}\) We performed SPECT imaging using a 12-detector single slice brain-dedicated scanner (Neurofocus 810, Strichmann Medical Equipment; Cleveland, OH) with a full-width at half-maximum resolution of 6.5 mm, throughout the 20 cm field-of-view (http://www.neurophysics.com).

**Image Analysis**

After attenuation correction and reconstruction in 3D mode (http://www.neurophysics.com), we selected regions of interest (ROIs) for midbrain (MID), DIENC and cerebellum (CER) by using validated templates (see Figure 2.3).\(^{25}\) One examiner (HGR), blinded for diagnosis positioned all ROIs in two series. Intra-class correlation coefficients were ≥0.98 for all ROIs. If the two series differed by >5%, scans were re-evaluated by a second investigator (JB). In the analyses we averaged the counts for the two series.

We assumed activity in CER to represent non-displaceable activity (non-specific binding and free radioactivity).\(^{26}\) We calculated the binding potential (BP) as the rate of specific to non-displaceable (ND) binding (BP\text{ND} = \frac{\text{activity}_{\text{ND}} - \text{activity}_{\text{free}}}{\text{activity}_{\text{ND}}}) for MID and DIENC.\(^{27}\) BP\text{ND} is proportional to transporter number under equilibrium conditions.

**Statistics**

General linear models were used to analyze differences in BP\text{ND} in MID and DIENC between depressed patients and HC using the following modeling strategy.

We first compared mean BP\text{ND} in MDD patients versus HC in univariable (‘crude’) models, only containing the main effect of diagnosis (categorical: MDD/HC). We then fitted multivariable models by adding variables to the model, which in the literature have been reported to influence BP\text{ND}. These variables included: gender (categorical: male/female), age (continuous), smoking (categorical: yes/no), season of scanning (categorical: “winter” October-March/ “summer” April-September).\(^{24}\) In addition, a number of specific two-way interactions were examined, again because they have been reported as significant in previous studies, which included: diagnosis*gender, diagnosis*smoking, gender*smoking (‘full multivariable models’). Three-way interactions were not examined because of the relatively small sample-size (from a statistical perspective). The Akaike’s Information Criterion (AIC) was used to judge whether the two-way interactions improved the model. If a two-way interaction did not improve the fit of the model, it was removed from the model in order to facilitate the interpretation of the model (‘reduced multivariable models’). Main effects always remained in the model, irrespective of their significance, in order to report their (lack of) impact. Diencephalon and midbrain data were analyzed separately, but the same set of variables and interaction terms were examined using the same modeling approach. If significant interactions were present, we performed post-hoc analyses in order to report the absolute differences in SERT availability in the involved subgroups. Differences between subgroups were analyzed within the framework of the multivariable model and tested for significance using the residual variance estimate of the model. For explorative analyses see the supplementary appendix.

We examined the association of BP\text{ND} with HDRS scores using linear regression models in patients, correcting for covariates used in the multivariable models. We used SPSS (version 15.0.1) for statistical procedures (www.spss.com). We expressed all means ±SD, except in Figures 7.1 and S7.1, where ±SEM was used for legibility.
Results

We studied 17 male and 32 female patients with MDD, versus 17 male and 32 female HC (Table 7.1). Patients smoked significantly more (n= 27; 55.1%) than HC (n= 11; 22.4%; $\chi^2= 11.2$; df= 2; $p= 0.004$). Among HC we included more Caucasians (n= 44; 89.8%) than among patients (n= 31; 63.3%; $\chi^2= 11.9$; df= 3; $p= 0.008$). We scanned 67% of patients and 55% of HC in winter ($\chi^2= 1.55$, df= 1, $p= 0.213$). In one female patient insufficient CER was scanned as a reference, and in three patients and one control MID slices were insufficient; these were omitted in the analyses. Of 15 patients who used antidepressants during their lives, 3 used antidepressants for the current episode. One patient used mirtazapine until 4 weeks before scanning, all others stopped antidepressants 6-132 months before scanning. Illicit drug abuse mainly involved cannabis. Lifetime MDMA-use occurred in none of the patients and in one female control (<10 tablets; last use 8 months before scanning).

### Table 7.1. Baseline characteristics of MDD patients and healthy controls (stratified by gender).

<table>
<thead>
<tr>
<th></th>
<th>MDD patients</th>
<th>Healthy controls (HC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male (n= 17)</td>
<td>Female (n= 32)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>43.2 ±7.8</td>
<td>41.8 ±8.6</td>
</tr>
<tr>
<td>Current cigarette smokers; n (%)</td>
<td>12 (70.6)</td>
<td>15 (46.9)</td>
</tr>
<tr>
<td>Alcohol use &gt;8 Units/ week; n (%)</td>
<td>3 (17.6)</td>
<td>4 (12.5)</td>
</tr>
<tr>
<td>Race: n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>12 (70.6)</td>
<td>19 (59.4)</td>
</tr>
<tr>
<td>Creole</td>
<td>2 (11.8)</td>
<td>7 (21.9)</td>
</tr>
<tr>
<td>Asian</td>
<td>2 (11.8)</td>
<td>5 (15.6)</td>
</tr>
<tr>
<td>MDD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severity‡</td>
<td>23.2 ±4.9</td>
<td>25.4 ±4.7</td>
</tr>
<tr>
<td>First episode; n (%)</td>
<td>9 (52.9)</td>
<td>18 (56.3)</td>
</tr>
<tr>
<td>Drug-naïve; n (%)</td>
<td>12 (70.6)</td>
<td>22 (68.8)</td>
</tr>
<tr>
<td>Melancholic; n (%)</td>
<td>12 (70.6)</td>
<td>24 (75.0)</td>
</tr>
<tr>
<td>Atypical; n (%)</td>
<td>0</td>
<td>3 (9.4)</td>
</tr>
<tr>
<td>Suicidal thoughts, plan or attempt; n (%)</td>
<td>7 (41.2)</td>
<td>10 (31.3)</td>
</tr>
<tr>
<td>Duration of episode: n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5 months</td>
<td>5 (29.4)</td>
<td>8 (25.0)</td>
</tr>
<tr>
<td>5 months – 2 years</td>
<td>8 (47.1)</td>
<td>21 (65.6)</td>
</tr>
<tr>
<td>&gt; 2 years</td>
<td>4 (23.5)</td>
<td>3 (9.4)</td>
</tr>
<tr>
<td>Age of first episode (years)</td>
<td>35.9 ±10.3</td>
<td>35.8 ±10.4</td>
</tr>
<tr>
<td>Co-morbidity: n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>4 (23.5)</td>
<td>7 (21.9)</td>
</tr>
<tr>
<td>Dysthymia</td>
<td>1 (5.9)</td>
<td>1 (3.1)</td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>1 (5.9)</td>
<td>3 (9.4)</td>
</tr>
<tr>
<td>Drug (alcohol, cannabis, benzodiazepines) abuse</td>
<td>4 (23.5)</td>
<td>3 (9.4)</td>
</tr>
<tr>
<td>SPECT scan in winter; n (%)</td>
<td>9 (52.9)</td>
<td>24 (75.0%)</td>
</tr>
</tbody>
</table>

Numbers represent means ± SD

1. Significantly more smoking in MDD patients than in HC ($\chi^2 = 11.0$; 1 df; $p< 0.01$)
2. Trend of lower alcohol use in MDD patients than in HC ($\chi^2 = 3.75$; 1 df; $p= 0.053$)
3. Severity measured by Hamilton Depression Rating Scale (17-items) in MDD patients and by Beck Depression Inventory in HC

SERT availability in MDD patients versus healthy controls (‘crude models’)

Mean MID BP$_{ND}$ in patients (0.62 ±0.22 [SD]) was not significantly different from HC (0.63 ±0.19; $F_{1.94}= 0.118$; $p= 0.733$). BP$_{ND}$ in DIENC was 1.15 (±0.24) in patients and 1.09 (±0.26) in HC ($F_{1.97}= 1.209$; $p= 0.274$).

Multivariable models of SERT availability in MDD patients versus healthy controls

**Midbrain**

For Mid, the full multivariable model (including diagnosis, gender, age, smoking, season of scanning, diagnosis*gender, diagnosis*smoking, and gender*smoking) was subsequently reduced by removing the non-significant interaction terms diagnosis*smoking, and...
gender*smoking (AIC-decrease = 2.478; Figure 7.1A). This reduced multivariable model had a significant diagnosis*gender interaction ($F_{1,87} = 4.039; p = 0.048$). In post-hoc comparisons, MDD males showed a trend of lower BP$_{ND}$ compared with male HC (difference $-0.124; t_{87} = -1.718; p = 0.089$), while female patients and HC did not differ (difference $0.047; t_{87} = 0.898; p = 0.372$). Furthermore, the main effect of season of scanning was significant ($F_{1,87} = 5.814; p = 0.018$). Scans performed in winter showed on average 18% higher BP$_{ND}$ than scans made in summer ($F_{2,87} = 3.248; p = 0.044$). The main effects of age and smoking were also included in this reduced model, but were not significant ($p = 0.227$ and $p = 0.582$, respectively; Figure 7.1A).

### Figure 7.1. BP$_{ND}$ values for midbrain and diencephalon.

![Graph A] A. Midbrain (n= 94), corrected for main effects of diagnosis ($p = 0.414$), gender ($p = 0.128$), season of scanning ($F_{1,87} = 5.814; p = 0.018$), smoking ($p = 0.582$), age ($p = 0.227$) and diagnosis*gender interaction ($F_{1,87} = 4.039; p = 0.048$). * Post-hoc differences $t_{88} = -1.718; p = 0.089$ between MDD males versus HC males.

B. Diencephalon (n= 97), corrected for main effects of diagnosis ($p = 0.476$), gender ($p = 0.277$), season of scanning ($p = 0.679$), smoking ($p = 0.223$), age ($p = 0.247$) and diagnosis*gender interaction ($F_{1,88} = 10.127; p = 0.002$), gender*smoking interaction ($F_{1,88} = 4.541; p = 0.036$) and diagnosis*smoking interaction ($p = 0.127$).

![Graph B] ** Post-hoc differences $t_{88} = -2.384; p = 0.019$ between smoking HC males versus non-smoking HC males. ** Post-hoc differences $t_{88} = -2.643; p < 0.01$ between smoking MDD males versus smoking non-smoking HC males, non-smoking MDD females versus non-smoking HC females.

### Diencephalon

For DIENC, the full multivariable model (including the same variables as MID) could not be reduced as all three interactions terms (diagnosis*smoking, diagnosis*gender and smoking*gender) improved the fit of the model as evaluated by the AIC (Figure 2B). This full multivariable model showed significant diagnosis*gender ($F_{1,88} = 10.227; p = 0.002$) and smoking*gender ($F_{1,88} = 4.541; p = 0.036$) interactions. The diagnosis*smoking interaction was not significant ($p = 0.127$). In post-hoc comparisons, smoking MDD males had significant lower BP$_{ND}$ compared with smoking male HC (difference $-0.304; t_{88} = -2.643; p = 0.016$). In non-smoking MDD males BP$_{ND}$ was numerically lower than in non-smoking male HC (difference $-0.140; t_{88} = -1.340; p = 0.184$). Contrary, non-smoking female patients had higher BP$_{ND}$ than non-smoking female HC (difference $0.221; t_{88} = 3.064; p = 0.003$) with almost no difference in BP$_{ND}$ between MDD and HC in smoking females (difference $0.057; t_{88} = 0.616; p = 0.539$). Furthermore, male smoking HC had significantly higher BP$_{ND}$ than non-smoking male HC (difference $0.279; t_{88} = -2.384; p = 0.019$), while in female HC BP$_{ND}$ was not affected by smoking (difference $0.032; t_{88} = -0.371; p = 0.712$). The main effects of season of scanning and age were also included in this model, but were not significant ($p = 0.679$ and $p = 0.247$, respectively; Figure 7.1B).

### Relation of SERT availability and severity of MDD

With linear regression models, we found no significant relation between HDRS-scores and BP$_{ND}$, neither in MID, nor in DIENC when accounting for gender, age, smoking and season of scanning.
Discussion

In the present – until now largest – study of MDD patients versus HC, we aimed to quantify SERT availability in MDD patients versus HC while accounting for covariates and interactions, and to correlate SERT availability with depression severity. We did not find significant differences in SERT availability in MID or DIENC in crude comparisons. However, a significant diagnosis*gender interaction existed in MID and DIENC, combined with a significant gender*smoking interaction in DIENC only. Depressed males, but not females, had lower MID SERT availability compared with HC. In DIENC depressed smoking males had significantly lower MID SERT availability compared with smoking male HC, while non-smoking female patients had higher SERT availability than female HC. Furthermore, the season of scanning influenced SERT availability in MID, with higher SERT availability in winter. We found no clinically relevant correlation of HDRS-scores with SERT availability.

Comparison with previous studies

Our results replicate earlier reports of similar SERT availability in MID and DIENC in MDD patients and HC. Other studies reported increased or decreased SERT availability in MDD patients versus HC. However, none of these studies – except two – investigated the effect of gender and no study corrected for season. Furthermore, we replicated a significant contribution of season of scanning on MID BPND.

The diagnosis*gender interaction in MID and DIENC is our most important finding. Contrary to Staley and colleagues, we found a different direction of this interaction in DIENC: significantly lower BPND in MDD males in MID (-17%) and DIENC (-18%), and higher BPND in MDD females in MID (+9%, non-significant) and DIENC (+13%, significant) than in HC. Staley et al. found 1% lower DIENC SERT availability in MDD males, and 22% lower DIENC SERT availability in MDD females. We found a gender*smoking interaction in DIENC (with highest BPND in smoking healthy males), while Staley found higher SERT availability in the brainstem (attributable to males). Our findings suggest that a failure to stratify for a diagnosis*gender interaction may obscure differences between patients and HC.

Methodological explanations for inconsistent findings

Despite technical differences between studies (scanning protocols, radioligands, image analyses), variation in the selection of HC (e.g. having relatives with psychiatric diagnoses) or patients (from different source populations) is the most probable explanation for inconsistent findings. Previous studies recruited patients in general psychiatric outpatient and university clinics. We recruited 65% of our patients from primary care settings. We adequately diagnosed patients by SCID, and required an HDRS >18 for inclusion. Thus, we recruited severe and often melancholic patients that were drug-free, with 69% of the patients drug-naive. Three studies included larger proportions of drug-naive patients. Like Parsey et al., we observed (non-significant -15%) lower MID SERT availability in drug-naive patients (results available on request). Additionally, some studies suggested that anxiety disorders influence SERT availability, and MDD with co-morbid anxiety may differ from ‘pure MDD’. However, this was not observed in our sample (results available on request).

Role of SERT in the pathogenesis of MDD

SERTs evacuate extracellular 5-HT from the synapse. Observed differences in SERTs between patients and HC may represent differences in the number of SERT containing neurons, in the number of SERTs per neuron or a combination of both.

Two major mechanisms for the role of SERT in MDD are hypothesized. First, increased SERT availability reduces 5-HT from the synapse more easily, which might lower 5-HT transmission, possibly leading to MDD. Second, as the brain might apply compensation mechanisms to retain homeostasis, a decreased 5-HT transmission by MDD may result in downregulation (decrease) of...
SERT in order to increase 5-HT transmission. A sequential occurrence of these two mechanisms could also be hypothesized: an initially increased SERT availability destabilizes (with or without an additional factor) and leads to MDD, which is followed by a decrease in SERT to compensate for decreased 5-HT transmission.

Differential effects of MDD on SERT availability between sexes may be explained via sex-hormones. Estrogen replacement after ovariectomy increased SERT mRNA and SERT availability in female rats and in hypothalamic regions of female macaques. Depressed women may have significantly higher 24h mean levels of diurnal estradiol rhythms, and may have higher testosterone levels compared with HC. Testosterone may increase SERT availability by conversion to estrogen by aromatase, which is especially available in DIENC. This could explain our finding of increased SERT availability in DIENC in females. In depressed men, the sex steroid testosterone is decreased, with 34-61% biochemical hypogonadism in depressed males compared with 6-14% in HC. This lack of testosterone in MDD may reduce SERT availability by reduced conversion to estrogen. Replacement of testosterone in castrated male rats increased SERT mRNA and SERT availability. Since we did not measure sex hormones, and Best et al. found no relation between menstrual cycle or sex hormones and SERT availability in HC, these explanations remain speculative and should be examined further.

We replicated a main effect of season demonstrated previously in DIENC in 12 healthy women and mesencephalon in 29 HC. Neumeister et al. observed decreased SERT availability in winter. In contrast, Buchert et al. found increased SERT availability in winter, which was also found in our study. Serotonin modulates the effects of photic input in the suprachiasmatic nucleus (SCN). The SCN imposes a circadian rhythm by affecting hormonal and autonomic output (reviewed by Buijs). Serotonin release in the SCN is highest during waking and activity. Because raphe neurons show regular high firing rates during waking and decreased firing during sleep, it could be hypothesized that during winter, with decreased daylight, more serotonergic activity is needed, which may be mediated via the raphe input into the SCN. Increased serotonergic activity (increased free synaptic serotonin) may result in a compensatory increase in SERTs. Nevertheless, the small size of the SCN (~0.27 mm³) by itself cannot explain higher SERT availability found by SPECT or PET.

Limitations of the present study

The cerebellum (especially the vermis) contains small amounts of SERT, which could result in an underestimation of BPND in patients and HC, expected to be 7% at most. Although a systematic underestimation of SERT due to DAT- or NET-rich areas in MID (substantia nigra and locus coeruleus, respectively) cannot be ruled out, we think this does not differentially affect patients or HC.

Second, lower levels of endogenous 5-HT (e.g. in MDD) could result in less competition with radioligands, increasing the specific binding measured. This was demonstrated in rhesus monkeys with [123I]β-CIT SPECT but not in humans. After tryptophan depletion (artificially reducing endogenous 5-HT) no differences in SERT availability were observed but the radioligand ([11C]DASB) in that study does not bind to the 5-HT recognition/translocation site, and may not be suitable to image such changes in extracellular 5-HT.

Third, we used previously validated ROIs instead of Magnetic Resonance Imaging for coregistration. Because these templates cover larger brain areas, BPND in small regions (raphe nuclei) cannot be determined. This potential measurement error (non-differential for patients and HC), might have increased variance in our measurements, despite very good intra-rater correlation coefficients. Additionally, we did not correct for non-uniform photon attenuation and partial volume effects in our gender-analyses. Greater skull thickness might underestimate BPND in males, and smaller MID and DIENC in females might suppress BPND compared with males. However, these factors are unlikely to explain the observed interactions.
Fourth, the 4-week washout of antidepressants (binding to SERT) may be too short.48 Because all but one patient stopped antidepressants ≥6 months before scanning, we think no substantial bias of the BPND assessment was introduced by competitive binding by traces of previous antidepressants.

Fifth, we allowed previous incidental use of illicit drugs (marijuana/cannabis n= 10, or MDMA n= 1) in our HC. Because heavy use of MDMA (>50 MDMA tablets) can damage serotonin neurons,49 we performed an additional analysis in which we excluded data from the MDMA-user in the HC group. However, this exclusion did not affect our results (results available on request). Sixth, we did not check personal or family history of psychiatric illnesses in HC, nor did we test for alcohol or drug abuse.

**Conclusion**

We showed lower SERT availability in MID and DIENC in depressed males and higher SERT availability in DIENC in depressed (non-smoking) females compared with HC. We replicated a seasonal influence on MID SERT availability, and found a gender*smoking interaction on DIENC SERT availability. This study points to complex effects of gender, smoking and season on the serotonergic system in the pathogenesis of MDD.

**Acknowledgements**

We wish to thank patients and healthy controls for their participation in the study. We thank general practitioners and psychiatric residents for appropriate referrals of patients. Staff members and technicians of the department of Nuclear Medicine provided indispensable help in making the SPECT-scans. We would like to acknowledge Dr. Maartje M. de Win, MD PhD for her help to recruit HC. Dr. Jules Lavalaye, MD PhD helped in the initial design of this study. Mrs. Michelle L. Miller revised the text linguistically. Dr. Ruud M. Buijs, PhD commented constructively on an earlier version of the manuscript. Henricus G. Ruhé received a grant from the Netherlands Organisation for Health Research and Development (ZonMw), program Mental Health, education of investigators in mental health (OOG; #100-002-002).

**Conflicts of interest**

None

**References**


5. Reivich M, Amsterdam JD, Brunswick DJ, Shiue CY. PET brain imaging with $[^{11}C](+)$McN5652 shows increased serotonin transporter availability in major depression. *J Affect Disord*. 2004; 82: 31-41.


Supplementary appendix: explorative analyses

METHODS

In addition to the analyses described in the main text, in exploratory analyses we examined whether potentially relevant other factors (alcohol use, lifetime antidepressant use and duration of depression episode) had an impact on BP_{ND} by adding them as main effects to the final multivariable model. Because our exploratory models used small subgroups for assessment of further confounding in multivariable models, we only present these analyses as hypothesis generating, exploratory models.

RESULTS

Midbrain

For MID differentiation of patients who never took antidepressants (drug-naive) versus those who did improved the reduced model (AIC-decrease= 1.399). Drug-naive male and female patients had non-significantly lower BP_{ND} than patients with a history of antidepressant use (data not shown).

Differentiation of the duration of the current episode, stratified as ≤2 years or >2 years ('chronic' depression) improved the reduced model (AIC-decrease= 7.289). Men depressed for more than 2 years had significantly higher BP_{ND} than men depressed ≤2 years (difference 0.333; t_{85} = -3.102; p= 0.003), men depressed ≤2 years had significantly lower BP_{ND} than HC (difference -0.204; t_{85} = -2.795; p= 0.006; Figure S7.1A). Differences between females were not significant.

Diencephalon

Differentiation between patients with or without a secondary diagnosis of alcohol abuse/dependence improved the model for DIENC (AIC-decrease= 0.602). Smoking, alcohol-abusing male patients had numerically higher BP_{ND} than smoking male patients not abusing alcohol (difference 0.717; t_{85} = -1.175; p= 0.064). Non-smoking, female, alcohol-abusing patients had significantly lower BP_{ND} than non-smoking, female patients not abusing alcohol (difference -0.388; t_{85} = 2.182; p= 0.032; Figure S7.1B).

Figure S7.1. BP_{ND} values for midbrain and diencephalon (exploratory models)

**Post-hoc differences for MDD males between: episodes ≤2 years versus >2 years (t_{85} = 3.102; p= 0.003), episodes ≤2 years versus HC (t_{85} = 2.795; p= 0.006).

**Post-hoc differences for MDD males between: episodes ≤2 years versus >2 years (t_{85} = 3.102; p= 0.003), episodes ≤2 years versus HC (t_{85} = 2.795; p= 0.006).

A. Midbrain (n= 94), corrected for main effects of duration of episode (F_{1,85} = 4.777; p= 0.011), gender (p= 0.294), age (p= 0.385), smoking (p= 0.636), season of scanning (F_{1,85} = 6.194; p= 0.015), and duration of episode*gender interaction (F_{2,85} = 3.953; p= 0.023).

B. Diencephalon (n= 97), corrected for main effects of alcohol abuse/dependence (p= 0.521), gender (p= 0.485), season of scanning (p= 0.511), age (p= 0.486), smoking (p= 0.067) and alcohol abuse/dependence*gender interaction (F_{2,85} = 7.236; p= 0.001), gender*smoking interaction (F_{2,85} = 7.112; p= 0.009) and alcohol abuse/dependence*smoking interaction (p= 0.099).

* Post-hoc differences t_{85} = 2.182; p= 0.032 between non-smoking MDD females without, versus with alcohol abuse/dependence.
<table>
<thead>
<tr>
<th>Author &amp; Date (reference)</th>
<th>Population (mean age)</th>
<th>Sex</th>
<th>N</th>
<th>SERT Imaging and ROI*</th>
<th>Results</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahonen 200414</td>
<td>Drug free MDD pts (34 ±7) Controls (35 ±7)</td>
<td>M + F ??</td>
<td>10</td>
<td>$[^{123}I]$ADAM SPECT, 10min, 5h, 7h MidBr, Thal, Caud, Putam, Pons, Cer</td>
<td>Nonsign. Differences of V3'' in MidBr (7% higher in patients); large interindividual variation in V3'' in MidBr. Other ROIs no sign. differences.</td>
<td>Abstract of preliminary results only; little information provided. Gender not reported. Apparently no matching of controls and patients. Recruitment of patients unspecified.</td>
</tr>
<tr>
<td>Catafau 200615</td>
<td>Drug-free MDD pts (36 ±11) Controls (36 ±11)</td>
<td>M + F</td>
<td>10</td>
<td>$[^{123}I]$ADAM SPECT, 4h MidBr, Thal, Striatum, Cer</td>
<td>Slightly lower but nonsign. differences in SERT availability in MidBr (-4%), Thal (-11%) in MDD patients vs controls.</td>
<td>Patients were drug-free for &gt;6m. Age matched historic control group, no matching for sex. Unequal male-female distribution in MDD vs. controls. No information on family history of controls. Recruitment of patients unspecified.</td>
</tr>
<tr>
<td>Dahström 200016</td>
<td>Drug-naive MDD pts (13.5 ±2.5) Drug-naive Non-MDD pts (12.2 ±2.9)</td>
<td>M + F</td>
<td>31</td>
<td>$[^{123}I]$-CIT SPECT, 1, 4, 24h Striatum (24h), PFC, Thal (4h), MidBr (4h), Occ</td>
<td>MidBr V3'' in patients sign. higher (8%) at 1hr compared to controls (difference not sign. at 4h, again sign. at 24h). Prefrontal and Thal differences not sign. different</td>
<td>Control-group is due to ethical considerations not depressed but a psychiatrically affected group of adolescents. MDD patients and controls are not matched for age and sex. No information on family history of controls. Recruitment of outpatients in university clinic.</td>
</tr>
<tr>
<td>Herold 200616</td>
<td>Drug-free MDD pts (42 ±12) Controls (36 ±13)</td>
<td>M + F</td>
<td>21</td>
<td>$[^{123}I]$-ADAM SPECT, 4h MidBr, Cer</td>
<td>MidBr V3'' in patients 2% higher (nonsign.). MDD males nonsign. lower V3'' than MDD females. No correlation with SERT and MDD severity rating.</td>
<td>All patients were drug-free for &gt;2m. No matching for age or sex. Unequal male-female distribution in MDD vs. controls. Study also reports on occupancy after 1 week of treatment with CIT 10 mg (n=13). Recruitment of patients unclear.</td>
</tr>
<tr>
<td>Ichimiya 200219</td>
<td>Drug-free MDD pts (44.1 ±13.5) or BD (41.7 ±8.8) Controls (42.3 ±14.5)</td>
<td>M + F</td>
<td>13</td>
<td>$<a href="+">^{11}C</a>$McN5652 PET MidBr, Thal, Cer</td>
<td>Sign. higher (22%) in Thal BP in complete patient group vs controls (in MDD only (23%), no change in MidBr</td>
<td>6 patients had BD; 10 patients had been treated with antidepressants &gt;6w before scanning (1 patient with BD 2w). Age matched (groupwise??) controls, not for sex. Recruitment of outpatients in university and general psychiatric clinics.</td>
</tr>
<tr>
<td>Joensuu 200713</td>
<td>Drug-naive MDD pts (38.8 ±8.6) Controls (36.6 ±8.9)</td>
<td>M + F</td>
<td>29</td>
<td>$[^{123}I]$nor β-CIT SPECT, 6, 24h MidBr, Cer</td>
<td>Sign. lower V3'' in patients (10%) in MidBr vs controls. No correlation with SERT and MDD severity rating.</td>
<td>Although mentioned as such, incomplete matching on gender and age. 14/29 pts had HDRS ≤18. Recruitment of outpatients for psychodynamic psychotherapy in university clinic. No sign. effects of gender (included patients mostly F) or season.</td>
</tr>
<tr>
<td>Lehto 200617</td>
<td>Drug-naive MDD pts (28.3 ±7) Controls (30.6 ±9.2)</td>
<td>M + F</td>
<td>29</td>
<td>$[^{123}I]$nor-β-CIT SPECT, 5 min, 6h, 24h MidBr, Striatum, Cer</td>
<td>Sign. lower SERT (-10%) in MidBr in MDD patients vs controls. No differences between melancholic, atypical, nondifferentiated MDD subtypes. Linear inverse correlation between SERT and atypical score. No correlation between SERT and HDRS.</td>
<td>Patients and controls were (groupwise??) matched for age and sex. No substantial rationale given for relation between SERT availability and atypical dimension of MDD. Recruitment of outpatients in university clinic.</td>
</tr>
</tbody>
</table>
Table S7.1. Previous studies measuring SERT availability in MDD patients versus healthy controls. (Continued)

<table>
<thead>
<tr>
<th>Author &amp; Date (reference)</th>
<th>Population (mean age)</th>
<th>N</th>
<th>SERT Imaging and ROI*</th>
<th>Results</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malison 1998; Kugaya 2004</td>
<td>Drug-free MDD pts (44 ±10) Controls (45 ±11) M + F</td>
<td>15</td>
<td>15</td>
<td>$^{[123I]}$β-CIT SPECT, 24h Brainstem, Striatum, Occ</td>
<td>Sign. reduction (-19%) in brainstem V3'' of MDD vs. controls, but not (-11%) in striatum. No correlation with SERT and MDD severity ratings</td>
</tr>
<tr>
<td>Meyer 2001</td>
<td>Drug-free MDD pts (37 ±8) Controls (7 ±7) M + F</td>
<td>13</td>
<td>13</td>
<td>$^{[1]}$C]DASP PET striatum, Cer</td>
<td>No sign difference in striatum V3'' of MDD vs controls. Sign. effect of age.</td>
</tr>
<tr>
<td>Meyer 2004a</td>
<td>Drug-free MDD pts (? ±?) Controls (? ±?) M + F</td>
<td>37</td>
<td>35</td>
<td>$^{[1]}$C]DASP PET striatum, Cer</td>
<td>No sign. differences in striatum V3'' of MDD vs controls.</td>
</tr>
<tr>
<td>Meyer 2004b</td>
<td>Drug free MDD pts (35 ±11) Controls (35 ±11) M + F</td>
<td>20</td>
<td>20</td>
<td>$^{[1]}$C]DASP PET Bilateral anteromedial PFC, dorsolateral PFC, Ant. Cing, Caudate, Putam, Thal, MidBr,Cer</td>
<td>No difference in SERT availability between MDD and controls for any region. Within patients (but not in controls) sign. correlations between increased Dysfunctional Attitude Scale (DAS) scores and increased SERT avail. Significant increase in SERT in patients vs. controls in all regions for 8 patients with high DAS scores (&gt;190) vs controls.</td>
</tr>
<tr>
<td>Newberg 2005</td>
<td>Drug-free MDD pts (38 ±?) Controls (37 ±?) M + F</td>
<td>7</td>
<td>6</td>
<td>$^{[1]}$C]McN5652 PET MidBr, Putam, Amyg, Thal, Hippoc, Ant. Cing, Cer</td>
<td>Sign. lower (-20%) BP' in all regions in MDD patients vs controls. Post-hoc assigned to amygdala and midbrain regions. Drug-naïve patients had sign. lower BP' in amygdala and midbrain vs non drug-naïve patients. No correlation with BP and depression severity ratings</td>
</tr>
<tr>
<td>Parsey 2006</td>
<td>Drug free MDD pts (38.0 ±13.4) Controls (38.8 ±15.9) M+F</td>
<td>25</td>
<td>43</td>
<td>$^{[1]}$C]McN5652 PET MidBr, Putam, Amyg, Thal, Hippoc, Ant. Cing, Cer</td>
<td>Sign. lower (-20%) BP' in all regions in MDD patients vs controls. Post-hoc assigned to amygdala and midbrain regions. Drug-naïve patients had sign. lower BP' in amygdala and midbrain vs non drug-naïve patients. No correlation with BP and depression severity ratings</td>
</tr>
</tbody>
</table>
Table S7.1. Previous studies measuring SERT availability in MDD patients versus healthy controls. (Continued)

<table>
<thead>
<tr>
<th>Author &amp; Date (reference)</th>
<th>Population (mean age)</th>
<th>N</th>
<th>SERT Imaging and ROI*</th>
<th>Results</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reivich 2004&lt;sup&gt;18&lt;/sup&gt;</td>
<td>Drug free MDD pts (22-56)&lt;br&gt;Controls (23-59) M+F</td>
<td>4</td>
<td>&lt;sup&gt;<a href="+">11C</a>&lt;/sup&gt;McN5652 PET Bilateral Frontal, Cing, Thal, Pons, Cer</td>
<td>Sign. larger BP in left frontal (+17%) and right cingulate (+24%) in patients vs controls. Thalasus and pons no sign. difference</td>
<td>MDD patients were drug free for ≥5 half-lives of the drug, ≥2w for MAO-I and ≥3w for fluoxetine. Patients were not matched for age or sex; no information on family history of controls. Wide age-range without possibility to correct for age-effects on SERT availability. Large SDs for thalamus or pons ROIs. Recruitment of outpatients.</td>
</tr>
<tr>
<td>Staley 2006&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Drug-free MDD pts (38.8 ±9.7)&lt;br&gt;Controls (38.9 ±10) M+F</td>
<td>32</td>
<td>&lt;sup&gt;[123I]&lt;/sup&gt;β-CIT SPECT, 24h Brainstem, Diencephalon, Striatum, Cer</td>
<td>Sign. reduction (-12%) in diencephalons V3'' of MDD vs controls, but not in brainstem. Interaction MDD by sex on diencephalon V3'': women -22% in MDD, men -1%.</td>
<td>Subjects were (group??) matched for sex, age and smoking status. No information on family history of controls. Recruitment of patients unspecified.</td>
</tr>
<tr>
<td>Willeit 2000&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Drug-free patients with SAD (30.5 ±8)&lt;br&gt;Controls (29.0 ±15.5) M+F</td>
<td>11</td>
<td>&lt;sup&gt;[123I]&lt;/sup&gt;β-CIT SPECT, 4, 24h Thal, Hypothalamus, MainBr, Pons, Cer</td>
<td>Sign. reduction (-15%) in Thalamus and Hypothalamus V3'' of MDD vs. controls, but not (-8%) in MidBr-Pons. No correlation with SERT and depression severity ratings</td>
<td>5 subjects had been treated with antidepressants ≥6m before scanning. Groupwise matching for age and sex. All differences in SPECT acquisitions 24h p.i., at 4h p.i. no significant differences found. Recruitment of outpatients of university clinic.</td>
</tr>
</tbody>
</table>

* Naming of identical regions as provided in the studies. Reference region in italics.

Abbreviations: Amyg = Amygdala, Ant. Cing = Anterior Cingulate, BD = Bipolar Disorder, BP = Binding Potential, Caud = Caudatus, Cer = cerebellum, CIT = citalopram, h = hour, Hippoc = Hippocampus, Hypothalamus, min = minute, m = month, MDD = Major Depressive Disorder, MidBr = Midbrain, Occ = Occipital lobe, PAR = paroxetine, PFC = Prefrontal Cortex, Putam = Putamen, SAD = Seasonal Affective Disorder, Thal = Thalamus, V3'' = BP, w = week
NEUROBIOLOGICAL EFFECTS OF PAROXETINE IN THE TREATMENT OF MAJOR DEPRESSIVE DISORDER

PART IV
SEROTONIN TRANSPORTER GENE PROMOTER POLYMORPHISMS MODIFY THE ASSOCIATION BETWEEN PAROXETINE SEROTONIN TRANSPORTER OCCUPANCY AND CLINICAL RESPONSE IN MAJOR DEPRESSIVE DISORDER

Pharmacogenetics and Genomics 2008; in press

Henricus G. Ruhé\textsuperscript{1}
Wendy Ooteman\textsuperscript{1}
Jan Booij\textsuperscript{2}
Martin C. Michel\textsuperscript{3}
Martina Moeton\textsuperscript{4}
Frank Baas\textsuperscript{4}
Aart H. Schene\textsuperscript{1}

\textsuperscript{1} Program for Mood Disorders, Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
\textsuperscript{2} Department of Nuclear Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
\textsuperscript{3} Department of Pharmacology and Pharmacotherapy, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
\textsuperscript{4} Department of Neurogenetics, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
Abstract

Background
In major depressive disorder (MDD), selective serotonin reuptake inhibitors (SSRIs) target the serotonin transporter (SERT). Their 30-50% response-rates are modified by SERT promoter polymorphisms (5-HTTLPR).

Aim
To quantify the relation between SERT occupancy and response, and whether 5-HTTLPR is a modifier.

Methods
Drug-free depressed outpatients (n= 49; both sexes; 25-55 years), received paroxetine 20 mg/day. We quantified SERT occupancy with \([^{123}\text{I}]\beta\)-CIT SPECT imaging at baseline and after 6 weeks; we genotyped 5-HTTLPR (S, L_G, L_A). Primary outcomes: percentage decrease in Hamilton Depression Rating Scale (HDRS_17) and response (≥50% decrease of HDRS_17).

Results
A significant positive relation between SERT occupancy and clinical response existed only in the L_A/L_A genotype (p< 0.002). Relative to paroxetine serum concentrations midbrain SERT occupancy was numerically higher for L_A/L_A compared with other genotypes, but this difference was not significant (p= 0.188).

Conclusions
Higher SERT occupancy is only associated with more clinical improvement in the L_A/L_A genotype. We hypothesize that the L_A/L_A carriers have a more dynamic serotonergic system, which appears more responsive to SSRIs.

ISRCTN register ISRCTN44111488
Introduction

Major depressive disorder (MDD) is often treated with antidepressants, most commonly selective serotonin reuptake inhibitors (SSRIs). SSRIs occupy the serotonin transporter (SERT) with high selectivity and affinity to block presynaptic serotonin reuptake. Unfortunately, clinical response rates for SSRIs are modest (30-50%) and difficult to predict. Unraveling the factors that modify clinical response may contribute to more effective SSRI treatment.

SERT occupancy can be estimated in-vivo using radioligands and SPECT or PET-imaging. Previous studies postulated that 80% SERT occupancy by SSRIs would be required for clinical response. However, when SERT occupancy-rates and clinical response were correlated, the same studies found no clear relationship between striatal SERT occupancy and clinical response after 4 weeks of SSRI treatment in MDD patients. Contrary, Kugaya et al. found that higher pretreatment SERT availability and greater SERT occupancy in diencephalon predicted better treatment response after 6 weeks of SSRIs. Recently, Zitterl et al. found a correlation between clomipramine diencephalon SERT occupancy and severity-scores in obsessive compulsive disorder. However, in these two studies SERT occupancies ranged between 23-61%.

Polymorphisms of the SERT gene promoter region (5-HTTLPR) are associated with its transcriptional activity and influences the rate of serotonin uptake. In human lymphoblasts, cells homozygous for the long (L) allele produce higher concentrations of SERT mRNA than cells containing one or two copies of the short (S) allele. Furthermore, serotonin uptake by the transporter is 2-fold higher in cells homozygous for the L-allele. In-vivo imaging inconsistently showed lower midbrain SERT availability for S-allele carrying healthy subjects compared with L-allele carriers. Individuals with an S-allele have increased anxiety related traits, elevated risk of depression after stress, increased amygdala reactivity, and increased adverse events after SSRI treatment, compared with subjects with an L-allele. An elegant, large MRI-study in healthy volunteers showed associations of the 5-HTTLPR S/S polymorphism with unfavorable alterations in anatomy and function of the amygdala-cingulate feedback circuit. This study points to an important role of the 5-HTTLPR-polymorphism in the development and functioning of emotional networks involved in MDD. A meta-analysis showed associations between the 5-HTTLPR-polymorphism and clinical SSRI efficacy within 4 weeks of treatment. Depressed patients without S-alleles had higher response rates to SSRIs. Thus, apart from developmental, functional, and stress reactivity effects on emotional networks, the 5-HTTLPR polymorphism also appears to influence the response to SSRIs.

Nowadays, the 5-HTTLPR-polymorphism is considered tri-allelic. The L-allele is subdivided as $L_C$ and $L_A$ by a common single nucleotide polymorphism (SNP; rs25531). The $L_C$ SNP creates a functional AP2 transcription factor binding site, behaving like an S-allele. The tri-allelic S:$L_A$:$L_C$ ratio is approximately 4:5:1 and can be reclassified into a modified bi-allelic genotype. In-vivo studies predominantly found higher SERT availability for $L_A/L_A$ carriers in healthy and MDD subjects. Recently, a significant association between the $L_A/L_A$ allele and citalopram adverse effects but not with response was found in MDD patients.

In summary, the relation between SERT occupancy and clinical response to SSRIs is unclear, a SERT polymorphism likely influences the development and function of the serotonergic system (e.g. SERT availability) and affects clinical response to SSRIs. Since it was not investigated whether the relation between SERT occupancy and clinical response is modified by the 5-HTTLPR-polymorphism, we aimed to quantify (1) the relation between SERT occupancy and clinical response, (2) pre-treatment SERT availability for different tri-allelic 5-HTTLPR genotypes and (3) the modification by 5-HTTLPR with respect to the relation between SSRI-occupancy and clinical response.
We hypothesized that (1) the relation between occupancy and clinical response is nonlinear with increased response-rates above 80% occupancy after 6 weeks of treatment, (2) we would find lower pretreatment SERT availability for the tri-allelic S'/S' 5-HTTLPR genotype and (3) because of the higher response-rates in patients with a L-allele, the association between occupancy and clinical response would be distinct for tri-allelic 5-HTTLPR genotypes. We studied these questions in a 6 week open trial of paroxetine 20 mg/day, in subjects who participated in the first phase of a randomized dose-escalation trial with $[^{123}I]$$\beta$-CIT SPECT assessment of SERT occupancy.34

**Methods**

**Participants**

Following approval by the institutional medical ethical committee and written informed consent, we recruited drug-free outpatients (25-55 years) from primary care, our outpatient department, and public psychiatric settings (October 2003 - August 2006) and included them in the study before antidepressants were started. Inclusion criteria were: MDD determined by the Structured Clinical Interview for DSM-IV (SCID),36 and a Hamilton Depression Rating Scale (HDRS$_{17}$) score above 18. Patients were drug-naïve or drug-free ($\geq$ 4 weeks and $\geq$ 5 half-lives of a previous antidepressant, when treated previously) and had used no more than one antidepressant treatment (other than paroxetine) at an effective dose for $\geq$ 6 weeks for the present MDD-episode. Exclusion criteria were pregnancy (or wish), bipolar disorder, psychotic features, neurological cognitive impairments (i.e. dementia), primary anxiety and/or substance abuse disorders and acute, severe suicidal ideation. We allowed secondary co-morbid anxiety and/or substance abuse to increase applicability of the findings of our main study.34

**SSRI treatment**

After baseline assessment, patients were treated open-label with paroxetine 20 mg/day for six weeks. When severe adverse effects occurred, dosages were reduced to 10 mg/day and again increased to 20 mg/day after one week. We supplied paroxetine in pill-boxes to improve treatment adherence. We checked adherence by pill-counts and medical history.38 Benzodiazepines (temazepam 10-20 mg/day or oxazepam 10-30 mg/day) were allowed if necessary.

**Questionnaires and measurements**

Primary clinical outcomes were the percentage decrease in HDRS$_{17}$-score, and the proportion of patients achieving clinical response ($\geq$50% decrease in HDRS$_{17}$). We administered questionnaires at study-entry and after 6 weeks of treatment. In addition, depressive symptoms were monitored at week 2 and 4 using the Maier and Bech subscales of the HDRS$_{17}$ and the self-rated Inventory for Depressive Symptomatology (IDS-SR$_{30}$) scores.41 Three trained investigators who administered the clinician-rated questionnaires (HDRS$_{17}$ and subscales) had good inter-rater agreement (intra-class correlation coefficient = 0.98).

**SPECT imaging and analysis**

We performed single photon emission computed tomography (SPECT) imaging for in-vivo assessment of SERT availability at study-entry and after 6 weeks between 2 to 10 pm as described previously.42 The acquisition started 230 ±18 (SD) minutes after intravenous injection of 100 MBq iodine-123-labeled $2\beta$-carbomethoxy-3$\beta$-(4-iodophenyl)-tropane ($[^{123}I]$$\beta$-CIT), when the radioligand is at equilibrium for SERT binding in brain areas expressing high densities of SERTs (i.e. midbrain and diencephalon).43 To prevent thyroid uptake of $[^{123}I]$, all subjects received oral potassium-
iodide solution. We performed SPECT imaging using a 12-detector single slice brain-dedicated scanner (Neurofocus 810, Strichmann Medical Equipment; Cleveland, OH) with a full-width at half-maximum resolution of 6.5 mm, throughout the 20 cm field-of-view (www.neurophysics.com).

After attenuation correction and reconstruction in 3D mode (www.neurophysics.com), we defined regions of interest (ROIs) for midbrain, diencephalon and cerebellum by using validated templates (see Figure 2.3).42,44 One investigator, blinded for scan session (study-entry/6 weeks), positioned all ROIs in two series. Intra-class correlation coefficients between series were >0.97 for all ROIs. If the two series differed by >5%, scans were re-evaluated by a second investigator. In the analyses the counts were averaged for the two series.

Using activity in the cerebellum (CER) as indicator of non-displaceable activity (non-specific binding and free radioactivity),44 we estimated the non-displaceable binding potential (BP_{ND}) of the radioligand to SERT by calculating the ratio of specific to non-specific binding per scan as BP_{ND} = \frac{(\text{activity}_{\text{total}} - \text{activity}_{\text{CER}})}{\text{activity}_{\text{CER}}} . BP_{ND} is proportional to SERT availability under equilibrium conditions.45 In a different study, we found high reproducibility of SERT imaging with [\text{123I}]\beta-CIT SPECT after repeated scanning of subjects, using the same camera and scanning-protocol (de Win et al., submitted). We calculated SERT occupancy at 6 weeks relative to the untreated SERT BP_{ND} (study-entry): OCC_{6 \, \text{weeks}} = \frac{(\text{BP}_{\text{ND}_{\text{study-entry}}} - \text{BP}_{\text{ND}_{\text{6 \, \text{weeks}}}})}{\text{BP}_{\text{ND}_{\text{study-entry}}}}.

**Paroxetine serum concentrations**

We collected blood for paroxetine serum concentrations (PSC; therapeutic range 10-75 µg/L) after 6 weeks, immediately before SPECT scanning. Serum was stored at -20° C until analysis. PSCs were determined using a validated High Pressure Liquid Chromatography-MS/MS method (available on request). The lower limit of quantification was 5µg/L, the lower limit of detection was 0.3 µg/L.

**Genotyping Procedures and Analysis**

Genomic deoxyribonucleic acid (DNA) was isolated out of blood using a filter-based method (QIAamp DNA Mini Kit, Qiagen Ltd, United Kingdom). The length of 5-HTTLPR28 was determined by gel electrophoresis. The region around the polymorphism was amplified by PCR using forward primer \text{gtgtaaacgacggccagtgcagcacaccaacctaat} and reverse primer \text{caggaaacagctatgaccagggagatcctgggaga} (M13 primer sequence in italics). The PCR reaction was performed in 10µl 1.5mM MgCl_{2}, 0.2µM forward and reverse primer, 0.1mM dNTP's, 0.5 Units Hotfire Polymerase (Solis Biodyne, Estonia), Buffer B (Solis Biodyne, Estonia) and 20ng genomic DNA. The lengths of the different alleles were short \approx 250 bp and long \geq 298 bp. Genotyping of the rs25531 SNP was done by sequencing (Sanger) using Big Dye Terminators (Applied Biosystems). The M13 forward primer \text{gtgtaaacgacggccagt} was used for sequencing. Reactions were performed containing 5ng of a forward primer, 5µl PCR product, BDT mix (Applied Biosystems) and 2.5xBDT buffer (Applied Biosystems). The length of the 5-HTTLPR polymorphism was confirmed by looking at the length of the sequenced PCR product. We reclassified the genotypes as S'/S' (S/S, Lc/Lc, Lc/LC, LC/LC), S'/LA (S/LA, LG/LA) and LA/LA.29 In post-hoc analyses we grouped S'/S' and S'/LA genotypes to contrast these with the ‘high-expression’ LA/LA genotype.

**Statistical Analysis**

For cross tabulations and differences between groups we used χ^{2} tests, Fisher’s exact test and ANOVA in SPSS for Windows v15.0.1.1 (www.spss.com). To investigate the relation between SERT occupancy and PSC, we modeled SERT occupancy (OCC) after 6 weeks in an E\text{\textsuperscript{max}} model as OCC = \frac{\text{PSC}_{\text{a}}}{(\text{PSC}_{\text{a}} + \text{b})} , in which \text{a} represents maximal SERT occupancy in the model (OCC_{\text{max}}) and \text{b} the PSC with 50\% SERT occupancy (EC_{50}).4-7;9 We calculated \text{a} and \text{b} by fitting a nonlinear regression model that minimizes the sum of squares of the residuals (GraphPad Prism v5.00, www.graphpad.com). To assess whether PSC-occupancy or occupancy-response curves were improved by sub-grouping (e.g. genetic subgroups or responders versus non-responders), we fitted either one curve, or separate curves and determined whether separate curves decreased
the Akaike Information Criterion (AIC; lower is better), which expresses the $-2 \log$-likelihood of the (nested) model penalized for the number of independent variables in the model. We furthermore verified these analyses by regression models including interaction terms for genotype.

To predict clinical response by SERT occupancy, we performed $\chi^2$ and ANOVA tests to estimate group differences in clinical response-rate and proportional decrease in HDRS$_{17}$, respectively, with SERT occupancy stratified as poor <50%, intermediate 51-79%, and high ≥80%, and linear and logistic regression analyses for continuous outcomes (proportional decrease in HDRS$_{17}$) and dichotomous outcomes (response). In these models occupancy was entered as a continuous and dummy-coded group variable (stratified SERT occupancy), while the models were corrected for age, baseline midbrain/diencephalon SERT availability and sex.

One responder and one non-responder were potentially non-adherent after 6 weeks (PSC <5µg/L), but were included in the analyses for three reasons. First, this estimation of medication adherence was only based on a single measurement of PSC over a 6-week treatment period, second, patients had empty pill-boxes and had claimed adherence, and third, potential non-adherence in other patients before the measurement of PSC could not be excluded either.

Table 8.1. Characteristics of patients, stratified by clinical response after 6 weeks of paroxetine 20 mg/day.

<table>
<thead>
<tr>
<th></th>
<th>Responders*(n= 10)</th>
<th>Non-responders*(n= 32)</th>
<th>p†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at baseline (years)</td>
<td>43.5 ±3.7</td>
<td>41.5 ±7.6</td>
<td>0.51</td>
</tr>
<tr>
<td>Female sex - n (%)</td>
<td>6 (60.0)</td>
<td>21 (65.6)</td>
<td>1.00</td>
</tr>
<tr>
<td>Current smoking - n (%)</td>
<td>5 (50.0)</td>
<td>17 (53.1)</td>
<td>1.00</td>
</tr>
<tr>
<td>Alcohol use: n (%)</td>
<td>7 (70.0)</td>
<td>30 (93.8)</td>
<td>0.08</td>
</tr>
<tr>
<td>≤ 7 Units/week</td>
<td>3 (30.0)</td>
<td>2 (6.2)</td>
<td></td>
</tr>
<tr>
<td>HDRS$_{17}$ at baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDD</td>
<td>23.8 ±3.7</td>
<td>25.0 ±4.7</td>
<td>0.47</td>
</tr>
<tr>
<td>IDS-SR$_{20}$ baseline</td>
<td>41.0 ±7.2</td>
<td>44.7 ±9.1</td>
<td>0.28</td>
</tr>
<tr>
<td>First episode - n (%)</td>
<td>4 (40.0)</td>
<td>18 (56.3)</td>
<td>0.48</td>
</tr>
<tr>
<td>No of episodes</td>
<td>2.0 ±1.1</td>
<td>1.9 ±1.8</td>
<td>0.92</td>
</tr>
<tr>
<td>Drug-naive - n (%)</td>
<td>7 (70.0)</td>
<td>22 (68.8)</td>
<td>1.00</td>
</tr>
<tr>
<td>Used AD in current episode - n (%)</td>
<td>0 (0.0)</td>
<td>3 (9.4)</td>
<td>1.00</td>
</tr>
<tr>
<td>Melancholic - n (%)</td>
<td>7 (87.5)</td>
<td>24 (96.0)</td>
<td>0.43</td>
</tr>
<tr>
<td>Duration of episode: n (%)</td>
<td></td>
<td></td>
<td>0.97</td>
</tr>
<tr>
<td>&lt;5 months duration</td>
<td>3 (30.0)</td>
<td>10 (31.3)</td>
<td></td>
</tr>
<tr>
<td>5 months – 2 years duration</td>
<td>6 (60.0)</td>
<td>18 (56.3)</td>
<td></td>
</tr>
<tr>
<td>≥ 2 years</td>
<td>1 (10.0)</td>
<td>4 (12.5)</td>
<td></td>
</tr>
<tr>
<td>Age of first episode (years)</td>
<td>33.3 ±11.2</td>
<td>36.4 ±9.6</td>
<td>0.40</td>
</tr>
<tr>
<td>Co-morbidity - n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>3 (30.0)</td>
<td>3 (9.4)</td>
<td>0.14</td>
</tr>
<tr>
<td>Dysthymia</td>
<td>0 (0.0)</td>
<td>1 (3.1)</td>
<td>1.00</td>
</tr>
<tr>
<td>Drug abuse / dependence‡</td>
<td>3 (30.0)</td>
<td>1 (3.1)</td>
<td>0.04</td>
</tr>
<tr>
<td>Alcohol abuse / dependence‡</td>
<td>2 (20.0)</td>
<td>1 (3.1)</td>
<td>0.14</td>
</tr>
<tr>
<td>SERT genotype</td>
<td></td>
<td></td>
<td>0.56</td>
</tr>
<tr>
<td>S'/S' (S/S, S/LG, Lc/Lc)</td>
<td>4 (40.0)</td>
<td>8 (25.0)</td>
<td></td>
</tr>
<tr>
<td>S'/L' (S/LA, Lc/La)</td>
<td>5 (50.0)</td>
<td>17 (53.1)</td>
<td></td>
</tr>
<tr>
<td>L'/L' (LA/LA)</td>
<td>1 (10.0)</td>
<td>7 (21.9)</td>
<td></td>
</tr>
<tr>
<td>SERT availability Bsl-scan (BPND)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midbrain</td>
<td>0.60 (0.24)</td>
<td>0.61 (0.18)</td>
<td>0.93</td>
</tr>
<tr>
<td>Diencephalon</td>
<td>1.16 (0.29)</td>
<td>1.15 (0.23)</td>
<td>0.89</td>
</tr>
</tbody>
</table>

* Responders defined as patients with ≥50% decrease in baseline HDRS$_{17}$-score
† p values for $\chi^2$ for categorical, Fisher’s exact test for dichotomous, and independent T-tests for continuous data
‡ cannabis, benzodiazepines
Results

Patients
From 177 patients assessed, 53 patients were excluded and 73 refused participation. Of the 51 patients included in the study, 7 patients dropped out due to paroxetine adverse effects (n= 5) or refusal of a second scan (n= 2), leaving 44 patients who were scanned at baseline and after six weeks of treatment. Of these 44 patients, 12 patients were treatment-responder after 6 weeks. Of responders, two SPECT scans could not be used for analysis due to technical reasons, leaving 42 patients for the final analysis. Benzodiazepines were prescribed in 11/42 (26%) patients.

At study-entry, no significant differences were found between responders and non-responders except for drug (cannabis/benzodiazepine) abuse or dependence with higher prevalence in responders compared with non-responders (Fisher’s Exact; p= 0.036; Table 8.1). None of the patients reported lifetime use of 3,4-methylenedioxy methamphetamine (MDMA).

SERT occupancy and clinical response
When SERT occupancy was stratified as poor (<50%), intermediate (51-79%) and high (≥80%), we found no significant relation between SERT occupancy and response-rate or the percentage decrease in HDRS<sub>17</sub> after 6 weeks of paroxetine treatment, neither in midbrain nor in diencephalon (p>0.05) (Table 8.2). Neither in linear nor in logistic regression models occupancy significantly predicted clinical response (p>0.05). The minimum occupancy rate at which response occurred in at least one patient was 29% in midbrain and 49% in diencephalon. We found no increase in response-rate for SERT occupancy >80%. Only midbrain SERT-occupancy was associated with PSC (p=0.02). PSC significantly predicted response (OR= 1.04 (95% CI= 1.00-1.09)). After exclusion of 2 potential non-adherent patients, the association of PSC with response became non-significant, other results did not change.

<table>
<thead>
<tr>
<th>Occupancy after 6 weeks paroxetine</th>
<th>Diencephalon n= 9</th>
<th>Midbrain n= 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean occupancy (%)</td>
<td>30.4 ±5.40</td>
<td>30.2 ±10.89</td>
</tr>
<tr>
<td>Mean PSC (µg/L)</td>
<td>18.2 ±6.27</td>
<td>18.2 ±6.27</td>
</tr>
<tr>
<td>Responders n (%)</td>
<td>4 (44.4)</td>
<td>4 (44.4)</td>
</tr>
<tr>
<td>% decrease HDRS&lt;sub&gt;17&lt;/sub&gt; †</td>
<td>25.7 ±4.66</td>
<td>26.1 ±6.88</td>
</tr>
<tr>
<td>Responders n (%)</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>% decrease HDRS&lt;sub&gt;17&lt;/sub&gt; †</td>
<td>28.9 ±7.43</td>
<td>30.2 ±10.89</td>
</tr>
<tr>
<td>Responders n (%)</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>% decrease HDRS&lt;sub&gt;17&lt;/sub&gt; †</td>
<td>29.4 ±7.25</td>
<td>30.2 ±10.89</td>
</tr>
<tr>
<td>Responders n (%)</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>% decrease HDRS&lt;sub&gt;17&lt;/sub&gt; †</td>
<td>30.2 ±10.89</td>
<td>30.2 ±10.89</td>
</tr>
</tbody>
</table>

Values represent means (±SEM) unless indicated different. N= 42, including 2 non-adherent patients. PSC = Paroxetine serum concentration * χ<sup>2</sup> for linear trend for response rates; ANOVA for % decrease † after 6 weeks of treatment

Modification of the relation between PSC, occupancy and response by genotype

PSC-occupancy curves in both the midbrain and diencephalon were curvilinear, with significantly different (non-stratified) curves for midbrain and diencephalon (determined by a decrease in AIC; figure available on request). The curves for responders versus non-responders were not significantly different in both brain regions. Stratification of the PSC-occupancy curves by SERT genotype indicated numerically higher OCC<sub>max</sub> and lower EC<sub>50</sub> in the L<sub>A</sub>/L<sub>A</sub> genotype (n= 8/42; 19%) compared with the S'/S' (n= 12/42; 29%) and S'/L<sub>A</sub> genotypes (n= 22/42; 52%), especially in midbrain (OCC<sub>max</sub> for L<sub>A</sub>/L<sub>A</sub>=99.6; p=0.188) but not in diencephalon (OCC<sub>max</sub> for L<sub>A</sub>/L<sub>A</sub>=79.3; p=1.0;
Figures 8.1A and B). The difference in AIC between one or separate curves indicated no significant improvement by including genotype in the models, both in midbrain (AIC increase 2.359) and diencephalon (AIC increase 4.037).

**SERT availability and SERT occupancy by genotype**

Mean pretreatment SERT availability and mean SERT occupancies after 6 weeks in midbrain or diencephalon did not significantly differ between genotypes (Table 8.3). Six of 8 patients (75%) with the L̄ₙ/Lₘ genotype reached midbrain occupancies ≥ 80% after 6 weeks compared with 12 of 31 patients (39%) with other genotypes (Fisher’s exact; p= 0.112).

After exclusion of 2 potentially non-adherent patients, mean SERT occupancies in midbrain were significantly different between genotypes (S′/S′ 80.9% ±5.80, S′/Lₘ 65.8% ±5.17, Lₘ/Lₘ 91.6% ±6.07; p= 0.017). Moreover, 6 of 7 patients (86%) with the Lₘ/Lₘ genotype reached midbrain occupancies ≥80% after 6 weeks compared with 12 of 30 patients (40%) with the other genotypes (Fisher’s exact; p= 0.042).

Figure 8.1. Paroxetine serum concentration and SERT occupancy by paroxetine, stratified by SERT genotype.

### Table 8.3. Ethnicity, clinical response and SERT occupancy by paroxetine stratified for SERT genotype.

<table>
<thead>
<tr>
<th>SERT genotype</th>
<th>S/S, S/Lₘ, Lₘ/Lₘ</th>
<th>S/Lₘ, Lₘ/Lₘ</th>
<th>Lₘ/Lₘ</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ethnicity - n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>7 (24.1)</td>
<td>15 (51.7)</td>
<td>7 (24.1)</td>
<td></td>
</tr>
<tr>
<td>Creole</td>
<td>3 (42.9)</td>
<td>3 (42.9)</td>
<td>1 (14.3)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>2 (33.3)</td>
<td>4 (66.7)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Clinical n = 12</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 6 % decrease HDRS₁₀₀</td>
<td>39.4±8.18</td>
<td>22.2±5.34</td>
<td>30.4±7.97</td>
<td>0.183</td>
</tr>
<tr>
<td><strong>Occupancy Diencephalon n = 22</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline BPₙ₀</td>
<td>101.7±3.94</td>
<td>116.2±6.01</td>
<td>122.9±8.84</td>
<td>0.353</td>
</tr>
<tr>
<td>Mean occupancy (%)</td>
<td>58.3±6.94</td>
<td>62.7±4.25</td>
<td>62.0±9.76</td>
<td>0.857</td>
</tr>
<tr>
<td>Occupancy ≥80% (%)</td>
<td>2 (16.7)</td>
<td>4 (18.2)</td>
<td>2 (25.0)</td>
<td>0.888</td>
</tr>
<tr>
<td><strong>Occupancy Midbrain n = 11</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline BPₙ₀</td>
<td>60.1±5.20</td>
<td>61.9±5.03</td>
<td>57.8±4.52</td>
<td>0.879</td>
</tr>
<tr>
<td>Mean occupancy (%)</td>
<td>73.6±9.04</td>
<td>65.8±5.17</td>
<td>81.3±11.56</td>
<td>0.381</td>
</tr>
<tr>
<td>Occupancy ≥80% (%)</td>
<td>5 (45.5)</td>
<td>7 (67.0)</td>
<td>6 (75.0)</td>
<td>0.159</td>
</tr>
</tbody>
</table>

Continuous values represent means (±SEM). N= 42, including 2 non-adherent patients.

* Post-hoc Lₘ/Lₘ vs. S/S, S/Lₘ, Lₘ/Lₘ, S/Lₘ, Lₘ/Lₘ Fisher’s exact p= 0.112

Panels represent PSC-occupancy curves for S′/S′ vs. S′/Lₘ vs. Lₘ/Lₘ in midbrain (A) and diencephalon (B).

Equation fitted: OCC₆_weeks = a + b(PSCᵇ)*

Differences between curves were not significant (lower AIC for 1 fitted curve vs. three fitted curves)

PSC= Paroxetine Serum Concentration, OCC₆_weeks = SERT occupancy after 6 weeks of paroxetine treatment.
SERT occupancy and clinical response by genotype

Despite no significant differences in mean proportional HDRS$_{17}$-decrease across genotypes (Table 8.3), we found a significant genotype*SERT occupancy interaction in diencephalon and midbrain (Figure 8.2). Higher diencephalon SERT occupancy was associated with larger proportional HDRS$_{17}$-decreases in L$_A$/L$_A$ carriers (ANOVA; $F_{2,5} = 27.643; p = 0.002$) (Figure 8.2B; corrected for age), and L$_A$/L$_A$ and L$_A$/S' carriers grouped together (ANOVA; $F_{1,28} = 8.380; p = 0.007$) (Figure 8.2D). The genotype*SERT occupancy interaction in diencephalon was significant (ANOVA; $F_{2,36} = 56.918; p < 0.001$). Introduction of baseline SERT availability in diencephalon (or midbrain) showed no significant improvement of the association between proportional HDRS$_{17}$ decrease and SERT occupancy (p>0.05). These results remained unchanged with absolute decrease in HDRS$_{17}$-scores instead of proportional HDRS$_{17}$-decreases. Performing the same analyses with the bi-allelic 5-HTTLPR classification found identical directions of the effects, but explained less of the variances (available on request).

**Figure 8.2.** SERT occupancy by paroxetine and proportional decrease in HDRS$_{17}$, stratified by SERT genotype.

Panels A and C: Midbrain, panels B and D: Diencephalon.

In panel A and C (midbrain) a significant relation between SERT occupancy and % decrease in HDRS$_{17}$ (determined by the slopes of regression lines) existed for the L$_A$/L$_A$ genotype when confounding by age was accounted for ($F_{1,28} = 56.92; p < 0.001$). In panels B (diencephalon) this relation was significant for the L$_A$/L$_A$ genotype ($F_{1,28} = 15.36; p = 0.008$), and confounded by age ($F_{2,5} = 27.64; p = 0.002$); for panel D this relation was significant for the S'/L$_A$ L$_A$/L$_A$ L$_A$/S' genotypes ($F_{1,28} = 8.380; p = 0.007$).

In panel A and C there was no significant interaction between the different genotype groups. In panel B and D a significant interaction existed between the genotype groups ($F_{2,36} = 5.273; p = 0.049$ and $F_{1,38} = 6.293; p = 0.017$ respectively).
Discussion

We quantified the relation between SERT occupancy in midbrain and diencephalon and clinical response after 6 weeks of paroxetine 20 mg/day. We found no significant relation between these occupancies and clinical response. Nor did we find significant differences in pretreatment SERT availability in patients with different 5-HTTLPR-genotypes. We found a significant modifying effect of the 5-HTTLPR genotype with respect to the association between SERT occupancy and the decrease in HDRS$_{17}$-score (proportional and absolute). In diencephalon, increased SERT occupancy was associated with larger decreases in HDRS$_{17}$ for L$_A$/L$_A$ and S'/L$_A$ carriers. In midbrain, increased SERT occupancy was associated with larger proportional decreases in HDRS$_{17}$ for L$_A$/L$_A$ carriers only. Age was a covariate in these associations. Unexpectedly, we found a trend of a modifying effect of the 5-HTTLPR genotype with respect to the association between paroxetine serum concentrations and occupancy.

Critique of methods

Some limitations might be addressed first. We observed a low 6 week response rate (24%), which may be attributable to 1) the inclusion of some patients with secondary co-morbid anxiety and/or substance abuse, who might show less and slower recovery,46 and 2) the relatively high pretreatment initial HDRS-scores. Nevertheless, the recent Sequential Treatment Alternatives to relieve Depression trial measured only 30% response rate after 6 weeks of citalopram as well.1

Second, we included patients who were ethnically heterogeneous (Table 8.3). This raises the possibility of spurious results secondary to population stratification, which can be a problem when analyzing non-functional markers which are in linkage disequilibrium with the functional variants. Since we have analyzed the relation between a functional DNA sequence variant and transporter occupancy we do not expect effects by the heterogeneity of our sample.

Third, as recruitment of MDD patients willing to undergo two SPECT scans is difficult, we could only successfully scan 42 patients twice. This led to small genotype-subgroups, and replication in larger samples will be necessary. Nevertheless, these 42 patients scanned twice while treated with the same antidepressant (and dose) is the largest patient sample to date, and allowed us to address the relation between SERT occupancy and clinical response, and its modification by the 5-HTTLPR-genotype for the first time. Furthermore, we had low dropout rates and good adherence.

Fourth, in this open label study we did not control for placebo response (estimated to be 30%). Placebo-responses may potentially obscure the relation between occupancy and response. Nevertheless, our major finding of a significant relation between SERT occupancy and proportional response in L$_A$/L$_A$-carriers was found despite such potential placebo-effects. Additionally, a placebo comparator would have estimated test-retest variability in SERT availability, but this variability is known to be much lower than the observed 60-70% SERT occupancy.

Fifth, we used $[^{123}I] \beta$-CIT for SPECT imaging, which is a non-selective radioligand, and also binds to dopamine transporters (DAT; e.g. midbrain substantia nigra).47 Nevertheless, uptake in midbrain and diencephalon is considered to reflect predominantly SERT,48 as these structures are rich of SERT relative to DAT. Therefore, although some additional DAT-binding might have concealed SERT occupancy, we think the change in $[^{123}I] \beta$-CIT-binding in diencephalon and midbrain mainly reflects SERT occupancy. Nevertheless, it would be challenging to replicate our study with selective radioligands like $[^{11}C]$DASB or $[^{123}I]$ADAM.

Finally, we did not further explore whether 5-HTTLPR (or other polymorphisms) also influenced transporter site kinetics. Rausch et al. showed that initial serotonin kinetic conditions of SERT in platelets could predict treatment response in MDD patients, which was modified by SERT polymorphism.49 Patients with an L-allele were more responsive to fluoxetine than patients with an S-allele, with the initial affinity constant $K_m$ and dose being significant covariates for treatment outcome. It is unlikely that $K_m$ is directly affected by the SERT promoter polymorphism that primarily regulates expression. Because the S and L variants actually represent 10 subtypes,50
other subtypes or an unknown other polymorphism could very well be associated with differences in $K_m$, which may better explain the inter-individual variation between PSC and SERT occupancy.\(^4\) Recently, Smeraldi et al.\(^{51}\) investigated the relation between clinical response to fluvoxamine and different L and S subtypes (identified by Nakamura et al.\(^{50}\)). They confirmed better responses in depressed patients bearing the L-allele, but also found significant differences in response among L-allele carriers according to the subtype of the L-allele, accounting for 0.6\% of the variance. Because these subtypes are relatively uncommon, it would be very interesting to specifically study these L-allele subtypes in future neuroimaging studies using a case-control design.

**SERT occupancy and clinical response**

Previous studies postulated that a SERT occupancy of at least 80\% would be associated with clinical response to SSRIs.\(^4\)\(^5\)\(^8\) Our data did not show significant associations between pretreatment SERT availability or SERT occupancy and clinical response to paroxetine, and neither suggested an 80\% SERT occupancy threshold with increased response rates. We could not replicate a relation between SERT occupancy and decrease of symptoms.\(^11\)\(^13\) Receiver operating characteristic curves relating SERT occupancy to response (available on request) showed that SERT occupancy had no diagnostic value in the prediction of response.

The relation between SERT occupancy and clinical response is modified by 5-HTTLPR genotype

We found that 5-HTTLPR-polymorphisms modify the relation between SERT occupancy and clinical response. Previous clinical studies and meta-analyses found superior treatment effects in L-allele carriers. Despite a low response rate in our sample, we found that in L\(_A\)/L\(_A\) carriers higher SERT occupancy was associated with more improvement. We put forward two possible explanations for this finding.

First, the L\(_A\)/L\(_A\) genotype is associated with a more than two-fold higher serotonin uptake in human lymphoblasts.\(^14\) If we assume that serotonergic cells maintain a certain tonus in neurotransmission, in L\(_A\)/L\(_A\) carriers more serotonin must be released as this is more effectively evacuated. When SERTs are blocked with paroxetine, this may increase serotonin neurotransmission more in L\(_A\)/L\(_A\) carriers than in other genotypes, which may result in larger postsynaptic effects.

Second, 5-HTTLPR also modifies the development and synaptic plasticity of neural networks critically involved in MDD.\(^14\) Our findings might thus point to differences in postsynaptic effects (evoked by increased serotonin neurotransmission) in neuronal networks that have been developed differently as a result of different SERT genotypes. Pezawas et al. showed that healthy subjects with an S/S polymorphism demonstrated relative uncoupling of the amygdala from the anterior perigenual cingulate.\(^26\) S/S carriers were also associated with increased anxiety traits, increased amygdala reactivity,\(^21\) decreased mood after tryptophan depletion\(^52\) and increased risk for MDD.\(^22\) In summary, the limbic-cortical network and the serotonergic innervations appear to be more flexible in non S/S carriers. Our results may be supportive of the hypothesis that in L\(_A\)/L\(_A\) carriers the significant association between higher SERT occupancy and increased reduction of symptoms could be indicative for a broader range for regulation of the serotonergic system. Higher SERT occupancy might then result in more effects of serotonergic antidepressants in L\(_A\)/L\(_A\) carriers.

Interestingly, Pollock et al. found quicker response in L/L carriers when treated with paroxetine for 12 weeks, but no difference when treated with the noradrenergic antidepressant nortriptyline.\(^53\) Additionally better treatment outcomes and fewer adverse events in S/S genotypes were found when treated with mirtazapine (a postsynaptic 5-HT\(_{2A}\), 5-HT\(_{2C}\)antagonist and noradrenergic agonist),\(^54\)\(^55\) which were contrary to the effects of paroxetine.\(^55\) Future pharmacogenomic studies should investigate whether non S'/S' carriers will benefit more from serotonergic drugs while S'/S' carriers may benefit from noradrenergic antidepressants. Furthermore, since the mechanisms of action of antidepressants may be influenced by polymorphisms of other genes, future research should investigate additive effects of multiple candidate gene polymorphism combinations.\(^56\)
Are paroxetine serum concentrations and SERT occupancy modified by 5-HTTLPR genotype?

Six of 8 patients (75%) with the L_A/L_A genotypereached midbrain occupancies ≥80% after 6 weeks compared with 12 of 31 patients (39%) with other genotypes, although this was statistically not significant (p = 0.112, Fisher’s exact). Which factor determines the maximum SERT occupancy remains unclear. 5-HTTLPR-polymorphisms primarily affect gene transcription, and as a consequence affect SERT expression. Therefore, we a priori did not expect differences in SERT occupancy for different genotypes, as occupancy is expected to be independent of available SERTs. As such, our results could represent an epiphenomenon: e.g. a different SNP in the SERT gene, in linkage disequilibrium with the 5-HTTLPR-polymorphism (e.g. rs2228673; www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=2228673) may have affected SERT occupancy. Contrary, the suggested modifying properties of the 5-HTTLPR-genotype for the relation between PSC and occupancy may also lack power, since the PSC-occupancy curves in diencephalon show a clear trend of modification by 5-HTTLPR. Therefore, this needs further exploration in larger samples.

A rationale for a modifying role of the 5-HTTLPR polymorphism in the occupancy of the SERT by paroxetine and clinical response could be postulated as follows. SERT imaging assesses the in-vivo binding to SERTs. After 6 weeks of treatment with paroxetine, the availability of SERTs to bind to the radiotracer is lower, presumably due to SERT blockade by the SSRI. However, it is not possible to assess whether in addition to blockade, secondary effects occur such as down-regulation of the SERT. This might also result in lowering of the availability of SERTs to bind with the radiotracer. Indeed, in rats down-regulation of SERTs after prolonged exposure to paroxetine has been reported.57-58 Said differently, we are unable to discriminate between direct (i.e., blockade of SERTs by paroxetine) and potential indirect pharmacological effects (i.e., down-regulation). Importantly, Benmansour et al., suggested that SSRI-induced down-regulation of the SERT may be a key component for the clinical response to SSRIs.57-58 Down-regulation of SERTs in rats was not caused by decreased gene transcription, but is presumably caused by increased (posttranslational) internalization of SERTs.57 Although speculative, the level of down-regulation of SERTs might be diverse for different 5-HTTLPR-genotypes, with faster or greater down-regulation in L_A/L_A versus other genotypes. This hypothesis could be studied in future imaging studies with antidepressant-exposed primates (having 5-HTTLPR-polymorphisms) in which abrupt discontinuation is possible.

Conclusion

In conclusion, we find that SERT genotype modifies the relation between SERT occupancy and clinical response, with more improvement at higher midbrain and diencephalon SERT occupancy in L_A/L_A carriers. Although this SERT promoter-polymorphism presumably does not influence SERT occupancy at given PSCs, our data may point to a more flexible serotonergic system in L_A/L_A carriers, which is more easily influenced by serotonergic antidepressants. This hypothesis could serve as a starting point for future pharmacogenetic studies. This might provide the necessary data to guide the choice of antidepressant treatment for individual patients, reducing the patient’s suffering, and lowering healthcare costs.

Acknowledgements

The authors wish to thank the patients in this study for their participation, and willingness to participate in the SPECT-study. We thank all participating general practitioners in the area of Amsterdam Oost and Zuidoost, Hoofddorp, Nieuw Vennep, and Abcoude for their inclusions and referrals for the study. Mrs. E. Miedema M.D. and Mrs. M.C. ten Doesschate M.D. were indispensable for their help in rating questionnaires. This study was financed by a grant from the Netherlands Organisation for Health Research and Development (ZonMw), program Mental Health, education of investigators in mental health to H.G. Ruhé (OOG; #100-002-002).
Conflicts of interest

None

References


Evidence why paroxetine dose-escalation is not effective in major depressive disorder: a randomized-controlled trial with assessment of serotonin transporter occupancy

Neuropsychopharmacology 2008; in press

Henricus G. Ruhé
Jan Booij
Henk C. v Weert
Johannes B. Reitsma
Eric J.F. Fransen
Martin C. Michel
Aart H. Schene

1 Program for Mood Disorders, Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
2 Department of Nuclear Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
3 Department of General Practice, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
4 Department of Clinical Epidemiology, Biostatistics & Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
5 Department of Pharmacy and Clinical Pharmacology, Onze Lieve Vrouwe Gasthuis, Amsterdam, the Netherlands
6 Department of Pharmacology and Pharmacotherapy, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
Abstract

Background
Dose-escalation is often used in depressed patients who fail to respond to standard doses of SSRIs, but clinical efficacy is equivocal.

Aim
To reassess the efficacy of paroxetine dose-escalation and quantify whether paroxetine dose-escalation increases occupancy of the serotonin transporter (SERT) more than placebo dose-escalation, in a randomized controlled trial.

Methods
We recruited 107 non-psychotic, unipolar depressed outpatients (18-70 yrs; Hamilton Depression Rating Scale (HDRS17) >18) from primary care and psychiatric outpatient departments. After 6 weeks open-label paroxetine 20 mg/day (T0), non-responding patients (HDRS17 decrease <50%; n= 60) were randomized to double-blind paroxetine (30-50 mg/day as tolerable) or placebo dose-escalation (paroxetine 20 mg/day + placebo). Patients were followed until 6 weeks after randomization (T1). Forty-nine patients, drug free at study-entry, underwent SPECT-scanning before treatment and were scanned repeatedly at T0 and T1. Paroxetine serum concentrations and SERT occupancy were determined at T0 and T1 (n= 32).

Results
We terminated the dose-escalation trial after an interim analysis. Thirty non-responding patients were randomized to paroxetine (46.7 ±5.5 mg/day), 27 to placebo dose-escalation. Response-rates were 10/30 (33.3%) and 10/27 (37.0%), respectively. Repeated measurement analyses showed no significant effect for treatment (p= 0.88, exceeding a priori stopping rules for futility [p>0.5]). Overall dropout was higher for placebo (26.7%) than paroxetine (3.3%; p= 0.03). Paroxetine dose-escalation increased paroxetine serum concentrations (p<0.001). SPECT measurements (12 patients randomized to paroxetine (46.9 ±4.8 mg) and 14 to placebo dose-escalation) showed no significant increase of midbrain SERT occupancy (2.5 ±26.4%, paroxetine; 3.1 ±25.8% placebo; p= 0.687) nor in diencephalon (p= 0.529).

Conclusions
Paroxetine dose-escalation in depressed patients has no clinical benefit over placebo dose-escalation. This is explained by the absence of significant increases of SERT occupancy by paroxetine dose-escalation, despite increased paroxetine serum concentrations.
(ISRCTN register nr. ISRCTN44111488)
Chapter 9

SERT occupancy by paroxetine dose-escalation in MDD

Introduction

Major depressive disorder (MDD) is often treated with antidepressants, particularly selective serotonin reuptake inhibitors (SSRIs). Unfortunately, response and remission rates are modest (30-50%), which require additional strategies to gain remission. Switching and augmentation have recently been evaluated. A third, and frequently applied option is dose-escalation, recommended in treatment guidelines and frequently used preceding other strategies. Only the recent NICE guideline is more reluctant in recommending dose-escalation. Although an individual patient may improve after dose-escalation, this could also represent a delayed drug response or reflect the natural course of the disease. Prolonged (up to 10 weeks), unaltered treatment with fluoxetine 20 mg/day improved the response rates of initial week 6 non-responders. Theoretically, the concept of dose-escalation assumes linear dose-response relationships which have not been proven for SSRIs. Therefore, the efficacy of dose-escalation of SSRIs has been questioned.

Previous studies did not show improved clinical effectiveness of dose-escalation, but had serious methodological weaknesses. All previous studies increased dosages probably too early (mostly after 3-4 weeks) and too abruptly, which may have obscured true dose-escalation effects by delayed effect of the standard doses and selective early drop-out of patients receiving true dose-escalation. Moreover, no study provided a rationale why dose-escalation was ineffective.

The primary molecular target of SSRIs is the serotonin transporter (SERT). Imaging techniques such as positron emission tomography (PET) and single photon emission computed tomography (SPECT) allow in-vivo labeling of SERT in the brain, which can be used to study their occupancy. To date, several imaging studies measured SERT occupancy after short or prolonged treatment with SSRIs (Table S9.1). Particularly, Meyer et al. showed 60-80% SERT occupancy after standard clinical doses of SSRIs, and demonstrated curvilinear dose-response relationships for SERT occupancy by SSRIs. However, high doses of SSRIs were rarely studied, and dose-escalation was never studied.

Taking into account previous methodological criticisms, and considering the molecular target of SSRIs, we have tested whether paroxetine dose-escalation increases SERT occupancy, and improves depressive symptoms more than placebo dose-escalation. We performed a 6 week, multicenter, randomized study in depressed patients not responding to 6 weeks of paroxetine at 20 mg/day. As a novel extension to previous clinical trials, and in order to elucidate the neurobiological basis for an expected lack of benefit of dose-escalation, we included a SPECT imaging approach. Herewith, we quantified whether paroxetine dose-escalation increased SERT occupancy more than placebo dose-escalation. This enabled us to relate clinical findings to the neurobiological correlate of SERT occupancy.

Methods

Participants

Following approval by the institutional ethical committee and written informed consent, we recruited outpatients (18-70 years) from primary care, our outpatient department, and public psychiatric settings between October 2003 and February 2007. Inclusion criteria were: MDD determined by the structured clinical interview for DSM-IV (SCID), and a Hamilton Depression Rating Scale (17 items; HDRS17) score above 18. All participants were drug-free or had undergone no more than one antidepressant treatment (other than paroxetine) at an effective dose for ≥ 6 weeks for the present MDD-episode. By the latter criterion, we avoided treatment resistance as potential bias for inefficacy of dose-escalation. Exclusion criteria, apart from pregnancy (or wish), were bipolar disorder, psychotic features, neurological cognitive impairments (i.e. dementia),
primary anxiety and/or substance abuse disorders and acute, severe suicidal ideation. Contrary, we allowed secondary co-morbid anxiety and/or substance abuse to increase applicability of our findings.

**Interventions**

Patients were treated by their referring physician or were referred to our outpatient department. After assessment at study-entry, all patients were treated open-label with paroxetine 20 mg/day for 6 weeks (see Figure 2.1). When severe adverse effects occurred, dosages were reduced to 10 mg/day and again increased to 20 mg/day after one week. We randomized all patients who did not achieve ≥50% decrease in HDRS17-score after 6 weeks, relative to study entry. They received a true paroxetine or a placebo dose-escalation added to paroxetine 20 mg/day. Dose-escalation was provided in blue capsules containing 10 mg paroxetine or placebo. Randomization was stratified for treatment setting (SPECT-group, outpatient department AMC, primary care, public psychiatry), gender and age. Within strata, we applied a minimization method to achieve a balanced distribution. We concealed allocation by using an independently operated computer program.

Dose-escalation consisted of incremental steps of one capsule every 5 days towards a maximum of 50 mg/day (20 mg + 3 capsules). Patients were allowed to increase at a slower pace (e.g. by 7 days) or stop further escalation (e.g. 20 mg + 2 capsules) according to adverse effects. No dosage adjustments were allowed during the last 3 weeks of the study. We checked adherence by pill-counts and anamnesis.

**Outcomes and measurements**

Primary clinical outcomes were HDRS17-scores, and the proportion of patients achieving response (≥50% decrease in HDRS17) or remission (HDRS17 ≤7). Secondary outcomes were total and specific (adverse effects / inefficacy) dropout rates, the Maier and Bech 6 item subscales of the HDRS17, the Inventory for Depressive Symptomatology self-rated (IDS-SR30) scores, the occurrence of adverse effects and health-related quality of life (MOS-SF36; physical and mental component scales standardized to a general Dutch population).

We administered questionnaires at study-entry, randomization (T0), and 6 weeks after randomization (T1). Depressive symptoms were also monitored at week 1, 2 and 4 using the Maier and Bech subscales and IDS-SR30 (see Figure 2.1). Three trained investigators administered clinician-rated questionnaires. Agreement between raters was good (intraclass correlation coefficient = 0.98). Raters and patients were blinded for treatment.

**Subgroup for SPECT imaging**

From all patients who entered the trial, we recruited patients who were drug-free (>4 weeks and ≥5 half-lives of a previous antidepressant) as potential candidates for SPECT imaging. These patients were asked to participate in the SPECT sub-study if their age was between 25-55 years to reduce variability in SERT measurements by age. Forty-nine patients could thus be recruited for a first SPECT scan. None of these patients reported past or present use of 3,4-methylenedioxymethamphetamine. We made a second scan in those patients who completed 6 weeks of paroxetine treatment (n= 44; including 12 responders), while only randomized non-responders (n= 32) were invited for a third scan at the end of the study. We treated SPECT patients at the AMC outpatient department. Medication was supplied in pillboxes.

**SPECT imaging and analysis**

We performed SPECT imaging at study-entry (baseline-scan), T0 and T1 (see Figure 2.1) between 2 to 10 pm according to previously described procedures. We made all scans 230 ±18 (SD) minutes after intravenous injection of approximately 100 MBq iodine-123-labeled 2β-carbomethoxy-3β-(4-iodophenyl)-tropane ([123I]β-CIT), when the radioligand is at equilibrium for SERT binding in brain areas expressing high densities of SERTs. To prevent thyroid uptake of [123I], all subjects received
oral potassium-iodide solution. We performed SPECT imaging using a 12-detector single slice brain-dedicated scanner (Neurofocus 810, Strichmann Medical Equipment; Cleveland, OH) with a full-width at half-maximum resolution of 6.5 mm, throughout the 20 cm field-of-view (http://www.neurophysics.com). Blood for paroxetine serum concentrations (PSC) was collected at T0 and T1 immediately before scanning. Serum was stored at -20° C until analysis. PSC were determined in May 2007 using a validated HPLC-MS/MS method (therapeutic range 10-75 µg/L; see appendix). The lower limit of quantification was 5 µg/L, the lower limit of detection was 0.3 µg/L.

After attenuation correction and reconstruction in 3D mode (http://www.neurophysics.com), we defined regions of interest (Rols) for midbrain, diencephalon and cerebellum by using validated templates (see Figure 2.3). One examiner, blinded for scan session (baseline-scan/T0/T1), positioned all Rols in two series. Intraclass correlation coefficients were >0.97 for all Rols. If the two series differed by >5%, scans were re-evaluated by a second investigator. In the analyses we averaged the counts for the two series.

Using activity in cerebellum as indicator of non-displaceable activity (non-specific binding and free radioactivity), we calculated specific to non-specific binding ratios per scan as

$$BP_{ND} = \frac{\text{Activity}_{ND} - \text{Activity}_{ND, baseline}}{\text{Activity}_{ND, baseline}}$$

BP_{ND} is proportional to transporter number under equilibrium conditions. In a different study, we found high reproducibility of SERT imaging with \[^{123}\text{I}\]\(\beta\)-CIT SPECT after repeated scanning of subjects, using the same camera and scanning-protocol (de Win et al., submitted). As primary outcomes, we calculated SERT occupancies at T0 or T1 relative to untreated baseline-scan SERT availability: $OCC_{T0 or T1} = \frac{BP_{ND,T0 or T1} - BP_{ND,baseline}}{BP_{ND,baseline}}$.

### Power and interim analysis

We performed a-priori power-calculations for two co-primary endpoints: (a) To detect a difference of ≥5-points in HDRS_{17} scores between paroxetine and placebo dose-escalation, while assuming a common standard deviation of 7 and using a one-tailed $\alpha = 0.025$ and $\beta = 0.05$, sample sizes of 60 per group were required; (b) For response rates (assumed to be 50% and 30% for paroxetine vs. placebo dose-escalation) a two-tailed $\alpha = 0.05$ and $\beta = 0.20$ required 110 participants per group. Because previous dose-escalation studies indicated no benefits relative to placebo dose-escalation, we planned an interim analysis after SPECT data for had been collected on at least 30 randomized patients in the SPECT subgroup. Stopping criteria, using the most informative continuous scores in a mixed model, were predetermined using the O'Brien and Fleming approach, were undisclosed while performing the interim analysis, and were $p<0.0026$ in case of superiority and $p>0.50$ for futility.

### Data analyses

Analyses were performed while blinded for treatment allocation. We based endpoint analyses on intention to treat (ITT), with last-observation carried forward (LOCF). To examine the effectiveness of paroxetine vs. placebo dose-escalation, we compared the proportion of patients with response, remission and drop-outs at the end of study using $\chi^2$ or Fisher’s exact test. We examined differences in mean continuous endpoints by ANCOVA with treatment as factor and value at randomization (T0) as covariate.

We used linear mixed models to assess differences in trends over time between groups in Maier, Bech and IDS-SR_{30} scores. Mean scores for these questionnaires were modeled as a function of the randomized group (paroxetine vs. placebo dose-escalation), score at randomization, and time since randomization (categorical, four levels). The interaction between time*group was added to the model to test whether trends over time were different between the two treatment groups. We used the Akaike Information Criteria to choose the best fitting variance/covariance structure (unstructured, compound symmetry or first-order auto-regressive) for each outcome parameter.

To examine changes in SERT occupancy between T0 and T1, we used ANCOVAs with treatment as factor, and SERT occupancy at T0 and age as covariates. In order to obtain maximum information of dose-escalation in these analyses, we excluded patients that were likely non-adherent to paroxetine at T0 or T1 (PSC < 5 µg/L). Thereafter, we plotted SERT occupancy against...
dose and PSC. We modeled dose-response in an $E_{\text{max}}$ model as $\text{OCC} = \frac{a \times \text{PSC}}{b + \text{PSC}}$, in which $a$ represents maximal SERT occupancy and $b$ the PSC with 50% SERT occupancy.\textsuperscript{18,20-22,27,28} We calculated $a$ and $b$ by fitting a nonlinear regression model that minimizes the sum of squares of the residuals. For quantification of differences in SERT occupancy between final responders and non-responders, we used ANCOVA models corrected for differences at T0, age and baseline-scan SERT availability in diencephalon.\textsuperscript{46} We performed all analyses in SPSS v15.0.1.1 (www.spss.com).

## Results

### Patient disposition

One-hundred and seven patients (mean age 43.8 ±9.8) started open-label paroxetine (Figure 9.1). The response-rate in the open phase was 27/107 (25.2%), and 60 non-responding patients were randomized for the double-blind phase. Randomization over the 2 treatment-arms resulted in comparable groups (Table 9.1). Fifty-one patients completed the 6 week randomization phase including 31 from the SPECT study. We obtained at least 1 post-randomization HDRS\textsubscript{17} score for 57 patients.

HDRS\textsubscript{17} scores at study-entry (-6 weeks) were comparable in T0 responders vs. non-responders (ANOVA, $F_{1,85} = 1.972$, $p = 0.164$). At T0, HDRS\textsubscript{17} scores (±SD) were 7.8 ±3.62 (66.6 ±14.3% decrease) in responders vs. 20.5 ±6.25 (17.7 ±20.3% decrease) in non-responders.

Figure 9.1. Recruitment and flow of participants.

* Three patients who refused dose-escalation after randomization, never ingested study drugs and refused further questionnaires, were excluded for endpoint analysis.

\textsuperscript{1} One SPECT patient dropped out early due to inefficacy, but for all SPECT-patients clinical data could be obtained. For SPECT analyses, 6 patients were excluded: 1 patient missed the T1 scan (placebo dose-escalation), 3 patients were likely non-adherent at T0 (paroxetine serum concentration <5µg/l; all paroxetine dose-escalation), and 2 were likely non-adherent at T1 (1 paroxetine dose-escalation, 1 placebo dose-escalation).
Table 9.1. Characteristics of non-responding MDD patients after 6 weeks of open treatment with paroxetine 20 mg/day (trial population).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>All patients (n= 60)</th>
<th>Paroxetine DE (n= 30)</th>
<th>Placebo DE (n= 30)</th>
<th>SPECT subgroup (n= 32)</th>
<th>Paroxetine DE (n= 16)</th>
<th>Placebo DE (n= 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Marital status - n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single (never married)</td>
<td>15 (51.7)</td>
<td>12 (40.0)</td>
<td>6 (37.5)</td>
<td>5 (31.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>8 (27.6)</td>
<td>6 (20.0)</td>
<td>7 (43.8)</td>
<td>4 (25.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divorced</td>
<td>4 (13.8)</td>
<td>11 (36.7)</td>
<td>2 (12.5)</td>
<td>7 (43.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>2 (6.9)</td>
<td>1 (3.3)</td>
<td>1 (6.3)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Educational level - n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>6 (20.0)</td>
<td>11 (36.7)</td>
<td>2 (12.5)</td>
<td>3 (18.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate</td>
<td>19 (63.3)</td>
<td>15 (50.0)</td>
<td>10 (62.5)</td>
<td>10 (62.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>5 (16.7)</td>
<td>4 (13.3)</td>
<td>4 (25.0)</td>
<td>3 (18.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>6 (20.0)</td>
<td>11 (36.7)</td>
<td>4 (25.0)</td>
<td>4 (25.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Income €/month (median; 25 &amp; 75 quartiles)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoking - n (%)</td>
<td>13 (43.3)</td>
<td>16 (53.3)</td>
<td>6 (37.5)</td>
<td>11 (68.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol use: n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 2 Units/week</td>
<td>20 (66.7)</td>
<td>21 (70.0)</td>
<td>10 (62.5)</td>
<td>11 (68.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-7 Units/ week</td>
<td>6 (20.0)</td>
<td>7 (23.3)</td>
<td>4 (25.0)</td>
<td>5 (31.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-21 Units/ week</td>
<td>2 (6.7)</td>
<td>1 (3.3)</td>
<td>1 (6.3)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;22 Units/ week</td>
<td>2 (6.7)</td>
<td>1 (3.3)</td>
<td>1 (6.3)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Race - n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>17 (56.7)</td>
<td>19 (63.3)</td>
<td>9 (56.2)</td>
<td>13 (81.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creole</td>
<td>4 (13.3)</td>
<td>7 (23.3)</td>
<td>2 (12.5)</td>
<td>3 (18.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>6 (20.0)</td>
<td>1 (3.3)</td>
<td>3 (18.8)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3 (10.0)</td>
<td>3 (10.0)</td>
<td>2 (12.5)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MDD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDRS &lt; 17 at study-entry (-6 weeks)</td>
<td>24.5 ±4.7</td>
<td>25.5 ±5.0</td>
<td>25.6 ±5.0</td>
<td>24.4 ±4.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDRS &lt; 17 (To)*</td>
<td>20.1 ±6.6</td>
<td>1.0 ±5.9</td>
<td>21.3 ±7.4</td>
<td>19.3 ±5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maier at (To)*</td>
<td>10.0 ±3.0</td>
<td>10.5 ±3.1</td>
<td>10.4 ±3.6</td>
<td>10.1 ±2.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bech (To)*</td>
<td>10.5 ±2.9</td>
<td>11.3 ±3.0</td>
<td>11.0 ±3.4</td>
<td>10.8 ±3.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDS-SR18 &lt; 17 (To)*</td>
<td>38.1 ±21.3</td>
<td>40.8 ±11.2</td>
<td>40.3 ±11.6</td>
<td>39.9 ±11.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First episode - n (%)</td>
<td>17 (56.7)</td>
<td>21 (70.0)</td>
<td>9 (56.3)</td>
<td>9 (56.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No of episodes</td>
<td>1.6 ±0.8</td>
<td>1.7 ±1.8</td>
<td>1.6 ±0.8</td>
<td>2.3 ±2.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug-naïve - n (%)</td>
<td>24 (80.0)</td>
<td>19 (63.3)</td>
<td>14 (87.5)</td>
<td>8 (50.0)†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used AD in current episode - n (%)</td>
<td>3 (10.0)</td>
<td>5 (16.7)</td>
<td>1 (6.3)</td>
<td>2 (12.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melancholic - n (%)</td>
<td>23 (76.7)</td>
<td>19 (63.3)</td>
<td>12 (75.0)</td>
<td>12 (75.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of episode: n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5 months</td>
<td>9 (30.0)</td>
<td>5 (16.7)</td>
<td>7 (43.8)</td>
<td>3 (18.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duration 5 months – 2 years</td>
<td>19 (63.3)</td>
<td>22 (73.3)</td>
<td>7 (43.8)</td>
<td>11 (68.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duration &gt; 2 years</td>
<td>2 (6.7)</td>
<td>3 (10.0)</td>
<td>2 (12.5)</td>
<td>2 (12.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of first episode (years)</td>
<td>37.1 ±9.1</td>
<td>38.1 ±11.8</td>
<td>38.4 ±10.8</td>
<td>34.4 ±10.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Co-morbidity – n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>5 (16.7)</td>
<td>7 (23.3)</td>
<td>0</td>
<td>4 (25.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysthymia</td>
<td>2 (6.7)</td>
<td>0</td>
<td>1 (6.3)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug (alcohol, cannabis, benzodiazepines) abuse/</td>
<td>2 (6.7)</td>
<td>1 (3.3)</td>
<td>1 (6.3)</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dependence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MOS-SF36</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical*</td>
<td>41.2 ±2.9</td>
<td>40.3 ±10.7</td>
<td>41.2 ±8.6</td>
<td>43.9 ±9.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental*</td>
<td>26.1 ±8.4</td>
<td>25.7 ±10.2</td>
<td>25.3 ±5.9</td>
<td>22.3 ±6.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SERT availability baseline-scan</strong> (BPND)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midbrain</td>
<td>N/A</td>
<td>N/A</td>
<td>0.553 ±0.119</td>
<td>0.657 ±0.217</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diencephalon</td>
<td>N/A</td>
<td>N/A</td>
<td>1.157 ±0.226</td>
<td>1.134 ±0.247</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SERT occupancy (% of BPND in Bl-scan)</strong>†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midbrain</td>
<td>N/A</td>
<td>N/A</td>
<td>73.2±17.1</td>
<td>82.2±18.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diencephalon</td>
<td>N/A</td>
<td>N/A</td>
<td>63.8 ±15.4</td>
<td>70.3 ±12.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Numbers represent means (±standard deviation) unless specified otherwise. * at randomization (To). † sign. difference between SPECT patients randomized to conditions (Fisher's exact p = 0.026).

† n = 29; excluding three patients who were likely nonadherent at To (PSC <5 μg/l). BPND = Binding Potential (non-displaceable), DE = dose-escalation, MDD = Major Depressive Disorder, PSC = Paroxetine serum concentration
Clinical effectiveness of paroxetine vs. placebo dose-escalation

During dose-escalation (To-T1), 1, 8 and 21 patients reached final doses of 30, 40 and 50 mg/day respectively. The placebo group escalated to a comparable number of capsules ($\chi^2 = 0.895$, df= 2, $p= 0.639$). Adherence based upon pill-counts was comparable between both groups (Fisher’s exact, $p= 0.492$).

Paroxetine dose-escalation did not yield better outcomes in depression severity and health-related quality of life compared to placebo dose-escalation (ITT; Table 9.2). The robustness of this finding was confirmed in the longitudinal analysis (mixed model). Changes over time in the Maier subscale and IDS-SR$_{30}$-scores (Figure 9.2), and Bech subscale-scores (available on request), were comparable between the two groups. Overall dropout was higher with placebo (26.7%) than with paroxetine dose-escalation (3.3%; $p= 0.03$). Paroxetine dose-escalation had significantly more adverse effects than placebo dose-escalation, but this did not result in higher discontinuation rates due to adverse effects (Table 59.2). Instead, adverse effects by paroxetine dose-escalation moderately decreased over time, suggestive of habituation.

Table 9.2. Depression and health-related quality of life scores after 6 weeks paroxetine vs. placebo dose-escalation (T1); all patients and SPECT subgroup.

<table>
<thead>
<tr>
<th></th>
<th>All patients (n = 57)</th>
<th>SPECT subgroup (n = 32)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Paroxetine DE (n = 30)</td>
<td>Placebo DE (n = 27)</td>
</tr>
<tr>
<td>Mean dosage mg/day</td>
<td>46.7 ±1.00</td>
<td>46.9 ±1.20</td>
</tr>
<tr>
<td>HDRS$_{17}$</td>
<td>16.1 ±1.22</td>
<td>15.3 ±1.28</td>
</tr>
<tr>
<td>Maier subscale</td>
<td>7.5 ±0.61</td>
<td>7.5 ±0.64</td>
</tr>
<tr>
<td>Bech subscale</td>
<td>8.1 ±0.63</td>
<td>8.1 ±0.66</td>
</tr>
<tr>
<td>Response $^*$ - n (%)</td>
<td>10 (33.3)</td>
<td>10 (37.0)</td>
</tr>
<tr>
<td>Remission $^*$ - n (%)</td>
<td>4 (13.3)</td>
<td>2 (7.4)</td>
</tr>
<tr>
<td>IDS-SR$_{30}$</td>
<td>34.8 ±1.83</td>
<td>32.5 ±2.05</td>
</tr>
<tr>
<td>MOS-SF36$^*$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>41.8 ±0.93</td>
<td>42.4 ±1.06</td>
</tr>
<tr>
<td>Mental</td>
<td>29.6 ±1.37</td>
<td>27.3 ±1.57</td>
</tr>
</tbody>
</table>

Scores at endpoint of the study are based on intention to treat, with last observation carried forward for early drop-outs. Values are means ±standard error, corrected for (mean) scores at randomization (To) (ANCOVA).

$^*$ ≥50% decrease in HDRS$_{17}$ with baseline score (6 weeks) as reference; Fisher’s exact test.$^1$ HDRS$_{17}$ ≤7; Fisher’s exact test.$^1$

Due to missing values: n = 29 paroxetine DE and n = 23 placebo DE for all patients.$^1$ Due to missing values: n = 39 paroxetine DE and n = 22 placebo DE for all patients. DE = dose-escalation.

Figure 9.2. Changes over time in Maier and IDS-SR scores after randomization.

Points represent mean Maier (A) and IDS-SR (B) scores (±SEM) adjusted for scores at randomization (To) for paroxetine (n= 30) and placebo (n= 27) dose-escalation. Mixed model analysis (Maier: n= 57, IDS-SR: n= 53): overall difference between paroxetine vs. placebo dose-escalation for Maier scores $F_{101,460} = 0.295$, p= 0.880, and for IDS-SR $F_{45,159} = 1.516; p= 0.213$. Due to missing values: n = 29 paroxetine DE and n = 23 placebo DE for all patients.$^1$

Mean IDS-SR score at week 4 differed significantly in favor of placebo dose-escalation ($t_{49,161} = 2.11; p= 0.040$).

DE = dose-escalation.
SERT occupancy and clinical response

Of 32 randomized patients in the SPECT subgroup, only 3 (9%) previously used mirtazapine or fluoxetine in the current episode (1 patient stopped mirtazapine 4 weeks prior to scanning, others stopped <2 months before scanning). One patient missed the T1 scan. Based upon PSC at randomization or T1, five patients in the SPECT-subgroup (4 with paroxetine and 1 with placebo dose-escalation; Fisher’s Exact: p = 0.172; see Figure 9.3) with PSC <5 µg/l were considered non-adherent, despite adherence according to pill-counts. We excluded these five patients for analyses of changes in SERT occupancy after dose-escalation, leaving 26 T1 scans analyzable for diencephalon occupancies after true or placebo dose-escalation. Paroxetine dose-escalation increased mean PSC from 36.2 to 154.3 µg/l (paired t-test; p<0.001), while mean PSC in placebo dose-escalation remained unchanged (Table S9.3). At randomization (T0), mean SERT occupancies (±SEM) for the paroxetine dose-escalation group were 76.2 ±4.70% in midbrain and 64.3 ±4.60% in diencephalon. For the placebo dose-escalation group these were 84.6 ±4.95% and 72.2 ±3.08%, respectively. Neither paroxetine nor placebo dose-escalation significantly increased SERT occupancy further (Figure 9.3A). Plotting PSC vs. SERT occupancy showed that PSCs > 50 µg/l were not associated with further increases of SERT occupancy in midbrain or diencephalon (Figure 9.3BC). Furthermore, individual changes in PSC (T0 to T1) were not significantly associated with changes in occupancy in midbrain (F1,24 = 0.101; p = 0.754) and diencephalon (F1,37 = 1.332; p = 0.259; Figure S9.1).

Figure 9.3. SERT occupancy during randomized dose-escalation of paroxetine. Changes over time and relation with paroxetine serum concentration.

A. Mean SERT occupancy (±SEM) for paroxetine dose-escalation and placebo dose-escalation at randomization (T0) and after 6 weeks of dose-escalation (T1). SERT occupancy was calculated as percentage of initial available SERTs (expressed as BPsD) at baseline (6-weeks) scans (see text). Changes in SERT occupancy between T0 and T1 for paroxetine dose-escalation and placebo dose-escalation were non-significant in ANCOVA models correcting for age and differences in T0 SERT occupancy (MIDBRAIN: F1,21 = 0.167; p = 0.687; DIENCEPHALON: F1,22 = 0.409; p = 0.529). For 1 patient insufficient midbrain was scanned at study entry to compute subsequent SERT occupancies.

B & C. Data for randomized patients used from both T0 (open circles) and T1 (diamonds). Dose - occupancy relationships are modeled as OCC = a + b*PSC. For Midbrain (B): a = 96.0 ±4.03 (SE), b = 2.65 ±1.39 (n= 30 at T0 and n= 27 at T1); for Diencephalon (C): a = 73.3 ±2.28, b = 3.92 ±1.07 (n= 32 at T0 and n= 29 at T1). Dashed lines represent 95% confidence interval of fitting.

DE = dose-escalation, OCC = occupancy, PAR = paroxetine, PLAC = placebo. PSC = paroxetine serum concentration.
We explored whether SERT occupancy was related to clinical response at T1 irrespective of paroxetine or placebo dose-escalation. Responders at T1 (n=12) had numerically higher SERT occupancy ($\pm$SEM) in midbrain (91.2 $\pm$5.8\%) and diencephalon (69.2 $\pm$2.8\%) than non-responders (n=14; 77.8 $\pm$5.1\% and 63.8 $\pm$2.6\%, respectively; ANCOVA: p=0.107 and p=0.178). These models used baseline-scan diencephalon SERT availability as covariate as this accounted for a major part of variance ($F_{1,21}=4.831$, p=0.039 and $F_{1,22}=10.407$, p=0.004 respectively). Contrary, T1 SERT occupancy in midbrain or diencephalon did not significantly predict the percentage decrease in HDRS$_{17}$ in linear regression, or response status in logistic regression (neither when corrected for baseline-scan SERT availability in diencephalon or age).

**Discussion**

In this randomized trial we examined clinical effectiveness of dose-escalation in MDD patients, who were non-responders to 6 weeks of 20 mg/day paroxetine, and explored potential underlying mechanisms. Despite markedly increased drug exposure, paroxetine dose-escalation to 30-50 mg/day did not improve depressive symptoms more than placebo dose-escalation, but was associated with more adverse effects. Concomitantly, increased paroxetine serum concentrations were not associated with substantially greater SERT occupancy, indicating that standard paroxetine doses (20 mg/day) already resulted in maximum SERT occupancy.

**Clinical outcomes of dose-escalation in non-responders**

The dose-response relationship for paroxetine was previously examined in fixed dose, parallel group designs. Twenty mg/day and higher paroxetine doses yielded similar clinical improvements.\(^7\) Similar findings were reported from parallel-group, fixed dose studies of other SSRIs.\(^8\) Accordingly, the usefulness of dose-escalation in non-responders to paroxetine and other SSRIs has been questioned.\(^8-10\) However, the underlying studies had methodological shortcomings, and dose-escalation remains a recommended standard approach for non-responders.\(^1;2\)

If dose-escalation is applied too early (before week 6), randomization of ‘late responders’ will likely dilute the difference between true and placebo dose-escalation, resulting in potentially false negative findings. While one study reported randomized dose-escalation of sertraline after 6 weeks of treatment, this was compromised by a non-randomized dose-increase 2 weeks prior to randomization.\(^17\)

The benefits of dose-escalation might be under-estimated if actively-treated patients drop-out early due to adverse effects\(^34\) or become more non-adherent. Our schedule for dose-escalation did not increase drop-out. Hypothetically, patients receiving a paroxetine dose-escalation might have interpreted the increased level of adverse effects as subjective clue of greater drug effects, encouraging them to persevere in the trial. At first sight a misbalance in adherence is suggested with 4 patients with paroxetine vs. 1 patient with placebo dose-escalation having low PSCs. However, this was not due to dose-escalation. As mentioned in Figure 9.1, in three patients, low PSCs at T0 already classified them as likely non-adherent, with only 1 paroxetine vs 1 placebo dose-escalation patient becoming likely non-adherent during dose-escalation (T0-T1). Therefore, we think that neither adverse effects nor non-adherence account for the observed inefficacy of dose-escalation.

Thus, the present study overcomes methodological limitations of previous SSRI dose-escalation studies by using a randomized, placebo-controlled, double-blind dose-escalation in non-responders to 6 weeks treatment with a standard dose of paroxetine. We also avoided treatment resistance as a factor for inefficacy of dose-escalation by inclusion of patients who received no more than one effective antidepressant trial for the current episode. Under these conditions paroxetine was not superior to placebo in dose-escalation. Moreover, our study was
more inclusive than most previous ones and thus may be more applicable to ‘real-world’ first-line antidepressant treatment. This may also explain why we observed lower response and remission rates than previous studies.

**Neurobiological effects of dose-escalation**

Our pharmacokinetic and imaging measurements were designed to explore why paroxetine dose-escalation would or would not improve treatment outcomes. Our imaging of SERT occupancy bypasses potential bias by inclusion of patients with ultrarapid drug-metabolism, which is often causally linked to non-response. Hypothetically the clinical selection of non-responders eligible for dose-escalation, might represent a selection of patients not reaching high levels of SERT occupancy.

Comparing SERT occupancies of different SSRI doses faces several methodological challenges. Firstly, the assessment of occupancy requires knowledge on the available number of SERTs. Due to inter-individual differences in available SERTs, only assessments with individual drug-free baseline-scans yield reliable data. Secondly, a given drug dose may yield a range of serum concentrations due to inter-patient pharmacokinetic differences. Hence, associations based upon serum concentrations are more reliable than those based upon administered dose. Finally, intra-individual comparisons of occupancy changes following dose-escalation are more powerful than those with historic data.

Against this background, Voineskos et al. recently reported high SERT occupancies in striatum (85%), thalamus (79%), and midbrain (98%) in 12 depressed patients exposed to > 4 weeks of venlafaxine 225-450 mg/day, sertraline 150-200 mg/day or citalopram 60-80 mg/day in a [11C]DASB PET study. They concluded that high doses significantly increased occupancy compared to an average of 80% SERT occupancy determined in previous studies with standard SSRI doses, which would favor the concept of dose-escalation. However, they did not determine occupancy relative to baseline scans of the same patients without medication, nor at standard doses. We performed drug-free study-entry scans, in addition, >90% of patients did not use antidepressants for the current episode of MDD. Furthermore, low therapeutic dosages of several SSRIs also yielded high SERT occupancy in most studies. The present study resolves this controversy by showing that 4-fold increases of PSC upon paroxetine dose-escalation did not significantly increase SERT occupancy. This offers an explanation for our findings: SERT occupancy is limited by a ceiling effect. A PSC achieved with a 20 mg/day paroxetine dose is sufficient to yield maximum SERT occupancy (Figures 9.3B,C). If low doses already yield maximum SERT occupancy, dose-escalation cannot be expected to increase treatment efficacy, which is in line with our clinical findings. Our results do not necessarily challenge the relationships between dose, SERT occupancy and clinical response but rather suggest that these relationships exist mainly at low and sub-therapeutic doses. Furthermore, the relationship between SERT occupancy and response might be confounded by other factors such as SERT gene polymorphisms.

In a recent study Owens et al. showed increased SERT occupancy with increasing paroxetine CR doses (12.5-75 mg/day) in an ex-vivo model using human transporter transfected cells, which might be at odds with our findings. However, validation of this ex-vivo method (in cultured cells) with concomitant in-vivo SPECT or PET SERT occupancy (the gold standard) is not yet available. Additionally, Zitterl et al. found a significant relation between SERT occupancy and clinical response in obsessive compulsive disorder treated with clomipramine (150 mg/day), but did not study the effects of dose-escalation in their study. Therefore, our study optimally quantifies the neurobiological effects of dose-escalation of antidepressants in patients.

**Critique of methods**

For logistic reasons, we used [123I]β-CIT for SPECT imaging, which is a non-selective radioligand, and also binds to dopamine transporters (DAT; e.g. substantia nigra) and norepinephrine transporter (NET; e.g. locus coeruleus). Furthermore, imaging studies indicated increased striatal DAT binding after treatment with paroxetine, especially when the occipital cortex was used as a reference. Nevertheless, uptake in midbrain and diencephalon is considered to reflect
predominantly SERT, as these structures are rich of SERT relative to DAT and NET. Therefore, although this non-selectivity might have concealed changes in SERT occupancies due to additional DAT or NET binding, we think our findings in diencephalon and midbrain mainly reflect SERT occupancy. Nevertheless, it would be challenging to replicate our study using a selective ligand for SERT like $[^{11}C]$DASB for PET or $[^{123}I]$ADAM for SPECT imaging.

Our study did not investigate secondary effects of paroxetine. Many adaptive pre- and post-synaptic effects of chronic administration of SSRIs have been documented, including neuroadaptive alterations in serotonin receptors and intracellular signalling pathways, as well as time-dependent effects on neurogenesis. These hypothetical additional effects of dose-escalation remain to be investigated. Nevertheless, neither the results of our trial, nor the findings in previous randomized controlled trials indicate that dose-escalation is an efficacious strategy for SSRI non-responders in MDD.

The fourfold increase of PSC after dose-escalation from 20 to 50 mg/day may question adherence of patients in the open phase of the study. However, paroxetine inhibits the cytochrome P$_{450}$ enzyme 2D6, also responsible for the metabolism of paroxetine. Therefore nonlinear increases of PSC reflect normal paroxetine pharmacokinetics.

Our study was discontinued after an interim analysis with relatively small patient numbers. However, the criteria for premature trial termination regarding futility had been pre-specified, making it highly unlikely that we overlooked clinically relevant differences. Moreover, the neurobiological parts of our study provide a rationale why even with much larger patient numbers no substantially different outcome can be expected. On the other hand, premature stopping reduced the power to examine whether subgroups of patients were more responsive to dose-escalation.

The present study was not designed to test the efficacy of paroxetine per se, as this is well established in patients with severe MDD (HDRS$_{17}>18$). Therefore, we did not include a pure placebo arm. Rather we investigated dose-escalation, and accordingly only included placebo during dose-escalation. This approach is similar to e.g. the STAR*D project, which interestingly reported similar response rates for open treatment with citalopram.

Conclusion

Previous studies had failed to demonstrate a clinical benefit of dose-escalation by SSRIs, but had methodological limitations. Addressing those limitations, our trial replicates that dose-escalation of paroxetine above the 20 mg/day standard dose has no additional clinical benefit. As a novel extension, we revealed the underlying neurobiological mechanism for this inefficacy: maximum SERT occupancy was already reached with the standard dose. Similarly high SERT occupancies reported with low doses of other SSRIs suggest that our conclusion may be applicable to the entire drug class. However, this does not exclude that dose-escalation has clinical benefits for antidepressants with additional molecular targets, e.g. the norepinephrine transporter, such as venlafaxine. This drug has shown dose-dependency of the clinical response in fixed-dose studies.

If dose-escalation is not promising for paroxetine and presumably other SSRIs, two clinical options remain for the treatment of non-responders to standard doses. These are either continuation of treatment until 10 weeks while waiting for a potential delayed response, or a change to a different and potentially more effective treatment strategy. Both strategies will further improve response-rates, but studies directly comparing these strategies have not yet been performed.
Acknowledgements

The authors wish to thank the patients in this study for their participation, and especially thank the patients that were willing to participate in the SPECT-study. We thank all participating general practitioners in the area of Amsterdam Oost and Zuidoost, Hoofddorp, Nieuw Vennep, and Abcoude for their inclusions and referrals for the study. Mrs. E. Miedema M.D. and Mrs. M.C. ten Doesschate M.D. were indispensable for their help in rating questionnaires. Mrs M. Haages managed randomization and maintained blinding. This study was financed by a grant from the Netherlands Organisation for Health Research and Development (ZonMw), program Mental Health, education of investigators in mental health to H.G. Ruhe (OOG: #100-002-002). We especially thank Professor M.E. Thase M.D. for his constructive comments on a previous version of the manuscript.

Conflicts of interest

All authors reported no biomedical financial interests or potential conflicts of interest.

References


Supplementary appendix: methods of paroxetine serum concentrations and results.

Paroxetine serum concentrations

PSC were determined in May 2007 using a previously validated HPLC-MS/MS method. A Thermo Finnigan (San Jose, CA, USA) Surveyor HPLC system was used, equipped with a Surveyor autosampler. Separation was carried out on a Agilent Eclipse XDB-CN column, 100 x 2.1 mm, particle size 3.5 µm (Agilent, Amstelveen, The Netherlands). A mixture of 60% (v/v) 2 mM ammonium acetate (pH 3.2) and 40% (v/v) acetonitrile was used as mobile phase. Paroxetine was extracted from serum by liquid-liquid extraction. Briefly, 500 µL of serum samples were mixed with sodium carbonate solution (0.5M). An internal standard solution (disopyramide 20 µg/l) was added and the mixture was vortexed for 5 seconds. Thereafter, 1-chlorobutane/ethylacetate mixture (60/40; v/v) was added and vortexed for 10 minutes. After centrifugation for 10 minutes (3000g), the upper layer was transferred to a second vial, and evaporated under a nitrogen atmosphere at 40 °C. After addition of methanol the solution was vortexed and transferred onto the HPLC-MS/MS system. Selected reaction monitoring (SRM) was used for drug quantification (SRM paroxetine 330.1/192.0 and SRM disopyramide 340.2/239.0). Calibration curves were constructed by adding known amounts of paroxetine to blank serum (5-75 µg/L range). The lower limit of quantification was 5 µg/L, which was 50% below the lower end of the therapeutic range of paroxetine in serum (10-75 µg/L range). The lower limit of detection was 0.3 µg/L.
Table S9.1. Studies of SERT occupancy after treatment with serotonergic antidepressants

<table>
<thead>
<tr>
<th>Author &amp; Date (reference)</th>
<th>Population (mean age)</th>
<th>N</th>
<th>Follow-up</th>
<th>Design Intervention / control</th>
<th>SERT Imaging and ROI</th>
<th>Results</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catafau 2006*28</td>
<td>Drug-free MDD pts (36 ±11) M + F</td>
<td>10</td>
<td>Bsl, 4-6 w</td>
<td>Paroxetine 20 mg</td>
<td>[123I]-ADAM SPECT, 4h MID, Thal, Striatum, Cer</td>
<td>In 9 patients a 2nd scan was made. Occupancies were: MID 66.4 ± 9.5%, Thal 63.0 ± 9.6%, Striat 61.3 ± 10.8%</td>
<td>Considerable variation in occupancy observed. Authors suggest that SERT occupancy is probably not related to SSRI response.</td>
</tr>
<tr>
<td>Cavanagh 2006*31</td>
<td>MDD pts treated</td>
<td>24</td>
<td>N/A mean AD use 26.3-12.8m</td>
<td>Variety of drugs used: monotherapy (n=17): venlafaxine (75-300 mg), SSRIs (20-60 mg), tricyclic (150 mg), Mirtazapine (30 mg); combination (n=7) of 2 antidepressants, addition of lithium, valproate, carbamazepine, T3 or antipsychotics; dosages unchanged for 22 weeks</td>
<td>[123I]-β-CIT SPECT, 3.3h BasG, Dienc, Occ</td>
<td>No occupancy percentages available. No significant difference in SERT residual activity between responders and non-responders. Wide range of SERT availability</td>
<td>Patients were not scanned before treatment (bsl), instead a comparison of binding potentials was made. Highly heterogeneous group considering drug treatment.</td>
</tr>
<tr>
<td>Erlandsson 2005*23</td>
<td>Healthy volunteers (32; 25-52) M</td>
<td>16</td>
<td>Bsl, 2-7d</td>
<td>Citalopram at different dosages (10-60 mg) for different durations (2-7 days). Volunteers were randomized to one of 7 different dosing regimens</td>
<td>[123I]-ADAM SPECT, 4h MID, Thal, Cer</td>
<td>No mean occupancy for separate dosing regimens given. Maximum MID occupancy 84%</td>
<td>Study aims at finding the best time frame (sigle scan protocol) for scanning with [123I]-ADAM. Considerable variation in relation occupancy and blood-levels of citalopram observed.</td>
</tr>
<tr>
<td>Herold 2006*29</td>
<td>Drug-free MDD pts (42 ±12) M + F</td>
<td>21</td>
<td>Bsl, 7d</td>
<td>Citalopram 10 mg, scans were made 6-7 hours after last application of oral dose</td>
<td>[123I]-ADAM SPECT, 4h MID, Cer</td>
<td>In 13 patients a 2nd scan was made. Mean MID occupancy was 61% (range 37-88%)</td>
<td>Considerable variation in occupancy observed. No correlation of occupancy and decrease in depression severity.</td>
</tr>
<tr>
<td>Hiltunen 1998*24</td>
<td>Healthy volunteers (30; 25-35) M + F</td>
<td>5</td>
<td>Bsl</td>
<td>1 patient received citalopram 30 mg 3hr prior to injection; 1 patient received citalopram 20 mg, 1 patient venlafaxine 37.5 mg 1hr after injection.</td>
<td>[123I]nor-β-CIT SPECT, o-24h a.o. aCG, MID, Thal, BasGang, Cer</td>
<td>After citalopram 30 mg 3hr prior to injection specific binding in the midbrain was 52% less than in untreated subjects. For venlafaxine and citalopram 20 mg no data given.</td>
<td>Study aims at establishing tracer-properties of [123I]nor-β-CIT. Specific binding of [123I]nor-β-CIT is 33% higher compared to [123I]-β-CIT (in other subjects).</td>
</tr>
<tr>
<td>Kent 2002*18</td>
<td>Patients with Social Fobia (35 ±11) M + F</td>
<td>5</td>
<td>Bsl, 3-6m</td>
<td>Paroxetine 20-40 mg (titrated by clinical response) (3-6m)</td>
<td>[11C]McN56 52 PET MID, Thal, Striat, Hip, Amyg, CingA, Cer</td>
<td>Occ by paroxetine: MID 98±3%, Thal 81±6%, Striat 75±7%, Hip 92±12%, Amyg 94±7%, CingA 81±22%</td>
<td>Higher occupancy of SERT than after acute challenge with higher doses (see 69). Unexpected decrease in Vt in cerebellum and white matter, indicative of decrease in non-specific distribution volume of ligand after chronic treatment with paroxetine.</td>
</tr>
</tbody>
</table>
Table S9.1. Studies of SERT occupancy after treatment with serotonergic antidepressants (Continued)

<table>
<thead>
<tr>
<th>Author &amp; Date (Reference)</th>
<th>Results</th>
<th>Population (mean age)</th>
<th>Sex</th>
<th>N</th>
<th>Follow-up</th>
<th>Design</th>
<th>Intervention / control</th>
<th>SERT Imaging and ROI</th>
<th>Results</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klein 2006&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Healthy volunteers (26.8 ±4.3)</td>
<td>M</td>
<td>25</td>
<td>Bsl, 6h</td>
<td>Citalopram (10 mg or 20 mg), escitalopram (5 mg, 10 mg, 20 mg). Volunteers were randomized to one of 5 different drug/dosing regimens</td>
<td>&lt;sup&gt;[123I]&lt;/sup&gt;-ADAM SPECT, 4h MID/Thal, Cer</td>
<td>Occupancy of 60 ±6%, 64 ±6% and 75 ±5% for 5, 10 and 20 mg escitalopram. For 10 and 20 mg citalopram occupancy 65 ±10% and 70 ±6%</td>
<td>Study aims to determine acute occupancy of SERT after single dose drug administration of citalopram or escitalopram</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klein 2007&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Healthy volunteers (28 ±7)</td>
<td>M</td>
<td>15</td>
<td>Bsl, 10d</td>
<td>Citalopram 20 mg, escitalopram 10 mg. Volunteers were randomized to one of the dosing groups</td>
<td>&lt;sup&gt;[123I]&lt;/sup&gt;-ADAM SPECT, 4h, 52h MID/Thal, Cer</td>
<td>Occupancy escitalopram 81.5 ±5.4%, citalopram 64.0 ±12.7%</td>
<td>Increased occupancy observed for escitalopram, also represented in ( E_{\text{max}} ) curves. ( E_{\text{max}} ) curves constructed with data for 4h and 52h post injection scans in 1 model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kugaya 2003, 2004&lt;sup&gt;46,48&lt;/sup&gt;</td>
<td>Volunteers (37.4 ±14.3) M + F 1. Patients with MDD (46.4 ±7.6) M + F 2. Patients with MDD (46.4 ±7.6) M + F</td>
<td></td>
<td>9</td>
<td>Bsl, 8d, 16d Bsl, 1-3w, 6w</td>
<td>Citalopram 40 mg (8d), Citalopram + Buproprion 100 mg (8-16d) Paroxetine 20 mg (6w)</td>
<td>&lt;sup&gt;[123I]&lt;/sup&gt;-β-CIT SPECT, 24h BrSt, Dienc, Cer</td>
<td>Occl by citalopram at 8d: BrSt 51.4%, Dienc 39.4%; no sign. change thereafter Occl by paroxetine at 1-3w: BrSt 36.5%, Dienc 29.6% &amp; at 6w BrSt 32.6%, Dienc 23.4%</td>
<td>Increased DAT-binding in Striatum during prolonged SSRI treatment. Bupropion (100-200 mg) only has no influence on DAT-binding in Striatum in study 2 higher Bsl Dienc SERT availability predicted better treatment response at week 6.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laasonen-Balk 2004&lt;sup&gt;70&lt;/sup&gt;</td>
<td>Patients with MDD (42.7 ±7) M + F</td>
<td></td>
<td>18</td>
<td>Bsl, 6m (±1m drug-free)</td>
<td>Open treatment with various antidepressants (n= 3) and/or benzodiazepines or supportive counselling</td>
<td>&lt;sup&gt;[123I]&lt;/sup&gt;-β-CIT SPECT, 1h &amp; 21-24h MID, Thal, Cer</td>
<td>Binding in MID was sign. increased in responders to treatment (corrected for age and sex)</td>
<td>Low depression severity at Bsl (HDRS 13.9 ±6.7); only three patients received antidepressants, and six benzodiazepines only. Increased SERT binding in responders suggests decreased midbrain density of SERT during MDD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meyer 2001&lt;sup&gt;19&lt;/sup&gt;</td>
<td>1. Patients with MDD, HDRS ≥16 Rx-free (32 ±8) M + F 2. Healthy volunteers (age-matched) M + F</td>
<td></td>
<td>12</td>
<td>Bsl, 4w Bsl, 4w</td>
<td>Open treatment with paroxetine 20 mg (n= 7), 10 mg (n=1) or citalopram 20 mg (n= 4). No treatment</td>
<td>&lt;sup&gt;[11C]&lt;/sup&gt;DASB PET, 0-1½ h Striat, Thal, CingA, PFC, Cer</td>
<td>1. Occ in Striat after paroxetine 20 mg: 83% ±5, and after citalopram 20 mg: 77% ±10. Occ in Thal after paroxetine 20 mg: 75-78% ±8, and after citalopram 20 mg: 65-70% ±15. Occ in CingA after paroxetine 20 mg: 76-77% ±15, and after citalopram 20 mg: 77-79% ±29.</td>
<td>No relationship was found between HDRS-score and occupancy level in any RoI. Striatal Occ increased with higher serum-levels of paroxetine, with app. 85% occ at serum levels of 28 μg/l. References to other studies that describe a hyperbolic relationship of serum level and occupancy. Test-retest in healthy subjects showed a mean difference of -3.7% ±3.7 (range -8 – 2%) over 4 weeks. Mean absolute difference was 10.9% ±2.9.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author &amp; Date (reference)</td>
<td>Results</td>
<td>Population (mean age)</td>
<td>Sex</td>
<td>N</td>
<td>Follow-up</td>
<td>Design Intervention / control</td>
<td>SERT Imaging and ROI</td>
<td>Results</td>
<td>Remarks</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------</td>
<td>-----------------------</td>
<td>-----</td>
<td>---</td>
<td>-----------</td>
<td>-------------------------------</td>
<td>---------------------</td>
<td>---------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Meyer 2004\textsuperscript{20}</td>
<td>1. Patients with MDD 2. Patients with MDD and co-morbid anxiety disorder 3. Healthy subjects (20-50) M + F</td>
<td>29 Bsl, 4w</td>
<td>16</td>
<td>37</td>
<td>MDD: citalopram 20-40 mg, fluoxetine 20 mg, sertraline 50-100 mg, paroxetine 20 mg, venlafaxine 75 mg MDD+anxiety: citalopram 40-60 mg, fluoxetine 40-60 mg, sertraline 150-200 mg, paroxetine 40-60 mg, venlafaxine 150-250 mg Healthy subjects: citalopram 1-10 mg, fluoxetine 1-10 mg, sertraline 5-25 mg, paroxetine 5-10 mg, venlafaxine 2.4-35.7 mg</td>
<td>[\text{[^1C]DASB PET, 0-1/2 h \text{Striat, Thal, CingA, PFC, Cuneus, MID, Cer}}]</td>
<td>Mean striatal occ was 81.4 ±7.2% for citalopram (20-40 mg), 76.2 ±7.5 for fluoxetine (20 mg), 85.0 ±7.0% for sertraline (50-100), 84.5 ±6.0% for paroxetine and 83.7 ±2.4 for venlafaxine (75 mg). Occ for Thal: 72.3 ±7.6%, 69.1 ±4.4%, 76.8 ±4.3%, 74.7 ±15.7% and 71.3 ±10.2% resp. Occ for MID: 87.5 ±7.7%, 82.3 ±9.3%, 91.8 ±14.0%, 93.4 ±7.7% and 91.0 ±7.7% resp.</td>
<td>12 patients also described in Meyer et al. 2001\textsuperscript{19} 2 MDD-patients received subtherapeutical doses in 'Healthy' group; in 5 patients of MDD group dosages were secondarly increased and also analysed (twice) in MDD + anxiety group. No relation between striatal occupancy and clinical remission or percentage change in Hamilton depression scores.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parsey 2006\textsuperscript{21}</td>
<td>Healthy volunteers (38 ±12) M</td>
<td>2 Bsl</td>
<td>17 Bsl, 4-6d</td>
<td>Sertraline 25 mg, 50 mg and 100 mg (4 days at designated dosage)</td>
<td>[\text{[^1C]McN56 PET, 0-2/5 h MID, Thal, Hip, Amyg, CingA, mTL, Occ, Cer}]</td>
<td>Occ by 60 mg paroxetine ‘pretreatment’ (n= 1): Amyg 64.8%, Hip 46.0%, Thal 38.4%, Midbr 83.9%, CingA 26.4%</td>
<td>Mainstudy (n= 6) aims to quantify characteristics and protocol for [^1C]McN5652.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pirker 1995\textsuperscript{22}</td>
<td>1. Patients with MDD (44 ±6-71) M + F 2. ‘Healthy’ controls (42.3 ±24-70) M + F</td>
<td>13 1 scan some where in treatment</td>
<td>11</td>
<td>27 1 scan</td>
<td>Citalopram 20 mg (n= 5), 40 mg (n= 6) and 60 mg (n= 1) for at least one week; one untreated patient.</td>
<td>[^{[123I]}\beta\text{-CIT SPECT, 2, 4, 16, 20, 24h Thal, hypoThal, MID, Pons, Cer}]</td>
<td>No difference in binding in Striat between patients and controls. Citalopram patients showed sign. decrease in Thal, hypoThal, MID compared to controls. No difference in binding between patients with citalopram 20 mg or 40 mg.</td>
<td>Study included patients with bipolar disorder (n= 1) and conversion disorder (n= 1). Control group included 4 patients with peripheral neurological disease. Citalopram-dosing was not blinded and higher dosages may be the result of an unclear selection process. Specific bound in Thal, hypoThal and Midbr region peaked at 4h; in Striat this was reached after 16-24h.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suhara 2003\textsuperscript{23}</td>
<td>1. Healthy volunteers (22.0 ±2.3) M 2. Patients with MDD (35.7 ±12.1)</td>
<td>27 1 scan</td>
<td>10</td>
<td>1. Clomipramine 5-50 mg, fluvoxamine 12.5-50 mg as single administration 2. Clomipramine 20-250 mg, fluvoxamine 25-200 mg as long-term administration</td>
<td>[^{[1C]}\text{McN56 PET, 0-1/2 h Thal}]</td>
<td>1. Clomipramine occ (\geq)83.9-100% at dosages 5-50 mg; fluvoxamine occ (\geq)7.7-87.7% at dosages 12.5-50 mg. 2. Clomipramine occ (\geq)61.3-100% at dosages 20-250 mg; fluvoxamine occ (\geq)76.6-93.6% at dosages 25-200 mg.</td>
<td>Hyperbolic relationship between oral dose, plasma concentration and occ. for clomipramine and fluvoxamine. Patient data suggest that the duration of treatment with fluvoxamine may be correlated with occ. This does not hold for clomipramine.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table S9.1. Studies of SERT occupancy after treatment with serotonergic antidepressants (Continued)

<table>
<thead>
<tr>
<th>Author &amp; Date (reference)</th>
<th>Results</th>
<th>Population (mean age)</th>
<th>Sex</th>
<th>N</th>
<th>Follow-up</th>
<th>Design</th>
<th>Intervention / control</th>
<th>SERT Imaging and ROI</th>
<th>Results</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shang 200772</td>
<td>Healthy controls (23.6 ± 6.3) M + F</td>
<td>8</td>
<td>Bsl, 9d</td>
<td>Venlafaxine 150 mg (4 days stable dose)</td>
<td>$[^{131}I]$-CIT SPECT, 23h Dience, MID, Striat, Cer</td>
<td>Occ in Dience: 52.5 ± 4.7%, in MID: 55.7 ± 4.5%</td>
<td>Study also investigates influence of venlafaxine on striatal DAT availability, which increases 10.7 ± 3.0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Takano 2006a73</td>
<td>Healthy controls (24.3 ± 4.8) M</td>
<td>6</td>
<td>Bsl, 5h, 26h, 53h</td>
<td>Fluvoxamine 50 mg once</td>
<td>$[^{11}C]$DASB PET, 0-1½ h PFC, Thal, Striat, Amyg, Hip, Cer</td>
<td>Mean occ of the 5 regions of 6 subjects were 72.9% ± 4.9% at 5 hours, 50.3% ± 11.0% at 26 hours, and 24.7% ± 15.3% at 53 hours. 5h occ: Thal 71.8 ± 3.5%, Amyg 71.6 ± 12.8%; 53h occ: Thal 24.2 ± 12.3%, Amyg 27.4 ± 18.2%</td>
<td>Study investigates time-course of occupancy of SERT after one (low) dose of fluvoxamine. In $E_{max}$ model dependent measurements at 3 timepoints are used as independent data-points</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Takano 2006b27</td>
<td>Healthy controls (24.1 ± 2.4) M</td>
<td>15</td>
<td>Bsl, 1. 6h, 25h, 35h ± 7d, 9d, 10d</td>
<td>1. single duloxetine dose (5, 20, 40, 60 mg) (n= 12) 2. duloxetine 60 mg for 7 days then stopped (n= 3)</td>
<td>$[^{11}C]$DASB PET, 0-1½ h Thal, Cer</td>
<td>1. Thal occ 35.3-86.5% at increasing dosages: at 6hrs occ: 43.6±8.8% (5 mg), 71.3±5.3% (20 mg), 80.6±4.8% (40 mg), 81.8±4.3% (60 mg) 2. Thal occ: 84.3±2.8% (7d), 71.9±2.6% (9d), 47.1±3.7% (11d)</td>
<td>Study concludes that more than 40 mg is required to attain 80% occupancy. Instead of plasma concentrations that decrease in a one-exponential way after stopping, occupancy appears to decrease linearly and slower than plasma concentrations.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tauscher 199974</td>
<td>Patient with BN and MDD (18) F</td>
<td>1</td>
<td>1 scan</td>
<td>Fluoxetine 60 mg. Until 1 week before scan addition with doxepine 75 mg.</td>
<td>$[^{13}I]$-CIT SPECT, 4h &amp; 21h Thal, hypoThal, Striat</td>
<td>Compared to controls an occ in Thal and hypoThal of app. 4% was expected by fluoxetine. Higher Striat DAT than in controls; in accordance with other MDD-patients</td>
<td>Methodologically poor. Results compared to a historic control group (n= 13, M + F; age 46 ±20; age range 24-71).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viinamäki 199875</td>
<td>1. Patients with PersDis (25 &amp; 34) M 2. Healthy controls (25±3 &amp; 34±1) M</td>
<td>10</td>
<td>Bsl, 1-5y</td>
<td>1. single scan</td>
<td>$[^{123}I]$-CIT SPECT, 4h &amp; 21h PFC, OC, MID, Thal, Striat</td>
<td>(Successfully) treated patient 1: increase in binding in MID: 34%, mPFC 43%, Thal 31%; with no changes in untreated patient 2; levels of binding post-treatment approach levels of controls</td>
<td>Patients were both abusing alcohol; patient 1 abstained during treatment; patient 2 did not. Patient 1 had low, and patient 2 had high HDRS-scores at baseline and follow-up. Controls used alcohol occasionally, no data on HDRS are given. Methodologically poor, study is described poorly.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voineskos 200730</td>
<td>Patients with MDD (36 ±9) Healthy controls (32 ±9) M+F</td>
<td>12</td>
<td>&gt;4w Bsl</td>
<td>Venlafaxine 225 mg, 45.0 mg, sertraline 150 mg, 200 mg, citalopram 60 mg, 80 mg</td>
<td>$[^{11}C]$DASB PET, 0-1½ h Striat, MID, Thal, PFC, ACg, Cer</td>
<td>Occ Striat: 85.8 ±3.4% (venlafaxine), 85.8 ±2.3% (sertraline) and 85.4±2.0% (citalopram) Occ MID: 99.5 ±4.1% (venlafaxine), 98.2 ±3.2% (sertraline) and 95.7±1.4% (citalopram) Occ Thal: 77.6 ±4.7% (venlafaxine), 76.3 ±7.9% (sertraline) and 82.2±3.6% (citalopram)</td>
<td>No baseline scans were performed in patients. Instead as reference baseline scans of controls were used to determine occupancy.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Amyg = Amygdala, aCG = anterior Cingulate Gyrus, BasGang = Basal Ganglia, BN = Bulimia Nervosa, BrSt = BrainStem, Bsl = Baseline, CingA = Cingulate Anterior, d = days, DAT = Dopamine Transporter, Dienc = Diencephalon, h = hours, Hip = Hippocampus, HDRS = Hamilton Depression Rating Scale, hypThal = Hypothalamus, m = months, MDD = Major depressive disorder, mTL = medial temporal lobe, MID = Midbrain, mPFC = medial PreFrontal Cortex, OC = Occipital Cortex, Occ = Occupancy, PersDis = Personality Disorder, PFC = PreFrontal Cortex, RoI = Regions of Interest, Striat = Striatum, Thal = Thalamus, w = weeks, y = years. * most studies used antidepressantsSERT occ. = SERT occupancy relative to non-specific binding (cerebellum)
Table S9.2. Drop-outs and emerging adverse effects reported more than once between T0 and T1 (ITT).

<table>
<thead>
<tr>
<th></th>
<th>Paroxetine DE (n = 30)</th>
<th>Placebo DE (n = 30)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Drop-outs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>1/30 (3.3%)</td>
<td>8/30 (26.7%)</td>
<td>0.026</td>
</tr>
<tr>
<td>Inefficacy</td>
<td>0/30 (0.0%)</td>
<td>3/30 (10.0%)</td>
<td>0.237</td>
</tr>
<tr>
<td>Adverse effects</td>
<td>0/30 (0.0%)</td>
<td>4/30 (13.3%)</td>
<td>0.112</td>
</tr>
<tr>
<td><strong>Adverse effect rates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% all</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Headache</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paroxetine DE (n = 30)</td>
<td>69.0</td>
<td>46.7 70.0 46.7 5/10 (50.0)</td>
<td>35.7</td>
</tr>
<tr>
<td>Placebo DE (n = 27)</td>
<td>44.8</td>
<td>36.7 50.0 43.3 2/3 (66.7)</td>
<td>28.6</td>
</tr>
<tr>
<td><strong>Sweating</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paroxetine DE (n = 30)</td>
<td>63.8</td>
<td>56.7 80.0 63.3 3/5 (60.0)</td>
<td>39.3</td>
</tr>
<tr>
<td>Placebo DE (n = 27)</td>
<td>63.8</td>
<td>66.7 63.4 50 1/5 (20.0)</td>
<td>64.3</td>
</tr>
<tr>
<td><strong>Dry mouth</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paroxetine DE (n = 30)</td>
<td>63.8</td>
<td>56.7 80.0 63.3 3/5 (60.0)</td>
<td>39.3</td>
</tr>
<tr>
<td>Placebo DE (n = 27)</td>
<td>63.8</td>
<td>66.7 63.4 50 1/5 (20.0)</td>
<td>64.3</td>
</tr>
<tr>
<td><strong>Orthostatic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>hypotension</strong></td>
<td>46.6</td>
<td>50.0 53.3 50.0 3/5 (60.0)</td>
<td>35.7</td>
</tr>
<tr>
<td><strong>Constipation</strong></td>
<td>44.8</td>
<td>36.7 50.0 43.3 2/3 (66.7)</td>
<td>28.6</td>
</tr>
<tr>
<td><strong>Libido</strong></td>
<td>41.4</td>
<td>36.7 56.7 56.7 2/2 (100.0)</td>
<td>28.6</td>
</tr>
<tr>
<td><strong>Agitation</strong></td>
<td>41.4</td>
<td>33.3 50.0 33.3 6/9 (66.7)</td>
<td>21.4</td>
</tr>
<tr>
<td><strong>Appetite</strong></td>
<td>39.7</td>
<td>46.7 46.7 43.3 3/7 (42.9)</td>
<td>35.7</td>
</tr>
<tr>
<td><strong>Blurry vision</strong></td>
<td>37.9</td>
<td>40.0 43.3 43.3 0/4 (0.0)</td>
<td>14.3</td>
</tr>
<tr>
<td><strong>Tremor</strong></td>
<td>36.2</td>
<td>36.7 46.7 40.0 4/7 (57.1)</td>
<td>32.1</td>
</tr>
<tr>
<td><strong>Dizziness</strong></td>
<td>34.5</td>
<td>36.7 46.7 26.7 8/11 (72.7)</td>
<td>28.6</td>
</tr>
<tr>
<td><strong>Concentration</strong></td>
<td>34.5</td>
<td>26.7 43.3 30.0 5/8 (62.5)</td>
<td>28.6</td>
</tr>
<tr>
<td><strong>Paraesthesia</strong></td>
<td>32.8</td>
<td>23.3 36.7 36.7 4/5 (80.0)</td>
<td>32.1</td>
</tr>
<tr>
<td><strong>Sleep disturbances</strong></td>
<td>31.0</td>
<td>23.3 43.3 43.3 8/9 (88.9)</td>
<td>35.7</td>
</tr>
<tr>
<td><strong>Nausea</strong></td>
<td>29.3</td>
<td>33.3 33.3 33.3 5/11 (45.5)</td>
<td>28.6</td>
</tr>
<tr>
<td><strong>Weakness</strong></td>
<td>29.3</td>
<td>16.7 36.7 26.7 3/3 (100)</td>
<td>17.9</td>
</tr>
<tr>
<td><strong>Diarrhea</strong></td>
<td>24.1</td>
<td>16.7 23.3 13.3 4/5 (80.0)</td>
<td>25.0</td>
</tr>
<tr>
<td><strong>Urinary retention</strong></td>
<td>24.1</td>
<td>13.3 36.7 23.3 2/3 (66.7)</td>
<td>14.3</td>
</tr>
<tr>
<td>Other sexual dysfunction</td>
<td>22.8</td>
<td>16.7 26.7 30.0 3/3 (100)</td>
<td>17.9</td>
</tr>
<tr>
<td><strong>Palpitations</strong></td>
<td>22.4</td>
<td>16.7 26.7 16.7 3/7 (42.9)</td>
<td>10.7</td>
</tr>
<tr>
<td><strong>Rash</strong></td>
<td>10.3</td>
<td>6.7 6.7 3.3 1/4 (25.0)</td>
<td>7.1</td>
</tr>
</tbody>
</table>

* Adverse effects are listed in order of overall frequency (mentioned more than once in Randomization Phase).
† Recurrence of adverse effect was determined if an adverse effect was initially present during the open phase, disappeared at the end of the open phase, and reemerged during the randomization phase. Columns represent reemerging cases (nominator) of all subjects with the adverse effect initially present, but in whom it disappeared at the end of the open phase (denominator). Values in brackets represent this rate as percentages.
‡ Sign. difference paroxetine DE vs. placebo DE (p<0.05; Fisher’s exact test).
§ Sign. difference paroxetine DE vs. placebo DE (p<0.01; Fisher’s exact test).
DE = dose-escalation, T0 = time of randomization at end of open phase, ≥2x RP = mentioned more than once in randomization phase, Rec. AE = Recurrence of adverse effect.

Table S9.3. Changes in paroxetine serum concentrations (μg/l) in the SPECT subgroup due to paroxetine or placebo dose-escalation.

<table>
<thead>
<tr>
<th></th>
<th>T0</th>
<th>T1</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paroxetine DE (n= 11)</td>
<td>36.2 (22.7-52.8)</td>
<td>154.3 (112.4-202.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Placebo DE (n= 15)</td>
<td>60.3 (41.0-83.3)</td>
<td>52.2 (32.9-75.8)</td>
<td>0.268</td>
</tr>
</tbody>
</table>

Values represent means with 95% confidence intervals, computed on square-root transformed paroxetine serum concentrations. Probable adherent patients only; in 1 patient (paroxetine dose-escalation) the measurement of paroxetine concentration was missing. The change in paroxetine serum concentrations after paroxetine dose-escalation vs. placebo dose-escalation was significant (ANCOVA correcting for paroxetine serum concentration at T0: F1,23= 59.938; p <0.001). * Significance of difference between T0 and T1; paired T-test. DE = dose-escalation.
Figure S9.1. Changes in paroxetine serum concentration versus changes in occupancy (randomization to endpoint).

Regression lines show linear relationship (with 95% CI) between increase in paroxetine serum concentration (PSC) and change in occupancy of midbrain (circles; beta= 0.03 ±0.08; $F_{1,24}= 0.101; p= 0.754$) and diencephalon (diamonds; beta= 0.07 ±0.06; $F_{1,27}= 1.332; p= 0.259$).
SUCCESSFUL PHARMACOLOGICAL TREATMENT OF MAJOR DEPRESSIVE DISORDER ATTENUATES AMYGDALA ACTIVATION TO NEGATIVE FACIAL EXPRESSIONS. AN fMRI STUDY

Submitted

Henricus G. Ruhé*
Jan Booij1
Dick J.Veltman3,4
Martin C. Michel5
Aart H. Schene1

1 Program for Mood Disorders, Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
2 Department of Nuclear Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
3 Department of Addiction Research, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
4 Department of Psychiatry, Vrije Universiteit, Amsterdam, the Netherlands
5 Department of Pharmacology and Pharmacotherapy, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
Abstract

Background
Despite modest response and remission rates to selective serotonin reuptake inhibitors (SSRIs), underlying neurobiological mechanisms of pharmacological treatment of major depressive disorder (MDD) are scarcely investigated in vivo. Amygdala function might be a biomarker for MDD treatment-response.

Aim
to quantify the changes in activation by emotional faces in amygdala and other limbic-subcortical-prefrontal structures after paroxetine treatment in MDD.

Methods
We performed an event-related functional magnetic resonance imaging study in SCID-positive, non-psychotic, unipolar depressed patients (n= 22; M+F; 25-55 years, HDRS, \( >18 \)) who were scanned at study-entry, and at 6 weeks (T0) and 12 weeks (T1) of paroxetine treatment. After 6 weeks open-label paroxetine 20mg/day (T0), 15 non-responding patients (HDRS, decrease <50\%) were randomized to double-blind paroxetine (30-50mg/day) or placebo dose-escalation (paroxetine 20mg/day + placebo). The paradigm consisted of sex-judgments of fearful, angry, happy and neutral faces, relative to scrambled (baseline) faces. Twenty-one age- (±2.5 yrs) and sex-matched healthy controls (HC; without 1st degree relative with, or lifetime psychiatric disorder) were scanned once as reference.

Results
At study-entry, MDD-patients showed increased ventral/limbic (amygdala and insula) and decreased dorsal cortical activations relative to HC. At T0 and T1, 5/20 and 13/20 patients were responders, respectively. Treatment response was associated with decreased activations in amygdala, right insula and orbitofrontal cortex (OFC), and increased activation in right dorsolateral prefrontal cortex (DLPFC) and nucleus accumbens. Amygdala activation was inversely correlated with increased pregenual cingulate and DLPFC activation as well as with improvement using the HDRS. Paroxetine dose-escalation was associated with decreased activity in ventral (limbic) and dorsal cortical regions, whereas activity of right hippocampus and left subgenual cingulate increased relative to placebo dose-escalation.

Conclusions
Decreased amygdala activation by negative faces is a biomarker for clinical response to antidepressants, but needs further exploration in other MDD-treatments. Paroxetine response-effects appear to be mediated by increased fronto-limbic control. (ISRCTN register nr. ISRCTN44111488)
Introduction

Major depressive disorder (MDD) is a highly prevalent and disabling disease,\(^1\) often treated with antidepressants, particularly selective serotonin reuptake inhibitors (SSRIs). Unfortunately, response and remission rates are only modest (30-50%). An additional strategy to achieve remission\(^2-4\) is dose-escalation\(^5\), of which the underlying neurobiological mechanism has scarcely been investigated in vivo. The development of non-invasive neuroimaging techniques may aid in clarifying this issue. For example, we recently showed that a true dose-escalation of paroxetine did not increase serotonin transporter occupancy (the molecular target of SSRIs) more than placebo-dose-escalation.\(^5\)

Previous studies of the etiopathogenesis of MDD have provided evidence for a dysfunctional limbic-subcortical-prefrontal network in MDD.\(^6,7\) A processing bias to perceive negative emotions was associated with increased activation of the ‘ventral’ compartment of this network (consisting of amygdala, insula, ventral striatum, ventral anterior cingulate gyrus and prefrontal cortex), responsible for the identification of the emotional significance of stimuli and the production of affective states.\(^7\) Activation of the (left) amygdala to negative facial expressions (sad, angry or fearful) has been consistently found increased in MDD compared with healthy subjects, and may be considered a biomarker for MDD.\(^8-11\)

Treatment with SSRIs may not fully reverse these differences between MDD and healthy subjects. Four studies have investigated changes in brain activation after 8 weeks of treatment with antidepressants in 8-19 patients per study.\(^10,12-14\) Another study in patients who remitted had a duration of 22 weeks (Table S10.1).\(^15\) These studies found an increase in activation in neocortical regions, including the cingulate gyrus (dorsal and anterior parts) after treatment with antidepressants.\(^10,13\) A decrease in amygdala activation, as originally found by Sheline et al.,\(^12\) was replicated in one study with sertraline (sad faces task),\(^10\) and in one study with bupropion (emotional oddball task).\(^14\) In addition, two randomized placebo-controlled studies in healthy controls found decreased activation in right\(^16\) or bilateral\(^17\) amygdala after 7 days of treatment with reboxetine and citalopram, respectively. However, as most of these treatment studies reported high response rates and one study explicitly excluded non-responders,\(^13\) it is unclear whether the observed changes over time are drug effects, or due to clinical response. This could be addressed in analyses of responders vs. non-responders. Furthermore, although dose-escalation is a recommended strategy for non-responders, studies employing fMRI to investigate dose-response relationships have not yet been published.

Therefore, the aim of this study was to evaluate changes within the limbic-subcortical-prefrontal network, including the amygdala, to (negative) facial expressions after paroxetine treatment of MDD. We specifically aimed to investigate clinical response versus drug effects, including dose-escalation effects. To this end, we performed a functional MRI study in 22 depressed patients who were scanned at study-entry, 6 weeks and 12 weeks of paroxetine treatment, and 22 matched controls scanned once.

We expected to observe increased activation of the amygdala in patients versus healthy controls at baseline, followed by an attenuation of amygdala hyperactivity together with an increase of activation in the neocortical areas (anterior cingulate and dorsomedial and -lateral prefrontal cortex (DMPFC/DLPFC)) after 12 weeks of treatment. We hypothesized that these changes would correlate with clinical response, but not with dose-escalation of paroxetine compared with placebo.
Chapter 10

Methods

Participants
Following approval by the institutional ethical committee and written informed consent, we recruited 22 outpatients (25-55 years) from our outpatient department (August 2005 - August 2006). Inclusion criteria were: MDD determined by the structured clinical interview for DSM-IV (SCID), and a Hamilton Depression Rating Scale (17 items; HDRS17) score above 18. Patients were drug-free (>4 weeks and ≥5 half-lives of a previous antidepressant) and had not used more than one antidepressant treatment (other than paroxetine) at an effective dose for ≥6 weeks for the present MDD-episode. Exclusion criteria, apart from pregnancy (or wish) and standard fMRI contraindications, were bipolar disorder, psychotic features, neurological impairments, primary anxiety and/or substance abuse disorders and acute, severe suicidal ideation.

We individually matched each patient by gender and age (±2.5 years) with a healthy control (HC), who had good physical health and had never used psychotropic medication. Exclusion criteria for HC were any lifetime psychiatric disorder according to the SCID (including abuse or addiction disorders), a Beck Depression Inventory (BDI) score >9, average alcohol use >4 units per day (preceding month) and a 1st-degree relative with psychiatric disorder(s).

Treatment schedule
After assessment at study-entry patients were treated with paroxetine 20 mg/day (supplied in pill-boxes) for 6 weeks. All patients who did not achieve ≥50% decrease in HDRS17-score after 6 weeks (T0; see Figure 2.1) were randomized (stratified for gender and age). Patients received a true paroxetine or a placebo dose-escalation added to paroxetine 20 mg/day. Methods of dose-escalation were described previously, we applied incremental steps of one capsule every 5 days (reaching 30-50 mg/day according to adverse effects). Dosages remained unchanged the last 3 weeks of the study. Adherence was checked by self-report, pill-counts and measured paroxetine serum concentrations.

Measurements
Primary clinical outcome was the proportion of patients achieving response (≥50% decrease in HDRS17). Depression severity was measured at study-entry, randomization (T0), and 6 weeks after randomization (T1), also using the Inventory for Depressive Symptomatology self-rated (IDS-SR30). At these time-points we planned fMRI-sessions of 50 minutes, each including a cognitive task (reported elsewhere), a structural scan and the facial expression task reported here. Agreement between trained raters was good (intra-class correlation coefficient = 0.98). Raters and patients were blinded for the randomized dose-escalation.

Facial expression task paradigm
We used an event-related emotional faces paradigm, which reliably activates the anterior medial temporal lobe including the amygdala. We presented four human face stimuli: angry, fearful, happy, and neutral human faces, and scrambled faces (with two arrows in the centre of the screen) as baseline condition. Each face stimulus condition consisted of 10 pictures; each picture was presented three times. Stimuli were randomized once and presented in identical order to all subjects, using the same task for each session. Stimuli were displayed for 2500ms with a variable interstimulus interval (400-600ms), to increase experimental power and to decrease expectancy effects. To control for overflow effects, we displayed a baseline stimulus after each one or two face pictures. Subjects were instructed to make gender judgements during presentation of face stimuli to control for attention differences. To familiarize participants, the task was first explained outside the scanner. No feedback was provided during the task.
fMRI imaging

We acquired fMRI scans in the afternoon/early evening using a 3Tesla Intera MRI scanner (Philips, Eindhoven, Netherlands). A 6-channel head-coil was used for radiofrequency reception, with the head fixated in the coil by foam pads to prevent movement-artefacts. Stimuli were generated by a Pentium PC and projected on a screen at the patient’s feet, visible through a mirror mounted on the headcoil. Stimulus onset was triggered by a pulse from the scanner. Two magnet compatible response boxes were used to record subject’s performance and reaction times (RTs).

At each session, we obtained a volumetric T1-weighted coronal scan (TE/TR= 4.6/9.63 msec, field of view= 24×24 cm, flip angle= 8°, number of excitations= 1, matrix= 256×256, 182 slices, slice thickness= 1.2 mm, interslice gap= 0 mm, scan time= 7 min) covering the entire brain volume, and 260 T2*-weighted axial echoplanar imaging (EPI) images sensitive to blood oxygen level dependent (BOLD) contrast (TE/TR= 35/2530.4 msec, field of view= 24×24 cm, flip angle= 90°, number of excitations= 1, matrix= 128×128, 36 ascending slices, slice thickness= 3 mm, interslice gap= 0.3 mm, scan time=10 min).

fMRI data analysis

Individual analysis

For all fMRI data-analyses we used SPM5 (Statistical Parametric Mapping; Wellcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk/spm/), operated under Matlab version 7.3.0.267 (2006b; the Mathworks, Natick, Massachusetts, USA). Standard preprocessing of scans consisted of correcting for slice-timing differences, head movements, coregistration to the structural scan, normalization to SPM/MNI standard space, and smoothing (8 mm full-width half-maximum Gaussian filter). Next, BOLD responses were modeled to affective facial expressions and baseline conditions for each voxel. For each subject, weighted contrasts were computed for simple main effects across all stimulus types combined (angry/anxious, happy, and neutral faces vs. baseline = ‘all faces’), and within stimulus type contrasts (angry/anxious vs. baseline = ‘negative faces’; happy vs. baseline = ‘happy faces’).

Group analysis

The contrast images obtained in individual analyses were entered into second level (random effects) analyses for between-group and within subject (i.e. T0–T1) comparisons (with interdependent measures and assuming equal variances). Main effects were identified at p<0.01 and FDR-corrected for multiple comparisons, with an extent threshold of three voxels. Planned contrasts were changes over time (study-entry – T0 – T1), responders versus non-responders, and effect of paroxetine dosage. Interactive effects of time/response*stimulus type (all faces, negative faces, happy faces; masked with the relevant main effect) were identified at p<0.001 (z >3.09), uncorrected, with an extent threshold of three voxels, but in the amygdala region at p<0.005 (z>2.57).

Statistics

We compared study-entry characteristics of patients and controls with independent samples t-tests for continuous and χ² or Fisher’s exact test for categorical variables. We used linear mixed models (compound symmetry variance/covariance structure) to assess differences in scan performance (reaction times and errors); between patients and healthy controls, and in patients in trend over time. Parameter estimates were extracted for bilateral amygdala from SPM5, and these values were used for graphical representation. With regression models, we quantified the association of parameter estimates with HDRS17-scores. We performed additional analyses in SPSS v15.0.1.1 (www.spss.com) and GraphPad Prism v5.00 (www.graphpad.com).
Results

Patients, controls and behavioural data

Table 10.1 displays demographic data and MDD ratings for patients and HC. HC had significantly higher education-levels. MDD-patients reported significantly higher state and trait anxiety than controls ($p < 0.001$), and showed slower RTs ($p = 0.045$) and more gender-judgement errors ($p = 0.032$). Three male patients withdrew from the study after the study-entry scan; 2 withdrew completely. One female HC did not attend her visit and could not be replaced. Due to excessive movements, two study-entry scans (one male, one female patient), two T0 scans (one male, one female patient) and three T1 scans (one male, two female patients) could not be analyzed. Thus, 41 study-entry scans (20 patients, 21 controls), 17 T0 and 16 T1 scans were available for analyses.

<table>
<thead>
<tr>
<th>Educational level n (%)</th>
<th>Patients (n= 22)</th>
<th>Healthy Controls (n= 22)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>5 (22.7)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Intermediate</td>
<td>12 (54.5)</td>
<td>6 (27.3)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>5 (22.7)</td>
<td>16 (72.7)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MDD-severity</th>
<th>Patients</th>
<th>Healthy Controls</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDRS$_{17}$</td>
<td>23.1 ±3.61</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>IDS-SR$_{30}$</td>
<td>42.9 ±7.69</td>
<td>4.6 ±3.87*</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>STAI-I</td>
<td>58.9 ±8.84</td>
<td>29.4 ±7.74*</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>STAI-II</td>
<td>62.7 ±9.25</td>
<td>29.8 ±11.14*</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Co-morbidity n (%)</th>
<th>Patients</th>
<th>Healthy Controls</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
<td>3 (13.6)</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Alcohol dependance</td>
<td>2 (9.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cannabis dependance</td>
<td>1 (4.5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scan task performance</th>
<th>Patients</th>
<th>Healthy Controls</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction Time (ms)</td>
<td>954.7 ±238.3</td>
<td>830.0 ±143.3</td>
<td>0.045</td>
</tr>
<tr>
<td>Correct response (%)</td>
<td>95.5 ±3.24</td>
<td>97.5 ±2.66</td>
<td>0.032</td>
</tr>
</tbody>
</table>

*1 female HC did not attend; no questionnaire and scan performance data available

Twenty patients completed the treatment protocol, 5/20 (25%) and 13/20 (65%) had a clinical response at week 6 and 12 respectively. No week 6 responders deteriorated afterwards (Table 10.2). A randomized dose-escalation was provided to 15 non-responding patients after 6 weeks of paroxetine treatment; for 11 of these repeated scans were analyzable. Clinical outcomes were neither numerically, nor statistically significantly different between true and placebo dose-escalation. RTs and errors did not change over time, and were not significantly different for responders versus non-responders nor associated with HDRS$_{17}$ scores.

Main effects at study-entry in patients and healthy controls

Combining the study-entry scans of patients and HCs (all faces contrast) showed robust activation of bilateral amygdala, fusiform gyrus, DLPFC, (anterior) insula, occipital cortices, and right orbitofrontal cortex (OFC; extending into the right anterior insula), parietal cortex and DMPFC. These effects were also found for negative faces, except for the right amygdala, left insula, and left DLPFC, which were not activated above threshold. With the happy faces contrast, we found main effects for bilateral fusiform gyrus, insula, occipital cortices, and right DLPFC, OFC (extended from insula), thalamus and parietal cortex (Table S10.2).
Amygdala deactivation in responders to paroxetine

Chapter 10

201

Table 10.2. Clinical response and task performance of patients during follow-up (n= 20).

<table>
<thead>
<tr>
<th></th>
<th>T0 (Randomization)</th>
<th>T1 (Endpoint)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDRS&lt;sub&gt;17&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All (n= 20)</td>
<td>15.7 ±5.14</td>
<td>11.1 ±5.31</td>
</tr>
<tr>
<td>To-Responders (n= 5)</td>
<td>9.8 ±3.11</td>
<td>7.2 ±1.48</td>
</tr>
<tr>
<td>True DE (n= 8)</td>
<td>18.0 ±6.13</td>
<td>13.5 ±6.70</td>
</tr>
<tr>
<td>Placebo DE (n= 7)</td>
<td>17.3 ±2.75</td>
<td>11.0 ±3.87</td>
</tr>
<tr>
<td>% decrease in HDRS&lt;sub&gt;17&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All (n= 20)</td>
<td>30.9 ±26.67</td>
<td>51.6 ±25.66</td>
</tr>
<tr>
<td>To-Responders (n= 5)</td>
<td>59.3 ±9.53</td>
<td>70.5 ±2.24</td>
</tr>
<tr>
<td>True DE (n= 8)</td>
<td>20.5 ±26.36</td>
<td>39.4 ±34.07</td>
</tr>
<tr>
<td>Placebo DE (n= 7)</td>
<td>22.0 ±21.01</td>
<td>52.0 ±15.09</td>
</tr>
<tr>
<td>% decrease in HDRS&lt;sub&gt;17&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responders (HDRS&lt;sub&gt;17&lt;/sub&gt; ≥50%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All (n= 20)</td>
<td>5 (25%)</td>
<td>13 (65%)</td>
</tr>
<tr>
<td>To-Responders (n= 5)</td>
<td>5 (100%)</td>
<td>5 (100%)</td>
</tr>
<tr>
<td>True DE (n= 8)</td>
<td>-</td>
<td>4 (50%)</td>
</tr>
<tr>
<td>Placebo DE (n= 7)</td>
<td>-</td>
<td>4 (57%)</td>
</tr>
<tr>
<td>IDS-SR&lt;sub&gt;30&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All (n= 20)</td>
<td>30.6 ±8.80</td>
<td>27.2 ±11.46</td>
</tr>
<tr>
<td>To-Responders (n= 5)</td>
<td>19.8 ±6.46</td>
<td>21.6 ±4.72</td>
</tr>
<tr>
<td>True DE (n= 8)</td>
<td>35.3 ±6.71</td>
<td>33.5 ±14.92</td>
</tr>
<tr>
<td>Placebo DE (n= 7)*</td>
<td>33.0 ±5.66</td>
<td>23.9 ±7.03</td>
</tr>
<tr>
<td>Scan task performance*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction Time (ms)</td>
<td>936.8 ±242.8</td>
<td>921.7 ±252.2</td>
</tr>
<tr>
<td>Correct response (%)</td>
<td>96.3 ±3.74</td>
<td>95.6 ±6.26</td>
</tr>
</tbody>
</table>

Values represent means ±SD. DE= dose-escalation. All differences between True DE and Placebo DE non-significant (ANOVA; p>0.14) * n= 18 at T0 and n= 17 at T1

Activations in the amygdala in MDD-patients versus HC

The all faces contrast showed higher activations in bilateral (extended) amygdala in MDD-patients than in HC. Higher (right) amygdala activities in patients were also found when we contrasted negative faces, but not when contrasting happy faces (Table 10.3A). In post-hoc analyses, baseline amygdala activations were not different between final treatment responders and non-responders (all contrasts).

Other activations of the limbic-subcortical-prefrontal network in MDD-patients versus HC

Contrasting all faces versus baseline showed higher activations in the left insula in MDD-patients compared with HC (Table 10.3A). For happy faces MDD-patients showed higher activation in the left subthalamic nucleus.

With the all faces contrast, we found lower activations in MDD-patients relative to HC in bilateral ventrolateral prefrontal cortex (VLPCF), posterior and medial dorsal cingulate, left DMPFC, bilateral DLPFC and fusiform gyrus (Table 10.3B). For negative faces, we found lower activations in MDD-patients in bilateral VLPCF, left posterior cingulate gyrus, left DMPFC, right DLPFC, and bilateral fusiform gyrus. With happy faces, we found lower activations in right VLPCF, right premotor cortex, and left fusiform gyrus in MDD-patients relative to HC.

In post-hoc analyses, final treatment responders showed higher activations at study-entry in the right pregenual (rostral) cingulate (MNI 14, 44, 3; k= 4; Z= 2.77; p= 0.003; negative faces), relative to final non-responders. Instead, non-responders showed higher study-entry activations in the subgenual cingulate (MNI 0, 26, -3; k= 11; Z= 3.90; p<0.001; negative faces).
Changes in amygdala activations after 6 and 12 weeks of paroxetine treatment

After 6 weeks of treatment (T0), for the all faces and negative faces contrasts, the activations in bilateral amygdala regions had not significantly decreased relative to the study-entry scan (Table S10.3A). In contrast, the right lateral amygdala showed a significant increase in activation after 6 weeks of treatment (all faces, negative faces and happy faces). Further exploration showed that week 6 treatment non-responders (n= 12) had increased amygdala activations for all faces (left z= 2.33; p= 0.01; right z= 2.52; p= 0.006) and negative faces (left z= 3.22; p= 0.001; right z= 2.87; p= 0.002) relative to week 6 responders.

After 12 weeks of treatment (T1), when we contrasted all faces, the right amygdala showed decreased activations (although subthreshold), relative to study-entry (Table 10.4A). Other contrasts did not reveal amygdala deactivations over time.

Table 10.3. Study-entry scans: MDD-patients versus HC.

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>L/R</th>
<th>x,y,z (MNI mm)</th>
<th>Cluster size (k)</th>
<th>Max. voxel Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. MDD &gt; HC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Faces</td>
<td>Amygdala</td>
<td>R</td>
<td>18 -2 -9</td>
<td>12</td>
<td>3.04</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-16 -4 -9</td>
<td>39</td>
<td>2.63</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>L</td>
<td>-26 4 12</td>
<td>30</td>
<td>3.13</td>
<td>0.001</td>
</tr>
<tr>
<td>Neg. Faces</td>
<td>Amygdala</td>
<td>R</td>
<td>18 -2 -9</td>
<td>13</td>
<td>2.99</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-16 -4 -9</td>
<td>16</td>
<td>2.46</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>L</td>
<td>-28 6 15</td>
<td>32</td>
<td>3.26</td>
<td>0.001</td>
</tr>
<tr>
<td>Hap. Faces</td>
<td>Subthalamic nucleus</td>
<td>L</td>
<td>-12 -6 -6</td>
<td>14</td>
<td>3.35</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>B. HC &gt; MDD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Faces</td>
<td>VLPFC</td>
<td>R</td>
<td>50 20 -6</td>
<td>222</td>
<td>3.95</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-34 22 -6</td>
<td>23</td>
<td>3.20</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>R</td>
<td>42 16 27</td>
<td>61</td>
<td>3.73</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>48 -2 51</td>
<td>16</td>
<td>3.31</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-54 16 0</td>
<td>25</td>
<td>3.48</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DMFPC</td>
<td>L</td>
<td>-4 10 63</td>
<td>76</td>
<td>3.49</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Fusiform gyrus</td>
<td>L</td>
<td>-42 54 -21</td>
<td>61</td>
<td>3.61</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cingulate, medial</td>
<td>L</td>
<td>-8 26 42</td>
<td>58</td>
<td>3.36</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>posterior</td>
<td>L</td>
<td>-6 20 51</td>
<td>23</td>
<td>3.59</td>
</tr>
<tr>
<td>Neg. Faces</td>
<td>DMFPC</td>
<td>L</td>
<td>-4 10 63</td>
<td>383</td>
<td>4.37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>VLPFC</td>
<td>R</td>
<td>50 20 -6</td>
<td>241</td>
<td>3.74</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-34 22 -6</td>
<td>41</td>
<td>3.09</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>R</td>
<td>42 16 27</td>
<td>76</td>
<td>4.25</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-42 54 -21</td>
<td>67</td>
<td>3.78</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Fusiform gyrus</td>
<td>R</td>
<td>46 -40 -24</td>
<td>49</td>
<td>3.10</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-16 38 -21</td>
<td>13</td>
<td>3.28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Sup. Temporal gyrus</td>
<td>R</td>
<td>58 56 -12</td>
<td>37</td>
<td>3.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cingulate, posterior</td>
<td>L</td>
<td>-12 22 48</td>
<td>39</td>
<td>3.41</td>
</tr>
<tr>
<td></td>
<td>Hap. Faces</td>
<td>VLPFC</td>
<td>R</td>
<td>54 32 3</td>
<td>81</td>
<td>3.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Premotor cortex</td>
<td>R</td>
<td>44 2 48</td>
<td>27</td>
<td>3.37</td>
</tr>
<tr>
<td></td>
<td>Sup. Temporal gyrus</td>
<td>R</td>
<td>54 48 12</td>
<td>34</td>
<td>3.51</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>48 36 6</td>
<td>27</td>
<td>3.20</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Fusiform gyrus</td>
<td>L</td>
<td>-40 54 -18</td>
<td>50</td>
<td>3.13</td>
<td>0.001</td>
</tr>
</tbody>
</table>

DLPFC = dorsolateral prefrontal cortex; HC = healthy control; MDD = major depressive disorder; MNI = Montreal neurological institute; VLPFC = ventrolateral prefrontal cortex.
Changes in activations in other brain areas after 6 and 12 weeks of paroxetine treatment

After 6 weeks of treatment (T0), relative to study-entry, we additionally found decreased activations in the left posterior hippocampus (all faces; Table S10.3A) and bilateral cuneus (all faces and negative faces). In contrast, increased activations were found in the right dorsal hippocampus, left posterior and medial cingulate gyrus, right pregenual cingulate gyrus, and left DMPFC (all faces contrast; Table S10.3B). For negative faces, activations of right dorsal hippocampus, right anterior and pregenual cingulate gyrus, bilateral cingulate gyrus, bilateral DLPFC, left DMPFC and left nucleus accumbens were increased. With the happy faces contrast, after 6 weeks no activations in any of these regions were found.

After 12 weeks of treatment (T1), relative to study-entry, we found decreased activations for the all faces contrast in left OFC, right posterior hippocampus, and right posterior cingulate gyrus (Table 10.4A). For negative faces, we found no significant decreases in activation for the negative faces. For happy faces we found a decrease in activation in the left insula, right dorsal hippocampus and right parietal cortex. Furthermore, after 12 weeks, relative to

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>L/R</th>
<th>x,y,z (MNI mm)</th>
<th>Cluster size (k)</th>
<th>Max. voxel Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Study entry &gt; T1</td>
<td>All Faces</td>
<td>R</td>
<td>20 0 -12</td>
<td>4</td>
<td>2.4</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>Amygdala</td>
<td>R</td>
<td>50 -48 18</td>
<td>101</td>
<td>3.67</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Hippocampus posterior</td>
<td>R</td>
<td>22 -36 0</td>
<td>34</td>
<td>3.30</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>L</td>
<td>-38 32 -12</td>
<td>9</td>
<td>3.15</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Cingulate, posterior</td>
<td>R</td>
<td>4 -36 48</td>
<td>4</td>
<td>3.09</td>
<td>0.001</td>
</tr>
<tr>
<td>Neg. Faces</td>
<td>Hip. Faces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>L</td>
<td>-34 30 -9</td>
<td>140</td>
<td>3.91</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>L</td>
<td>-48 22 18</td>
<td>32</td>
<td>3.58</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Motor cortex</td>
<td>R</td>
<td>10 -12 63</td>
<td>16</td>
<td>3.25</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>L</td>
<td>36 12 -3</td>
<td>12</td>
<td>3.33</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>B. T1 &gt; Study entry</td>
<td>All Faces</td>
<td>R</td>
<td>28 6 48</td>
<td>194</td>
<td>6.02</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>L -28 4 51</td>
<td>21</td>
<td>4.15</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>L</td>
<td>-38 2 48</td>
<td>23</td>
<td>3.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>L -46 22 18</td>
<td>32</td>
<td>3.58</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L -36 34 30</td>
<td>7</td>
<td>3.10</td>
<td>0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motor cortex</td>
<td>R</td>
<td>10 -12 63</td>
<td>16</td>
<td>3.25</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>L -4 0 60</td>
<td>97</td>
<td>3.19</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neg. Faces</td>
<td>Premotor Cortex</td>
<td>R</td>
<td>28 6 48</td>
<td>487</td>
<td>5.14</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>L -28 4 51</td>
<td>107</td>
<td>4.82</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>R</td>
<td>42 20 21</td>
<td>30</td>
<td>3.66</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>R 60 -20 30</td>
<td>65</td>
<td>3.63</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L -46 20 24</td>
<td>114</td>
<td>3.82</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L -50 8 0</td>
<td>7</td>
<td>3.45</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L -46 8 30</td>
<td>35</td>
<td>3.44</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Motor cortex</td>
<td>L</td>
<td>-26 -16 57</td>
<td>39</td>
<td>3.67</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>R 12 -4 66</td>
<td>20</td>
<td>3.31</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cingulate, posterior</td>
<td>R</td>
<td>6 -12 30</td>
<td>106</td>
<td>3.53</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>L</td>
<td>-26 -6 51</td>
<td>30</td>
<td>3.84</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Hippocampus</td>
<td>R</td>
<td>24 -16 -18</td>
<td>19</td>
<td>3.23</td>
<td>0.001</td>
</tr>
</tbody>
</table>

*Intraparietal sulcus and supramarginal gyrus respectively. Abbreviations: see also Table 10.3. OFC= orbitofrontal cortex.
study-entry, we found increased activations in left DLPFC, and bilateral premotor and motor cortices for the all faces and negative faces contrasts, and in the right hippocampus and left DLFPC for the happy faces contrast (Table 10.4B).

**Changes in amygdala activations in relation to clinical response**

When we compared responders and non-responders after 6 weeks and 12 weeks (full factorial model), we found higher right bilateral amygdala activations in non-responders relative to responders for the all faces contrast (Table 10.5A). With the negative faces contrast, higher activations in non-responders were significant in bilateral amygdala ($p= 0.001$; Figure 10.1). Moreover, the amygdala negative faces contrast estimates were significantly associated with HDRS$_{17}$-scores, with higher activations in more depressed patients. No differential effects between left and right amygdala were found.

In post-hoc analyses, left amygdala activations correlated positively with the contralateral amygdala, bilateral OFC and right subgenual anterior cingulate (Table S10.3). Right amygdala activation correlated positively with the contralateral amygdala, bilateral OFC and right DLPFC. Left amygdala signal was inversely correlated with the pregenual anterior cingulate, right amygdala signal was inversely correlated with the left DLPFC and the left nucleus accumbens.

### Table 10.5. Full-factorial model of 6-weeks and 12-weeks responders versus non-responders.

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>L/R</th>
<th>$x,y,z$ (MNI mm)</th>
<th>Cluster size (k)</th>
<th>Max. voxel Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Non-Responders &gt; Responders</td>
<td>All Faces Amygdala</td>
<td>R</td>
<td>18 -2 -18</td>
<td>30</td>
<td>3.24</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>18 -10 -21</td>
<td>24</td>
<td>2.22</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OFC lateral</td>
<td>R</td>
<td>28 16 -18</td>
<td>98</td>
<td>3.80</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>posterior</td>
<td>R</td>
<td>10 12 -18</td>
<td>98</td>
<td>3.03</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>40 38 -12</td>
<td>39</td>
<td>3.56</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>R</td>
<td>42 -10 3</td>
<td>60</td>
<td>3.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Brainstem (L. coeruleus)</td>
<td>R</td>
<td>6 -16 -12</td>
<td>30</td>
<td>3.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Neg. Faces Amygdala</td>
<td>R</td>
<td>18 -2 -18</td>
<td>30</td>
<td>3.20</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>22 -2 -18</td>
<td>24</td>
<td>3.11</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>R</td>
<td>40 38 -12</td>
<td>33</td>
<td>3.48</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>posterior</td>
<td>R</td>
<td>10 12 -18</td>
<td>*</td>
<td>2.90</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>R</td>
<td>42 -10 3</td>
<td>49</td>
<td>3.33</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>B. Responders &gt; Non-Responders</td>
<td>All Faces DLPFC</td>
<td>R</td>
<td>20 20 63</td>
<td>5</td>
<td>3.71</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>N.Accumbens</td>
<td>L</td>
<td>-10 10 -9</td>
<td>22</td>
<td>3.21</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Neg. Faces  N.Accumbens</td>
<td>L</td>
<td>8 12 -9</td>
<td>40</td>
<td>3.69</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Parietal cortex</td>
<td>L</td>
<td>-46 -52 45</td>
<td>21</td>
<td>3.49</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cerebellum</td>
<td>R</td>
<td>26 -66 -24</td>
<td>11</td>
<td>3.67</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>8 -74 -15</td>
<td>78</td>
<td>3.64</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hippocampus, dorsal</td>
<td>L</td>
<td>-24 -32 -6</td>
<td>15</td>
<td>3.38</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cingulate, medial</td>
<td>LR</td>
<td>4 16 42</td>
<td>41</td>
<td>3.29</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>L</td>
<td>-34 24 6</td>
<td>30</td>
<td>3.17</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Hypothalamus</td>
<td>R</td>
<td>8 -10 3</td>
<td>15</td>
<td>3.11</td>
<td>0.001</td>
</tr>
</tbody>
</table>

* in one OFC cluster. Abbreviations: see Table 10.3 and 10.4.
Figure 10.1. Amygdala response relative to clinical response (HDRS$_{17}$ ≥50% decrease) after 6 and 12 weeks of treatment with paroxetine.

A. SPM-results showing design matrix, contrast and activation in right amygdala (MNI 18, -2, -18).
B. Activations of left (MNI -22, 2, -18; dark) and right (MNI 18, 2, -18; white) amygdala stratified for non-response and response at 6 and 12 weeks. Significant difference between non-responders and responders ($F_{1,32} = 4.591$; $p < 0.0001$; $r^2 = 0.19$; line) and right (0.38 ± 0.14 [SE]; $F_{1,31} = 6.929$; $p = 0.0131$; $r^2 = 0.18$; dashed line) amygdala. Regression lines for left and right amygdala are not significantly different ($F_{1,62} = 0.0683$; $p = 0.79$).

C. Plot of contrast estimate by HDRS 17-score (centralized to mean). Significant positive correlations for left (0.33 ± 0.12 [SE]; $F_{1,31} = 7.076$; $p = 0.0123$; $r^2 = 0.19$; line) and right (0.38 ± 0.14 [SE]; $F_{1,31} = 6.929$; $p = 0.0131$; $r^2 = 0.18$; dashed line) amygdala. Regression lines for left and right amygdala are not significantly different ($F_{1,66} = 0.0683$; $p = 0.79$).

Changes in activations in other brain areas in relation to clinical response
Non-responders also showed significant ($p < 0.001$) higher activations in right OFC and right insula (all faces and negative faces), brainstem (all faces), relative to non-responders after 6 and 12 weeks of treatment (Table 10.5A). In contrast, treatment responders showed higher activations in right DLPFC, and left nucleus accumbens (all faces and negative faces; Table 10.5B). Furthermore, with the happy faces contrast, responders had higher activations in the left dorsal hippocampus, bilateral medial cingulate gyrus, left insula and right hypothalamus.

Effects of true dose-escalation versus placebo dose-escalation
We investigated the interaction of randomized dose-escalation by time in full factorial models. Relative to placebo dose-escalation, 6 weeks of true dose-escalation decreased activations in left OFC, bilateral insula, right VMPFC, left ventrolateral thalamus, right anterior cingulate, left DMPFC and motor cortex, right parietal cortex, right putamen and left globus pallidus (all faces; Table S10.5A). With negative faces decreased activations after dose-escalation were found in right insula, bilateral OFC, left ventrolateral thalamus, right anterior cingulate, left DMPFC and bilateral
motor cortex. With happy faces decreased activations were found in right insula, bilateral OFC, right VMPFC, bilateral anterior cingulate, left pregenual cingulate, bilateral parietal cortex and right putamen.

Relative to placebo dose-escalation, 6 weeks of true dose-escalation increased activations in right hippocampus, left subgenual cingulate, bilateral medial temporal gyrus, left parietal and DLPFC and right cerebellum (all faces; Table S10.5B). With the negative faces contrast, dose-escalation increased activations in right hippocampus, left parahippocampal gyrus, left subgenual cingulate, left medial temporal gyrus and bilateral parietal cortex and cerebellum. With happy faces 6 weeks of dose-escalation increased activations in left parahippocampal gyrus, left medial temporal gyrus and left DLPFC.

Discussion

This fMRI study evaluated the changes in amygdala activation to (negative) facial expressions after paroxetine treatment of MDD, as well as changes in other brain regions within the limbic-subcortical-prefrontal network. Depressed patients were scanned at study-entry, 6 weeks and 12 weeks of paroxetine treatment. To our knowledge, this is the first fMRI study investigating SSRI-treatment with a randomized, placebo-controlled dose-escalation.

MDD-patients showed higher ventral (amygdala and insula) and decreased dorsal activations relative to HC at study entry. After 6 weeks of paroxetine treatment, when only 5/20 patients were classified as responders, we found increased ventral (right amygdala) and dorsal activations (posterior, medial and pregenual cingulate cortex, left DMPFC and DLPFC). In contrast, after 12 weeks of treatment (with 13/20 responders), we found decreased activations in ventral (left OFC) and increased activations in the dorsal compartment, all relative to study-entry. Treatment response was associated with lower activations in amygdala, right OFC and insula, and higher activations in right DLPFC and nucleus accumbens. The placebo-controlled randomized dose-escalation resulted in decreased activations in a number of ventral regions, but not amygdala, while right hippocampus and left subgenual cingulate activations increased, relative to placebo dose-escalation. These results suggest that the attenuation of amygdala activation during SSRI-treatment is not a direct pharmacological response, but appears to be a biomarker for treatment response.

Brain activations in MDD and changes following antidepressant treatment

Our study replicates an increased reactivity of the ventral ‘limbic’ structures in MDD-patients (e.g. amygdala and insula), with decreased activations in dorsal prefrontalareas (PFC and cingulate).8-11 Furthermore, like previous treatment studies10,13-15 paroxetine treatment over 12 weeks increased dorsal cortical function, and reduced activations in the ventral structures. However, these effects were more prominent when we distinguished treatment responders and non-responders.

In the present study, we investigated the role of the amygdala as a biomarker for antidepressant treatment. Previous reports suggested that the SSRIs sertraline and fluoxetine decreased abnormal amygdala activations in response to negative stimuli in MDD-patients,10,12 as was found for bupropion.14 Two studies in healthy volunteers, treated for 7 days with citalopram and the noradrenaline uptake inhibitor reboxetine also found an attenuation of amygdala response to negative faces.16,17 Although we replicate this attenuation, our data suggest an important difference regarding the interpretation previous findings. Specifically, our results indicate that abnormal amygdala activation is decreased in association with treatment response, but is not a direct pharmacological effect, as was also indicated by the absence of amygdala-effects after dose-escalation. Because the response-rates in previous MDD fMRI-studies were high; 10/11 (91%),12 13/19 (68%),13,14 and 6/8 (75%)14 this response-related effect may have been missed in previous group analyses.
As expected, in our study amygdala activation at week 6 and 12 was correlated with other ventral structures (contralateral amygdala, OFC, subgenual cingulate) and inversely correlated to pregenual (rostral) cingulate cortex and DLPFC activity. This highlights an inhibitive connection between the amygdala and the pregenual cingulate, a region that plays a crucial role in integrating the function of dorsal and ventral prefrontal compartments.\textsuperscript{6,29-31} Previous pre-treatment, resting state, fluorodeoxyglucose positron emission tomography (PET) studies have shown that increased metabolism in the pregenual cingulate predicted better treatment response (reviewed by Dougherty and Rauch).\textsuperscript{29} This finding has been replicated in two fMRI studies with negative faces/pictures,\textsuperscript{10,33} and in our post-hoc comparisons of study-entry scans. Moreover, we also found increased activations in the subgenual cingulate in final treatment non-responders,\textsuperscript{32,33} suggesting that activity of pregenual and subgenual cingulate areas may serve as additional biomarkers for (future) treatment response.

The present study also included a paroxetine versus placebo dose-escalation comparison in a secondary randomization of treatment non-responders after 6 weeks. Relative to placebo, higher paroxetine doses was associated with decreased activity in several regions, including left OFC, bilateral insula, right VMPFC and left ventrolateral thalamus. In contrast, dose-escalation vs. placebo resulted in increased activity of the (para)hippocampal cortex and, unexpectedly, the subgenual cingulate cortex. Two explanations can be put forward to account for these results. First, higher doses of paroxetine may further decrease ventral (limbic) hyperactivation and increase dorsal prefrontal activation, in line with the effects seen at standard doses.\textsuperscript{10,12,14} However, the observed increased activity of the subgenual cingulate after dose-escalation is at odds with this hypothesis, because persistent or higher activation in the subgenual cingulate was previously associated with non-response.\textsuperscript{32,33} Also within this hypothesis, the increased activations in the hippocampus might represent increased hippocampal neurogenesis after prolonged antidepressant administration, as has been found in rodents,\textsuperscript{34,35} and which is presumably associated with second-messenger effects of SSRIs resulting in increased levels of brain derived neurotrophic factor (BDNF).\textsuperscript{36} In contrast, Mayberg et al. previously reported decreased hippocampal metabolism in fluoxetine-responders, and no change in non-responders.\textsuperscript{33} However, these studies used fluorodeoxyglucose resting-state PET, comparing changes between week 1 and week 6, so that direct comparisons are problematic.

An alternative explanation might be that dose-escalation does not lead to further improvement MDD in patients, nor affect its neurobiological substrate. This would be in line with the absence of unequivocal improvement after dose-escalation found in previous clinical trials.\textsuperscript{5,37-39} Following this line of argument, the increased activations in the subgenual cingulate may be interpreted as supportive to the finding of Licht et al.,\textsuperscript{40} who reported that patients with increased sertraline doses (200 mg/day) had a poorer response-rate relative to placebo dose-escalation (sertraline 100 mg/day).

In our study, correlation analyses showed that decreased amygdala activity in responders was associated with increased rostral cingulate and DLPFC activity, suggesting that antidepressant response is associated with increased cognitive control by the dorsal compartment as well as a reduction of the hyperactivation in the ventral compartment. Recently, Chen et al. investigated functional connectivity of the amygdala, and reported increased functional coupling with the frontal and (pregenual) cingulate cortex (and striatum and thalamus) after 8 weeks of sertraline treatment.\textsuperscript{31} Although assessed differently, we also found associations of the amygdala with these regions. Therefore, it may be hypothesized that SSRI-exposure over time improves dorsal prefrontal regulation of abnormal limbic activity.\textsuperscript{30,41,42} However, the SSRI-effects on pregenual cingulate-amygdala coupling remain controversial,\textsuperscript{31,41} and may be modified by patient selection and other variables (e.g. polymorphisms of the serotonin transporter promoter region and BDNF).\textsuperscript{31,43,44}

In treatment responders, we found increased activation not only of dorsal prefrontal areas but also of the left nucleus accumbens. Activation of the nucleus accumbens is associated with anticipation of reward.\textsuperscript{45} In contrast, expected loss might deactivate the nucleus accumbens. Although a recent pilot-study did not reveal differences in nucleus accumbens activation between MDD patients and HC during a monetary reward task,\textsuperscript{46} it might still be that during MDD either the
章節 10

第 10 章

期望獎勵的減少，或損失的反應增加，相較於較輕的抑鬱狀態。這可能是假設性的解釋，對大腦基底核的活動增加（或減少的失活）在患者的臨床反應中所見的增強。此結果值得進一步研究報酬處理在 MDD。

限制

與 previous studies 相似，10;12;13 我們選擇不包括安慰劑治療組，而是包括一個隨機的安慰劑控制治療週期。47

若以往的 PET 研究，47

因此，正如最近提出的， future fMRI 研究應該優先包括一個完整的安慰劑組。31

第二，我們一次掃描健康控制組，所以差異習慣化效果不能被評估。10;13

不過，我們的患者樣本顯示了雙相反應，初試期時的 amygdala 活動的增加與極高的反應率 adj

我們認為這不太可能；首先，一個患有共病焦慮和酒精濫用的患者因掃描時過度運動而無法包括在分析中，其次，酒精和大麻的使用在 6 周的帕卡西特治療後停藥。

第三，我們的隨機治療週期在小樣本大小的背景下需要謹慎評估。因此，需要在更大的樣本大小中進行重複。

第四，我們的樣本包含共病的次級焦慮症（3/22）或藥物濫用症（3/22），理論上可能影響我們的發現。我們認為這種可能不高；首先，一個患有共病焦慮和酒精濫用的患者因掃描時過度運動而無法包括在分析中，其次，酒精和大麻的使用在 6 周的帕卡西特治療後停藥。

結論

在 MDD 患者中，我們研究了 amygdala 活動和其它大腦區域在帕卡西特治療後的變化，並應用了一個隨機的安慰劑控制治療週期在治療反應者後 6

在 HC 者，amygdala 的活動和雙側 amygdala 的活動在治療反應者中是不同的。這些結果支持 amygdala 活動透過 fMRI 為治療反應的生物標記，並指出前額-基底核的控制是帕卡西特藥物反應的效果。未來的研究需要澄清自发的、心理治療和/或安慰劑反應者是否也有相似的 amygdala 活動抑制。

致謝

作者感謝參與此 fMRI 研究的患者。E. Miedema M.D. 是不可或缺的治療師並協助掃描。A. Nederveen, PhD, 和 M. de Ruiter, PhD 被稱為他們的貢獻到掃描和分析。M. Haages 管理隨機化和維持盲測。此研究由荷蘭國家健康研究和發展基金會（ZonMw）的項目 Mental Health, 教育於心理健康（OOG; #100-002-002）贊助，以及荷蘭大腦基金會（14F06.45）資助。
Conflicts of interest

None

References

<table>
<thead>
<tr>
<th>Author &amp; Date (reference)</th>
<th>Population (mean age)</th>
<th>N</th>
<th>Design fMRI paradigm</th>
<th>Roles studied / reported</th>
<th>Results</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davidson et al. (2003)</td>
<td>MDD Pts Rx-free ≥7 weeks, HDRS ≥18</td>
<td>12</td>
<td>8 week, VLX 225 mg (open), 3 scans Bsl, 2wk, 8wk; 1.5T Negative, positive, neutral IAPS evoking high arousal.</td>
<td>L.Ins, L.CG(A), Visual Cortex, Amyg</td>
<td>Activation in L.Ins decreased in Pts vs HC (neg. vs neutral); after 2 weeks Pts = HC; after 8 weeks Pts increased activation vs HC (n.s.). Activation L.CG(A) decreased in Pts vs HC (neg. vs neutral); after 8 weeks Pts increased activation vs HC (n.s.) Bsl Amyg activity in Pts = HC; Bsl DLPFC activity in Pts &lt; HC (L) and Pts = HC (R).</td>
<td>Recruitment by advertisement, assessment not specified. Higher pre-treatment CG(A) activation predicted improvement. Start stimulus and scanning manually. No placebo. Responders only: of originally 14 MDD Pts 2 non-responders were excluded.</td>
</tr>
<tr>
<td>Fu et al. (2004)</td>
<td>MDD Pts Rx-free ≥4 weeks, HDRS ≥18 (42.8 ±6.7) Matched HC (43.2 ±8.8)</td>
<td>19</td>
<td>8 week FLX 20 mg (open), 2 scans Bsl, 8wk; 1.5T Sad Ekman Faces of variable intensity: study of activation and dynamic range</td>
<td>HCM, PHC, Amyg, Hypothal, Thal, Ins, CG(A+D), VStriat, NC, InfPC, CER</td>
<td>Activation in all regions increased in pts vs. HC at bsl. Limbic and subcortical activations decreased, dynamic range in neocortical regions increased, in pts after 8 weeks. Decrease in HDRS – reduction dynamic range in CG(A) and CER.</td>
<td>Patients recruited by advertisement, assessment by SCID, no axis 1 co-morbidity. No placebo, 13/19 week 8 responders; 9/19 remission; HDRS responders 5.8 ±2.7; non-responders 15.0 ±4.9²⁸</td>
</tr>
<tr>
<td>Fu et al. (2007)</td>
<td>MDD Pts Rx-free ≥4 weeks, HDRS ≥18 (42.8 ±6.7) Matched HC (43.2 ±8.8)</td>
<td>19</td>
<td>8 week FLX 20 mg (open), 2 scans Bsl, 8wk; 1.5T Happy Ekman Faces of variable intensity: study of activation and dynamic range</td>
<td>extrastriatal cortex, LingG, CER, Cun, CG(P), Put, Thal, HCM, PreC</td>
<td>Activation in extrastriatal cortex, CER, LingG, Cun, CG(P) lower in pts vs HC at bsl. Activation increased in pts after 8wks. Pts have smaller dynamic range in Put, Thal, PreC, CG(P) vs HC at baseline and wk 8.</td>
<td>See under Fu et al. (2004)</td>
</tr>
<tr>
<td>Harmer et al. (2006)</td>
<td>HC without (lifetime) psychiatric disorder</td>
<td>24</td>
<td>7 days double blind randomised: PL or CIT 20 mg (matched on gender, age and IQ), 1 scan at 7 days; 1.5T Masked (17ms + 167ms neutral) fearful, happy or neutral Ekman faces</td>
<td>Amyg (L+R)</td>
<td>Amyg(L+R) and Med.PFC activity significantly lower after fearful faces under CIT vs PL. No sign. effects of CIT vs PL on happy faces.</td>
<td>Recruitment not specified, assessment not specified. No baseline scans. No patients. Decreased Med.PFC activity may be contradictory to decreased Med.PFC activity in MDD. Explained as decrease in HC due to CIT inhibition of Amyg</td>
</tr>
<tr>
<td>Norbury et al. (2007)</td>
<td>HC without (lifetime) psychiatric disorder</td>
<td>24</td>
<td>7 days double blind randomised: PL or REB 8mg, 1 scan at 7 days; 1.5T Masked (17ms + 83ms neutral) and overt (200ms) fearful, happy or neutral Ekman faces</td>
<td>Amyg (L+R)</td>
<td>Masked Amyg(R) activity significantly lower after fearful faces under REB vs PL; no sign differences in overt stimulus presentation or Amyg(L). Increased activity under REB vs PL in FG(R) after happy faces</td>
<td>Recruitment not specified, assessment not specified. No baseline scans. No patients. Response to masked faces might represent effect of REB on automatic aspects of processing.</td>
</tr>
</tbody>
</table>
### Table S10.1 Previous fMRI studies of affective stimuli (faces/IAPS) in MDD during treatment with antidepressants (Continued)

<table>
<thead>
<tr>
<th>Author &amp; Date (reference)</th>
<th>Population (mean age)</th>
<th>N</th>
<th>Design fMRI paradigm</th>
<th>Roles studied / reported</th>
<th>Results</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Robertson et al. (2007)</strong>&lt;sup&gt;14&lt;/sup&gt;</td>
<td>MDD Pts Rx-free ≥7 Wks, HDRS ≥15 (41.4 ±7) M+F</td>
<td>10</td>
<td>8 week BUP 150-450 mg (open), 2 scans Bsl, 8 wk; 1.5T Neg. IAPS (warfare/violence) in Emotional Oddball Task</td>
<td>OFC, VMFPC, CG(A,P), InfPFC, FG, DMPFC, Amyg, HC</td>
<td>Reduced activation in OFC(R), VMPF(C,R), CG(A,R), InfPFC(R), HC+Amygdala(R), NC(R), FG(R), DMPFC(L), CG(P,L) after 8 weeks of BUP Increased activation in InfPFC(L), FG(L). Decrease in activation in OFC(R), InfPFC(R), FG(R) and Amyg (L) correlated with improvement in HDRS</td>
<td>Recruitment not specified. Scans of 8 subjects completing follow-up used for analyses. No placebo, no controls. After BUP, 2/8 of subjects showed no improvement, 4/8 reached HDRS ≤7. 3/8 pts withdrew before 8 weeks, scans made at withdrawal.</td>
</tr>
<tr>
<td><strong>Schaefer et al. (2006)</strong>&lt;sup&gt;15&lt;/sup&gt;</td>
<td>MDD or dysthymia Pts Rx-free ≥4 wks, HDRS ≥22 (35.9 ±?) HC without 1&lt;sup&gt;st&lt;/sup&gt; degree relatives with psych. Disord. (28.2 ±?) M+F</td>
<td>9</td>
<td>Average 22 week VLX 150-300 mg (open), 2 scans Bsl, after remission (4 weeks HDRS ≤10); 1.5T Pos. IAPS (social interaction, single person, human faces, people, erotica, nonhuman appetitive) in passive viewing</td>
<td>Inf., Med., Sup.PFC, Ins, Med.DNT, FG(R)</td>
<td>For Faces vs people/nonhuman appetitive: In patients sign increased activity at 2&lt;sup&gt;nd&lt;/sup&gt; scan in Inf.,Med.,Sup.PFC(L), Ins(L,R), Med.DNT(R), FG(R), while in controls these regions showed increased activity at Bsl scan only.</td>
<td>Recruitment by advertisement, assessment by SCID. 2&lt;sup&gt;nd&lt;/sup&gt; scan after remission only. 6/9 patients reached remission. Two other contrasts: social interaction and erotica also presented in manuscript. In experiments for faces except for Inf and FG(R) a significant laterality was observed. No interaction with clinical improvement or VLX dose found.</td>
</tr>
<tr>
<td><strong>Sheline et al. (2001)</strong>&lt;sup&gt;12&lt;/sup&gt;</td>
<td>MDD Pts Rx-free ≥4 wks, HDRS ≥8 (40.3 ±?) Matched HC HDRS &lt;8 (39.8 ±?) M + F</td>
<td>11</td>
<td>8 week SER ~100 mg (open), 2 scans Bsl, 8wk; 1.5T Masked (40ms) fearful, neutral or happy followed by neutral (160ms) Ekman Faces</td>
<td>Amyg (L+R)</td>
<td>Bsl: L.Amyg activation Pts &gt; HC; R.Amyg Pts&lt;HC (n.s.) (fearful/happy vs neutral) Pts: after SER, L+R Amyg activation after all or fearful faces sign. Decreased. After SER L+R Amyg activity in Pts = HC</td>
<td>Recruitment by advertisement, assessment by DSM-IV criteria. After SER mean HDRS decreased from 23.3 to 9.7. Only 1 non-responder and 2 partial responders. No correlation between L.Amyg signal intensity change and HDRS or anxiety</td>
</tr>
</tbody>
</table>

**Abbreviations:** Amyg= Amygdala, BUP= bupropion, CER= Cerebellum, CG= Cingulate gyrus (A-anterior, D-dorsal, P-posterior), CIT= Citalopram, Cun= Cuneus, DMPFC= Dorsal Medial Prefrontal Cortex, FLX= fluoxetine, FG= Fusiform Gyrus, HC= Healthy Control, HCm= Hippocampus, HDRS= Hamilton Depression Rating Scale, Hypothal= Hypothalamus, InfC= Inferior Frontal Cortex, InfPC= Inferior Parietal Cortex, Ins= Insula, L= left, LingG= Lingual gyrus, MDD= Major Depressive Disorder, Med.DNT= Medial Dorsal Nucleus of Thalamus, NC= Nucleus Caudatus, PFC= Prefrontal Cortex, PHC= Parahippocampal gyrus, PL= placebo, PreC= Precuneus, Pts= patients, Put= Putamen, R= right, REB= Reboxetine, SER= sertraline, Thal= Thalamus, VLX= venlafaxine, VStriat= Ventral striatum.
### Table S10.2  Main effects in study-entry scans: MDD-patients and HC combined.

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>L/R</th>
<th>x,y,z (MNI mm)</th>
<th>Cluster size (k)</th>
<th>Max. voxel Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MDD + HC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>All Faces</strong></td>
<td>Amygdala</td>
<td>R</td>
<td>20 -4 -12</td>
<td>12</td>
<td>3.73</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-18 -4 -15</td>
<td>8</td>
<td>4.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Fusiform gyrus</td>
<td>R</td>
<td>40 -52 -21</td>
<td>160</td>
<td>6.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-42 -56 -24</td>
<td>99</td>
<td>5.34</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>R</td>
<td>46 16 30</td>
<td>405</td>
<td>5.94</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-40 24 21</td>
<td>15</td>
<td>3.61</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Insula (Ant.)</td>
<td>R</td>
<td>36 30 -3</td>
<td>249</td>
<td>5.21</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-34 30 -6</td>
<td>119</td>
<td>4.49</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>R</td>
<td>32 34 -12</td>
<td></td>
<td>4.77</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Occipital</td>
<td>R</td>
<td>12 -94 6</td>
<td>88</td>
<td>4.86</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-14 -98 15</td>
<td>64</td>
<td>4.99</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>R</td>
<td>48 -50 6</td>
<td>126</td>
<td>4.72</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>48 -68 12</td>
<td>34</td>
<td>4.03</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DMPFC</td>
<td>R</td>
<td>4 22 45</td>
<td>35</td>
<td>3.91</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cuneus</td>
<td>R</td>
<td>4 -64 33</td>
<td>15</td>
<td>3.84</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Neg. Faces</strong></td>
<td>Amygdala</td>
<td>L</td>
<td>-18 -4 -15</td>
<td>8</td>
<td>3.95</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Fusiform gyrus</td>
<td>R</td>
<td>40 -52 -21</td>
<td>165</td>
<td>6.28</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-42 -56 -24</td>
<td>82</td>
<td>4.61</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>R</td>
<td>42 16 27</td>
<td>339</td>
<td>5.65</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>R</td>
<td>34 -36 -12</td>
<td>107</td>
<td>4.37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Insula (Ant.)</td>
<td>R</td>
<td>36 30 -3</td>
<td>*</td>
<td>4.35</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Occipital</td>
<td>R</td>
<td>14 -96 9</td>
<td>45</td>
<td>4.31</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-12 -98 15</td>
<td>56</td>
<td>5.24</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>R</td>
<td>48 -50 6</td>
<td>61</td>
<td>4.25</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>50 -68 12</td>
<td>9</td>
<td>3.73</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DMPFC</td>
<td>R</td>
<td>4 22 45</td>
<td>11</td>
<td>3.66</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cuneus</td>
<td>R</td>
<td>4 -60 33</td>
<td>10</td>
<td>3.93</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Hap. Faces</strong></td>
<td>DLPFC</td>
<td>R</td>
<td>46 16 30</td>
<td>173</td>
<td>5.60</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Fusiform gyrus</td>
<td>R</td>
<td>38 -56 -21</td>
<td>90</td>
<td>5.36</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-44 -56 -24</td>
<td>61</td>
<td>4.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Insula (Ant.)</td>
<td>R</td>
<td>36 22 3</td>
<td>163</td>
<td>5.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-34 24 0</td>
<td>72</td>
<td>4.44</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>R</td>
<td>32 34 -12</td>
<td>*</td>
<td>4.70</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Occipital</td>
<td>R</td>
<td>16 -94 6</td>
<td>38</td>
<td>4.41</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-10 -96 9</td>
<td>17</td>
<td>3.98</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Thalamus</td>
<td>R</td>
<td>4 -14 6</td>
<td>20</td>
<td>4.04</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>R</td>
<td>52 -50 6</td>
<td>8</td>
<td>3.80</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Clusters are FDR (p<0.05)-corrected with extend threshold of 3 voxels.
* cluster size not reported, one region extended from right anterior insula to right OFC
Abbreviations: see also Table 10.3 and 10.4. DMPFC= dorsomedial prefrontal cortex

---

Amygdala deactivation in responders to paroxetine
Table S10.3 Changes after 6 weeks of treatment (T₀) compared with study-entry.

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>L/R</th>
<th>x,y,z (MNI mm)</th>
<th>Cluster size (k)</th>
<th>Max. voxel Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Study-entry &gt; T₀</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Faces</td>
<td>Amygdala</td>
<td>R</td>
<td>14 -12</td>
<td>11</td>
<td>1.92</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-16 0 -15</td>
<td>12</td>
<td>1.97</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>Cuneus</td>
<td>LR</td>
<td>0 -66 12</td>
<td>67</td>
<td>4.27</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Hippocampus, posterior</td>
<td>L</td>
<td>-18 -48 9</td>
<td>7</td>
<td>3.60</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>28 -36 0</td>
<td>31</td>
<td>3.27</td>
<td>0.001</td>
</tr>
<tr>
<td>Neg. Faces</td>
<td>Amygdala</td>
<td>L</td>
<td>-14 2 -12</td>
<td>11</td>
<td>2.25</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>Cuneus</td>
<td>L</td>
<td>-2 -66 12</td>
<td>51</td>
<td>3.97</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Sup.Temporal Sulcus</td>
<td>L</td>
<td>-48 -46 6</td>
<td>11</td>
<td>3.48</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>L</td>
<td>-36 -16 -18</td>
<td>31</td>
<td>3.36</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. T₀ &gt; study-entry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Faces</td>
<td>Amygdala</td>
<td>R</td>
<td>28 -2 -15</td>
<td>28</td>
<td>3.22</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Cingulate, posterior</td>
<td>L</td>
<td>-12 -18 48</td>
<td>21</td>
<td>4.13</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>medial</td>
<td>L</td>
<td>2 4 33</td>
<td>70</td>
<td>3.37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>pregenual</td>
<td>R</td>
<td>4 36 3</td>
<td>15</td>
<td>3.41</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Sup.Temporal gyrus</td>
<td>R</td>
<td>62 -8 -9</td>
<td>8</td>
<td>3.79</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Hippocampus, dorsal</td>
<td>R</td>
<td>28 -28 -12</td>
<td>31</td>
<td>3.74</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DMFPC</td>
<td>L</td>
<td>-10 22 60</td>
<td>32</td>
<td>3.39</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Inf. Temporal Gyrus</td>
<td>L</td>
<td>-44 -4 -39</td>
<td>14</td>
<td>3.28</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neg. Faces</td>
<td>Amygdala</td>
<td>R</td>
<td>28 0 -15</td>
<td>47</td>
<td>3.47</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DMFPC</td>
<td>L</td>
<td>-4 22 60</td>
<td>288</td>
<td>4.56</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cingulate, anterior</td>
<td>R</td>
<td>2 30 33</td>
<td>133</td>
<td>3.21</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>medial</td>
<td>LR</td>
<td>2 2 36</td>
<td>69</td>
<td>3.27</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>pregenual</td>
<td>R</td>
<td>6 34 3</td>
<td>5</td>
<td>3.43</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>VLPFC</td>
<td>L</td>
<td>-24 58 30</td>
<td>82</td>
<td>3.89</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>L</td>
<td>-50 26 -3</td>
<td>16</td>
<td>3.65</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-28 44 42</td>
<td>55</td>
<td>3.51</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-30 22 54</td>
<td>14</td>
<td>3.28</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>32 -2 54</td>
<td>12</td>
<td>3.22</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>24 42 48</td>
<td>136</td>
<td>3.09</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Hippocampus, dorsal</td>
<td>R</td>
<td>30 -28 -9</td>
<td>48</td>
<td>3.53</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cerebellum</td>
<td>L</td>
<td>-22 -34 -27</td>
<td>15</td>
<td>3.21</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Sup. Temporal gyrus</td>
<td>R</td>
<td>62 -8 -9</td>
<td>7</td>
<td>3.17</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>N.Accumbens</td>
<td>L</td>
<td>-4 8 12</td>
<td>9</td>
<td>3.13</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hap. Faces</td>
<td>Cerebellum</td>
<td>L</td>
<td>-24 -32 -24</td>
<td>23</td>
<td>3.78</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Sup.Temporal gyrus</td>
<td>R</td>
<td>62 -8 -9</td>
<td>3</td>
<td>3.57</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Amygdala</td>
<td>R</td>
<td>22 -12 -18</td>
<td>8</td>
<td>3.22</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-18 -10 -18</td>
<td>3</td>
<td>2.47</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Abbreviations: see Table 10.3, 10.4 and S10.2.
### Table S10.4 Positive and negative correlations with amygdala activation (negative faces).

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>L/R</th>
<th>x,y,z (MNI mm)</th>
<th>Cluster size (k)</th>
<th>Max. voxel Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Left Amygdala (MNI -22, 2, -18) activation</td>
<td>Hippocampus</td>
<td>R</td>
<td>26 -18 -12</td>
<td>102</td>
<td>4.24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Amygdala</td>
<td>R</td>
<td>22  4 -18</td>
<td>127</td>
<td>5.09</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>R</td>
<td>24  14 -21</td>
<td>33</td>
<td>3.77</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Parahippocampal gyrus</td>
<td>L</td>
<td>-30 -34 -18</td>
<td>98</td>
<td>4.25</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>R</td>
<td>38 -6 -12</td>
<td>18</td>
<td>3.60</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cingulate, subgenual</td>
<td>R</td>
<td>6 14 -12</td>
<td>10</td>
<td>3.59</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>R</td>
<td>52 -54 -15</td>
<td>37</td>
<td>3.88</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Precuneus</td>
<td>R</td>
<td>4 -58 21</td>
<td>4</td>
<td>2.78</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Cingulate, pregenual</td>
<td>R</td>
<td>4 -50 0</td>
<td>14</td>
<td>2.73</td>
<td>0.001</td>
</tr>
<tr>
<td>B. Right Amygdala (MNI 18, -2, -18) activation</td>
<td>Hippocampus</td>
<td>R</td>
<td>24 -22 -12</td>
<td>17</td>
<td>4.56</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Amygdala</td>
<td>L</td>
<td>-18 -4 -21</td>
<td>65</td>
<td>4.49</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>L</td>
<td>34 14 -12</td>
<td>18</td>
<td>3.77</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>R</td>
<td>58 14 36</td>
<td>14</td>
<td>3.70</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Precuneus</td>
<td>L</td>
<td>-22 16 57</td>
<td>6</td>
<td>3.16</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Accumbens</td>
<td>L</td>
<td>-10 12 -9</td>
<td>3</td>
<td>2.68</td>
<td>0.004</td>
</tr>
</tbody>
</table>

* activations for p<0.001 with extend voxel size 10
† in one cluster with right amygdala
‡ activations for p<0.01 with extend voxel size 3
Abbreviations: see Table 10.3, 10.4 and S10.2.

### Table S10.5 Changes between T0 and T1 for true dose-escalation (paroxetine 30-50 mg) relative to placebo dose-escalation (paroxetine 20 mg/day).

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>L/R</th>
<th>x,y,z (MNI mm)</th>
<th>Cluster size (k)</th>
<th>Max. voxel Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Decrease in activation by DE relative to placebo</td>
<td>DMPFC</td>
<td>L</td>
<td>-6 14 51</td>
<td>333</td>
<td>4.27</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>L 24 16 51</td>
<td>96</td>
<td>4.31</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R 26 14 63</td>
<td>7</td>
<td>3.10</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thalamus, ventrolateral</td>
<td>L</td>
<td>-8 -18 3</td>
<td>96</td>
<td>4.26</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Motor cortex</td>
<td>R</td>
<td>26 -2 63</td>
<td>138</td>
<td>4.19</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>L -4 -6 66</td>
<td>56</td>
<td>3.97</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>L</td>
<td>-38 -28 -3</td>
<td>226</td>
<td>3.93</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>R</td>
<td>44 -6 6</td>
<td>100</td>
<td>3.91</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>R 36 22 3</td>
<td>110</td>
<td>3.71</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L -42 -10 3</td>
<td>42</td>
<td>3.65</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Putamen</td>
<td>R</td>
<td>28 0 -3</td>
<td>24</td>
<td>3.84</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>R</td>
<td>62 -36 39</td>
<td>86</td>
<td>3.69</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>R 66 -32 18</td>
<td>60</td>
<td>3.64</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R 56 -52 39</td>
<td>99</td>
<td>3.48</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cingulate, anterior</td>
<td>R</td>
<td>14 32 18</td>
<td>36</td>
<td>3.30</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Precuneus</td>
<td>R</td>
<td>12 -60 21</td>
<td>42</td>
<td>3.42</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>VMPFC</td>
<td>R</td>
<td>16 58 33</td>
<td>37</td>
<td>3.28</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Sup.Temporal Gyrus</td>
<td>L</td>
<td>-42 -32 12</td>
<td>19</td>
<td>3.15</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>G. Pallidus</td>
<td>L</td>
<td>-10 0 -3</td>
<td>10</td>
<td>3.14</td>
<td>0.001</td>
</tr>
</tbody>
</table>
### Table S10.5
Changes between T0 and T1 for true dose-escalation (paroxetine 30-50 mg) relative to placebo dose-escalation (paroxetine 20 mg/day). (Continued)

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>L/R</th>
<th>x,y,z (MNI mm)</th>
<th>Cluster size (k)</th>
<th>Max. voxel Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neg. Faces</td>
<td>Motor cortex</td>
<td>LR</td>
<td>-4 -6 66</td>
<td>62</td>
<td>4.42</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>L</td>
<td>-18 6 57</td>
<td>101</td>
<td>4.22</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DMFPC</td>
<td>L</td>
<td>-6 14 51</td>
<td>*</td>
<td>3.72</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Thalamus, ventrolateral</td>
<td>L</td>
<td>-6 -18 0</td>
<td>60</td>
<td>3.80</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Insula</td>
<td>R</td>
<td>36 22 3</td>
<td>72</td>
<td>3.74</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>44 -6 6</td>
<td>41</td>
<td>3.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>R</td>
<td>38 34 6</td>
<td>†</td>
<td>3.46</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-36 30 -3</td>
<td>61</td>
<td>3.55</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cingulate, anterior</td>
<td>R</td>
<td>12 30 21</td>
<td>14</td>
<td>3.28</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Precuneus</td>
<td>R</td>
<td>12-60 21</td>
<td>9</td>
<td>3.24</td>
<td>0.001</td>
</tr>
<tr>
<td>Hap. Faces</td>
<td>Insula</td>
<td>R</td>
<td>36 12 0</td>
<td>207</td>
<td>4.21</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>32 16 -21</td>
<td>40</td>
<td>3.67</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Putamen</td>
<td>R</td>
<td>22 12 3</td>
<td>134</td>
<td>3.20</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Parietal</td>
<td>L</td>
<td>-60 -36 18</td>
<td>109</td>
<td>4.08</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>62 -28 42</td>
<td>166</td>
<td>3.10</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>26 -10 60</td>
<td>10</td>
<td>3.55</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>22-26 66</td>
<td>67</td>
<td>3.21</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Cingulate, anterior</td>
<td>L</td>
<td>-6 32 18</td>
<td>218</td>
<td>3.17</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>12 42 24</td>
<td>180†</td>
<td>3.59</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>pregenuval</td>
<td>L</td>
<td>-8 46 -6</td>
<td>127</td>
<td>3.90</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>VMPFC</td>
<td>R</td>
<td>22 58 18</td>
<td>180</td>
<td>3.91</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>OFC</td>
<td>L</td>
<td>-28 42 -15</td>
<td>‡</td>
<td>3.00</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>22 46 -15</td>
<td>136</td>
<td>3.43</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>16 -16 -15</td>
<td>23</td>
<td>3.23</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Premotor Cortex</td>
<td>R</td>
<td>4 0 54</td>
<td>14</td>
<td>3.21</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>R</td>
<td>26 38 45</td>
<td>19</td>
<td>3.15</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-24 20 54</td>
<td>67</td>
<td>3.10</td>
<td>0.001</td>
</tr>
</tbody>
</table>

B. Increase in activation by DE relative to placebo

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>L/R</th>
<th>x,y,z (MNI mm)</th>
<th>Cluster size (k)</th>
<th>Max. voxel Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neg. Faces</td>
<td>Med.Temporal Gyrus</td>
<td>L</td>
<td>-52 -6 -12</td>
<td>32</td>
<td>4.02</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>60 -34 -3</td>
<td>19</td>
<td>3.44</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Hippocampus</td>
<td>R</td>
<td>24 -14 -18</td>
<td>69</td>
<td>3.85</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Cingulate, subgenual</td>
<td>L</td>
<td>-2 22 6</td>
<td>48</td>
<td>3.65</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>24 -42 -6</td>
<td>25</td>
<td>3.63</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>DLPFC</td>
<td>L</td>
<td>-54 24 15</td>
<td>45</td>
<td>3.16</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Cerebellum</td>
<td>R</td>
<td>12 -46 -15</td>
<td>23</td>
<td>3.14</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-40 -34 66</td>
<td>9</td>
<td>3.19</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LR</td>
<td>-6 -42 -15</td>
<td>79</td>
<td>3.39</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>26 -62 -12</td>
<td>19</td>
<td>3.10</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-60 -42 -6</td>
<td>43</td>
<td>4.51</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-28 -32 -12</td>
<td>12</td>
<td>3.55</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>-58 20 21</td>
<td>43</td>
<td>3.01</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

† in one cluster with left DLPFC
‡ in one cluster with right VMPFC
¶ in one cluster with left pregenual cingulate

Abbreviations: see also Table 10.3, 10.4 and S10.2. VMPFC= ventromedial prefrontal cortex.
CHAPTER 11

EFFECT OF THE SELECTIVE SEROTONIN REUPTAKE INHIBITOR PAROXETINE ON PLATELET FUNCTION IS MODIFIED BY A SLC6A4 SEROTONIN TRANSPORTER POLYMORPHISM

Journal of Thrombosis and Haemostasis 2008; in press

Nadia Abdelmalik1,2*
Henricus G. Ruhé*
Kurdo Barwari2
Erik-Jan van den Dool3
Joost C.M. Meijers2
Saskia Middeldorp2,4,5
Harry R. Büller2
Aart H. Schene1
Pieter W. Kamphuisen2

* Both authors contributed equally to the manuscript

1 Program for Mood Disorders, Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
2 Department of Vascular Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
3 Department of Clinical Chemistry, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands
4 Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands
5 Department of General Internal Medicine and Endocrinology, Leiden University Medical Center, Leiden, the Netherlands
Abstract

Background
Selective serotonin reuptake inhibitors (SSRIs) have been associated with an increased bleeding tendency.

Aim
To prospectively quantify the dose-response effects of paroxetine and the influence of the serotonin transporter gene (SLC6A4) promoter polymorphism (5-HTTLPR) on platelet function.

Methods
Nineteen drug-free psychiatric outpatients (44.5 ±10.8 years), were tested before and after 6 weeks of paroxetine treatment (20 mg/day). Based on clinical symptoms, paroxetine dosages were increased (40-50 mg/day) for 6 more weeks in 11 patients. Parameters related to platelet function were assessed by bleeding time, platelet function analyzer (PFA), platelet serotonin, platelet factor 4 (PF4), β-thromboglobulin (β-TG), and aggregation tests.

Results
Paroxetine 20 mg/day increased mean bleeding time by 1.2 minutes (95% confidence interval (95% CI) -0.2-2.7) and reduced median platelet serotonin level (463 ng/10^9 platelets; Inter Quartile Range (IQR) 361-666), and platelet β-TG concentration (3.1 IU/10^6 platelets; IQR 0.3-6.0). Other platelet parameters did not change significantly. Serial platelet aggregation tests did not become abnormal. Paroxetine dose-escalation did not further influence platelet function. However, 5-HTTLPR polymorphisms modified these effects: in LA/LA-carriers, bleeding times did not change (-0.2 minutes; (95% CI -0.6 to 0.9)), while bleeding times significantly increased in <2LA-allele carriers (2.3 minutes (95% CI 0.5 to 4.07); p= 0.032). Platelet serotonin decreases were larger in patients without LA-alleles (868 ng/10^9 platelets; (IQR 585 to 1213)) than in ≥1 LA-allele carriers (457 ng/10^9 platelets; (IQR 392 to 598); p= 0.035). PFA closure time and PF4 increased significantly in patients without LA-alleles.

Conclusions
Paroxetine 20 mg/day does not increase overall bleeding time, but impairs platelet function by decreasing the levels of platelet serotonin and platelet β-TG. These paroxetine effects appear to be mediated by 5-HTTLPR, with most pronounced effects in patients without LA-alleles.
Chapter 11

Paroxetine induced platelet function changes

Introduction

Selective serotonin reuptake inhibitors (SSRIs) belong to the most widely prescribed classes of drugs and are often used for the treatment of depression and anxiety disorders.\(^1\) However, because of their extensive use, the impact of relatively infrequent side effects is of potential clinical importance. One of these harmful side effects is an increased bleeding tendency.

Over the years, numerous case reports\(^2\)\(^-\)\(^6\) and case control studies\(^7\)\(^-\)\(^\text{11}\) have suggested that SSRI use is associated with an increased risk of bleeding. The bleeding pattern varied from minor bleeding to spontaneous gastro-intestinal bleeding and excessive perioperative hemorrhage. One study showed that SSRI users who underwent orthopedic surgery had a nearly four-fold increased risk to require blood transfusion.\(^2\) In addition, large population based case-control studies have consistently shown an increased incidence of upper gastrointestinal bleeding associated with SSRI use,\(^2\) especially when co-medicated with nonsteroidal anti-inflammatory drugs, aspirin\(^8\)\(^-\)\(^\text{10}\) or coumarins.\(^8\)

The underlying pathophysiological mechanisms of the increased bleeding tendency of SSRIs were previously studied, and significant changes in platelet function related to SSRI use were consistently reported. These changes included a decreased aggregation with epinephrine\(^14\)\(^-\)\(^\text{15}\) and collagen,\(^15\)\(^-\)\(^\text{16}\) a prolongation of the PFA closure time,\(^17\) and a decrease in plasma beta-thromboglobulin (\(\beta\)-TG)\(^8\) and plasma platelet factor 4 (PF4).\(^18\)\(^-\)\(^19\)

Furthermore, SSRIs block serotonin transporters (SERTs), which increases serotonergic transmission between neurons and results in antidepressant effects. However, SSRIs also block the SERTs of blood platelets, causing decreased platelet uptake of serotonin.\(^20\) Serotonin mediates vasoconstriction, platelet aggregation, and platelet activation after vessel injury, and since platelets cannot synthesize serotonin, treatment with an SSRI could lead to depletion of platelet serotonin and impair hemostasis.\(^20\)

Polymorphisms of the SERT gene (SLC6A4) promoter region (5-HTTLPR) are associated with the transcriptional activity of the SERT gene and the rate of serotonin uptake.\(^21\) The 5-HTTLPR polymorphism is located approximately 1 kb upstream of the transcription initiation site and is composed of 16 repeat elements. Human lymphoblasts homozygous for the long (L) 5-HTTLPR allele produce higher concentrations of SERT mRNA than cells containing one or two copies of the short (S) allele. Furthermore, the rate of serotonin uptake by the transporter is more than twofold higher in cells homozygous for the L allele.\(^21\) In Caucasians the L allele is found more frequently (57%) than the S allele (43%), with a 5-HTTLPR genotype distribution of 32% LL, 49% LS and 19% SS. However, other populations show different prevalences, especially Asians, who have more than two-fold higher SS genotype frequencies.\(^22\) Nowadays, the 5-HTTLPR polymorphism is considered tri-allelic.\(^23\) The L allele can be subdivided in an L\(_L\) and an L\(_A\) variant by a common SNP (rs25531), which creates a functional transcription factor binding site (L\(_C\)), behaving like a short allele. This tri-allelic classification is probably more reliable to find associations between 5-HTTLPR polymorphisms and phenotypes (e.g. SSRI adverse effects).\(^24\) In Caucasians the S:L\(_A\):L\(_L\) ratio is approximately 8:10:2, while in black patients this is 5:10:5.\(^23\)

Although an association between SSRI use and platelet dysfunction is undoubted,\(^14\)\(^7\)\(^-\)\(^\text{18}\)\(^-\)\(^\text{25}\)\(^-\)\(^\text{26}\) it remains uncertain whether a dose-response relationship exists between SSRI use and platelet function.\(^20\) In addition, since functional polymorphisms of the serotonin transporter influence the rate of serotonin uptake, they may also mediate the effects on hemostasis.\(^27\)

In this study, we evaluated the effect of standard and increasing dosages of the SSRI paroxetine on the platelet function of SSRI free patients. In addition, we assessed whether this effect is influenced by the 5-HTTLPR/rs25531 polymorphism.
Methods

Participants
After approval by the local ethics committee and written informed consent, we recruited 19 adult psychiatric outpatients (18-75 years) from February 2006 until February 2007. The inclusion criterion was indication of SSRI use (for mood or anxiety disorders) as determined by the treating physician. We excluded patients who used aspirin, NSAIDs (in the preceding 48 hours) or antidepressants (less than 4 weeks prior to inclusion), and patients who had a history of increased bleeding tendency.

Treatment and timing of measurements
After baseline assessment, patients were treated with paroxetine 20 mg/day for 6 weeks. If at 6 weeks the clinical symptoms of the patients had not improved by 50%, physicians could increase the dose (dose-escalation) to a maximum of 50 mg/day over the next 6 weeks, as recommended by current psychiatric treatment guidelines. As a measure to detect increased bleeding tendency, in all patients the bleeding severity score (Tosetto et al.) was assessed at baseline, after 6 weeks, and after 12 weeks of treatment (if a dose-escalation was prescribed). Platelet parameters were assessed at the same time points by technicians that were unaware of the treatment regimen.

Measurements of platelet parameters

Blood samples
Blood was collected in open 10 mL tubes containing 1 ml 3.2% sodium citrate by venapuncture from the antecubital vein using a 19-gauge needle. Within 60 minutes after collection, platelet rich plasma (PRP) was prepared by centrifugation at 190 g at room temperature for 10 minutes without braking. PRP was separated from the packed cells; after 30 minutes aggregation studies were started. PECT tubes for 4 ml blood containing 0.4 ml of a mixture of 94 nmol/L Prostaglandin E1, 90 mmol/L EDTA, 0.63 mmol/L sodium carbonate and 10 mmol/L theophilline were used for the determination of β-TG and PF4 in plasma. Blood samples were collected in PECT tubes to prevent in vitro release of β-TG and PF4 from platelets. K3EDTA 7.5% anti-coagulant tubes were used for the determination of platelet count.

Platelet parameters
Bleeding time was measured with a standard incision using the Surgicutt device (Surgicutt Adult, ITC Edison, USA). Platelet function was measured with the PFA-100 (Siemens Healthcare Diagnostics, Marburg Germany) using epinephrine (EPI) and ADP containing cartridges. Platelet aggregation in platelet-rich-plasma with ADP, ristocetin, collagen and arachidonic acid was analyzed qualitatively with a standardized aggregometer (model 540-vs, Chrono-Log Corporation, Havertown, USA).

The amount of β-TG and PF4 in platelets and plasma was measured using enzyme-linked immunosorbent assays (Asserachrom β-TG and PF4, Diagnostica Stago, Roche). β-TG and PF4 in plasma were determined in PECT samples. β-TG and PF4 in platelets were determined after pelleting the platelets from PRP. The platelets were then destroyed by a combination of Triton (2% Triton X-100) and sonication for 15 seconds on ice (microtip, Branson, amplitude 50%). After 5 minutes of centrifugation at 13,000 rpm, the supernatant was used for analysis.

Serotonin
In the sonicated supernatant, the amount of serotonin was measured in an acidified (HClO4) environment by fluorimetry (Fluostar Galaxy, BMG Offenburg, Germany).
Paroxetine induced platelet function changes

Paroxetine serum concentrations

We collected blood samples for measurement of paroxetine serum concentrations (PSC) after 6 and 12 weeks of paroxetine treatment from 19 and 11 patients, respectively. PSC were measured using a validated HPLC-MS/MS method (details available on request). The lower limit of quantification was 5 µg/L, which was at the lower end of the therapeutic range of paroxetine in serum (5-75 µg/L). The lower limit of detection was 0.3 µg/L.

Genotyping procedures and analysis

Genotyping was performed as described earlier.23 In brief, genomic deoxyribonucleic acid (DNA) was isolated and the length of the 5-HTTLPR polymorphism was determined by gel electrophoresis. The region around the polymorphism was amplified by PCR. Genotyping of the rs25531 SNP was done by sequencing. The length of the 5-HTTLPR polymorphism was confirmed by looking at the length of the sequenced PCR product. Taking into account this SNP, we reclassified the tri-allelic genotypes into a modified bi-allelic classification: S'/S' (S/S, L_C/L_S, L_C/L_C = non L_A), S'/L_A (S/L_A, L_C/L_A = 1 L_A) and L_A/L_A (≥ 2 L_A). Because it is unknown which allele is dominant (heterozygosity), we either grouped S'/S' and S'/ L_A as <2 L_A alleles versus ≥ 1 L_A or contrasted ≥ 1 L_A vs. non L_A genotypes in our analyses.

Statistical analyses

We first computed means for the platelet parameters at baseline, after 6 and 12 weeks of paroxetine administration. We compared patients who received a further dose-escalation from week 6 onwards with patients who remained at 20 mg/day with χ² tests for categorical data and independent T-tests for continuous data. We compared the changes in platelet parameters over time with paired T-tests (baseline vs. 6 weeks and 6 vs. 12 weeks). In case of a non-normal distribution (assessed by histograms and Shapiro Wilks’ test) we present medians with interquartile ranges (IQR) and used nonparametric Wilcoxon signed ranks tests for paired data. For the platelet parameters that appeared to change significantly after 6 weeks of treatment, we further explored the changes in linear mixed models; we examined the effects of paroxetine dosage by the significance of a dose*time interaction.

In addition, we investigated the relation of PSC with changes in relevant platelet parameters in linear regression models using all observations (after 6 and 12 weeks) at once. We introduced the 5-HTTLPR polymorphism to explore the genetic influence on these relations. Because of our sample size and non-normal distribution of bleeding times and platelet parameters, we tested differences between genotypes with nonparametric Mann-Whitney tests. All analyses were performed in SPSS v15.0.1.

Results

Patients

Nineteen subjects who were diagnosed with either a mood disorder (n = 13), an anxiety disorder (n = 2) or both (n = 4) participated in this study. The mean age was 44.5 years (SD 10.8) and 11 (58%) were female. Co-medication consisted of depakine (for epilepsy; n = 1), levothyroxine (n = 1) and omeprazol (n = 1).

All subjects received a standard dose of paroxetine 20 mg/day for 6 weeks (mean PSC 25.6 µg/L (95% CI 14.3 to 36.9)). In 11 subjects who failed to show sufficient clinical improvement after this period, paroxetine dosages were increased to 40 (n = 4) or 50 mg/day (n = 7) for another 6 weeks (mean PSC 107.2 µg/L (95% CI 42.2 to 172.2)). The group that only received standard dosage was comparable with the increased dosage group in type of psychiatric disorder, age and sex (Table 11.1).
Chapter 11

Chapter 11

Bleeding severity scores

At baseline, the bleeding severity score was low (<3) in all subjects. Six weeks at paroxetine 20 mg/day, or a subsequent dose-escalation for 6 weeks thereafter did not affect the bleeding score.

Platelet parameters

Six weeks of treatment with paroxetine non-significantly increased the mean bleeding time by 1.2 minutes (95% CI -0.2 to 2.7; Paired t-test; p= 0.083; Table 11.2). Furthermore, paroxetine resulted in a significant reduction of median platelet serotonin level of 463 ng/10⁹ platelets (IQR 361 to 666; Wilcoxon signed ranks test; p<0.001), and median platelet ß-TG level of 3.1 international units (IU)/10⁶ platelets (IQR 0.3 to 6.0; Wilcoxon signed ranks test; p= 0.016). Other changes in platelet parameters were not statistically significant, nor did the aggregation tests reveal qualitative abnormalities.

Dose-escalation of paroxetine to 40-50 mg/day for another 6 weeks did not lead to further reduction of platelet serotonin or platelet ß-TG levels or to a further increase of the bleeding time (Table 11.3). The absence of further changes in platelet ß-TG and serotonin levels were confirmed by non-significant dose*time interactions in mixed models.

There was no correlation between plasma paroxetine levels and changes in platelet function tests, including the change in platelet serotonin and platelet ß-TG levels. Platelet aggregation tests with ADP, ristocetin, collagen and arachidonic acid were not influenced by paroxetine administration (data not shown).

Table 11.1. Baseline characteristics of patients starting treatment with paroxetine 20 mg/day.

<table>
<thead>
<tr>
<th></th>
<th>All (n= 19)</th>
<th>Dose not increased after 6 weeks (n= 8)</th>
<th>Dose-escalation 40-50 mg after 6 weeks (n= 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age ±SD</td>
<td>44.5 ±10.8</td>
<td>45.6 ± 13.2</td>
<td>43.6 ± 9.2</td>
</tr>
<tr>
<td>M/F ratio</td>
<td>8/11</td>
<td>4/4</td>
<td>4/7</td>
</tr>
<tr>
<td>Psychiatric diagnosis</td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>Mood disorder†</td>
<td>13 (68.4%)</td>
<td>5 (62.5%)</td>
<td>8 (72.7%)</td>
</tr>
<tr>
<td>Anxiety disorder‡</td>
<td>2 (10.5%)</td>
<td>1 (12.5%)</td>
<td>1 (9.1%)</td>
</tr>
<tr>
<td>Mood + anxiety disorder</td>
<td>4 (21.1%)</td>
<td>2 (25.0%)</td>
<td>2 (18.2%)</td>
</tr>
<tr>
<td>Substance abuse‡</td>
<td>4 (21.1%)</td>
<td>2 (25%)</td>
<td>2 (18.2%)</td>
</tr>
<tr>
<td>Eating disorder</td>
<td>1 (5.3%)</td>
<td>1 (12.5%)</td>
<td>-</td>
</tr>
</tbody>
</table>

* Unipolar depression (n= 16; 89.5%) and/or dysthymia (n= 3; 15.8%)
† Panic disorder with or without agoraphobia (n= 4; 21.1%), Post traumatic stress disorder (n= 2; 10.5%)
‡ Cannabis abuse (n= 1; 5.3%), Alcohol and cannabis abuse (n= 2; 10.5%), Benzodiazepine abuse (n= 1; 5.3%)

Table 11.2. Changes in platelet function after 6 weeks of treatment with paroxetine 20 mg/day (total n= 19).

<table>
<thead>
<tr>
<th></th>
<th>Baseline (mean ±SD)</th>
<th>6 weeks (mean ±SD)</th>
<th>n*</th>
<th>Difference baseline vs. 6 weeks (mean 95% CI)</th>
<th>p†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platelets (x10⁹/l)</td>
<td>252 ±58</td>
<td>249 ±59</td>
<td>19</td>
<td>-3.2 (-14.0 – 7.6)</td>
<td>0.547</td>
</tr>
<tr>
<td>Bleeding time (min)</td>
<td>4.3 ±1.1</td>
<td>5.3 ±2.0</td>
<td>13</td>
<td>1.2 (-0.2 – 2.7)</td>
<td>0.083</td>
</tr>
<tr>
<td>PFA-ADP (sec)</td>
<td>85.7 ±18.4</td>
<td>88.9 ±18.9</td>
<td>19</td>
<td>3.2 (-13.5 – 2.2)</td>
<td>0.530</td>
</tr>
<tr>
<td>PFA-epinephrine (sec)</td>
<td>112.4 ±26.4</td>
<td>119.5 ±33.9</td>
<td>19</td>
<td>7.1 (-22.0 – 17.8)</td>
<td>0.329</td>
</tr>
<tr>
<td>(median IQR)</td>
<td>(median IQR)</td>
<td>(median IQR)</td>
<td>(median IQR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PF4 platelets‡ (IU/10⁶ platelets)</td>
<td>15.9 (14.7 – 16.7)</td>
<td>15.9 (13.2 – 17.9)</td>
<td>17</td>
<td>-0.2 (-2.3 – 2.7)</td>
<td>0.925</td>
</tr>
<tr>
<td>PF4 plasma‡ (IU/mL)</td>
<td>6.9 (4.0 – 12.1)</td>
<td>5.8 (3.4 – 17.6)</td>
<td>18</td>
<td>-0.7 (-6.8 – 8.9)</td>
<td>0.983</td>
</tr>
<tr>
<td>ß-TG platelets‡ (IU/10⁹ platelets)</td>
<td>33.5 (30.4 – 36.4)</td>
<td>30.7 (28.1 – 34.2)</td>
<td>17</td>
<td>-3.1 (-6.0 – 0.0)</td>
<td>0.016</td>
</tr>
<tr>
<td>ß-TG plasma‡ (IU/mL)</td>
<td>35.8 (22.4 – 39.7)</td>
<td>28.0 (20.3 – 64.7)</td>
<td>18</td>
<td>-2.2 (-16.4 – 22.0)</td>
<td>0.879</td>
</tr>
<tr>
<td>Platelet serotonin‡ (ng/10⁹ platelets)</td>
<td>667 (489 – 758)</td>
<td>128 (80 – 141)</td>
<td>16</td>
<td>-463 (-666 – 361)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

* n for pairs without missing values
† p-values for paired t-tests for normally distributed data and paired Wilcoxon signed ranks test for non-normally distributed data
‡ due to non-normal distribution medians and interquartile ranges (IQR) are shown.
Paroxetine induced platelet function changes

### Effects of the 5-HTTLPR polymorphism

Genotype was analyzed in eighteen patients, of whom 3 had the S'/S', 9 the S'/L_A, and 6 the L_A/L_A genotype, respectively. For these 18 patients, 26 observations after 6 and 12 weeks were available when combining data from these time points. In patients with the 2L_A genotype, paroxetine treatment did not increase bleeding time (-0.2 minutes; 95% CI -0.6 to 0.9), while in patients carrying the <2L_A genotype, bleeding time increased by 2.3 minutes (95%CI 0.5 to 4.1, Mann-Whitney test; p= 0.032). In a linear regression model the <2LA genotype significantly predicted the change in bleeding time (p<0.001). This difference in bleeding time between the two genotype groups was not explained by a decrease in platelet serotonin, dosage or PSC (linear regression; pchange> 0.37).

The median decrease in platelet serotonin levels of patients with paroxetine treatment was not different between the <2L_A and 2L_A genotypes (<2L_A: 587 ng/10^9 platelets; IQR 393 to 841; 2L_A: 457 ng/10^9 platelets; IQR 373 to 582; Mann-Whitney test; p= 0.032). In a linear regression model the <2L_A genotype significantly predicted the change in bleeding time (p<0.001). This difference in bleeding time between the two genotype groups was not explained by a decrease in platelet serotonin, dosage or PSC (linear regression; pchange> 0.37).

The median decrease in platelet serotonin concentrations of patients with paroxetine treatment was not different between the <2L_A and 2L_A genotypes (<2L_A: 587 ng/10^9 platelets; IQR 393 to 841; 2L_A: 457 ng/10^9 platelets; IQR 373 to 582; Mann-Whitney test; p= 0.032). In a linear regression model the <2L_A genotype significantly predicted the change in bleeding time (p<0.001). This difference in bleeding time between the two genotype groups was not explained by a decrease in platelet serotonin, dosage or PSC (linear regression; pchange> 0.37).

### Table 11.3. Changes in platelet parameters in patients who received a secondary dose-escalation (40-50 mg) paroxetine after 6 weeks (n= 11).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Difference baseline vs. 6 weeks (mean ±SD)</th>
<th>6 weeks (mean ±SD)</th>
<th>12 weeks (mean ±SD)</th>
<th>n</th>
<th>Difference 6 vs.12 weeks (mean 95% CI)</th>
<th>p†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platelets (x10^9/l)</td>
<td>-9.9 (-23.5 - 3.7)</td>
<td>242 ±61</td>
<td>249 ±53</td>
<td>11</td>
<td>7.0 (-13.8 - 27.8)</td>
<td>0.470</td>
</tr>
<tr>
<td>Bleeding time (min)</td>
<td>1.3 (-0.1 - 3.1)</td>
<td>5.6 ±2.3</td>
<td>5.8 ±2.3</td>
<td>9</td>
<td>0.4 (-1.1 - 2.0)</td>
<td>0.523</td>
</tr>
<tr>
<td>PFA-ADP (sec)</td>
<td>-1.4 (-15.8 - 13.1)</td>
<td>0.0 ±16.4</td>
<td>91.7 ±15.3</td>
<td>11</td>
<td>3.7 (-8.2 - 15.6)</td>
<td>0.502</td>
</tr>
<tr>
<td>PFA-epinephrine (sec)</td>
<td>9.5 (14.7 - 33.6)</td>
<td>122.6 ±39.7</td>
<td>124.5 ±30.0</td>
<td>10</td>
<td>0.9 (19.5 - 21.3)</td>
<td>0.923</td>
</tr>
<tr>
<td>(median IQR)</td>
<td>(median IQR)</td>
<td>(median IQR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PF4 platelets (IU/10^9 platelets)</td>
<td>-0.2 (-2.3 - 2.7)</td>
<td>16.0 (12.5 - 18.4)</td>
<td>15.0 (14.3 - 17.0)</td>
<td>10</td>
<td>-0.8 (-2.2 - 2.2)</td>
<td>0.799</td>
</tr>
<tr>
<td>PF4 plasma (IU/mL)</td>
<td>0.1 (-6.8 - 8.9)</td>
<td>6.4 (4.3 - 18.2)</td>
<td>5.2 (3.7 - 12.2)</td>
<td>10</td>
<td>-2.6 (-13.4 - 9.9)</td>
<td>0.575</td>
</tr>
<tr>
<td>β-TG platelets (IU/10^6 platelets)</td>
<td>-3.7 (-9.4 - -0.3)</td>
<td>31.2 (27.4 - 35.3)</td>
<td>32.8 (28.6 - 35.4)</td>
<td>10</td>
<td>1.0 (-2.2 - 2.5)</td>
<td>0.575</td>
</tr>
<tr>
<td>β-TG plasma (IU/mL)</td>
<td>0.9 (-15.5 - 22.0)</td>
<td>28.0 (20.9 - 65.2)</td>
<td>22.7 (19.3 - 38.9)</td>
<td>10</td>
<td>-5.6 (-39.6 - -0.1)</td>
<td>0.074</td>
</tr>
<tr>
<td>Platelet serotonin (ng/10^9 platelets)</td>
<td>-499 (-780 - -366)</td>
<td>118 (74 - 143)</td>
<td>100 (86 - 145)</td>
<td>10</td>
<td>1.50 (-23.50 – 14.00)</td>
<td>0.799</td>
</tr>
</tbody>
</table>

† n for pairs week 6 – week 12 without missing values

### Table 11.4. Changes in platelet parameters after 6 and 12 weeks of paroxetine treatment, relative to baseline, stratified for genotype.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>No L_A/L_A</th>
<th>≥1 L_A</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platelet serotonin (ng/10^9 platelets)</td>
<td>-868 (-1213 – -585)</td>
<td>-457 (-598 – -392)</td>
<td>0.035</td>
</tr>
<tr>
<td>PFA-ADP (sec)</td>
<td>23.0 (17.0 - 48.0)</td>
<td>-0.5 (-15.5 - 5.0)</td>
<td>0.001</td>
</tr>
<tr>
<td>PFA-epinephrine (sec)</td>
<td>47.0 (26.5 - 63.5)</td>
<td>4 (-16.0 - 19.0)</td>
<td>0.001</td>
</tr>
<tr>
<td>PF4 platelets (IU/10^6 platelets)</td>
<td>3.2 (0.7 - 4.7)</td>
<td>-0.4 (-1.5 - 1.7)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

| p-values for paired t-tests for normally distributed data and paired Wilcoxon signed ranks test for non-normally distributed data
| due to non-normal distribution medians and interquartile ranges (IQR) are shown.
| Wilcoxon signed ranks test p< 0.051

Table 11.4. Changes in platelet parameters after 6 and 12 weeks of paroxetine treatment, relative to baseline, stratified for genotype.

Values represent changes of platelet parameters of combined observations after 6 and 12 weeks of paroxetine treatment (relative to baseline; 29 observations for 18 patients). Negative numbers indicate a decrease of the parameter.

* Mann-Whitney non-parametric test
Discussion

This study confirms that treatment with the SSRI paroxetine at 20 mg/day decreases platelet serotonin and platelet β-TG levels, and increases bleeding time. Paroxetine treatment did not affect other platelet parameters or bleeding tendency. Dose-escalation of paroxetine to 40-50 mg/day did not further change platelet parameters. Additionally, the 5-HTTLPR polymorphism of the patients strongly affected the effect of paroxetine on bleeding time, platelet serotonin and other platelet parameters or bleeding tendency. Dose-escalation of paroxetine to 40-50 mg/day did not change platelet serotonin level, and increased PFA-ADP, PFA-EPI and platelet PF4. Especially patients with the S'/S' genotype had a greater reduction in platelet serotonin level, and increased PFA-ADP, PFA-EPI and platelet PF4.

Previous studies did not investigate dose-escalation. Because of the large inter-individual difference of plasma SSRI levels at stable doses, proper investigation of a dose-response relation requires quantification of within-subject changes, instead of comparisons of (small) groups. By doing so, paroxetine at 20 mg/day already maximally prolonged bleeding time and decreased platelet serotonin, which was not altered by consecutive dose-escalation (to 40-50 mg/day). This is in accordance with the recent finding that SERT occupancy in the brain is not increased by dose-escalation.

Of interest are two in vitro studies that reported dose-response relations for sertraline and venlafaxine. Serebruany et al. reported a significantly prolonged in vitro PFA-ADP and PFA-EPI with sertraline concentrations that were claimed to mimic the plasma levels observed in patients using sertraline 50, 100 or 200 mg/day. Increased platelet aggregation with ADP, arachidonic acid, epinephrine and collagen was found in vitro, but with venlafaxine...
concentrations 1000-fold higher than usual plasma concentrations of treated patients. However, these in vitro studies incubated samples with antidepressants immediately before assessing platelet function. Normally platelet serotonin levels decrease in vivo only after >7 days, since the circulating platelet pool must be renewed before effects of SSRIs on platelet function are measurable. This raises concern about the validity of in vitro experiments.

The effect of the SERT polymorphism on platelet function

The effect of paroxetine on platelet function highly depended on the SERT polymorphism of the patients. In mice, the influence of the serotonin transporter can be studied in strains without (knockout Slc6a4<sup>−/−</sup>), with 1 (Slc6a4<sup>+/−</sup>) or with 2 alleles (Slc6a4<sup>+/+</sup>) of the SERT gene. In knockout mice the serotonin level in peripheral tissues (e.g. platelets) is <10% compared with non-knockouts. In Slc6a4<sup>−/−</sup>-mice (best resembling humans with S/S genotypes), peripheral serotonin levels were unchanged compared with Slc6a4<sup>+/−</sup> mice. Thus, in Slc6a4<sup>−/−</sup>-mice overall serotonin homeostasis can be retained, despite a 4 to 5 fold decrease in serotonin reuptake. Furthermore, a recent platelet aggregation study in Slc6a4<sup>−/−</sup>-mice showed 80% reduction in ADP-induced aggregation, not observed in Slc6a4<sup>+/−</sup>-mice, while results of this study also indicated a more prominent role of the requirement of ongoing SERT transport of serotonin in maintaining platelet aggregation. In other words, apart from the direct effect of SSRIs reducing serotonin granules in platelets, blockade of the SERT will also impair the functioning of the SERT at the platelet surface, which is required for the aggregation process itself. This might explain our findings: despite compromised serotonin reuptake in S'/S' subjects, platelets manage to achieve adequate, normal serotonin levels under normal circumstances. However, blocking the SERT by paroxetine might affect this compromised transporter system more than in patients with the most effective L<sub>A</sub>/L<sub>A</sub> genotype, resulting in larger decreases of serotonin, reduced aggregation capacity (increased PFA-ADP and PFA-EPI closure times), and longer bleeding times.

A recent study failed to find an effect of 5-HTTLPR polymorphisms on PFA-closure time. Consistent with our findings, this study reported no significant differences in bleeding tendency, but unfortunately platelet serotonin levels were not measured. Furthermore, in this study the bi-allelic variant of 5-HTTLPR was genotyped, while the reclassification of the ‘tri-allelic’ A to G SNP in the long allele as an S'-variant is probably more precise to detect differences.

Limitations

Due to the small number of patients, the results of dose-escalation and genotype must be interpreted with some restraint. However, despite the modest sample size of our study, the effects of the genetic variants on platelet serotonin and bleeding time were demonstrated in conservative nonparametric tests. We furthermore consider our findings as valid, because our bleeding time abnormalities and platelet serotonin findings for subgroups of SERT genotypes clearly coincide. We did not test changes in platelet parameters over time in control patients without SSRI, nor did we investigate dose-escalation in a randomized design. Additionally, our sample size did not allow further exploration of large versus small changes in platelet parameters within genotype groups.

Clinical relevance

Two potential clinical implications emerge. First, dose-escalation of paroxetine above 20 mg/day does not further impair platelet function, as it is already maximally impaired at a standard dose of 20 mg/day. In other words, increased bleeding tendency associated with SSRI use will occur irrespective of the administered dose. Second, and most important, the impairment of platelet function appears to depend on the common functional SERT promoter polymorphism. Although we only investigated paroxetine, we think this finding, if replicated, can be extrapolated to other SSRIs. The absence of an increased bleeding tendency or abnormal bleeding severity score may represent low sensitivity of the bleeding severity score to detect changes. However, this may also suggest that the observed decrease in platelet function is not clinically relevant in patients.
without trauma or surgery, unless a preexisting platelet abnormality is present. However, because of their enhanced antiplatelet response to SSRIs, patients with a \(<2^L_A\) genotype may be at increased risk of bleeding complications during surgery or while using concomitant anticoagulants.

With a low prevalence of clinically important bleeding complications associated with SSRI use, determination of the SERT polymorphism in all patients who start with an SSRI will not be cost-effective. However, for patients with a previous severe bleeding episode (either with or without SSRI use) genotyping may be considered. When an \(S'/S'\) genotype is found, switching to another, non-serotonergic antidepressant might be advisable. Generally, non-serotonergic antidepressants (i.e. mirtazapine, maprotiline, doxepin or bupropion) are associated with fewer hospitalizations due to bleeding complications than intermediate serotonin reuptake inhibitors (e.g. venlafaxine and amitriptyline; Odds ratio 1.9, 95% CI 1.1 to 3.5), and high degree serotonin reuptake inhibitors (fluoxetine, sertraline, clomipramine and paroxetine; Odds ratio 2.6, 95% CI 1.4-4.8), although calculated odds ratios differed for individual antidepressants. The decision to switch must then depend on the risks (e.g. oncoming surgery) and benefits (clinical response) for individual patients.

**Conclusion**

This study shows that the SSRI paroxetine already maximally impairs platelet function at 20 mg/day, with no overall effect on bleeding time, but significant decreases in platelet serotonin and \(\beta\)-TG levels. Dose-escalation does not further influence hemostasis. In addition, the effect of paroxetine on platelet function appears to be largely mediated by the common functional 5-HTTLPR SERT polymorphism, with patients with the \(S'/S'\) polymorphism having the largest hemostatic impairment with larger decreases in serotonin, and increases in PFA closure time and platelet PF4.

**Acknowledgements**

We wish to thank patients for their participation in the study. We thank general practitioners and psychiatric residents for appropriate referrals of patients.

**Conflicts of interests**

All authors reported no biomedical financial interests or potential conflicts of interest. This study was partly funded by a grant from the Netherlands Organization for Health Research and Development (ZonMw), program Mental Health, education of investigators in mental health (OOG; #100-002-002) to Henricus G. Ruhé.
References


Paroxetine induced platelet function changes
PART V

SUMMARY AND GENERAL DISCUSSION
SUMMARY, CONCLUSIONS AND GENERAL DISCUSSION
Summary

In this chapter, a summary of the conclusions of our studies and the discussion of these findings will be presented. Possible limitations and strengths of the thesis will be discussed and directions for future research will be presented.

Summary of this thesis and conclusions

This thesis addressed several questions:

1. **Is a short, easy to use clinician rated questionnaires as effective and precise as the routine Hamilton depression rating scale (HDRS)?**
   
   In chapter 3, we reanalyzed the treatment-outcomes of two antidepressant-psychotherapy trials. In line with previous reports, we found that the effect-sizes of two 6-item sub-scales of the HDRS (17-items) – the Maier and Bech sub-scales – were comparable to the original HDRS in the measurement of depression severity, and the sensitivity to measure changes. Furthermore, this comparability was stable across the full range of response to treatment, across both pharmacotherapy and psychotherapy, and for patients with different baseline severities of their depression. With an item response theory approach, we calculated a conversion table linking HDRS-scores and Maier and Bech scores, and determined cut-off points for remission for these subscales compared with conventional HDRS definitions. With these subscales clinicians can measure depression severity and clinical response more efficient than with the original HDRS.

2. **What is the evidence for dose-escalation as a strategy for non-response to a first SSRI?**
   
   In chapter 4, we presented a systematic review of the evidence for the dose response relationship of selective serotonin reuptake inhibitors (SSRIs) in major depressive disorder (MDD). In our literature-search, we identified 8 dose-escalation studies that increased dosages after at least 3 weeks of a standard dosage. Furthermore, 3 systematic reviews were published by then, which included three or four of the eight identified dose-escalation studies. Only one of the dose-escalation studies approached stringent methodological criteria. We found no evidence for increased efficacy of dose-escalation within the first 4 weeks of treatment. Dose-escalation after 6 weeks appeared less effective than continuing the same dose. We found some, but limited evidence for efficacy of dose-escalation after 8 weeks, particularly in partial responders. This effect was seen within 4 weeks after dose-escalation. Irrespective of efficacy, dose-escalation unequivocally increased side-effects, but effects on drop-out rates due to side effects were less straightforward. We therefore concluded that the available evidence for dose-escalation neither unequivocally confirmed its efficacy, nor deemed it ineffective. This conclusion was the starting point for the DELPHI-study, described in the chapters 2 and 7 to 10.

3. **What is the evidence for switching antidepressants as a strategy for non-response to a first SSRI?**
   
   In chapter 5, we presented a systematic review of the evidence for switching after failure of a first SSRI in MDD. In addition to our literature-search we included four studies released after these searches, of which three studies from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial. We identified eight randomized controlled trials (RCTs) and 23 open switch studies. Studies were of variable methodological quality and carried out in heterogeneous populations. The STAR*D results largely increased the amount, and quality of the available evidence, but did not show differential class effects to guide switching. Three randomized switch-studies investigating the switch from SSRI to another SSRI or SNRI (venlafaxine) were eligible for a meta-analysis. This meta-analysis showed that in the most favorable analysis, venlafaxine was slightly more effective as a second antidepressant,
compared with other SSRIs (number needed to treat (NNT) = 10 (95% confidence interval (95% CI) 6.3-33.3). We therefore concluded that after the failure of a first SSRI switching is open to all antidepressant classes (except irreversible MAO-inhibitors), without clear recommendations other than those that apply for the selection of initial treatment.

4. **Does the depletion of monoamine (5-HT and NA/DA) systems lower mood in humans, and is this lowering of mood different across different populations?**

In chapter 6, we presented a systematic review of monoamine depletion studies reporting mood effects of depletion. As an extension of previous systematic reviews of monoamine depletion studies, we pooled the results of 53 small-sized depletion studies with an adapted pooling technique (modified from conventional meta-analyses of RCTs to handle the statistically paired cross-over designs, and including an adjustment for small sample bias). Pooling is important because it quantifies the balance of positive versus negative studies.

We pooled 45 acute tryptophan depletion (ATD) studies and 8 acute phenylalanine/tyrosine depletion (APTD) studies, acutely lowering serotonergic and dopaminergic/norepinephrinergic neurotransmission, respectively. Serotonin or dopamine/norepinephrine depletion did not decrease mood in healthy controls, but slightly lowered mood in healthy controls with a family history of MDD. In drug-free patients with MDD in remission, a moderate mood decrease was found for ATD, without an effect of APTD. ATD induced relapse in patients with MDD in remission who used serotonergic antidepressants. MDD patients did not experience mood deterioration after ATD while depressed.

We concluded that simple, direct correlations of serotonin or dopamine/norepinephrine levels in the brain and mood do not exist. Because the serotonin or dopamine/norepinephrine depletion induced by ATD or APTD most clearly decreases mood in vulnerable individuals (patients who are in remission from MDD and healthy controls with a positive family history for MDD), we concluded that the monoamine systems are important systems in the vulnerability to become depressed. Moreover, the changes in brain metabolism in remitted patients who relapse after ATD or AMPT suggest that the serotonergic and norepinephrinergic systems give input to a final common pathway which needs further research to be clarified.

5. **Do MDD-patients and healthy controls differ in the number of central serotonin transporters, and is the amount of available SERTs correlated with depression severity?**

In chapter 7, we measured the availabilities of the central serotonin transporters (SERTs) in baseline SPECT scans of the DELPHI-SPECT participants, and compared these with SERT availabilities in age- and sex-matched healthy controls. Because of earlier reports, we investigated potential covariates as smoking behaviour, age, season of scanning and possible interactions with gender in multivariate models. We found interaction effects of smoking and diagnosis with gender for diencephalon SERT availability. For midbrain SERT availability, an interaction of gender with diagnosis was found, while season of scanning was a covariate. These interaction effects were expressed as lower SERT availability in midbrain and diencephalon in depressed males, and higher SERT availability in diencephalon in depressed (non-smoking) females compared with healthy controls. These findings point to complex effects of gender, smoking and season of scanning on the serotonergic system in healthy controls and patients with MDD, which must be considered in future studies comparing SERT availabilities in MDD patients versus healthy controls.

6. **Does a common genetic polymorphism of the promoter region of the serotonin transporter gene (SLC6A4) modify the association between the SERT occupancy by paroxetine and the clinical response?**

In chapter 8, we studied the relationship between clinical response after 6 weeks of paroxetine treatment and SERT occupancy (which is the change in SERT availability after 6 weeks of treatment relative to the pre-treatment SERT availability). Because clinical response to SSRIs is likely associated with polymorphisms of the serotonin transporter gene promoter region (5-HTTLPR), we also aimed to investigate whether this relation of...
response and SERT occupancy was modified by the 5-HTTLPR genotype. We obtained study-entry and week 6 SPECT scans for 44 patients treated with paroxetine 20 mg/day, of which 42 scans were analyzable (10 responders and 32 non-responders).

We found that for all patients, SERT occupancy was not associated with clinical response (expressed as the proportional decrease in HDRS relative to study entry). However, when we grouped patients by 5-HTTLPR polymorphism, we found that in patients with the (favorable) L_A/L_A variant a significant relation between SERT occupancy and clinical response (absolute and proportional change in HDRS) existed. A previous study found a better amygdala-cingulate coupling in subjects with an L-allele relative to the unfavorable S/S variant, probably reflecting differences in the development of the serotonergic system induced by 5-HTTLPR. As an explanation of our findings, we therefore hypothesized that patients with the 5-HTTLPR L_A/L_A genotype have a more flexible serotonergic system, which is more easily influenced by serotonergic antidepressants.

7 Is dose-escalation of paroxetine an effective clinical strategy for non-response in MDD? In chapter 9, we studied the strategy of dose-escalation which was found equivocally efficacious in the systematic review, as described in chapter 4. In this study, we addressed the methodological flaws found in previous dose-escalation trials, but moreover, we also measured whether paroxetine dose-escalation increased SERT occupancy more than placebo dose-escalation. After inclusion of 49 SPECT patients (with successful follow-up of 31 patients participating in the randomized placebo-controlled dose-escalation SPECT study), we performed an interim analysis including all randomized patients (n= 57), and tested differences between placebo and true dose-escalation with pre-defined cut-off values for futility and superiority. We found no clinical benefits of true dose-escalation compared with placebo, and so the trial was stopped because of futility.

Most important, our data also showed that true dose-escalation of paroxetine did not increase SERT occupancy more than placebo dose-escalation, probably because a plateau was reached at standard doses already. We concluded, that in line with previous, though more equivocal evidence, dose-escalation is not beneficial in MDD. Therefore, two clinical options remain for the treatment of MDD-patients who do not respond to standard doses: either continuation of treatment until 10 weeks while waiting for a potential delayed response, or a switch to a different and potentially more effective treatment strategy (e.g. another SSRI or SNRI or TCA, or psychotherapy, or a combination).

8 Does treatment with paroxetine normalize amygdala hyperactivation in MDD? In chapter 10, we describe the (negative) faces task of the DELPHI-fMRI sub-study. We studied the effects of paroxetine treatment on the hyperactivation of the amygdala (and other brain areas) during a depressive episode. We specifically investigated: 1) whether amygdala-activation by (negative) facial expressions differed from healthy controls, 2) whether amygdala activation changed after 6 and 12 weeks of treatment, 3) whether the amygdala activation merely changed by paroxetine treatment or in relation with clinical response, and 4) whether dose-escalation of paroxetine in week 6 non-responders affected activations, compared with placebo-dose-escalation. Apart from the amygdala, we also investigated these questions for other brain areas. We performed fMRI scans in 22 MDD patients and 21 age- and sex-matched healthy controls (controls were scanned once).

Compared with healthy controls, we found increased activations in bilateral (extended) amygdala and left insula in MDD-patients. In contrast with previous studies, we found an increase in right amygdala activation after 6 weeks of treatment, which was attributed to the high number (12/20) of treatment non-responders. This increased amygdala-activation was reduced thereafter until week 12. When we analyzed the differences in week 6 and 12 treatment responders versus week 6 and 12 non-responders, bilateral amygdala activation was reduced in responders versus non-responders. Amygdala activations in week 6 and 12 were associated with HDRS-scores. Furthermore, non-responders had higher activations in
right orbitofrontal cortex (OFC) and insula, while treatment responders showed higher 
activation in right dorsolateral prefrontal cortex (DLPFC) and left nucleus accumbens. These 
differences were non-existent at study-entry.

Although randomized groups became small, changes over time in the true dose-escalation 
group showed decreased activations in ventral and dorsal regions, while activations of the 
right hippocampus and left subgenual cingulate increased, when compared with changes 
observed after placebo dose-escalation. We concluded that in line with previous treatment 
studies46-59 paroxetine treatment over 12 weeks reduced activations in the amygdala and other 
(ventral) emotion-generating structures, and increased cortical function (in the dorsal 
regulatory structures; see chapter 1, pages 18 and 21), especially in treatment responders. 
These findings may point to increased fronto-limbic control as a mechanism of paroxetine 
drug-response effects, which was supported by findings of other groups.60;61

9 What are the changes in hemostasis and blood platelet parameters when patients are treated 
with paroxetine, and are these changes modified by dose-escalation or a genetic 
polymorphism of the promoter region of the serotonin transporter gene?

In chapter 11, we present the results of a study investigating an infrequent, but dangerous 
 adverse effect of SSRIs: increased bleeding tendency.62 In this study, we measured platelet 
parameters and coagulation, while we applied a secondary (non-randomized) dose-escalation 
paroxetine treatment non-responders. We additionally investigated whether 5-HTTLPR 
polymorphisms modified these effects on platelet parameters. We found that a standard dose 
of paroxetine (20 mg/day) already significantly decreased platelet serotonin levels and 
platelet β-thromboglobulin (β-TG), without further decreases after dose-escalation. 
Moreover, we found that 5-HTTLPR polymorphisms modified these effects: compared with 
LA/LA-carriers, bleeding time significantly increased in <2LA-allele carriers, and platelet 
serotonin decrease was larger in patients without LA-alleles. Furthermore, the platelet 
f Function analyzer closure time significantly increased in patients without LA-alleles. Although 
the observed bleeding tendency presumably is not clinically relevant in patients without 
trauma or surgery (unless a preexisting platelet abnormality is present),63 these findings are 
applicable to patients with a previous severe bleeding episode. For them genotyping may be 
considered; when an S'/S' genotype is found, switching to another, non-serotonergic 
antidepressant might be advisable.

Clinical relevance

Antidepressants have comparable efficacy, however only 50% of the MDD-patients respond to the 
first antidepressant trial given, while fewer achieve full remission of symptoms.54;65 When the 
miracle doesn’t happen, there are 5 strategies for non-response: prolongation of the initial trial, 
dose-escalation, switching to another drug, augmentation with another drug or a combination of 
antidepressants. This thesis addresses switching strategies23 and dose-escalation.11;51

Switching

Despite a small benefit of a switch to venlafaxine compared with a second SSRI (in the most 
favorable analysis only), we concluded that switching-options after a first SSRI are open to all 
antidepressant classes, without clear recommendations other than those that apply for the 
selection of initial treatment. One more RCT comparing venlafaxine and citalopram,66 and three 
open studies, one with venlafaxine (with randomization to different doses)67, one with a 
venlafaxine-mirtazapine combination strategy,68 and one with duloxetine69 appeared thereafter.

In addition, another meta-analysis was published, comparing a non-SSRI switch (venlafaxine 
(3 studies) or mirtazapine (1 study) or bupropion (1 study)) versus a second SSRI.70 This meta-
analysis included 2 different studies (abstracts from symposia)71;72 compared with our meta-
analysis, and excluded one study with questionable methodology, like we did in our most
favorable analysis.²⁴ Their conclusion was that the NNT for remission was 22 (confidence interval not given), modestly in favor of a non-SSRI switch, which is far below the standard for clinical relevance (NNT ≤ 10) as suggested by the United Kingdom’s National Institute of Clinical Excellence.²⁵²⁶ When they considered response as an outcome, they found no significant benefit, comparable to our results.²⁷ In the final manuscript of the previous abstract,²⁸ Lenox-Smith and Jiang reported increased efficacy of venlafaxine over citalopram (flexible doses) in a subgroup of patients with severe MDD (HDRS₁₉ > 31), who participated in a multicenter RCT of SSRI non-responders.²⁹ However, this difference was only significant for the continuous scores of the HDRS₁₉ (difference approximately 4.7 points; SD not given), and not for the remission rates. Therefore, we think that only at the level of public health, the difference between switching-strategies might be interesting, but not at the level of individual patients.

**Dose-escalation**

In our systematic review of dose-escalation it was concluded that the available evidence neither unequivocally confirmed its efficacy, nor deemed it ineffective. This was in agreement with another review, performed by an independent group, and published just before ours.³⁰ The results of the dose-escalation RCT in this thesis first of all addressed the methodological shortcomings of previous studies, and was stopped for futility after an interim analysis.³¹ Even though the numbers of patients randomized to placebo or true dose-escalation were modest (n= 27 and n= 30, respectively), the changes in HDRS-scores and subscales were far from significantly different, while the level of this significance was far above the a priori cut-off for futility (which was undisclosed at the time of analysis). This finding aligns with the dose-escalation study by Licht et al.³² Unfortunately, the premature termination of the trial precluded planned secondary analyses to distinguish subgroups who still might benefit form dose-escalation (e.g. partial responders; defined as 25-50% reduction of HDRS).³³⁻³⁴ In our sample the subgroup of partial responders consisted of 24 patients (from 60 randomized; 40%), and did not show any indication of beneficial treatment-effects for true dose-escalation. Nevertheless this observation cannot replace true subgroup analyses in larger samples.

Our sub-study of the changes in SERT occupancy during the randomized, placebo-controlled dose-escalation phase further corroborates our clinical findings. We therefore conclude that our systematic review, our dose-escalation study and our imaging study provide substantial evidence for clinicians, that dose-escalation is not an effective strategy for depressed patients who do not respond to 6 weeks of a first SSRI at a standard dose.

In a recent meta-analysis of 9 placebo-controlled, fixed-dose, dose-finding studies of SSRIs,³⁵ Papakostas et al. found a modest but statistically significant difference of 4% higher response rates in patients who started with higher than standard doses of citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine and sertraline (p= 0.04). This 4% difference in response rates corresponds with a NNT of 25, which is clinically not impressive, nor relevant. In their meta-analysis, Papakostas et al. could not investigate whether this difference was attributable to a single SSRI. However, when inspecting their data, all paroxetine trials did not show benefits of higher paroxetine doses compared with standard 20 mg/day doses, with relative risks of 1.00 (95% CI 0.67⁻1.50), 1.02 (95% CI 0.69⁻1.52) and 0.94 (95% CI 0.65⁻1.35). In contrast, the rate of discontinuation due to adverse events was also higher in the patients treated in the above standard doses group (16.5% compared with 9.8%). Therefore, to our opinion, it remains also questionable whether higher doses of SSRIs at the initiation of treatment indeed are advantageous to MDD-patients. Instead, they could give patients more adverse effects and preclude prolonged exposure to tolerable standard doses. Finally, but most important for this thesis, these findings presumably hardly apply to paroxetine.

**Other findings of clinical relevance**

Besides the switching and dose-escalation strategies, two other relevant points from this thesis emerge. First, we showed that the Maier and Bech subscales of the HDRS₁₇ are equivalent to the full HDRS₁₇.³⁶ Others even suggested that these subscales might measure the core-symptoms and
severity of MDD more specifically, and their use might even be more efficient,23 because of the reduction of noise by less responsive items.26 We therefore propose that the Maier or Bech-subscals should be used in the clinical assessment of (treatment of) MDD patients, instead of, or besides self-rating scales like the Inventory for depressive symptoms self-rated (IDS-SR) or Beck depression inventory (BDI). This will be of great importance in patients who cannot reliably fill out self-rated questionnaires, like in MDD with psychotic features or MDD with severe cognitive dysfunction. Of course, providing a shorter, easy to use clinician rated scale will not translate in direct implementation of one of these subscales, but may at least reduce the time burden of rating scales for physicians.

Second, our findings regarding the increased bleeding tendency in patients merit attention.62 Increased bleeding tendency, especially in combination with non-steroidal anti-inflammatory drugs (NSAIDs) was observed before.77-82 This even led to the recommendation to use TCAs when concomitant NSAIDs should be used in the guideline for Dutch general practitioners.83 However, the serotonergic TCA clomipramine, presumably has comparable effects. Because of the relatively low incidence of dangerous gastro-intestinal bleeding complications in general (4.3/1000 SSRI treatment years; 14.5/1000 SSRI+NSAID/aspirin treatment years; relative to 1.2/1000/year in untreated humans),80 the increased bleeding tendency by SSRIs is of no clinical relevance in patients without trauma, surgery or a previous severe bleeding episode.63 However, in patients with a previous severe bleeding episode, or in patients who need elective surgery, and maybe also in patients regularly using NSAIDs, genotyping of the 5-HTTLPR may be considered. When an S'/S' genotype is found, switching to another, non-serotonergic antidepressant might be advisable. For which of the aforementioned subgroups this genotyping will be cost-effective, remains to be investigated.

Neurobiology of the treatment of major depressive disorder

With our studies we also aimed to investigate the neurobiology of paroxetine treatment effects in MDD in vivo. The enhancement of serotonergic neurotransmission by paroxetine is undoubted, caused by a reduced serotonin uptake after SERT occupancy as the primary mechanism. Therefore, the most important question which we addressed was whether dose-escalation of paroxetine would increase SERT occupancy compared with placebo.

SERT occupancy after short or prolonged treatment with SSRIs was investigated before 84-96 even at variable and higher doses.86;96 However, without a pre-assessment of SERT availability before treatment96 and without randomization to a placebo-controlled dose-escalation,86 one can never exclude the possibility that non-responders after 6 weeks of a standard dose consist of a selected sample who had lower SERT occupancy. This hypothetically lower SERT occupancy could then be considered explanatory for non-response, and might increase by subsequent dose-escalation compared with placebo dose-escalation.

Our SPECT measurements in the subgroup of the DELPHI study showed no overall relation between SERT occupancy and clinical response in the open phase of treatment with paroxetine over 6 weeks. Furthermore, we found no evidence for the postulated 80% SERT occupancy as a requisite for clinical response.45 In addition, the second, randomized, placebo-controlled dose-escalation phase of DELPHI showed no differences in changes in SERT occupancy between placebo and true dose-escalation for 6 more weeks.51 With this latter finding, we are the first group that studied sequential dose-escalation in non-responders, controlling for placebo-effects and measurement-bias by repeated SPECT scans. Interestingly, we did find an association between SERT occupancy and clinical response in the subgroup of patients with the (tri-allelic) L_A/L_A polymorphism of the 5-HTTLPR. Because the tri-allelic L_A:L_C ratio is approximately 5:1,97 our finding aligns with data from previous treatment studies which found increased response rates in subjects with the bi-allelic L/L 5-HTTLPR polymorphism.46;47
However, the mechanism why increased SERT occupancy is associated with increased clinical response in LA/LA carrying patients, cannot be explained by higher SERT occupancies in LA/LA patients.45 Probably, the explanation must be sought in differences in the function of the serotonergic system (and its development) mediated by 5-HTTLPR polymorphisms. From previous lines of research, it may be concluded that the S-allele predisposes to a more reactive arousal system, and that the L-allele accounts for more flexibility and/or the possibility to react and adapt better to external changes imposed on the subject.98 Furthermore an elegant fMRI-study showed that both the anatomy and functional connections between limbic, subcortical and cortical regions are mediated by 5-HTTLPR polymorphisms, with S/S carriers showing relative uncoupling of the amygdala-pregenual cingulate connectivity.48

The biochemical consequences of chronic administration of SSRIs are thought to include an increase in extracellular levels of serotonin followed by neuroadaptive alterations in serotonin receptors and postsynaptic intracellular signalling pathways (see figure 1.2, page 20), as well as time-dependent effects on neurogenesis.99;100 Not surprisingly, animal studies demonstrated that in mice knocked out for the murine Slc6a4 SERT transporter (Slc6a4 +/-), fluoxetine and fluvoxamine are ineffective on behavioural and other measures, while mice with at least one copy of the allele (Slc6a4S or Slc6a4L) showed the typical antidepressant effects.99 This is pivotal evidence for the working mechanism of SSRIs, for which SERTs are apparently required to be present. The Slc6a4S mice have ~50% less expression and transport function and can be viewed as a model for humans with the S/S genotype.101 Therefore, the findings of almost identical effects of SSRIs in Slc6a4S or Slc6a4L mice102 do not support the findings of differential effects for 5-HTTLPR polymorphisms from meta-analyses of human studies.46;47 However, the 5-HTTLPR polymorphism is only present in humans and higher order primates (e.g. macaques). To our knowledge, no studies in primates investigated the mediating effects of 5-HTTLPR polymorphisms on the relation between SERT occupancy and behavioural effects. Furthermore, we are unaware of studies investigating dose-response relationships of SSRIs in animals. Finally, although a robust increase in anxiety-like behaviour was found in Slc6a4S/mice compared with other genotypes, so far no studies were performed to investigate the differences in brain development and functional connections for mice with a Slc6a4S, Slc6a4L or Slc6a4L genotype.103

Serotonin has an important role in the regulation of neuronal information processing by constraining (instead of inhibiting) the reactivity of the brain to internal and external stressors.104 In general, serotonin appears 1) to prevent an overshoot of other dynamic systems (e.g. constraining the release of dopamine after rewarding stimuli), and 2) to control the sensitivity of the system to perturbation by new elements entering the system (e.g. constraining a natural propensity to switch to alternate behaviours; this can be provoked by serotonin depletion). This constraint can be experimentally challenged by acute tryptophan depletion (ATD), which results in increased perception of aversive stimuli, increased food intake, increased sexual, aggressive, depression-like and anxiety-like behaviours, but also decreased sensitivity to cues of punishment, and less flexibility to change behaviour in rats.104

Besides the cognitive effects,98 our meta-analysis quantified the mild mood lowering effects by ATD in humans, which were especially seen in healthy controls with depressed relatives, and patients who had remitted from a previous MDD-episode.32 Because mood-effects can be considered as more complex behaviour, these may be more difficult to provoke in relatively short-term ATD studies compared with longer-lasting studies in animals. Nevertheless, the large body of research on the serotonergic system has now changed the categorical model –decreased serotonin is specific for depression – into a model in which alterations in the serotonergic neurotransmission can be seen as a biological risk-factor, which interacts with innate and/or external factors. This so-called serotonergic vulnerability is now seen as neither a sufficient, nor necessary factor to develop and maintain MDD, but in combination with other factors increases the risk for MDD.98
Our fMRI-study,\textsuperscript{54} corroborates the hypothesis that paroxetine treatment interferes with the (maladaptive) functioning of various interconnected regions in the limbic-subcortical-cortical network.\textsuperscript{105,106} Like others, we found increased ventral\textsuperscript{*} activations (e.g. in the amygdala and insula), and decreased dorsal\textsuperscript{*} activations (anterior cingulate, DMPFC, DLPFC) in MDD-patients relative to healthy controls. We furthermore found that the ventral hyperactivation and dorsal hypoactivation was normalized in treatment responders, but not in non-responders. Recently, Chen et al. demonstrated increased functional coupling of the amygdala with the frontal and (pregenual) cingulate cortex (and striatum and thalamus) after 8 weeks of sertraline treatment.\textsuperscript{61} Although we did not investigate this coupling per se, we found that the amygdala responses in our patients were correlated with pregenual cingulate and frontal regions.

Furthermore, in responders we found increased activation in the nucleus accumbens, pointing to the involvement of the limbic-cortico-striato-pallido-thalamic network, which indirectly suggests the involvement of changes in dopaminergic neurotransmission after paroxetine treatment.\textsuperscript{107} Two coinciding explanations can be given for this increased activation of the nucleus accumbens. First, in rodents, chronic treatment with antidepressants (among others SSRIs) potentiated dopamine transmission, due to an increased sensitivity of postsynaptic dopamine receptors and possibly also due to a decreased sensitivity of presynaptic (inhibitive) dopamine autoreceptors, preferentially in the limbic system.\textsuperscript{108} Second, again in rodents, decreases in serotonin are associated with a release of normally inhibited behaviour, including impaired suppression of behaviour inducing punishment.\textsuperscript{104} Activation of the nucleus accumbens is associated with anticipation of reward,\textsuperscript{109} while the nucleus accumbens is deactivated by loss or punishment. Therefore, it might be postulated that the increase in serotonergic neurotransmission by SSRIs, reduces a lack of constraint on dopaminergic neurotransmission in the nucleus accumbens, with subsequently less punishment perceived. This would then result in less deactivation of the nucleus accumbens. The fact that this was demonstrated in treatment responders relative to non-responders might suggest that in non-responders this enhanced dopaminergic constraint might not be achieved.

As interesting, but contradictive finding, we did not find any indication that dose-escalation increased SERT occupancy, but at the other hand, we found indications that a true dose-escalation decreased hyperactivations in the ventral compartment (OFC, insula, VMPFC, ventrolateral thalamus), and increased activation in the dorsal compartment (hippocampus, DLPFC, and parietal cortex). Also we counter intuitively found an increased activation in the subgenual cingulate after dose-escalation.\textsuperscript{54} As explained, we could unfortunately not disentangle response- and drug-effects. If the decreases in the ventral, and the increases in the dorsal compartment could have been attributed to responders to dose-escalation, while the increased activation in the subgenual cingulate was associated with non-responders (despite dose-escalation), this would have made our findings less surprising.

Limitations

Limitations of the studies in this thesis were discussed in each chapter. However, some must be recapitulated in this general discussion.

At the moment of publication of this thesis, the systematic reviews presented in the chapters 4, 5, and 6 might have become outdated, since the literature searches were performed until October 2006\textsuperscript{32} or February 2005.\textsuperscript{11,23} Although we did not repeat the searches until July 2008, for the dose-escalation and switching reviews, we commented on additional studies that we encountered thereafter in this general discussion. We did not review additional monoamine depletion studies, as this should have resulted in new meta-analyses.

\textsuperscript{*} For explanation of the ventral and dorsal compartments of the limbic-subcortical-cortical network, see chapter 1, page 18.
Furthermore, in contrast with other authors, we did not perform a meta-analysis for the dose-escalation studies, and only for the venlafaxine versus second SSRI switch studies. At this point, we considered the dose-escalation studies with various moments of randomization after the initiation of treatment too heterogeneous to pool. For switching, we decided not to pool (by then unpublished) data on various dual acting antidepressants versus second SSRIs.

In the SPECT-studies, we used $^{[123]}\beta$-CIT for SPECT imaging, which is a non-selective radioligand, and also binds to the dopamine transporter (DAT; e.g. substantia nigra) and norepinephrine transporter (NET; e.g. locus coeruleus). Nowadays, selective SERT ligands like $^{[11C]}$DASB for PET or $^{[123]}$ADAM for SPECT are available. However, previous imaging study in primates, showed that $^{[123]}\beta$-CIT uptake in midbrain and diencephalon predominantly reflects SERT, as these structures are rich of SERT relative to DAT and NET. Although this non-selectivity might have slightly concealed changes in SERT occupancies due to additional DAT- or NET-binding, we think our findings in diencephalon and midbrain mainly reflect SERT occupancy. Furthermore, if one would consider non-selectivity as relevant, the use of SPECT-scans at randomization as a reference for changes in occupancy will further reduce the magnitude of this non-selectivity bias, because non-selectivity will be measured systematically between T0 and T1. Despite this rationale, in future studies, a selective ligand for SERT ($^{[11C]}$DASB or $^{[123]}$ADAM) should be considered. Furthermore, it would be a challenge to replicate our study using one of these selective ligands.

Although the region-of-interest (RoI) approach to determine non-displaceable binding potential (BPND) and SERT occupancies is a valid and widely accepted method, we did not co-register individual SPECT images with individual structural MRI scans. This might have reduced signal-to-noise variance in our measurements. However, because we applied a randomized, double-blind dose-escalation, and analyzed the SPECT scans while blinded for the moment in treatment (T0 or T1) and intervention, we think that this RoI approach only affected measurement-bias, and did not introduce differential bias between the intervention groups.

Scientific studies are powered to investigate major questions, but often additional questions will be explored. In the case-control study of SERT availability of MDD-patients versus controls, only small subgroups remained for effect-modifying variables, which might have increased the chance of detecting spurious results. We therefore presented these differences as exploratory analyses only. For the SERT occupancy-response relation by genotype interaction, we also compared small groups, but nevertheless found significant results, which we consider valid. Nevertheless, replication in a larger sample would strengthen these results. Because of the interim-analysis, we stopped our trial for futility. Thereafter, our modest sample size of the randomized dose-escalation for example precluded proper subgroup analyses. Our fMRI-study, which was relatively large for imaging studies found robust effects. However, the comparison of the dose-escalation versus placebo dose-escalation in our fMRI study was hindered by a modest number of patients in the true dose-escalation group. This precluded the investigation of the 3-way interaction intervention by time by response, which could have properly answered which effects were drug-induced. Therefore, as indicated, these results must be considered with some restraint. Finally, in the hemostasis study, we studied 19 patients, but did so in a within-subject design, which has more power to detect changes. Again the number of subjects receiving a dose-escalation was modest, which might have precluded changes in other platelet parameters with smaller effect-sizes.

A general limitation is that our neurobiological investigations did not investigate adaptive neuronal effects caused by paroxetine. We either investigated the pre-synaptic effects of dose-escalation on SERT occupancy, or the changes in the final common pathway: the changes in activations of the limbic-subcortical-cortical network. Many adaptive pre- and post-synaptic effects of SSRIs have been documented, mostly in animal studies, but also in humans. We did not investigate any of these specific secondary effects of SSRIs. The radiation burden for scientific research did not allow additional radioligand scans of other receptors (e.g. 5-HT$_2$A or 5-HT$_1$A) or other monoamines (e.g. endogenous dopamine release). Furthermore, additional challenges (e.g. gepirone to assess 5-HT$_1$A receptor sensitivity) were considered too burdensome in these patients. We did not perform animal studies to investigate dose-escalation,
for example in microdialysis experiments or studies investigating SERT occupancy, SERT down-regulation, or other secondary effects of serotonergic neurotransmission on receptors and/or other pathways.

Nevertheless these limitations, one should bear in mind that our, and previous clinical studies did not support dose-escalation as a rational strategy for SSRI non-responders in MDD. Therefore, the merits of additional dose-escalation studies would be to provide fundamental insights in working mechanisms.

Finally, this thesis might also make clinicians wonder why initial, high doses of SSRI have been found effective in (some) fixed-dose, dose-finding studies in anxiety disorders, and especially in obsessive compulsive disorder (OCD). Of course, anxiety disorders are beyond the scope of this thesis, but this question remains intriguing and might warrant an approach as we did for MDD.

Strengths

First of all, a major strength of this thesis is its clinical starting point. All psychiatrists recognize the problem of non-response to a first SSRI, and most of them increase the dose as a first strategy, followed by switching. As such, this thesis covers an important clinical topic, which could be very relevant for future treatment guidelines.

Our systematic reviews were conducted in accordance with the stringent guidelines of the Cochrane Collaboration. We performed ‘sensitive’ literature searches in several databases, decreasing the chance of missing valuable studies. All studies were critically appraised and abstracted. We only performed meta-analyses when clinical heterogeneity allowed pooling, a methodological requirement which is often violated by referring to a random effects approach. In the obvious heterogenous populations investigated in the monoamine depletion review, we achieved better homogeneity by stratification. In this latter meta-analyses, we acknowledged the statistical problems of pooling crossover studies (having dependent measurements) and adapted the standard statistical approach to solve this problem.

The empirical studies in this thesis were designed carefully and the methodology of the dose-escalation trial was based on a thorough review of the previous dose-escalation studies. We therefore randomized patients after sufficient time to expect not many further delayed clinical responses. Also, we gradually increased doses which obviously prevented early drop-out due to intolerable adverse effects. Most importantly, our SERT occupancy SPECT protocol was unique, and provided the rationale why dose-escalation was ineffective.

With the fMRI-study, which could be started after a second grant was obtained, we looked beyond SERT occupancy to investigate the effects of paroxetine treatment on emotion-regulation. Although the dose-escalation results of this study are not unambiguous, this study is only the sixth fMRI-study investigating the treatment-effects of SSRIs. The fact that we scanned patients three times, and that many patients initially were treatment non-responders allowed us to investigate the distinction in brain activations between drug-effects and clinical response, which is an important addition to the existing literature.

Finally, we also looked at peripheral effects of SSRIs, far beyond the brain, but also considering the inhibition of SERTs by paroxetine in association with hemostasis. Increased bleeding tendency is an infrequent but potentially dangerous adverse effect. We showed that even standard doses of paroxetine might affect hemostasis, and suggested subgroups who are eligible for secondary prevention by genetic screening of the 5-HTTLPR.
Future research

Future studies with data from DELPHI

Several lines of research on data from the DELPHI-study still need to be worked out. We collected salivary samples at study-entry, T0 and T1 (see chapter 2, page 36). We will investigate the changes in cortisol by treatment over time, and might be able to associate these changes with clinical response or other clinical or imaging (e.g. SERT) variables. We therefore will collaborate with the group of Kirschbaum and Rohleder. In addition, we collected blood-samples at study-entry, T0 and T1, to determine ex-vivo SERT and norepinephrine inhibition by paroxetine. Finally, we determined omega-3 and omega-6 polyunsaturated fatty acids, again at study-entry, T0 and T1, which will be analyzed in relation with clinical response to paroxetine, and additionally in relation with SERT availability and occupancy.

Furthermore, we are currently working on a sub-study which investigates the association of midbrain and amygdala SERT occupancy with the changes in amygdala-activation after (negative) facial expressions in the patients who underwent both SPECT and fMRI scans.

Future clinical studies

The findings presented in this thesis may be the starting point of new research.

- For better evidence-based recommendations about when to choose between switching, augmentation, combination, or psychotherapeutic strategies as a next step, comparisons of these strategies relative to each other are necessary. This will require future algorithm-based switch studies, including psychotherapy and augmentation or combination strategies to improve our knowledge to guide treatment for SSRI non-responders. In these studies, one arm should be to continue treatment (e.g. for another 6-8 weeks, depending on the timing of switching) as a ‘placebo’ control. Of course the somehow disappointing results from STAR*D should be reconsidered when such a major endeavor is set up again.

- As a smaller enterprise, the relative benefits of switching after 6 weeks versus a continuation of the same treatment for another 6 weeks would provide important guidance to clinicians. Quitkin et al. showed that MDD-patients treated with fluoxetine 20 mg/day who were non-responders at week 6, showed remission rates at week 12 between 31%-41%, indicative of a substantial delayed remission. However studies directly comparing such prolonged treatment versus switching have not yet been performed in MDD.

- The controversial evidence advocating initial, high doses of SSRI in panic disorder, posttraumatic stress disorder, OCD and other anxiety disorders necessitates a well-performed meta-analysis of fixed-dose, dose-finding studies in anxiety disorders, stratified by disorder and drug. In addition, dose-escalation studies in these disorders would be interesting, especially when combined with the investigation of neurobiological mechanisms by neuroimaging. For example, the effects of a randomized, placebo-controlled dose-escalation on dopaminergic D2-like receptors, or dopamine transporters in OCD would be interesting.

Future neurobiological studies

In order to develop more successful treatment approaches, we need to understand better which effects of treatment are required for recovery, and which biomarkers are associated with treatment response. Therefore, more fundamental, neurobiological studies should aim at several gaps in the evidence.

- In order to elucidate SSRI working-mechanisms, the increased coupling between amygdala and frontal and cingulate cortices must be investigated with other SSRIs. Secondly, it should be investigated whether 5-HTTLPR polymorphisms mediate this coupling, or the increase after SSRIs, or both. If the magnitude of coupling would be mediated by 5-HTTLPR polymorphisms, this would also corroborate our findings for the association between SERT occupancy with clinical response in L/L carriers.
- In order to study secondary effects of SSRI treatment, neuroimaging approaches to quantify the effects on 5-HT\textsubscript{1A} receptors during treatment and after dose-escalation are warranted in patients. Although a desensitization of 5-HT\textsubscript{1A} was demonstrated for chronic SSRI treatment in animals,\textsuperscript{118,153} the question remains whether this desensitization is also associated with 5-HT\textsubscript{1A} downregulation and/or whether dose-response effects occur. This study should ideally be performed in combination with gepirone-challenges at different moments during treatment to assess 5-HT\textsubscript{1A} receptor sensitivity.\textsuperscript{125}

- In order to investigate secondary effects of serotonin, e.g. on endogenous dopamine release in the nucleus accumbens and other brain regions, microdialysis experiments in animals, chronically treated with SSRIs, will further investigate dose-escalation and time dependent changes in SSRI treatment.

- The hypothesis of the involvement of dopamine in mood and MDD must be investigated.\textsuperscript{107} Serotonergic antidepressants indirectly influence the dopaminergic system. Responders to amitriptyline showed decreased \([^{123}I]IBZM\) SPECT binding to striatal dopamine D\textsubscript{2} receptors, relative to pretreatment.\textsuperscript{154} This decrease correlated with a decrease in HDRS-score. This might represent an increased tonic dopamine release after amitriptyline, but may also be attributed to improved psychomotor activity in responders. Nevertheless, a failure to achieve alterations in the dopaminergic system by antidepressants, might explain non-response or residual symptoms (sleep disturbances, diminished pleasure, loss of interest, fatigue, and decreased motivation), which are often observed in clinical practice.\textsuperscript{107} Therefore, investigation of the dopaminergic effects of SSRIs, in combination with or versus behavioural approaches (re-activation) in combined fMRI and SPECT/PET studies in (retarded) MDD-patients are necessary.

- The effects of SSRIs on neurogenesis also merit further research. After prolonged antidepressant administration in rodents, increased hippocampal neurogenesis was found.\textsuperscript{121,123} The second-messenger effects of SSRIs upregulate cAMP responsive element binding protein (CREB) in the neuron’s nucleus, which consequently regulates CREB-directed gene transcription.\textsuperscript{197,155} To date, it is unclear which specific genes transcription are affected by CREB. Therefore the exact role of second messenger systems (and CREB in particular) in response to SSRIs remain to be elucidated.\textsuperscript{117} One of the genes that is (positively) influenced by CREB is the brain derived neurotrophic factor (BDNF) gene. BDNF is a plethoric factor, which regulates neuronal survival, migration, phenotypic differentiation, axonal and dendritic growth and synapse formation. As such BDNF has a role in cognitive functions (e.g. in memory and hippocampal plasticity).\textsuperscript{100} BDNF and serotonin are closely interlinked: BDNF promotes production of serotonin, upregulates serotonin uptake/release, and modifies firing rates of neurons in the raphe nuclei, while serotonin upregulates BDNF. This stimulating feedback mechanism may be involved in the selection of useful synapses, which could be hypothesized to be a requisite for clinical response.\textsuperscript{119} Few studies investigated the long-term changes in BDNF measured in venous blood of MDD patients treated with antidepressants.\textsuperscript{156-158} Although different effects for different drugs were seen, generally an increase in BDNF after 6 months was found, significantly associated with a decrease in HDRS. Nevertheless, the role of BDNF as biomarker for antidepressant response is equivocal. Furthermore, due to dilution, peripheral measures may not be adequate, so sampling from an internal jugular vein might be better.\textsuperscript{159} Therefore, this line of research might need development of the methods to measure BDNF (and CREB) more locally in the brain of humans.
Conclusion

This thesis addresses the pharmacological treatment of Major Depressive Disorder (MDD), and focuses on what to do when patients do not respond to a standard dose of an antidepressant. In two systematic reviews, we summarized the evidence for various switch strategies and dose-escalation, and concluded that for both strategies insufficient evidence existed to unambiguously guide clinicians. We thereafter performed a randomized controlled trial in which we studied dose-escalation in depressed patients who had not responded to 6 weeks of paroxetine (20 mg/day), and combined clinical measurements with SPECT and fMRI neuroimaging in subgroups of the participants (DELPHI-study). With the DELPHI-study, we show that dose-escalation of paroxetine has no clinical benefit over placebo dose-escalation, and furthermore quantify the absence of a pharmacological effect on the target of paroxetine: the SERT.

Another systematic review and four studies in this thesis, investigate the role of monoamines and SERT in the etiology of MDD and the neurobiological effects of pharmacological treatment with paroxetine. From these studies we conclude that MDD is not necessarily caused by a malfunction of the serotonergic system perse. MDD may better be considered as a disturbance of the emotional and cognitive functioning in the limbic-subcortical-cortical network. The increase of serotonergic neurotransmission by paroxetine (and likely other antidepressants), appears to improve the constraint of neuronal information processing, with less emotional activation and more cognitive control in treatment responders relative to non-responders. The findings of this thesis can be the starting point for further investigation of biomarkers for treatment response, in order to develop better understanding of MDD and a more effective treatment approaches.

References


54. Ruhe HG, Booij J, Veltman DJ, Michel MC, Schene AH. Successful treatment of major depressive disorder by paroxetine attenuates amygdala activation to negative facial expressions. An fMRI study. [Submitted].


107. Dunlop BW, Nemeroff CB. The role of dopamine in the pathophysiology of depression. Arch Gen Psychiatry. 2007; 64: 327-337.


Chapter 12


In deze Nederlandse samenvatting wordt een overzicht gegeven van een aantal basiskenmerken en een aantal problemen in de farmacologische behandeling van depressie. Deze problemen stonden aan de basis van de studies beschreven in dit proefschrift. Hiervoor is een onderzoeksproject opgezet met verschillende substudies (beschreven in hoofdstuk 2). Dit onderzoek had de overkoepelende naam DELPHI-studie, wat staat voor ‘Dose-escalation Legitimate? Pharmacology and Imaging studies in depression’. In het Nederlands: ‘Is dosisverhoging legitiem? Farmacologische en neuroimaging studies bij depressie.

Na de achtergrond wordt een samenvatting gegeven van de belangrijkste resultaten van de studies, de klinische relevantie ervan, gevolgd door een afsluitende conclusie.

Achtergrond en problemen in de behandeling van depressie

Major Depressive Disorder (MDD), in het Nederlands depressieve stoornis, verder aangeduid met depressie, is een veelvoorkomende en invaliderende ziekte die in potentie recidiveert of verwordt tot een chronische aandoening. Depressie staat momenteel wereldwijd op de tweede plaats van invaliderende ziekten en komt per jaar bij ±25,5% van de volwassenen voor. Ongeveer 12-14% van alle mannen en 22-24% van alle vrouwen wordt gedurende hun leven getroffen door depressie.\textsuperscript{52} Depressie gaat gepaard met hoge directe kosten betreffende de behandeling en indirecte kosten als gevolg van verminderde arbeidsproductiviteit en het verlies van kwaliteit van leven. Naast verschillende vormen van psychotherapie wordt farmacotherapie met antidepressiva vaak aangewend als behandeling.\textsuperscript{3-8}

Antidepressiva kunnen worden ingedeeld in vier groepen of ‘klassen’: Selectieve serotonine heropname remmers (SSRIs), tricyclische antidepressiva (TCAs), monoamine oxidase remmers (MAOIs) en een restgroep, waaronder de zogenaamde ‘dual action’ antidepressiva. SSRIs, ‘dual-action’ antidepressiva en TCAs worden het meest gebruikt.\textsuperscript{9} De meeste antidepressiva beïnvloeden de balans van verschillende boodschapperstoffen in het brein, de ‘neurotransmitters’ serotonine (5-HT) en noradrenaline (zogenaamde monoaminen). De boodschapperstoffen zijn noodzakelijk voor de communicatie tussen zenuwcellen. Communicatie vindt plaats door het uitscheiden van de boodschapperstoffen in de ruimte tussen de zenuwcellen: de synaps. De meeste antidepressiva blokkeren een belangrijk onderdeeltje van de zenuwcell: de serotonine (of noradrenaline) heropname transporter (respectievelijk SERT of NAT), waardoor de hoeveelheid serotonine en/of noradrenaline in de synaps wordt verhoogd. Op deze manier verhogen de meeste antidepressiva de serotonerge en/of noradrenerge signaaloverdracht of neurotransmissie. Als dit echter nader wordt bekeken, hebben verschillende groepen antidepressiva ook specifieke, aanvullende effecten op de zenuwcellen; zogenaamde pre- en post-synaptische effecten (zie hoofdstuk 1, pagina 19).\textsuperscript{10,11} In reactie op de verhoogde serotonerge en/of noradrenerge signaaloverdracht worden ‘second messenger’ systemen geactiveerd, die men op dit moment slechts gedeeltelijk begrijpt.\textsuperscript{12}

Het brein is ook een systeem van met elkaar verbonden kernen of schakelcentra, ook wel neuronale netwerken genoemd. Behandeling met antidepressiva blijkt aanzienlijke veranderingen in de functie van dergelijke neuronale netwerken te veroorzaken (zie hoofdstuk 1, pagina 21).\textsuperscript{13,15} Tot op heden is echter onvoldoende duidelijk welke veranderingen noodzakelijk en specifiek zijn om een depressie te genezen. Het lijkt erop dat de effecten van een toegenomen signaaloverdracht gemedieerd worden door de genetische eigenschappen van patiënten (bijvoorbeeld het 5-HTTLPR polymorphisme\textsuperscript{16}) (zie hoofdstuk 1, pagina 22). Er is namelijk een verband tussen genetische eigenschappen van patiënten en de effecten van behandeling met antidepressiva.\textsuperscript{17} Ook blijken de ontwikkeling van het brein en verbindingen in de neuronale netwerken genetisch bepaald te zijn.\textsuperscript{18}
Antidepressiva hebben een vergelijkbare effectiviteit; ongeveer 50% van de depressieve patiënten reageert op het eerste voorgeschreven antidepressivum. Bij slechts een kleiner percentage van de patiënten wordt volledige remissie van de depressieve klachten bereikt. Daarom is onvoldoende verbetering op een eerste antidepressivum een groot klinisch probleem.

Ten eerste wordt een behandelaar geconfronteerd met de vraag hoe de verandering van symptomen goed en efficiënt te meten. Om op juiste gronden klinische beslissingen te nemen, is de beschikbaarheid en implementatie van valide, korte en eenvoudige meetinstrumenten belangrijk. Ten tweede is het de vraag wat het optimale moment is om de gestarte antidepressieve behandeling te evalueren en zonodig te wijzigen: aangezien antidepressiva enige weken tijd nodig hebben om een verbetering te bewerkstelligen, moet dit niet te vroeg, maar ook niet te laat plaatsvinden.

Als een patiënt geen verbetering (veelal gedefinieerd als 50% verbetering van de klachten; ‘respons’) ervaart op de gestarte SSRI, zijn er vijf farmacologische strategieën: 1) nog enkele weken ongewijzigd voortzetten van de aanvankelijke behandeling, 2) dosisverhoging, 3) het veranderen of ‘switchen’ van het antidepressivum, 4) het toevoegen van een middel dat de werking van het antidepressivum versterkt (augmentatie) en 5) het combineren van verschillende antidepressiva. Over al deze strategieën is de wetenschappelijke kennis beperkt. Om goede ‘evidence-based’ behandelalgoritmen te ontwikkelen zijn goede systematische literatuuroverzichten noodzakelijk. Naast het overzicht over de effectiviteit van behandelingen, verduidelijken deze overzichten ook waar zich lacunes in de kennis over de behandeling met antidepressiva bevinden. Deze lacunes kunnen in nieuwe onderzoeksprojecten worden bestudeerd.

De verlate start van de werking van antidepressiva, de beperkte werkzaamheid van antidepressiva en de verscheidenheid aan farmacologische effecten in het brein, roepen de meer fundamentele vraag op wat het pathofysiologische model voor depressie is. Meer specifiek kan men zich afvragen: wat is precies de onderbouwing van het uitgangspunt van de werking van de meeste antidepressiva: de monoamine hypothese en de betrokkenheid van het serotonerge systeem als een oorzakelijke verklaring van het ontstaan van depressie?

**Samenvatting en conclusies van de studies in dit proefschrift**

Dit proefschrift behandelt verschillende vragen:

1. **Is een korte, door de behandelaar in te vullen vragenlijst even effectief en precies als de gebruikelijke Hamilton depressieschaal (HDRS)?**
   In hoofdstuk 3 heranalyseren wij de behandeluikomen van twee antidepressiva/psycotherapie onderzoeken met als doel de evaluatie van de uitkomsten bepaald door verschillende subschalen van de oorspronkelijke HDRS. Conform andere studies vonden wij dat de twee 6-items bevattende subschalen van Maier en Bech overeenkomen met de originele 17-items bevattende HDRS. Dit gold voor het meten van de ernst van depressie en de gevoeligheid om veranderingen in de tijd te meten. Wij berekenen een conversietable van oorspronkelijke HDRS-scores naar de Maier en Bech scores (en vice versa). Ook bepaalden we afkappunten voor remissie bij deze subschalen. Wij concluderen dat met deze subschalen cliënten de ernst en respons van patiënten efficiënter kunnen meten dan met de oorspronkelijke HDRS.

2. **Wat is het wetenschappelijke bewijs voor dosisverhoging als strategie voor patiënten die niet verbeterden op een eerste SSRI?**
   In hoofdstuk 4 beschrijven wij een systematisch literatuuroverzicht van het wetenschappelijke bewijs voor een dosis-respons relatie voor SSRIs bij de behandeling van depressie. Wij concludeerden dat er géén wetenschappelijk bewijs bestaat dat dosisverhoging van SSRIs effectief is in de eerste 4 weken van de behandeling met een standaarddoserings. Verder lijkt dosisverhoging na 6 weken behandeling met een standaarddosering zelfs minder effectief
dan het ongewijzigd doorgebruiken van de oorspronkelijke dosering. Dosisverhoging gaat gepaard met een toename van de bijwerkingen zonder dat patiënten de behandeling staken vanwege deze bijwerkingen. Vanwege methodologische tekortkomingen concludeerden wij desondanks dat het wetenschappelijke bewijs voor dosisverhoging noch een duidelijk toegenomen werkzaamheid, noch een duidelijke onwerkzaamheid aantoont. Deze conclusie vormde het startpunt voor de DELPHI-studie, beschreven in de hoofdstukken 2 en 7 tot en met 10.

3 Wat is het wetenschappelijke bewijs voor het ‘switchen’ van antidepressiva als strategie voor patiënten die niet verbeterden op een eerste SSRI?
In hoofdstuk 5 beschrijven wij een systematisch literatuuroverzicht waarin het wetenschappelijke bewijs voor het wisselen (‘switchen’) van antidepressiva in het geval dat er onvoldoende verbetering van de depressieve klachten optreedt bij het eerste antidepressivum. Voor drie gerandomiseerde switch-studies die het switchen naar een tweede SSRI vergeleken met een switch naar een SNRI (venlafaxine) konden wij een meta-analyse uitvoeren. Deze meta-analyse liet alleen in de meest rooskleurige analyse een beperkte meerwaarde van het switchen naar venlafaxine als tweede antidepressivum zien. Wij concludeerden dat bij onvoldoende respons op een eerste SSRI, het ‘switchen’ naar elk ander middel kan plaatsvinden, zonder duidelijke verschillen in werkzaamheid en dus vergelijkbaar is met de keuze van het eerste middel.

4 Verlaagt de depletie van monoaminerge systemen de stemming van mensen en is deze stemmingsdaling verschillend voor verschillende subgroepen van mensen?
In hoofdstuk 6 beschrijven wij een systematisch literatuuroverzicht van het wetenschappelijke bewijs dat zogenaamde monoamine depletie studies de stemming van onderzochte proefpersonen verminderen. Hierbij voerden wij een meta-analyse uit, zodat de resultaten van 53 kleine depletie studies konden worden samengevat (‘gepooled’). Wij vonden dat serotonerge, noch noradrenerge/dopaminerge depletie een stemmingsdaling geeft in gezonde controlepersonen, terwijl depletie de stemming van gezonde controlepersonen met een eerstegraads familiaal lid dat depressief is geweest enigszins vermindert. Bij medicatievrije patiënten die na een depressie een volledige remissie van hun klachten hebben bereikt vonden we een matige stemmingsdaling. Monoamine depletie veroorzaakt een terugval bij patiënten die na een depressie een volledige remissie van hun klachten hebben bereikt, echter alleen als deze patiënten een serotonerg werkend antidepressivum gebruiken. Bij depressieve patiënten vindt geen verdere stemmingsdaling plaats na depletie.

5 Verschillen depressieve patiënten en gezonde controlepersonen in het aantal serotonine transporters en is het aantal serotonine transporters gecorreleerd met de ernst van de depressie?
In hoofdstuk 7 onderzoeken we het aantal serotonine transporters (SERTs) in het brein van depressieve patiënten, zoals gemeten in de eerste (baseline) SPECT scans van de deelnemers aan de DELPHI-SPECT studie. Wij vonden diverse interactie-effecten tussen geslacht en ziektestatus met betrekking tot het aantal SERTs in zowel het diencephalon ((hypo)-thalamusregio) als het midbrain (de hersenstam), met daarnaast interactie-effecten tussen roken en ziektestatus (alleen in het diencephalon). Deze interactie-effecten resulteerden in een lager aantal SERTs bij depressieve mannen maar juist meer SERTs bij depressieve vrouwen (ten opzichte van gezonde controlepersonen). Daarnaast bleek het seizoen waarin de scan gemaakt wordt een duidelijke verandering van het aantal SERTs in het midbrain te geven. Deze bevindingen wijzen op complexe effecten van geslacht, roken en het seizoen waarin de scan gemaakt wordt op het aantal SERTs bij depressieve patiënten en gezonde controles. Deze effecten moeten in toekomstig onderzoek van het aantal SERTs bij depressieve patiënten versus gezonde controlepersonen worden meegewogen in de analyses.
6 Verandert een veelvoorkomend genetisch polymorfisme van de promotor regio van het gen voor de serotoninereceptor (SLC6A4) de associatie tussen de SERT bezettingsgraad door paroxetine en de klinische respons?

In hoofdstuk 8 onderzoeken wij de relatie tussen klinische respons na 6 weken paroxetine gebruik en de mate van bezetting van het farmacologische doelwit van paroxetine: de serotonine transporter. De mate van bezetting, ofwel bezettingsgraad van de SERT is hierin de verandering van de hoeveelheid gemeten SERTs na 6 weken behandeling ten opzichte van de hoeveelheid gemeten SERTs vóór behandeling (baseline).38 Daarbij onderzoeken we of de relatie tussen klinische respons en bezettingsgraad wordt beïnvloed door genetische varianten in het SERT-gen, de 5-HTTLPR polymorfismen.

Wij vonden dat voor alle patiënten samen, de SERT bezettingsgraad niet geassocieerd is met de klinische respons. Echter, nadat we de patiënten onderverdeelden op basis van hun 5-HTTLPR polymorfisme, vonden we alleen bij patiënten met het (gunstige) L/A genotype een significante relatie tussen SERT bezettingsgraad en klinische respons. Wij verklaren onze bevindingen met de hypothese dat patiënten met het 5-HTTLPR L/A genotype mogelijk een meer flexibel serotonerg systeem bezitten,18 dat beter beïnvloed kan worden door serotonerge antidepressiva.

7 Is dosisverhoging van paroxetine een effectieve klinische strategie bij onvoldoende respons bij de behandeling van depressie?

In hoofdstuk 9 onderzoeken wij dosisverhoging als strategie voor patiënten die onvoldoende verbeteren op een standaarddosering paroxetine gedurende 6 weken in een gerandomiseerde klinische studie (RCT). Dit onderzoek was een gevolg van de twijfels over de effectiviteit van dosisverhoging (zie ook hoofdstuk 4).32 Als meest belangrijke aanvulling onderzoeken we in een subgroep van de patiënten of een echte dosisverhoging van paroxetine de SERT bezettingsgraad meer doet toenemen dan een placebo-verhoging.40

In een interim analyse analyseerden wij 57 gerandomiseerde patiënten en vonden dat er geen verschil in werkzaamheid bestaat tussen een echte en een placebo dosisverhoging van paroxetine. De studie werd hierop afgebroken in verband met futiliteit. Als belangrijkste bevinding laat ons onderzoek zien dat een echte dosisverhoging van paroxetine geen grotere toename van de SERT bezettingsgraad geeft dan een placebo dosisverhoging. Dit is te verklaren doordat waarschijnlijk al bij een standaarddosering een plateau van SERT bezetting wordt bereikt. We concluderen, in overeenstemming met eerdere, echter meer twijfelachtige gegevens,4447 dat dosisverhoging geen meerwaarde heeft in de behandeling van depressie.

8 Wordt de hyperactivatie van de amygdala bij depressie door een behandeling met paroxetine gemoduleerd?

In hoofdstuk 10 onderzoeken we de emotionele respons van het brein op een (negatieve) gezichtentak in de DELPHI-fMRI substudie.43 We onderzoeken hierin de effecten van behandeling met paroxetine op de vaak vastgestelde hyperactivatie van de amygdala1 tijdens een depressie. Ten opzichte van gezonde controlepersonen vonden we bij depressieve patiënten een toegenomen activiteit in beide amygdala’s en de daaraan grenzende regio. Wij vonden aanvankelijk een toename van de activatie van de rechter amygdala na 6 weken behandeling, die kon worden toegeschreven aan de patiënten (12/20) die onvoldoende verbeterden (non-responders). Deze toegenomen amygdala-activatie nam af in de periode tot week 12. Vooral bij de patiënten die in week 6 en/of week 12 ≥50% verbetering van hun klachten ervaren, nam de amygdala-activatie beiderzijds af, terwijl de amygdala-activatie in week 6 en 12 geassocieerd was met de HDRS-score. Patiënten met onvoldoende respons hadden meer activatie in andere emotie ‘genererende’ (ventrale; zie hoofdstuk 1, pagina 18) gebieden, terwijl responderende patiënten meer activatie toonde in emotie ‘regulerende’ (dorsale; zie hoofdstuk 1, pagina 18) gebieden. We concluderen dat, in overeenstemming met eerdere studies,4447 12 weken behandeling met paroxetine de amygdala en andere emotie

* De amygdala is een amandelvormig onderdeel van de hersenen, dat zich zowel in de linker als rechter hersenhelft bevindt, en onder andere betrokken is bij automatische reacties op angstige situaties.
genererende gebieden minder actief doet zijn en dat de corticale, regulerende functie verbeterd. Dit gebeurt vooral bij patiënten die een klinische respons ervaren. Deze bevindingen wijzen op een toegenomen frontaal-limbische controle als werkingmechanisme voor paroxetine geïnduceerde behandel-effecten (zie hoofdstuk 1, pagina 21).

9 Hoe veranderen de haemostase en stollingsparameters in bloedplaatjes gedurende een behandeling met paroxetine en worden deze veranderingen gemedieerd door dosisverhoging of een genetisch polymorfisme van de promotor regio van het serotonine transporter gen?

In hoofdstuk 11 beschrijven we de resultaten van een weinig voorkomende, maar potentiële gevaarlijke bijwerking van SSRIs: toegenomen bloedingsneiging.48 We vonden dat een standaard dosis paroxetine (20 mg/dag) al een significante daling van serotonine en β-thromboglobuline ($\beta$-TG) in de plaatjes veroorzaakt, zonder dat dit verder afneemt bij dosisverhoging. Bovendien vonden we dat het 5-HTTLPR polymorfisme deze effecten medieert: ten opzichte van $L_A/L_A$-dragers, neemt de bloedingstijd significant toe in patiënten met $<2L_A$-allelen en hadden patiënten zonder een $L_A$-allel, sterkere veranderingen in plaatjes aggregatie gerelateerde parameters. Deze bevindingen zijn van belang voor patiënten met een eerdere ernstige bloeding; bij hen kan genotypering overwogen worden, waarbij bij een $S'/S'$ genotype het starten van, of het switchen naar een niet-serotonerg antidepressivum moet worden overwogen.

Klinische relevantie

Switchen

Ondanks een klein voordeel van een eventuele switch naar venlafaxine in plaats van een tweede SSRI (slechts in de meest gunstige analyse), concluderen wij dat de switch-opties na een eerste SSRI identiek zijn aan de keuze voor het eerste middel. Op basis van de literatuur kunnen geen duidelijke aanbevelingen worden gegeven voor een keuze na een eerste SSRI. Sinds ons systematische literatuuroverzicht verschenen er nog één nieuwe RCT die switchen naar venlafaxine vergeleken met switch naar citalopram,49 en drie open-label studies.50-52

Recent werd er een andere meta-analyse gepubliceerd, waarin een niet-SSRI switch (venlafaxine, mirtazapine of buproprion) werd vergeleken met switch naar een tweede SSRI.53 In deze meta-analyse includeerden auteurs 2 extra studies (abstracts van symposia)54,55 en excluderden zij (net als wij in onze meest rookkleurige analyse54) een methodologisch twijfelachtige studie. Deze meta-analyse concludeert dat er bij switch naar venlafaxine een statistisch significant, doch klinisch niet relevant voordeel bestaat in het aantal patiënten dat een remissie van de klachten bereikt.56 Voor het uitkomstcriterium respons werd geen significant verschil vastgesteld.53 Onze conclusie blijft met deze aanvullende bevindingen ongewijzigd; er lijkt voor individuele patiënten geen duidelijk verschil tussen de verschillende switch-opties te bestaan.

Dosisverhoging

Na ons systematische literatuuroverzicht concluderen wij dat er geen overtuigend bewijs ten faveure of tegen dosisverhoging bestaat. Een onafhankelijke onderzoeksgroep concludeerde hetzelfde op basis van een ander overzichtsartikel.41 De resultaten van onze dosisverhogingsstudie (hoofdstuk 9) verdisconteerde de methodologische beperkingen van eerdere studies en werd vanwege futiliteit afgebroken na een interim-analyse.40 Hoewel het aantal gerandomiseerde patiënten beperkt was, waren de verschillen tussen echte en placebo dosisverhoging zodanig klein (en niet significant) dat de studie voortijdig moest worden beëindigd vanwege futiliteit. Dit houdt in dat het gevonden verschil in HDRS-scores kleiner was dan een vooraf berekend noodzakelijk verschil om nog enige kans op een daadwerkelijk verschil te verwachten, indien het onderzoek voltooid zou worden conform de oorspronkelijke planning.
Zodoende beschermd de deze interim-analysepatiënten om blootgesteld te worden aan een interventie waarvan op dat moment bijna 100% zeker was dat er uiteindelijk geen meerwaarde van zou worden vastgesteld. Deze uitkomst is in overeenstemming met het eerdere onderzoek van Licht et al.\textsuperscript{57} Helaas maakt het voortijdig staken van onze studie het onmogelijk om eventuele subgroepen te identificeren waarin dosisverhoging wellicht toch effectief is (bijvoorbeeld de patiënten die gedeeltelijk (partieel) verbeterden).\textsuperscript{58,59}

De substudie waarin wij de verandering in de SERT bezettingsgraad bepaalden bij patiënten die een dosisverhoging kregen, versterkt de hierboven beschreven klinische bevindingen. Wij concluderen op basis van ons systematisch literatuuroverzicht, onze RCT en onze imaging studie, dat dosisverhoging van SSRIs geen effectieve strategie is voor depressieve patiënten die na 6 weken onvoldoende verbetering ervaren op een standaarddosering.

Een recente meta-analyse van 9 placebo-gecontroleerde, ‘fixed-dose’ studies, gericht op het vinden van de optimale dosering (‘dose-finding’) van SSRIs,\textsuperscript{60} concludeerde dat er een gering, maar statistisch significant voordeel was van hogere doseringen van citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine en/of sertraline, gegeven vanaf het begin van de behandeling (4% meer patiënten met een respons; \( p = 0.04 \)). Het kleine verschil wordt echter als klinisch niet relevant beoordeeld. De auteurs konden niet aangeven dat de significantie van de resultaten aan één van de SSRIs kon worden toegeschreven. Een nadere inspectie van hun gegevens laat in ieder geval geen enkel voordeel van hogere doseringen paroxetine zien, maar wel een duidelijk grotere uitval vanwege bijwerkingen (16.5% ten opzichte van 9.8% bij standaarddoseringen). Daarom blijft het volgens ons dubieus of hogere doseringen SSRIs vanaf het begin van de behandeling zinvol zijn bij depressieve patiënten. Het zou zelfs zo kunnen zijn dat hogere doseringen mensen eerder doen stoppen vanwege bijwerkingen, zodat een adequate, langerdurende blootstelling aan (qua bijwerkingen acceptabele) standaarddoseringen voorkomen wordt.

**Andere klinisch relevante resultaten**

Twee andere relevante bevindingen komen in dit proefschrift naar voren. Ten eerste blijkt dat de Maier en Bech subschalen resultaten geven die overeenkomen met de volledige HDRS\textsuperscript{17,22} Andere auteurs suggereren dat deze subschalen beter de kernsymptomen van depressie meten en dat het gebruik van subschalen zelfs efficiëntere metingen geeft\textsuperscript{23,24} vanwege de afname in variatie.\textsuperscript{61} De Maier of Bech subschalen kunnen dus gebruikt worden in de klinische beoordeling van (de behandeling van) depressieve patiënten, naast zelfinvullijsten als de ‘Inventory for depressive symptomatology self-rated’ (IDS-SR) of de ‘Beck depression inventory’ (BDI). Dit is vooral belangrijk bij mensen die zelfinvullijsten niet betrouwbaar kunnen invullen (bijvoorbeeld bij psychotische kenmerken of bij ernstige cognitieve beperkingen).

Ten tweede verdient de verhoogde bloedingsneiging bij SSRI-gebruik meer aandacht.\textsuperscript{48} Dit was al bekend vooral in combinatie met NSAIDs.\textsuperscript{62-67} De standaard van het Nederlands Huisarts Genootschap adviseert een TCA als de patiënt tevens NSAIDs gebruikt,\textsuperscript{4} hetgeen volgens ons niet geheel correct is, aangezien een serotonerg werkend TCA als clomipramine vermoedelijk vergelijkbare problemen geeft. Vanwege de relatief lage incidentie van potentieel levensbedreigende gastro-intestinale bloedingen (4,3/1000 SSRI behandeljaren; 14,5/1000 SSRI+NSAID/aspirine behandeljaren; ten opzichte van 1,2/1000/jaar in onbehandelde volwassenen),\textsuperscript{65} lijkt de toegenomen bloedingsneiging bij SSRIs irrelevant bij patiënten die geen eerdere ernstige bloedingsincidenten hebben doorgemaakt, of een trauma of operatie meemaken. Echter, bij patiënten met anamnestisch een eerder ernstig bloedingsincident, of bij patiënten die een electieve operatie zullen ondervinden en wellicht ook bij patiënten die regelmatig NSAIDs gebruiken, is genotypering van het S-HTTPLR een overweging. Bij het vaststellen van het S'/S' genotype zou dan een switch naar een niet-serotonerg werkend antidepressivum kunnen worden geadviseerd. De (kosten-) effectiviteit van deze benaderingswijze dient nog wel te worden onderzocht.
Neurobiologie van de behandeling van depressie

Met ons onderzoek onderzochten we ook de neurobiologie van depressiebehandeling met paroxetine in vivo. De algemene beschouwing omtrent dit deel van het proefschrift vraagt aanzienlijke uitleg en enige voorkennis, welke gedetailleerder uiteen wordt gezet in de Engelstalige ‘general discussion’.

Met de SPECT-scans gedurende de gerandomiseerde, dubbelblinde dosisverhogings fase laten wij als eersten ondubbelzinnig zien dat er bij dosisverhoging geen sprake is van een stijging van de SERT-bezettingsgraad. Dit kon op basis van eventuele vertekening in eerder onderzoek niet zo eenduidig worden geconcludeerd. Bovendien wijzen onze gegevens op een matige relatie tussen bezettingsgraad en klinische respons, in tegenstelling tot een eerder gepostuleerde 80% bezettingsgraad die hiervoor vereist zou zijn. De feiten dat de relatie tussen bezetting en klinische respons alleen lijkt te bestaan voor dragers van het L_A/L_A genotype van het 5-HTTLPR gen en dat datzelfde genotype de anatomie en connectiviteit van emotiegerelateerde netwerken beïnvloedt, wijzen erop dat de (heilzame) effecten van een verhoogde serotonerge signaloverdracht waarschijnlijk afhankelijk zijn van de ontwikkeling van het serotonerge systeem.

Een toename van serotonine geeft in het algemeen een beter controle of beteugeling van de informatieverwerking in het brein, met 1) een vermindering van de kans op doorschieten in een reactie van andere systemen (bijvoorbeeld het dopaminesysteem) en 2) een vermindere gevoeligheid voor verstorende in- en externe invloeden op het brein. Deze effecten worden met depletie van serotonine omgekeerd beïnvloed. Daarnaast beïnvloedt serotonine de aanmaak van ‘brain derived neurotrophic factor’ (BDNF), een eiwit dat betrokken is bij neuroplasticiteit. Onze en andere fMRI-studies kunnen in dat licht worden geïnterpreteerd. Bij responders verbeterde de cognitieve controle en de functionele hyperactivatie af. Uit een eerder onderzoek bleek dat behandeling met sertraline een toename veroorzaakte in de functionele koppeling tussen de amygdala en de frontale en (pregenuale) cingulaire cortex, respectievelijk emotionele en controlerende, cognitieve onderdelen van het limbische-subcortico-corticale netwerk. Mogelijk is deze toegenomen functionele koppeling een gevolg van de verhoogde serotonerge neurotransmissie, mogelijk het gevolg van een toename van BDNF, of beide. Dit dient nog verder onderzocht te worden.

Beperkingen van het onderzoek

In ieder hoofdstuk en de Engelstalige ‘general discussion’ zijn de mogelijke beperkingen van de studies van dit proefschrift afzonderlijk beschreven. De belangrijkste hiervan worden in deze Nederlandse samenvatting aangestipt.

In het case-control onderzoek naar de beschikbaarheid van SERTs bij depressieve patiënten versus gezonde proefpersonen bleven kleine subgroepen over bij het onderzoeken van effect-mediërende variabelen. Daarom hebben wij deze verschillen als exploratieve analyses gepresenteerd. Ten aanzien van de associatie tussen SERT bezettingsgraad en klinische respons vergeleken wij ook kleine groepen, maar vonden wij statistisch sterk significante verbanden die wij daarom als valide beschouwen. Om toevallige bevindingen uit te sluiten is replicatie van deze bevindingen wenselijk. Na de interim-analyse staakten wij onze vergelijkende studie vanwege futiliteit, wat verdere analyse van subgroepen niet toeliet. De fMRI-studie was een voor neuroimaging vrij grote studie, waarin wij robuuste effecten vonden. Echter ook hierbij was het aantal patiënten in de dosisverhogingsgroep beperkt voor het onderzoeken van drie-weg interacties tussen dosisverhoging, tijd en respons.

Ons neurobiologische onderzoek onderzocht geen van de specifieke secundaire effecten van SSRIs. Er zijn vele adaptieve pre- en post-synaptische effecten van SSRIs bekend,\textsuperscript{10,83,90-97} deze werden vooral bij knaagdieren onderzocht.\textsuperscript{98} De maximale stralingsbelasting voor wetenschappelijk onderzoek liet geen ander scanonderzoek met radioliganden toe, evenmin onderzochten wij de effecten van dosisverhoging in diermodellen. Vooropgesteld moet echter blijven, dat de klinische effecten zowel in ons, als in eerder onderzoek geen effectiviteit van dosisverhoging laten zien. Aanvullend dierexperimenteel onderzoek zou daarom vooral meer fundamentele kennis over het werkingsmechanisme van SSRIs en de neurobiologie van dosisverhoging geven.

Tenslotte kunnen clinici zich bij het lezen van het onderzoek in dit proefschrift af gaan vragen waarom initieel hogere doseringen van SSRIs effectief zijn gebleken in (sommige) ‘fixed-dose, dose-finding’ studies bij patiënten met een angststoornis,\textsuperscript{99-102} en vooral bij de obsessief compulsieve stoornis (OCD).\textsuperscript{103-107} Dit vormt een intrigerend vraagstuk dat buiten het kader van dit proefschrift valt, maar een vergelijkbare aanpak als in ons onderzoek rechtvaardigt.

**Methodologisch sterke kenmerken van het onderzoek**

Dit proefschrift is gebaseerd op een relevant klinisch vraagstuk. Alle behandelaars kennen het probleem dat patiënten onvoldoende verbeteren op een eerste SSRI en dosisverhoging is bijna altijd de eerste stap, gevolgd door een switch van antidepressivum.\textsuperscript{108-112} Dit proefschrift kan daarom relevant zijn voor toekomstige richtlijnen.

Onze systematische literatuuroverzichten werden uitgevoerd conform de strenge richtlijnen van de Cochrane Collaboration.\textsuperscript{113} Wij voerden ‘sensitieve’ zoekacties van de literatuur, zodat de kans op het missen van waardevolle studies geminimaliseerd werd. Bovendien werden alle gevonden studies kritisch beoordeeld en samengevat, waarbij wij alleen een meta-analyse uitvoerden als de klinische homogeniteit dit toeliet. In onze meta-analyse van de monoamine deprietiestudies hanteerden wij bovendien het statistische probleem van het ‘poolen’ van ‘cross-over’ studies.\textsuperscript{37}

De empirische studies uit dit proefschrift werden zorgvuldig opgezet, waarbij de methodologische problemen van eerdere dosisverhogingsstudies werden voorkomen.\textsuperscript{32,114} Hiertoe randomiseerden wij patiënten pas na 6 weken behandeling, zodat de kans op een verlate respons werd verminderd. Ook leidde ons dosisverhogingsschema niet tot een toename van de uitval (door bijwerkingen) in de dosisverhogingsgroep. Tenslotte was onze SPECT-studie die het effect op de SERT bezettingsgraad kwantificeerde uniek en gaf dit een duidelijke rationale waarom dosisverhoging niet werkzaam is.

Met het uitvoeren van de fMRI-studieonderzochten we de effecten op de emotieregulatie als gevolg was van de behandeling met paroxetine. Dit was de zesde fMRI studie die de behandel effecten van SSRIs bij depressie onderzocht. Het feit dat wij 3 scans maakten en dat veel patiënten initieel niet voldoende verbeterden gaf ons de mogelijkheid om het verschil in hersenactivatie tussen responders en non-responders bij paroxetinebehandeling te onderzoeken. Dit is een belangrijke aanvulling op de bestaande literatuur.
Verder keken wij ook naar de perifere effecten van paroxetine, buiten het brein, maar met hetzelfde mechanisme: de blokkade van de opname van serotonine in de plaatjes en de effecten daarvan op de bloedstolling. Toegenomen bloedingsneiging is een infrequent, maar potentieel gevaarlijk complicatie. Wij lieten zien, dat ook de standaarddoserings van paroxetine invloed kan hebben op de haemostase, vooral in de subgroep van patiënten die drager zijn van het S'/S' polymorfisme van het 5-HTTLPR gen. Deze groep is mogelijk ook gebaat bij secundaire preventie van toekomstige bloedingsproblemen bij het gaan of blijven gebruiken van SSRIs.

Toekomstig onderzoek

De studies in dit proefschrift stimuleren om nieuwe studies en systematische literatuuroverzichten te starten. Een aantal onderdelen van de DELPHI-studie dienden nog uitgewerkt te worden, terwijl andere vraagstukken met deze studies onbeantwoord zullen blijven. In de Engelstalige ‘general discussion’ wordt een aanzet gegeven voor toekomstige studies, die zich met name zullen moeten richten op de vergelijking van verschillende strategiën voor onvoldoende respons, en het meer fundamentele begrip van de dwarsverbanden tussen verschillende neurobiologische systemen. Hierbij valt te denken aan een RCT waarin switchen na 6 weken wordt vergeleken met 6 weken doorbehandelen, onderzoek bij depressieve patiënten naar de beïnvloeding van het dopaminerg systeem door SSRIs, en de effecten van antidepressiva op de neurogenese.

Conclusie

Dit proefschrift heeft betrekking op de behandeling van depressie met antidepressiva en vooral op het vraagstuk wat te doen als patiënten niet verbeteren op een standaarddosing. Op basis van twee systematische literatuuroverzichten concluderen wij dat er onvoldoende wetenschappelijk bewijs bestaat voor een duidelijke aanbeveling over het switchen naar een ander antidepressivum en het verhogen van de dosering. In de hierop volgende gerandomiseerde placebo-gecontroleerde studie (de DELPHI-studie) onderzochten wij de klinische werkzaamheid van dosisverhogingen bij depressieve patiënten die onvoldoende waren verbeterd na 6 weken paroxetine gebruik (20 mg/dag). Daarnaast onderzochten we in een subgroep de neurobiologische effecten van dosisverhoging in SPECT en fMRI neuroimaging studies. Met de DELPHI-studie tonen we aan dat dosisverhoging bij paroxetine geen meerwaarde heeft ten opzichte van een placebo-verhoging. Bovendien kwantificeren we dat dosisverhoging de blokkade van het farmacologische doel van paroxetine, de SERT, niet doet toenemen.

Eveneens onderzochten wij in dit proefschrift de rol van monoaminen en de SERT in de etiologie van depressie en de neurobiologische effecten van behandeling met paroxetine. Op basis van deze studies concluderen wij dat depressie niet noodzakelijkerwijs veroorzaakt wordt door een slecht werkend serotonerg systeem. Depressie kan beter worden gezien als een verstoring van de balans tussen het emotionele en cognitieve functioneren van het limbisch-subcortico-corticale netwerk. De toename van de serotonergeneurotransmissie als gevolg van behandeling met paroxetine (en waarschijnlijk andere SSRIs) lijkt te liggen in een betere beteugeling van de neuronale verwerking van informatie. Patiënten die verbeteren na de behandeling laten een afname van de emotionele activiteit (bijv. in de amygdala) en een toename van de cognitieve controle zien. De bevindingen van dit onderzoek kunnen dienen als uitgangspunt voor het verder ontwikkelen van biomarkers voor behandel-uitkomsten, met als doel het ontwikkelen van een beter begrip van het ziektebeeld depressie en effectievere behandelingen.
Chapter 13

Referenties


43. Ruhe HG, Boos J, Veltman DJ, Michel MC, Schene AH. Successful treatment of major depressive disorder by paroxetine attenuates amygdala activation to negative facial expressions. An fMRI study. 2008; [Submitted].


104. Figue M, Denys D. New pharmacotherapeutic approaches to obsessive compulsive disorder. CNS Spectr. 2008; [In press].


ABBREVIATIONS, COLOR FIGURES, PUBLICATIONS, ACKNOWLEDGEMENT AND CURRICULUM VITAE
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[123I]ß-CIT</td>
<td>methyl 3ß-(4-iodophenyl) tropane-2 ß –carboxylate</td>
</tr>
<tr>
<td>5-HT</td>
<td>serotonin</td>
</tr>
<tr>
<td>5-HTTLPR</td>
<td>Serotonin transporter promoter polymorphism</td>
</tr>
<tr>
<td>95% CI</td>
<td>95% confidence interval</td>
</tr>
<tr>
<td>B-TG</td>
<td>β-thromboglobulin</td>
</tr>
<tr>
<td>AHCPR</td>
<td>Agency of healthcare policy and research</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike’s information criterion</td>
</tr>
<tr>
<td>AMC</td>
<td>Academic Medical Center</td>
</tr>
<tr>
<td>AMPT</td>
<td>Alpha-methyl-para-tyrosine</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance model</td>
</tr>
<tr>
<td>ATD</td>
<td>Acute tryptophan depletion</td>
</tr>
<tr>
<td>APTD</td>
<td>Acute phenylalanine/tyrosine depletion</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>BDNF</td>
<td>Brain derived neurotrophic factor</td>
</tr>
<tr>
<td>BDI</td>
<td>Beck depression inventory</td>
</tr>
<tr>
<td>BOLD</td>
<td>Blood oxygen level dependent</td>
</tr>
<tr>
<td>BPND</td>
<td>Binding potential (non-displacable)</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CCDAN</td>
<td>Cochrane collaboration depression anxiety and neurosis group</td>
</tr>
<tr>
<td>CER</td>
<td>Cerebellum</td>
</tr>
<tr>
<td>CGI-I/S</td>
<td>Clinical global impression scale (improvement/severity)</td>
</tr>
<tr>
<td>COMT</td>
<td>Catechol-O-methyltransferase</td>
</tr>
<tr>
<td>CI95</td>
<td>95% Confidence interval</td>
</tr>
<tr>
<td>CREB</td>
<td>cAMP response element binding protein</td>
</tr>
<tr>
<td>CYP P450</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>DA</td>
<td>Dopamine</td>
</tr>
<tr>
<td>DAT</td>
<td>Dopamine transporter</td>
</tr>
<tr>
<td>DE</td>
<td>Dose-escalation</td>
</tr>
<tr>
<td>DELPHI</td>
<td>Dose-escalation legitimate? Pharmacology and imaging studies in depression</td>
</tr>
<tr>
<td>DIENC</td>
<td>Diencephalon</td>
</tr>
<tr>
<td>DLPFC</td>
<td>Dorsolateral prefrontal cortex</td>
</tr>
<tr>
<td>DMPFC</td>
<td>Dorsomedial prefrontal cortex</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DLPC</td>
<td>Dorsolateral prefrontal cortex</td>
</tr>
<tr>
<td>DSM-IV</td>
<td>Diagnostic and statistical manual, 4th edition</td>
</tr>
<tr>
<td>E-S</td>
<td>Effect size</td>
</tr>
<tr>
<td>EEG</td>
<td>Electric encephalogram</td>
</tr>
<tr>
<td>EMEA</td>
<td>European agency for the evaluation of medicinal products</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and drug administration</td>
</tr>
<tr>
<td>fMRI</td>
<td>functional Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>HC</td>
<td>Healthy controls</td>
</tr>
<tr>
<td>HDRS</td>
<td>Hamilton depression rating scale</td>
</tr>
<tr>
<td>IDS-C/SR</td>
<td>Inventory for depressive symptoms – clinician / self-rated</td>
</tr>
<tr>
<td>IQR</td>
<td>Inter Quartile Range</td>
</tr>
<tr>
<td>IRT</td>
<td>Item response theory</td>
</tr>
<tr>
<td>ITT</td>
<td>Intention to treat</td>
</tr>
<tr>
<td>IU</td>
<td>International units</td>
</tr>
<tr>
<td>LOCF</td>
<td>Last observation was carried forward</td>
</tr>
<tr>
<td>LoE</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>MAACL</td>
<td>Multiple affect adjective checklist</td>
</tr>
<tr>
<td>MADRS</td>
<td>Montgomery Åsberg depression rating scale</td>
</tr>
<tr>
<td>MAO(-I)</td>
<td>Monoamine oxidase (inhibitor)</td>
</tr>
<tr>
<td>MHRA (British)</td>
<td>Medicines and healthcare products regulatory agency</td>
</tr>
<tr>
<td>MID</td>
<td>Midbrain</td>
</tr>
<tr>
<td>MDD</td>
<td>Major depressive disorder</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>MDMA</td>
<td>3,4-methylenedioxy methamphetamine (XTC)</td>
</tr>
<tr>
<td>MNI</td>
<td>Montreal Neurological Institute</td>
</tr>
<tr>
<td>MOS-SF36</td>
<td>Medical outcome study short form (measuring health related quality of life)</td>
</tr>
<tr>
<td>MRI, sMRI/fMRI</td>
<td>Magnetic resonance imaging, structural/functional</td>
</tr>
<tr>
<td>N/A</td>
<td>Not appropriate</td>
</tr>
<tr>
<td>NA/NE</td>
<td>Noradrenaline (=norepinephrine)</td>
</tr>
<tr>
<td>NAT/NET</td>
<td>Noradrenaline transporter</td>
</tr>
<tr>
<td>NERI</td>
<td>Norepinephrine reuptake inhibitors</td>
</tr>
<tr>
<td>NNT/NNH</td>
<td>Number needed to treat/ harm</td>
</tr>
<tr>
<td>NSAID</td>
<td>Non-steroidal anti-inflammatory drug</td>
</tr>
<tr>
<td>OCC</td>
<td>Occupancy</td>
</tr>
<tr>
<td>OCD</td>
<td>Obsessive compulsive disorder</td>
</tr>
<tr>
<td>OFC</td>
<td>Orbitofrontal cortex</td>
</tr>
<tr>
<td>PCPA</td>
<td>Para-chlorophenylalanine</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>PFA</td>
<td>Platelet function analyzer</td>
</tr>
<tr>
<td>PF4</td>
<td>Platelet factor 4</td>
</tr>
<tr>
<td>PFC</td>
<td>Prefrontal cortex</td>
</tr>
<tr>
<td>POMS</td>
<td>Profile of Mood States</td>
</tr>
<tr>
<td>PRP</td>
<td>Platelet rich plasma</td>
</tr>
<tr>
<td>PSC</td>
<td>Paroxetine serum concentrations</td>
</tr>
<tr>
<td>PUFA</td>
<td>Poly unsaturated fatty acids</td>
</tr>
<tr>
<td>QIDS-SR</td>
<td>Quick-inventory for depressive symptoms</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized controlled trial</td>
</tr>
<tr>
<td>RIMA</td>
<td>Reversible inhibitor of monoamine oxidase A</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver operator characteristic (curve)</td>
</tr>
<tr>
<td>RoI</td>
<td>Regions of interest</td>
</tr>
<tr>
<td>RR</td>
<td>Relative risk</td>
</tr>
<tr>
<td>RT</td>
<td>Reaction time</td>
</tr>
<tr>
<td>SCID</td>
<td>Structured clinical interview for DSM-IV</td>
</tr>
<tr>
<td>SCL</td>
<td>Symptoms check list (90 items)</td>
</tr>
<tr>
<td>SCN</td>
<td>Suprachiasmatic nucleus</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SE(M)</td>
<td>Standard error (of the mean)</td>
</tr>
<tr>
<td>SERT</td>
<td>Serotonin transporter</td>
</tr>
<tr>
<td>SLC6A4</td>
<td>Serotonin transporter gene</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SNRI</td>
<td>Serotonin and norepinephrine reuptake inhibitor</td>
</tr>
<tr>
<td>SPECT</td>
<td>Single photon emission computed tomography</td>
</tr>
<tr>
<td>SPM5</td>
<td>Statistical Parametric Mapping version 5</td>
</tr>
<tr>
<td>SPSP</td>
<td>Short psychodynamic supportive psychotherapy</td>
</tr>
<tr>
<td>SSR1</td>
<td>Selective serotonin reuptake inhibitor</td>
</tr>
<tr>
<td>STAR*D</td>
<td>Sequenced treatment alternatives to relieve depression</td>
</tr>
<tr>
<td>TCA</td>
<td>Tricyclic antidepressant</td>
</tr>
<tr>
<td>TE</td>
<td>Echo time</td>
</tr>
<tr>
<td>TR</td>
<td>Repetition time</td>
</tr>
<tr>
<td>TRD</td>
<td>Treatment-resistant depression</td>
</tr>
<tr>
<td>VA(M)S</td>
<td>Visual analogue (mood)scales</td>
</tr>
<tr>
<td>VLPFC</td>
<td>Ventrolateral prefrontal cortex</td>
</tr>
<tr>
<td>VMPFC</td>
<td>Ventromedial prefrontal cortex</td>
</tr>
</tbody>
</table>
Figure 1.2. Transporters, receptors and second messenger systems involved in the effects of antidepressants.
Figure adapted from Belmaker and Agam.\textsuperscript{79} The left half of the presynaptic neuron represents a serotonergic neuron, the right half a norepinephrinergic neuron.

In the presynaptic neuron, serotonin is synthesized from tryptophan by tryptophan hydroxylase and stored in vesicles. Likewise, norepinephrine is synthesized from tyrosine by tyrosine hydroxylase. These vesicles merge with the cell membrane when the neuron is depolarized, thereby releasing their contents into the synaptic cleft.

After release, serotonin and norepinephrine are transported back into the presynaptic neuron by serotonin and norepinephrine transporters. Furthermore, serotonin and norepinephrine are catabolized by the monoamine-oxidase A (MAO-A) enzyme.

In the synaptic cleft, serotonin and norepinephrine affect both the pre- and post-synaptic neuron. The pre-synaptic $5\text{-HT}_1A$ and $5\text{-HT}_1B$ auto-receptors decrease serotonin release by inhibitory feedback; the $\alpha_2$-adrenergic receptor does the same for the release of norepinephrine.

Post-synaptically, serotonin and norepinephrine bind to G-protein-coupled monoamine receptors (MARs): the cyclicAMP (cAMP)-coupled receptor, which activates protein kinase A (PKA), and the Phosphatidylinositol (PI)-coupled receptor, which activates phospholipase C (PLC) which thereafter form inositol triphosphate (IP$_3$) and diacylglycerol (DAG). IP$_3$ and DAG activate protein kinase C (PKC). Both PKA and PKC finally activate cAMP responsive element binding (CREB) protein, which stimulates DNA transcription. For example, this might result in the production of brain derived neurotrophic factor (BDNF).
DELPHI-SPECT and DELPHI-fMRI represent sub-studies nested within the total study. For these 2 sub-studies drug-free patients were recruited and treated in the Academic Medical Center.
Figure 2.1. Design of the DELPHI study.

6 weeks PAR (20mg)

Inclusion

Clinical visits, questionnaire measurements

6 weeks (PAR DE)

6 weeks (placebo DE)

Randomization of non-responders

Weeks: -6 -4 -2 0 1 2 4 6

Entry

T₀ T₁

Bsl SPECT

2nd SPECT

3rd SPECT

Bsl fMRI

2nd fMRI

3rd fMRI
Figure 2.3. Regions of interest (RoI) for Midbrain, Cerebellum and Diencephalon.
Example of SPECT images after 3D reconstruction. Templates with fixed ROIs are shown in green.

A. Midbrain (circle) and cerebellum. B. Striatum (for demarcation midbrain-diencephalon) and diencephalon (circle). ROIs were positioned by hand and situated based on anatomy and maximum concentration of activity/ml in the ROI.
Negative faces contrast. Week 6 and 12 scans and response status were combined in a full factorial model.

A. SPM-results showing design matrix, contrast and activation in right amygdala (MNI 18,-2,-18).

B. Activations of left (MNI -22,2,-18; blue) and right (MNI 18,-2,-18; red) amygdala stratified for non-response and response at 6 and 12 weeks. Significant difference between non-responders and responders ($F_{1,32}= 4.591; p<0.0001$); no significant lateralization ($F_{1,32}= 0.1517; p= 0.70$).

C. Plot of contrast estimate by HDRS17-score (centralized to mean). Significant positive correlations for left ($0.33 \pm 0.12 \text{[SE]}; F_{1,31}= 7.076; p= 0.0123; r^2= 0.19$; shown in blue) and right ($0.38 \pm 0.14 \text{[SE]}; F_{1,31}= 6.929; p= 0.0131; r^2= 0.18$; shown in red) amygdala. Regression lines for left and right amygdala are not significantly different ($F_{1,62}= 0.0683; p= 0.79$).
**Figure 10.1.** Amygdala response relative to clinical response (HDRS17 ≥50% decrease) after 6 and 12 weeks of treatment with paroxetine.
Publications


Ruhe HG, Booij J, Veltman DJ, Michel MC, Schene AH. Successful treatment of major depressive disorder by paroxetine attenuates amygdala activation to negative facial expressions. An fMRI study. 2008; [Submitted].


Published abstracts


Ruhé HG, Huyser J, Swinkels JA, Schene AH. Failure to respond to the first SSRI in major depressive disorder: an evidence-based guideline for timing of treatment changes, increasing the dose and switching. [Abstract]. *J Affect Disord.* 2004; **78** Suppl 1:S89.
DANKWOORD
Het beste bewaar je tot het laatste. In de afgelopen 5 jaar heb ik met veel plezier met vele mensen samengewerkt zonder wie dit proefschrift onmogelijk zou zijn geweest. In mijn gedachten heb ik dit dankwoord al vele malen geschreven.

In de eerste plaats is deze studie alleen mogelijk geworden dankzij heel veel patiënten, die ondanks hun depressie bereid waren verwezen te worden naar het AMC. Ik ben u allen buitengewoon erkentelijk. Zonder deze deelname was het onderzoek nooit geslaagd. Alle patiënten die naast de ‘gewone’ dosisverhogingsstudie ook nog meededen aan het SPECT en/of fMRI onderzoek wil ik dubbel bedanken. U kwam op moeilijke tijden, soms meerdere keren per week terug voor deze scans, soms ook als de somberheid nog niet zo was verbeterd.

Oké, de gezonde proefpersonen die bereid waren SPECT en fMRI scans te ondergaan wil ik hartelijk bedanken. Ik hoop uw interesse in wetenschappelijk onderzoek te hebben gestimuleerd met mijn uitleg over wat we onderzochten. Dankzij uw bereidheid is het mogelijk resultaten te vergelijken met een niet-zieke toestand.

Graag wil ik ook alle huisartsen, psychiater en sociaal psychiatrisch verpleegkundigen bedanken die in de periode dat wij op zoek waren naar deelnemers voor ons onderzoek trouw bleven verwijzen. Wellicht dat wij in de toekomst opnieuw met jullie samen kunnen werken.


Dr. Jan Booij, mijn co-promotor, dank ook voor jouw denkwerk, enthousiasme en gelijkwaardige begeleiding tijdens het uitvoeren van dit onderzoek. Jij regelde de infrastructuur in de samenwerking met de XTC-studie, de scan-faciliteiten, de uitleg over de analyses van de scans en mijn verdere scholing op het gebied van de nucleaire geneeskunde. Met jou is het altijd prettig samenwerken geweest, laten we dit vooral voortzetten.

De leden van de promotiecommissie hebben als eersten het product van mijn werk mogen lezen en beoordelen: Prof.dr. Damaan Denys, dank voor de faciliteiten die door de afdeling Nucleaire geneeskunde werden geschapen om dit onderzoek te kunnen doen. Dank voor het beoordelen van het manuscript, maar ook mijn dank voor de interesse in mijn onderzoek, als wij elkaar op F2 ontmoetten.

Prof.dr. C.G. Kruse, mijn hartelijke dank voor uw beoordeling van mijn manuscript. Ik waardeer het enorm dat u ondanks uw drukke reisschema aanwezig kon zijn bij de verdediging.

Prof.dr. Willem Nolen, biologische psychiater van het eerste uur en daarbij ervaren clinicus op het gebied van de stemmingsstoornissen. Hartelijk dank voor jouw interesse in mijn activiteiten, waar je al sinds 2001 bij betrokken bent. Dat jij voor mijn commissie zou worden gevraagd was voor mij vanzelfsprekend, ik vind het een eer dat je mijn manuscript zo spoedig van een oordeel voorzag. Het proefschrift is af, de studies zijn (bijna) gepubliceerd, en we moeten voort, op zoek naar nieuwe strategieën in de behandeling van depressie.
Dr. Robbert-Jan Verkes, ook jou wil ik hartelijk danken voor je bereidheid vragen te stellen bij mijn verdediging. Dank voor de plezierige samenwerking tijdens het schrijven van het hoofdstuk in het Handboek Depressie, dat smaakte naar meer.

Jochanan Huyser, tot drie keer toe was jij op de jou toevertrouwde wijze - strategisch aan de zijlijn - een beïnvloeder van formaat. Je nam mij eind 1996 aan als AGNIO. Vervolgens bracht jij me de beginselen van depressiebehandeling bij. Je kwam, kort na een jaar zoeken naar een ‘kristallisatiepunt’, met het aanbod om het richtlijnproject te gaan doen. Ik heb genoten van je stimulerende enthousiasme tijdens dit project, en denk met bewondering terug aan jouw kunde in het voorzitten van de commissievergaderingen. Toen we collegae psychiatriers waren, heb jij mijn keuze om het aanbod om het richtlijnproject te gaan doen. Ik heb genoten van je stimulerende enthousiasme tijdens dit project, en een kleine suggestie richting gegeven. Dan k, nu jij het AMC hebt verlaten, de dagelijkse leiding van ‘jouw’ behandelcentrum Stemmingstoornissen over heb mogen nemen beschouw ik als een grote eer.

Voor alle praktische en logistieke support was een telefoontje naar Marianne Haages meestal voldoende. Marianne, fantastisch hoe jij al sinds het richtlijnproject enthousiast hebt meegewerkt, en daarvoor ben ik je duizendmaal dank verschuldigd. Met jouw invulling krijgt de profielschets voor een ervaren secretaris/secretaresse een onbereikbare dimensie.

Mijn collegae van de afdeling stemmingstoornissen, Guido Nabarro, Prof.dr. Jan Swinkels, Mirjam van Zon, alle verpleegkundigen, assistenten in opleiding, ergotherapeuten, maatschappelijk werkvers, psychomotore therapeuten en psychologen, jullie hebben allemaal bijgedragen op het toest stand komen van dit boek. Ook de ex-collegae Nicole Eussen, Arjen Ronner, Nicole de Boer en Pierre Delfgauw horen in dit rijke huis. Dank voor jullie interesse, waarneming, vragen, samenwerking op het gebied van onderzoek en ingewikkelde casuïstiek, het verdragen van mijn afwezigheid, maar vooral jullie degelijke vakkennis en enthousiasme in de behandeling van onze depressieve patiënten. Dagelijks laten jullie zien waar het bij behandelen van depressieve, zieke mensen echt om gaat: betrekkenheid met gepaste distantie, veelvuldig proberen en herhalen, gedoseerde activatie, soms toegeeflijk zijn maar zonder hierbij de eigen grenzen te overschrijden, gecombineerd met een professionele aandacht voor stagnatie en alarmsignalen. Ik vind het, nogmaals, een eer om met jullie te werken en hoop daar mee door te kunnen gaan.

Een speciaal woord van dank gaat uit naar Mascha ten Doesschate en Elizabeth Miedema, die veel bijdroegen aan de opzet en de dataverzameling voor de DELPHI-studie. Mascha, veel succes met de afronding van jouw promotie. Elizabeth, gezamenlijk werkten wij anderhalf jaar aan de inclusie. Dank voor je toewijding en bereidheid om logistieke zaken omtrent de patiëntenzorg van mij over te nemen. Andere leden van onze onderzoeksgroep: Karin, Kim, leke, Anja, Wendy, Hanneke, Maarten, Monique, Judith en sinds kort Maaike en Mathilda, dank voor de prettige samenwerking, jullie bijdragen aan het onderzoeksoverleg, de gezamenlijke projecten die zijn afgerond, nog lopen of gaan komen. Ik verheug me nu al op een volgend bezoek aan ISAD. Nada, Michiel, Koen en Max, dank voor jullie input bij mijn projecten. Maaike, sinds kort toegetreden als AGIKO, jij startte jouw promotie in een tijd dat ik nog druk was/ben met de afronding van dit proefschrift. Ik bewonder je degelijkheid waarmee je complexe zaken uitdiept en kijk uit naar de verdere ontwikkeling van onze biomarker plannen.

Alle collegae van andere afdelingen ben ik dankbaar voor de prettige contacten, expertise en toegankelijkheid. Dr. Henk van Weert, met jou heb ik veel werk verzet om ons verwijzersnetwerk op te bouwen. Dank voor je betrokkenheid, je inzet en je hulp om DELPHI tot een succes te maken. Prof.dr. Dick Veltman, jij leerde mij in meerdere middagen één-op één onderwijs het werken met SPM5. Het lijkt me goed hiermee door te gaan en vooral samen nieuwe fMRI studies te gaan doen. Prof.dr. Martin Michel, jij besteedde als co-auteur bijzonder veel aandacht aan mijn manuscripten. Dank voor je kritische commentaren, het wordt er inderdaad alleen maar beter van. Dr. Hans Reitsma, “gelijk hebben en krijgen is niet altijd hetzelfde”, dank voor al jouw hulp bij ook voor een gewone epidemioloog ingewikkelde statistiek; ik heb veel van je geleerd. Prof.dr. Koos Zwinderman, dat geldt ook voor jou. Dr. Maartje de Win, aanvankelijk kon ik voor de SPECT-scans meeliften op jouw XTC-studie. Dank hiervoor en voor jouw ordentelijke planning waar altijd op te bouwen was. De medewerkers van de afdeling Nucleaire geneeskunde, dank voor het verdragen van al mijn ‘geregel’ om de scans op de juiste momenten te kunnen maken, en dank
voor jullie hulp bij het scannen in de avonduren. Dr. Rob Scholten, ik ken je vanaf mijn eerste serieuze onderzoeksinspanningen op het EMGO. Ondanks dat ik door jou gedegen aan de tand werd gevoeld tijdens mijn epidemiologie-examen moet ik je nog steeds lastig vallen voor meta-analytische problemen, waarvan jij zelfs verzuchtte dat ze zo ingewikkeld zijn. Dank je voor je prettige, relativerende en snelle hulp. Ook Dr. Astrid Goossens van de KEB wil ik bedanken voor haar enthousiasme tijdens het richtlijnproject. Dr. Aart Nederveen, dank voor je rustige, gedegen hulp bij het instellen van de MRI-scanner en alle antwoorden op andere vragen die ik je heb gesteld. Raschel en Sandra, dank voor de eerste hulp bij MRI-problemen. Tessa, jij was nooit te beroerd om extra te komen scannen als MRI-piloot. Michiel de Ruiter, veel dank voor het delen van jouw kennis van SPM en fMRI, ook als ik je op vakantie of in de supermarkt belde. Zsuzsika, dank voor je expertise bij het voorbewerken van de ToL, en succes met je eigen studies. Eric Franssen, dank voor de snelle bepalingen van de paroxetine-spiegels en je commentaren op eerdere versies van het artikel. Remco Knol, dank voor je prachtige illustratie (figuur 1.2) en veel succes de dag na mij. Peter Remijnse, we hadden een goede inclusie luchtrecht, moeten we vaker doen. Ik ben zeer benieuwd naar jouw vorderingen. Prof.dr. Frank Baas, Marja Jakobs en Martina Moeton, dank voor jullie hulp bij de genetische bepalingen, en de interpretatie ervan. Nadia, Pieter-Willem, Saskia, Joost, Eric-Jan, Kurdo en Prof.dr. Harry Büller, wat een mooi project is onze stollings-studie geworden, dank voor de samenwerking. Bep de Lange, sinds anderhalf jaar mijn buurvrouw, dank voor je zorg, je interesse en natuurlijk de schutbladen op momenten dat het echt nodig is. Prof.dr. Frans de Jonghe, Prof.dr. Jack Dekker en Jaap Peen van de Mentrum Depressie groep, dank voor het gezamenlijk werken aan het HDRS-artikel, meten is weten, en efficiënt meten is sneller weten.

Eugenie Oosterhuis, u verdient een aparte vermelding in dit dankwoord. Het bereiken van een plateau na een stevige beklimming wil niet zeggen dat je er niet alsnog af kan vallen; wel weet je beter waar de gevaarlijke randen zitten. Dank dus opnieuw.

Leden van de intervisiegroep, Ethy, Agnes, Joop, Saskia en Hiske, gelukkig dat ik van jullie ‘mag blijven’ als ik gepromoveerd ben. Dank voor jullie begripvolle, constructieve doch stevige probleemanalyses, waaruit nieuwe oplossingen opborrelden. Gelukkig is het ook erg gezellig om bij elkaar te komen.


Ook mijn (schoon-)familie wil ik bedanken voor de support van de laatste maanden en natuurlijk de jaren erna. Pap en mam, van jullie heb ik heel veel meegekregen, maar absoluut ook een zeker fanatisme, dat onontbeerlijk is voor een promotie. Dank daarvoor, en voor de investeringen die jullie in de afgelopen 40 jaar op alle terreinen in mij gedaan hebben. Dat de weg eigenlijk belangrijker is dan het doel, staat soms in schril contrast met ons fanatisme; laten we dus nog maar lang onderweg zijn, opdat we dat beter leren. Toine en Quinter, wat bof ik met jullie als broers, en dat gevoel wordt alleen maar sterker. Jullie hebben allebei een geweldige vrouw in onze familie geïntroduceerd. Toine en Andrea, geweldig dat jullie voor deze gelegenheid hierheen willen komen. Quinter en Miriam, bij jullie is het altijd gezellig. Laten we elkaar vaak zien, ondanks alle opleiding, promotie en andere drukte. Miriam, heel veel succes ook met jouw laatste loodjes. Toon en Johanna, dank voor al die support bij dit proefschrift, praktisch, logistiek en geestelijk. Jullie voelden op de een of andere manier vaak goed aan wanneer een handje hulp gewenst was. Het is heerlijk dat ik jullie met Katelijne heb leren kennen.

Jan Willem, beste vriend sinds de Bhagwan-discotheek in 1986. Wat hebben wij al veel met elkaar meegemaakt en wat ben ik blij dat jij bij deze gelegenheid mijn paranimf bent. Quinter, je blijft mijn broertje, ook jij bent mijn ideale paranimf. Ik wil jullie beiden speciaal bedanken voor het overnemen van een aantal belangrijke klussen rondom de promotie, waarbij jullie ook nog even de redactie van de laatste hoofdstukken deden. Dat heeft mij veel ruimte gegeven om me op de laatste loodjes te concentreren.
Het beste bewaar je tot het laatste. Laura en Merle, wat een feest dat jullie bij ons zijn geboren; we zouden nooit meer zonder jullie kunnen. Ik weet zeker dat ik het meest trots op jullie zal zijn, als jullie op woensdag op de eerste rij zitten. Dank voor de tientallen tekeningen, de zaterdagochtenden dat we met de lego speelden (en nog gaan spelen), jullie verhalen en ongelofelijke creativiteit. Hiermee hebben jullie mij geholpen om te blijven zien dat je soms gewoon even moet gaan spelen. Laura, jouw zelfvertrouwen is een lust om te zien, evenals je lach als je weer iets grappigs herkent. Merle, ik zou nog wel honderd jaar met jouw gezang wakker willen worden. En…. gelukkig zijn jullie veel sneller dan ik in het maken van boeken!

Katelijne, dank voor de mogelijkheid die jij biedt om mij op de wetenschap te richten en al het andere dat je voor me betekent. Hoewel je mij ook op gezette tijden, op meer of minder diplomatiek wijze, naar het nut van mijn exercities vraagt, heb ook jij de afgelopen jaren veel in dit proefschrift geïnvesteerd. Ik realiseer me maar al te goed dat jij mij op meerdere momenten behoedt voor doordraven, wat vermoedelijk in mijn genen zit, maar niet minder mijn eigen verantwoordelijkheid is. Ik ga je niet beloven dat het allemaal anders gaat worden na 5 November, daarvoor ben ik te nieuwsgierig, heb ik teveel plezier in dit werk en hebben we niets aan loze beloftes. Het vinden van een juiste, iets evenwichtiger balans moet wel weer kunnen. Gelukkig zal ik maar eenmaal promoveren. Kus.

Eric Ruhe

Amstelveen, September 2008
Eric Ruhé werd geboren op 18 November 1968 in Amsterdam, als oudste van drie zoons. Toen hij vijf was verhuisde het gezin naar Overijssel, waar zij tien jaar op het platteland woonden. Hierna verhuisden zij naar Den Haag, waar hij in 1986 het eindexamen VWO behaalde.

Vanaf 1986 studeerde hij geneeskunde aan de Universiteit van Amsterdam, waar hij in 1991 het doctoraal examen behaalde. Hierna vond een reis naar Kenia en Tanzania plaats waarbij hij zijn eerste kennismaking met wetenschappelijk onderzoek had bij het uitvoeren van de resistentiebepaling van de malaria parasiet voor de middelen chloroquine en amodiaquine via de Afrikaanse flying docters (AMREF).


In deze periode volgde hij de postacademische opleiding tot klinisch-epidemioloog.


Eric Ruhé was born in Amsterdam on November 18th, 1968. He is the eldest of three sons. At the age of five, the family moved to Overijssel, a rural part of the Netherlands, where they lived for ten years. The family then moved to The Hague, where he completed his secondary school (VWO).

From 1986 until 1991, he studied medicine at the University of Amsterdam. His first encounter with scientific research was in Kenya and Tanzania, where he participated in the determination of resistance of plasmodium Falciparum against chloroquine and amodiaquine, a project conducted for the African flying doctors service (AMREF).

He started his internships in 1992 and received his medical degree in 1995, after spending another 6 months in Tanzania, performing a special internship in tropical medicine. He thereafter served his civil service at the Vrije Universiteit, at the institute for extramural medicine (EMGO). He developed a mortality and morbidity registration for the Hoorn Studie, which was until then a cross-sectional study. This mortality and morbidity registration is still operational until now. During this civil service, he received his masters in clinical epidemiology.

Since 1997, he has been working in psychiatry, first as a non-resident (AGNIO) at the program for Mood Disorders of the Academic Medical Center (AMC). From 1998 until 2002, he specialized in psychiatry at the AMC and Mentrum. During his residency he combined his epidemiological and psychiatric knowledge, and performed a clinical practice guideline project, funded by the AMC. This local guideline was also incorporated in the Dutch multidisciplinary guideline for depression which was developed at the same time. Three publications of this thesis result form the guideline project. Thereafter two grant applications were written and awarded, which enabled the other studies of this thesis. Since his registration as a psychiatrist in 2003, he has been working as a specialized clinical psychiatrist at the program for Mood Disorders at the AMC.

He shares his life with Katelijne Schmeink, they were married in 1998. Together they have two wonderful daughters: Laura (2000) and Merle (2002).