Summary and General Discussion
This chapter starts with a short review of the studies included in this thesis. This thesis covers two major subjects: the efficacy of combination treatment in major depressive disorder (MDD) and the prediction of outcome in MDD. Accordingly, this chapter is divided into two main parts. The first part includes methodological problems associated with efficacy research on combination treatment (chapter 2), an overview of earlier studies on this subject (chapter 3), a description of background and practice of interpersonal psychotherapy (IPT) (chapter 4) and the outcome of a randomized clinical trial comparing combination treatment to three other treatments (chapter 5). The second part of this thesis consists of a study on the ability to reliably assess personality factors during an episode of MDD (chapter 6), a study of the factors predictive of outcome, with a special emphasis on personality factors (chapter 7) and a study of the relationship between belonging to an ethnic minority and outcome (chapter 8). Per study a short review of the main findings is presented, followed, if relevant, by a short update of the literature and ending with discussion and conclusions. The final paragraph is devoted to a discussion of the relevance of the findings from a theoretical, clinical and public health viewpoint. This discussion results in some suggestions for further research.

**PART I: EFFICACY OF COMBINATION TREATMENT IN MDD**

**Summary of results**

**Chapter 2** is an overview of the methodological and conceptual problems encountered in the study of combination treatment for depression. Nine recommendations were given to improve studies in this field.

a) The form of psychotherapy used has been shown to be effective in earlier studies;

b) The primary outcome measure should be an international recognized measure (HAMD, MADRS or IDS). As secondary measures, scales measuring interpersonal relationships, health costs and other psycho-social parameters should be used.

c) When performing combination therapy, the delivery of pharmacotherapy and psychotherapy should be done by two persons or be delivered by the same person throughout the whole study.

d) The ‘statistical power’ of the study should be large enough to allow conclusions to be drawn.

e) Interpretation of the effects requires comparison of several treatment conditions.

f) The study should be performed by a group with expertise in pharmacotherapy as well as in psychotherapy.
g) Selection of patients should be kept to a minimum. Patients who have received prior treatment, patients with chronic depression and patients with comorbid personality disorders should all be included.

h) Some of the advantages found for combination treatment became evident in the follow up after the trial. Especially a reduction of the risk of relapse has been found to be an advantage for combination therapy over medication only. Studies with a sufficient long follow up period are needed.

i) In prior studies the drop out in medication cells was especially high. One of the modern antidepressants may well have an advantage in this area.

Since our publication on methodological issues (Blom et al., 2000) a paper by Westen, Novotny et al. (2004) on the validity of RCT for psychotherapy has been published. To summarize this article very briefly, Westen et al. argue against overly relying on the outcome of RCT’s for psychotherapy practice. Many of the problems stated in their paper are in fact addressed in our study. For instance the problem of co-occurrent personality problems is addressed in Part II of our thesis.

**Chapter 3** is an overview of the current literature on combination treatment. In our original publication (Blom, et al., 2000) 18 studies were found that compared a combination of psychotherapy and medication for the treatment of MDD to either psychotherapy alone, medication alone or both. The main conclusion of our review was that combination treatment did not have an advantage over single treatment in acute depression.

Four reviews on the relative efficacy of combination treatment have since been published (Friedman et al., 2004; Hollon et al., 2005; Otto et al., 2005; Pampallona et al., 2004). All four find a small but consistent advantage for combination treatment. Contrary to our review (Blom et al., 2000), all reviews cited both included studies of chronic (including patients with dysthymia) and acute depressed patients. At this point it is not known whether the distinction between acute versus chronic depressed patients is only in the duration of the disorder or that the difference lies in distinctive etiological and pathophysiological pathways. There is some evidence in support of the latter (Akiskal, King et al., 1981). Until this problem has been further elucidated, it does not seem to be correct to combine these two groups in one review. In the paper by Pampallona et al (2004), the two largest studies included only chronic depressed or dysthymic patients (Browne et al., 2002; Keller et al., 2000). In fact, these two studies combined had almost the same total number of patients as all studies on acute MDD taken together. The total ‘weight’ of these two studies (number of chronic patients included) is thus very large in this review and conclusions about the relative value of combination treatment for acute depression can hardly be drawn from this review.
Friedman et al (2004) and Hollon (2005) in their reviews both reach the same conclusion that “combination treatment can have advantages for selected patients groups”. Essentially they both mean, for chronic depressed patients.

The conclusion regarding the efficacy of combined treatment for acute MDD, that we reached in 2000, is essentially the same if one considers the most recent literature.

**Chapter 4** gives a description of the technique and background of interpersonal psychotherapy (IPT). On theoretical grounds IPT has credibility because it directly influences the way depressed patients cope with interpersonal stress and its subsequent consequences.

A short summary of the expanding literature on life events and the causal pathways by which life events can lead to depression in patients genetically susceptible for MDD is included.

IPT is used in the treatment of MDD, but it has also been applied as the treatment of other disorders such as, post-traumatic stress disorder (Bleiberg et al., 2005), borderline personality disorder (Markowitz et al., 2006) and bipolar disorder (Frank, 2005). Since the publication of the papers on which chapter 2 was based, one review (de Mello et al., 2005) on IPT in the treatment of depression was published. In general, the conclusion of this review was that IPT is comparable in effectiveness to medication (in the treatment of MDD) and more effective in comparison to pill-placebo. Also, IPT may have a greater efficacy than CBT. However, this conclusion is based on a limited number of studies.

In our study we did not include patients with dysthymia (patients with chronic major depression were included) because of our finding in a pilot study (Blom et al., 1996) that patients with dysthymia did not seem to respond to IPT. This finding has recently been replicated (Markowitz et al., 2005) underscoring our decision.

**Chapter 5** describes the main outcome of the Randomised Controlled Trial. In our study we were primarily interested in the question whether combination therapy would be more effective compared to either medication or IPT. In designing the trial we explicitly included a psychotherapy/placebo condition. Our primary reason for this was to study whether a possible advantage for the combination treatment over single treatment, should not be attributed to either increased time spent with the patient or other factors such as enhancement of compliance.

For the study, 355 patients were screened initially of whom 207 were randomized over four treatment conditions: IPT, nefazodone, the combination of these two and the combination of IPT and placebo. Between randomization and start of the trial, 14 patients refused treatment or did not show up for the next appointment. Finally 193 patients entered the study. Patients were treated for 12 to 16 weeks depending on the occurrence of missing sessions of IPT in the first 12 weeks. Of the 193 patients who entered the study, 138 (71.5%) completed the full trial. We did not find any
difference in drop out rate between conditions. A higher drop out rate is sometimes seen in medication only conditions (e.g. de Jonghe et al, 2001) but this was not the case in the present study.

Using a multilevel analysis we found a clear advantage of combination therapy over the medication only condition on the MADRS. Repeating these analyses with HAMD scores as outcome measure these results could not be replicated.

When we looked at remittance (HAMD < 9) as outcome: again the combination of IPT and medication was significantly more effective than medication alone, but not than IPT alone. No other significant differences between conditions were found.

Finding inconsistencies between the MADRS and the HAMD is not entirely surprising. The MADRS was specifically designed to be more sensitive to change as compared to the HAMD (Montgomery & Asberg, 1979). Furthermore the HAMD has three items concerning sleep (item 4, 5 and 6) compared to the MADRS who has only one sleep-related item. Since nefazodone has a strong sleep enhancing effect, this could favor medication over psychotherapy.

With the CGI as outcome measure we found IPT to be more effective compared to nefazodone alone, but as the CGI may be more vulnerable to bias (Gaudiano & Herbert, 2005) this finding must be interpreted with caution.
Discussion

Methodological issues
We succeeded in implementing almost all of the recommendations from chapter 2 in our own study:

a. An empirically validated form of psychotherapy: The form of psychotherapy used has been shown to be effective in earlier studies. By choosing IPT as the comparator psychotherapy we insured ourselves of a widely recognized effective form of psychotherapy. One also “invented” elsewhere, so as not to create a “home-team advantage (see chapter 3).

b. Use of well known rating scales: We used three widely known and used rating scales for the measurement of depression severity, the HAMD, MADRS and BDI.

c. Execution of therapy: The delivery of pharmacotherapy and psychotherapy should be divided over two persons or be delivered by the same person throughout the whole study. Throughout the study, pharmacotherapy was delivered by experienced physicians, following a standard protocol. Time with patients in the pharmacotherapy arm of the study was restricted so as to maximize differences between the pharmacotherapy and psychotherapy condition. One of the critiques of earlier studies (e.g. Elkin et al., 1989) was that if for instance physicians would spend 30 minutes weekly with their patients, a supportive relationship would form and one would then compare one form of combination therapy (medication plus supportive therapy) to another combination therapy (psychotherapy plus medication).

In our study the psychotherapy and pharmacotherapy were delivered by two persons throughout the study. A psychiatrist could perform psychotherapy, but in these instances another physician would deliver the pharmacotherapy.

d. Sufficient statistical power: We calculated that with a sample size of 62 per treatment group this would provide 90% power to detect a difference in response rate of 25% (i.e. 50% response in both single groups versus 75% response in the combination treatment). With an $\alpha$ of 0.05 we set a target of 250 participants. To detect a more modest difference of 20%, the statistical power would be 74%. However, inclusion into the trial was stopped when we reached a total of 207 included patients.

Stopping was for practical reasons: because of a major merge between mental health care organizations in The Hague, many of the participating therapists and psychiatrists either moved to another location or left the organization altogether. So eventually, with approximately 50 included patients in each treatment condition, the statistical power to detect a 25% difference in response rate with alpha of .05 was 84.7%.

e. The total number of treatment conditions can not be extended without limit. In our study we did not choose for more than four conditions for obvious reasons: the number of patients needed would far exceed our possibilities. Another reason to limit the number of conditions was that we did not
want to include a no treatment or placebo-only condition. With effective treatments available we did not consider this ethically correct.

f. The study should be carried out by experts in both pharmacotherapy and psychotherapy. Before the start of this study, several medication trials were carried out in the organisation(s) where the present study was executed (Hoencamp, Haffmans et al., 2000; Hoencamp & Haffmans, 1991; Hoencamp, Haffmans et al., 1994a; Hoencamp, Haffmans et al., 1994b; Nolen, Hoencamp et al., 1993). Expertise in psychotherapy and IPT in particular in the present study was acquired before the start of the trial as described in chapter 5.

g. Exclusion of patients should be kept to a minimum: of the 355 patients screened for this study, the majority (207 or 58%) was included in the trial. This compares favourably to the percentage of 35 as reported for other treatment studies of depression (Westen et al., 2004). Although our trial was an efficacy trial, exclusion criteria were kept to a minimum. Patients with comorbid personality disorder and other axis I disorders could be included. We therefore believe that the findings of our study can be generalized to standard outpatient practice in secondary and tertiary care centres.

h. Follow up. Patients who were remitted at the end of the trial were followed up to a year. Patients who didn’t remit were given another treatment and exited the study. Because the number of patients remitted was small, no further analysis of these data was made. An attempt was made to contact all patients who had participated in the study. We did however not succeed in recovering a sufficient number of patients to make a reliable analysis possible.

i. Drop out should be kept as low as possible. By selecting a modern antidepressant for our study (nefazodone) we had hoped to reduce attrition. Although attrition was not high (28.5%), it was neither very low. We can conclude that we only partially succeeded in our objective. However, minimizing contact with the therapist in the pharmacotherapy arm of the study did not result in an increase in attrition. In fact, attrition in the pharmacotherapy arm was comparable to that in the combination arms of the study.

At the outset of our study we hypothesized that the attrition in the two combination arms of the study would be small. It has been suggested that one of the advantages of combination therapy could lie in its positive effect on attrition (Paykel, 1995). In our study we found attrition in the combination arms similar to that found in both single treatments.

Attrition also did not differ between the medication and the psychotherapy conditions. Reasons for drop out did of course differ between conditions with “no-show” being more frequent in the psychotherapy condition and “side effects” in the medication condition.

Since attrition as reported in other outpatient studies is almost never much lower than as observed in our study, it is well possible that the figure found in this study is a reflection of the “real world setting” i.e. a “natural bottom level”.
Outcome
Comparing our findings to the literature in general is not easy due to various methodological and patient recruitment factors. For instance the statistical method used in our study (i.e. multilevel analysis) is not often found in other studies despite the clear advantages (Gibbons et al., 1993). Keller et al. (2000) used a very similar design (although without inclusion of a pill-placebo + psychotherapy condition) in the treatment of chronic depressed patients. The result of this study shows a better general outcome and there was a clear advantage for the combination of psychotherapy and medication over both conditions alone. The reasons for these differences could be several as pointed out in chapter 5. The larger amount of time spent with patients, a different form of psychotherapy and the type of patients included, all could contribute to the differences in outcome. As has been pointed out by others (Rubin, 2000) the lack of a psychotherapy-pill-placebo condition makes it hard to definitely evaluate the findings of this study.

The most important difference with our own study is however that the Keller et al. (2000) study only included patients with chronic forms of depression whereas our own study was aimed at acute depression. Nonetheless 21.7% of our patients had a duration of the index episode at baseline of longer than 12 months and 4.6% even had a duration longer than 24 months.

Studies combining IPT and medication in acute MDD are very rare. An early study by DiMascio et al. (1979) showed an advantage of combination over both medication and IPT only. This study did not contain a psychotherapy + pill-placebo condition, so that it remains an open question if the advantage found for the combination is a specific treatment effect or the result of enhanced non-specific effects of the combination. A study by Reynolds et al. (1999) found the combination to be superior over medication alone and over the combination of IPT + pill-placebo. This study was however hampered by the fact that in the IPT + pill-placebo condition treatment had to be discontinued if patients failed to improve by week 8. As has been shown before, patients receiving IPT seem to improve relatively late as compared to medication patients (DiMascio et al., 1979; Klein & Ross, 1993; Watkins, Leber et al., 1993) and it could well be that patients were discontinued before the effect of IPT could fully set in.

Our study confirms that IPT alone is at least as effective as medication. This has been consistently shown in the treatment of acute depression (Hollon et al., 2005). Adding medication to IPT does not improve final outcome. A possible advantage for combination treatment could be the pace of change. If one looks at figure 3 in chapter 5, it seems that major improvement of symptoms was reached in the nefazodone condition in the first 6 weeks of treatment. In the second half of the study almost no improvement was seen. Although not as clear, the reverse seems to be true for the IPT condition, with a relatively small improvement in the first half and more substantial gains made in the second half of treatment. The same phenomenon is found in the Keller et al. (2000) study. However, this rather impressionistic interpretation of differences in the pace of change between
pharmacotherapy and psychological treatment has to be interpreted with great caution because they are not supported by a statistically significant group x time interaction effect and further studies with sufficient statistical power are needed to validate the possible existence of different change trajectories between pharmacotherapy and psychological treatment. If these patterns would be confirmed, it would favour combination treatment which combines two advantages: relatively fast improvement due to medication in the early stages of treatment and further improvement in the later stage when psychotherapy kicks in. For some patients, with for instance severe sleeping problems, suicidality and/or difficulties in daily life, a rapid decrease of symptoms would be especially important. It would also entail that both parts of the combination treatment should preferably start at the same time. Starting medication first and adding psychotherapy a few weeks later, would nullify this advantage. This is contrary to what is sometimes said about sequential treatment (Frank, Kupfer et al., 2007).
PART II: PREDICTION OF OUTCOME

Chapter 6 deals with the question whether personality factors can be reliably measured during an episode of MDD. Personality factors were assessed using the NEO-FFI. This instrument measures five dimensions of personality (Neuroticism (Ne), Extraversion (E), Openness to experience (O), Altruism (A) and Conscientiousness (C)).

It has long been argued that personality factors can not be reliably assessed during an episode of major depression. In fact, in an earlier publication we argued the same (Griens et al., 2002). In this study we have shown that this may not be true. We found that the structure of the Big Five personality factors remained stable during treatment for MDD. Participants did, however change over time: Ne and A scores changed compared to baseline. Although changes were significant, the absolute change was very small. Of note is that participants differed substantially from normative data obtained in the general population (Costa & McCrae, 1992). Since we did not perform a prospective study we do not know whether this reflects a vulnerability for depression prior to the onset of the episode or a ‘scar’ effect (Akiskal et al., 1983). In one of the few prospective studies (Ormel et al., 2004) Ne was shown to be high before the onset of depression, even higher during an episode of MDD and return to the original level upon recovery. This could well mean that all above mean scores on one or several of the personality traits of the big Five constitute a risk factor for subsequent episodes of MDD.

We found that changes in depression scores on the HAMD were relatively independent from changes in personality traits across the duration of the study. In our study we both examined a self-rating scale (the BDI) and an expert-rated scale (HAMD). We found important differences in correlations between changes as measured with both scales. Overlap between the change score on the HAMD and the NEO-FFI is not strong, with HAMD scores predicting 18.8% of variance in NEO-FFI scores of which only Extraversion (E) has a significant relationship with depression change score. Of note is that Neuroticism (Ne) does not have a significant relationship with depression change scores on the HAMD. Using the BDI, on the other hand, Ne does have a large and significant relationship with depression change scores (on the BDI) explaining 30.0% of the variance. Vice versa Ne explains 17.1% of variance in depression change scores. It is not clear what the exact cause of the overlap between Ne change scores and change scores on the BDI is. It could be that scores overlap because both are assessed with self-rating scales. Moreover, there could be an overlap in item content. This in turn would imply that the BDI measures, at least partly, trait rather than state features (Groth-Marnat, 1990). This interpretation is supported by the fact that at baseline
Ne and the HAMD show a smaller correlation ($r = 0.206, p=0.008$) compared to Ne and BDI at baseline ($r = 0.320, p < 0.0001$). Interestingly, the MADRS did not show a statistically significant correlation with Ne at baseline ($r = 0.094, p = 0.238$).

Differential or rank-order stability was also studied. Stability coefficients were high. Differences between personality traits were very small. This is noteworthy since Ne is often found to be less stable than the four other traits (Robins et al., 2001).

As found before (Clark et al., 2003; De Fruyt et al., 2006) personality traits as assessed by the NEO-FFI did not change in any important way during treatment, although therapy was effective in reducing depressive symptomatology.

Individual-level continuity was also found in our study. By far the majority (86.1%) of participants stayed stable on all five personality factors. A reliable change on one personality factor was shown by 6.5% and only 1.4% of participants showed a reliable change on two factors. Of the five factors, Ne was the least stable, and O the most stable factor.

In the same line, ipsative continuity was also found to be high. Only variability in shape was greater than expected in 12.5% of participants. Variability in scatter and elevation were either not significant or greater than expected in only a small minority of participants (1.9%). Again this is in accordance with the earlier study by de Fruyt et al (2006).

In chapter 7 we report on the different potential predictors of remittance and symptomatic improvement, with a special emphasis on personality dimensions. Univariate analysis showed a significant relationship of outcome with severity, duration of index episode, and use of medical services (UMS). None of the personality variables was predictive of outcome in the univariate analyses. A hierarchical logistic regression was carried out to test for significant predictors of remittance. Subsequently a multiple regression analysis was used to investigate variables predictive of changes on the HAMD as dependent variable. Both types of regression analyses showed that these disease-related variables each uniquely predicted outcome, but that the five personality factors did not contribute to the prediction model. Neither did we find personality factors to be differential predictors.

The main conclusion of this chapter is that, contrary to expectation, personality factors as assessed by the NEO-FFI do not predict response in the acute treatment of MDD. Only severity, duration of the current episode and to a lesser extent, use of medical services, significantly predict final outcome.

Chapter 8 involved a post-hoc analysis of ethnic minority patients (EMP) included in the trial. In writing up our findings, we were impressed by the paucity of studies comparing the treatment response of EMP to non-EMP. The literature on treatment of MDD in EMP is substantial, but to our
knowledge no RCT’s have been carried out in this area. This is a surprising finding, considering the widely acknowledged mental health problems among these patients.

In our study EMP start with significantly higher mean baseline score on both the HAMD and the MADRS. At endpoint again, mean scores on both instruments were significantly higher compared to non-EMP. Both groups showed a similar rate of improvement. Attrition rates however were significantly higher in EMP compared to non-EMP.

We expected that EMP would have a more ‘somatic’ presentation of symptoms and that this would reflect itself in the scores on the HAMD. This expectation was not confirmed. EMP did score significantly higher on only three items: ‘mood’ (item 1) and two of the three ‘sleep’ items. When using a Bonferroni correction, only one item (“early sleep”) remained significantly higher in EMP compared to non-EMP.

As expected, attrition among EMP was significantly higher compared to non-EMP. High attrition among EMP has been found before and a variety of reasons have been suggested for this problem. In our study, reasons for attrition were collected and EMP were notably more prone to drop out of treatment because of no show.

Importantly, we found that participation of EMP in a randomized clinical trial is feasible.

**DISCUSSION**

The main findings of this second part of the thesis can be summarized as follows: Personality factors can be reliably assessed during an episode of major depression. These personality factors do not predict short term outcome. Belonging to an ethnic minority group predicts high attrition in our study. The rate of improvement of ethnic minority patients is however the same as for ethnic Dutch patients.

Measurement of personality factors can be relevant for clinicians in predicting long term vulnerability to the occurrence of relapse in MDD. We have shown in our study that personality factors can be reliably assessed by the NEO-FFI during an episode of MDD.

Our results also point out that personality factors and depression can best be seen as separate constructs. An interesting finding, warranting further study, is the observation that change on the BDI shows a much higher overlap with change scores on the Neuroticism scale as compared to change scores on the HAMD. While both the BDI and the NEO-FFI are self-rating scales and the HAMD is an independently assessed scale, the convergence between BDI and NEO-FFI could be based upon similarity in the content of the items and lack of convergence between HAM-D and NEO-FFI could be caused by differences in interpretation of items.

Personality factors did not predict outcome. This finding has been observed before (Clark, 2005; Clark et al, 2003). Patients in our study generally had high scores on all personality dimensions,
foremost Neuroticism (Ne). This is in accordance with studies by both Ormel et al., (2004) and Costa et al., (2005), who both found that personality dimensions during an episode of major depression are accentuated. Ormel et al (2004) also found that Ne scores were elevated prior to the onset of a depressive episode. This can be seen as an indication that high Ne points to a vulnerability for the onset of MDD (Enns & Cox, 1997).

It may well be that Ne does predict the long term outcome, especially the risk of relapse, in MDD. For the short term this would imply that patients with an abnormal personality profile, as measured by the NEO-FFI should be treated as any other patient. This is an important finding because clinicians tend to think that patients with personality problems are harder to treat and respond less to therapy. In our study this has not been confirmed: we did not find any relationship between personality scores (as measured by the NEO-FFI) and outcome.

The question remains whether personality disorders, not factors, do predict outcome in acute depression. Three recent reviews (Mulder, 2002; Newton-Howes et al., 2006; Reich, 2003) on this subject give a mixed view, with two studies (Reich, 2003; Newton-Howes et al., 2006) concluding that comorbid personality disorder does negatively predict outcome in MDD and Mulder (2002) concluding that essentially because of a lack of well-controlled trials no clear conclusion on this topic could be drawn.

For the clinician our findings can be a small blessing: he or she should not be overly concerned by degree of personality pathology in his or her patients. If a diagnosis of MDD exists treatment should be applied regardless of the severity of any personality pathology.

As part of this thesis we performed a post-hoc analysis of the subgroup of patients from ethnic minority groups. First, we believe that more study needs to be done for this group of patients. With all the restrictions of a post-hoc study weighted, we do feel that some important lessons can be learned. We found that patients from ethnic minorities were generally willing and able to participate in a randomized clinical trial. This should lead to trials specifically aimed at these patient groups.

The analysis has further taught us that the interpretation of results obtained in EMP should be especially careful. For instance outcome can be interpreted in two ways: in absolute terms outcome for EMP was worse compared to non-EMP: endpoint HAMD was significantly higher and only few EMP were in remission by the end of the trial. However, no differences in relative change over time between both groups were found. It is our opinion that the latter finding is of more importance than the former. Finding a higher score on both the HAMD and the MADRS at baseline and at the end of treatment could well be due to cultural influences. This would imply that cut-off scores currently used should be different for EMP. Future study should further elucidate this.

Attrition was significantly higher in EMP. Strategies for the improvement of adherence to either pharmacotherapy and psychotherapy need to be developed. One course of action would be to see whether clinicians with the same ethnical background realize a lower attrition rate. A second
strategy sometimes deployed (Grote, Bledsoe et al., 2004), is to make care more accessible to these patients.

In our study we used IPT for the first time in this patient group. In the rate of change, EMP in the IPT condition did again not differ from non-EMP. This could be a potentially major finding because, according to practitioners in the field and one earlier study, the application of CBT in EMP seemed to be particularly difficult (Markowitz et al., 2000). From our experience in this trial we feel that several adaptations of IPT are needed to make it a more effective treatment for EMP. It is beyond the scope of this thesis, but adaptations such as length and frequency of sessions, introduction of a special focus for EMP (migration) and changes in the pace and wording of interventions are needed. Ideally such an adaptation of the IPT manual should be studied in clinical trials including enough patients from different ethnic groups so as to make empirically based clinical judgments possible.

**Concluding remarks and recommendations for future research**

**Clinical implications**

We have treated nearly 200 patients with a variety of treatments: psychotherapy, medication and both. At the end of the study we found that the combination treatment worked better as measured with the MADRS (but not with the HAMD) compared to medication only but not to psychotherapy only. Also, remittance as measured with the HAMD was significantly higher in the combination treatment as compared to nefazodone alone, but not to IPT alone.

**Consequence of the study for interpersonal psychotherapy.**

It is clear from this study that IPT is a good option in the treatment of MDD. Neither medication nor the combination of psychotherapy and medication had an advantage on any of the outcome measures over IPT alone. Psychotherapy has none of the disadvantages associated with medication: side effects, possibility of teratogenic damage, tolerance nor the problem of withdrawal. This noted we also found that patients with more severe forms of depression and with a longer duration of the index episode faired poorly. Our study showed that neither IPT, nor medication nor the combination improved the outlook for these patients. This is in line with for instance Parker, Parker et al. (2006) who state that IPT may be especially suited for the treatment of non-chronic depression. IPT is not a universal treatment suited for all types of depression, so much is clear from our study and we agree with Parker et al. (2006) in this. We disagree however in his finding that combination treatment does not offer advantages over medication. In our study, the reverse is rather true.

From this study we can not learn which treatment should be offered for chronic depression.

Recently, interest has heightened in a new form of integrative psychotherapy, Cognitive Behavior-
Analysis System of Psychotherapy (CBASP) specially designed for the treatment of chronic depression and dysthymia. This is a promising development, but it should be kept in mind that CBASP was studied in only one randomized clinical trial (Keller et al., 2000) and that in this trial CBASP was compared to nefazodone with equal efficacy. Put another way, there was no difference in outcome between CBASP and the same medication (nefazodone) as we used in our trial. Another promising venue for these patients is the development of more intensive specialized treatments, such as combinations of different forms of psychotherapy, medication and for instance running therapy (Blom, 2006).

All proposed forms of treatment (CBASP, intensive therapy) are more or less treatments of a limited duration. One option could also be an extension of therapy. The Health Council of the Netherlands (Gezondheidsraad) in its advice to the government on the position of long-term psychotherapy, explicitly recommends the study of long-term psychotherapy for chronic depression (Gezondheidsraad, 2001). The alternatives that have been developed for chronic depression are slightly longer in duration but primarily more intensive. It is our opinion that these should be studied first.

Final comments and recommendations for future research

In the last few years, several large studies, including our own, have shown that the treatment of MDD is a lot more daunting than optimists stated before. Very recently the main results of the STAR*D study have been published (Rush et al., 2006). In this large study, encompassing 3671 patients with MDD, patients were treated according to a complex algorithm. All patients received citalopram as first line treatment. After 8 weeks, 36.8% of patients were remitted. Those not remitted subsequently received four sequential steps involving switching, augmentation and combination strategies. CBT was part of the study but few patients were referred to this condition (Thase, Friedman et al., 2007).

The cumulative rate of recovery at the end of the study was 67%. However, using a much stricter outcome criterion: sustained recovery over the duration of the trial, only 43% of the patients were remitted by the end of the study.

Bringing together these findings with the findings from our own study, the following clinical implications and suggestions for further research can be given:

1. Psychotherapy should be the therapy of choice, over medication in the treatment of patients with MDD of moderate severity and with a duration of the index-episode of less than 12 months.
2. The combination of severe symptoms and a longer duration of the index episode has a poor prognosis. Almost none of the patients with a duration of the index episode longer than 12 months and a severe depression (baseline HAMD > 20) was remitted at the end of the study. It does not seem fair to offer these patients, as first line treatment, therapy consisting of short term psychotherapy, medication or both. For these patients more effective forms of treatment have to be developed and studied. Special emphasis should be on the study of intensive forms of treatment within a limited time span.

3. Duration and severity predict for a large part outcome in acute treatment of MDD. This finding can be used to further a staging method in MDD. In medicine, clinical characteristics of an illness are used to stage the illness and guide therapy. Cancer therapy is the most common example, but the staging method is used in other areas of medicine as well. It has been rarely used in psychiatry (Fava & Kellner, 1993; McGorry, Hickie et al., 2006). Staging, using factors such as the number of previous episodes, severity, duration and comorbidity can help to guide therapy. For this to work, specific forms of depression such as seasonal affective disorder, bipolar depression, psychotic depression should be distinguished because specific treatments have been tried and found successful for these groups. For the large remaining group of MDD stages could be construed, each needing its own treatment approach. A further advantage of staging could be that it would facilitate discussion between different care providers. This principle needs to be tested empirically.

4. Change in depression on the BDI and change in Neuroticism showed a much higher overlap as compared to change on the HAMD and change in Ne (chapter 6). In future studies the reasons for this overlap should be studied further. One possible reason for this finding could lie in the way the instrument is administered (self rating versus expert rating). Another reason could be that especially the BDI measures in part trait related symptoms, meaning that the BDI and the Ne factor of the NEO-FFI measure, at least in part, measure the same phenomena. We recommend that self-rated instruments and expert-rated instruments that are essentially similar (e.g. the Inventory of Depressive Symptoms (Rush et al., 1996) be used to study this question further.

5. Neuroticism, in fact all personality factors, was not found to be predictive of short term outcome. For the clinician this entails that patients with and without high scores on any personality factor can be treated as every other patient. Since measurement of personality factors during an episode is possible and valid, the role of personality factors in the prediction of relapse and long term prognosis should and can be studied. It is striking that even at the end of treatment, Neuroticism scores were generally still very high. This could be a reflection of an enduring propensity for relapse.
6. IPT should be adapted for patients from ethnic minorities and be studied further. For the large group of patients from different ethnical backgrounds it will be necessary to adapt IPT for the special needs of this population. We do however believe that the model of IPT holds special promise for these patients.

7. To be able to test for instance the proposed principle of staging (see #3) large numbers of patients are necessary. At the same time, the STAR*D trial has made it clear that expertise in treating depressed patients, especially those not responding to standard treatment, needs able physicians (Nolen, van den Broek et al., 2007). Large numbers of patients coupled with sophisticated research entails that researchers should move away from single studies to a more continuous model of research. For this, close cooperation between academic and specialized treatment centres is imperative. The clinician performing research outside academic centres is facing a difficult task. Besides this, after a study is done, the whole ‘machinery’ of the study is often dismantled and expertise is lost. By cooperation resources and large number of patients are brought together. Studies should be carried out on a continuous basis with different research questions introduced over time.