Chapter 5

Treatment of unipolar psychotic depression
A randomized, double-blind study comparing imipramine, venlafaxine, and venlafaxine plus quetiapine

Jaap Wijkstra MD,1 Huibert Burger MD PhD,2–3 Walter W van den Broek MD PhD,4 Tom K Birkenhâger MD PhD,4 Joost GE Janzing MD PhD,6 Marco PM Boks MD PhD,1 Jan A Bruijn MD PhD,4,5 Marc LM van der Loos MD,7 Leonie MT Breteler MD,8 Guillaume MGI Ramaekers MD MSc PhD,9 Robbert J Verkes MD PhD,6 Willem A Nolen MD PhD.10

1 Rudolf Magnus Institute of Neuroscience, Department of Psychiatry, University Medical Center Utrecht, Utrecht, The Netherlands
2 Interdisciplinary Center for Psychiatric Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
3 Julius Center for Health Sciences and Primary care, University Medical Center Utrecht, Utrecht, The Netherlands
4 Department of Psychiatry, Erasmus Medical Center, Rotterdam, The Netherlands
5 GGZ Regio Breda, Breda, The Netherlands
6 Department of Psychiatry, Radboud University Nijmegen Medical Center, The Netherlands
7 Department of Psychiatry, Isala Klinieken, Location Sophia, Zwolle, The Netherlands
8 Department of Psychiatry, St. Antonius-Mesos Hospital, Utrecht, The Netherlands
9 Department of Psychiatry, Waterland Hospital, Purmerend, The Netherlands
10 Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Funding/Support
The study was supported by grants from AstraZeneca and Wyeth Pharmaceuticals, both of which also provided the study medication.

Abstract

Objective: It remains unclear whether unipolar psychotic depression should be treated with an antidepressant and an antipsychotic or with an antidepressant alone.

Method: In a multi-center RCT, 122 patients (18-65 years) with DSM-IV-TR psychotic major depression and HAM-D-17≥18 were randomized to 7 weeks imipramine (plasma-levels 200-300 µg/L), venlafaxine (375 mg/day) or venlafaxine-quetiapine (375 mg/day, 600 mg/day). Primary outcome was response on HAM-D-17. Secondary outcomes were response on CGI and remission (HAM-D-17).

Results: Venlafaxine-quetiapine was more effective than venlafaxine with no significant differences between venlafaxine-quetiapine and imipramine, or between imipramine and venlafaxine. Secondary outcomes followed the same pattern.

Conclusion: That unipolar psychotic depression should be treated with a combination of an antidepressant and an antipsychotic and not with an antidepressant alone, can be considered evidence based with regard to venlafaxine-quetiapine versus venlafaxine monotherapy. Whether this is also the case for imipramine monotherapy is likely, but cannot be concluded from the data.

Trial registration: http://www.controlled-trials.com/isrctn/trial ISRCTN36607067
Introduction

Compared to non-psychotic depression, psychotic depression (major depressive disorder (MDD) with psychotic features) is more severe, more incapacitating, has a lower likelihood of placebo response, a longer duration of episodes, and more recurrence of psychotic features in subsequent episodes. Up to 20% of patients with a MDD episode meet criteria for psychotic depression.

Only 10 randomized controlled trials (RCTs) have addressed pharmacological treatment in psychotic depression. In a recent meta-analysis of these studies the combination of an antidepressant and an antipsychotic was found significantly more effective than an antipsychotic alone, but not than an antidepressant alone. Nevertheless, in many guidelines it is considered evidence based to treat psychotic depression with a combination of an antidepressant and an antipsychotic. However, in a recent review Mahli et al. are more careful in their recommendations stating that in the treatment of psychotic depression “the combination may be more effective, but not all studies support this”. To our opinion, treatment with an antidepressant and antipsychotic is, because of the greater risk of adverse effects, only justified when there is good evidence that combination treatment is superior to monotherapy.

The present study compares the efficacy and safety of the combination of an antidepressant and an antipsychotic with that of two antidepressants given as monotherapy. Because placebo response in this severely ill patients is unlikely, we considered placebo treatment in these severely ill patients undesirable. Moreover, we anticipated problems to get approval from our ethical review board for including a placebo-arm and – if we would have got approval – with recruitment.

The tricyclic antidepressant (TCA) imipramine was chosen as the reference treatment, based on good effect in a previous study. The serotonin and noradrenaline reuptake inhibitor venlafaxine was chosen as alternative monotherapy, based on its reported superior effect among depressed inpatients compared to placebo or to the selective serotonin reuptake inhibitor fluoxetine. As combination treatment, we selected venlafaxine plus the atypical antipsychotic quetiapine. We preferred an atypical antipsychotic to one of the classical antipsychotics because of better tolerability, especially regarding extra pyramidal symptoms.

Aims of the study

Aim of the study is to find an answer to the question whether patients with unipolar psychotic depression should be treated with the combination of an antidepressant and an antipsychotic or with an antidepressant alone.
Method

Patients

Hospitalized patients aged 18 to 65 years could be included if they met DSM-IV-TR criteria for a unipolar major depressive episode with psychotic features with a score ≥18 on the Hamilton Rating Scale for Depression-17 items (HAM-D)\(^{18}\) both at the screening visit and at baseline, i.e. the day prior to start of medication. Exclusion criteria were an acute indication for electroconvulsive therapy (ECT); mental retardation; alcohol or substance abuse or dependence within 3 months of enrollment; any serious somatic illness; somatic medication affecting mood; contraindications for study medication; adequate treatment of the current episode with imipramine (≥4 weeks with adequate plasma levels) or venlafaxine (≥4 weeks ≥ 300mg/day).

Study Design

This study was an investigator initiated (WAN), double-blind RCT with eight participating centers in the Netherlands. Recruitment took place between June 2002 and June 2007. The study was approved by the ethical review board of the University Medical Center Utrecht, and by the local review boards of the participating centers, and performed according to the rules of Good Clinical Practice. All patients, or their legal relatives in case of incapacity, gave written informed consent prior to enrollment. Before inclusion, all patients were without psychotropic medication for at least 4 days. After inclusion, patients were randomized 1:1:1 to 7 weeks double-blind treatment with imipramine, venlafaxine, or venlafaxine plus quetiapine, while stratifying for center, using randomly permuted blocks of size six. Randomization was executed centrally (UMC Utrecht) by using a computer-generated randomization list. The code could be opened during the study period of 7 weeks in case of medical emergency. After this period, the code was broken in order to inform the clinician how to proceed treatment, but only after two weeks when all data of that patient were monitored and entered into the database. Another argument for unblinding was that it was organizationally impossible to keep patients on double-blind medication throughout the total study duration.

Patients received a predefined number of capsules/tablets (6 capsules imipramine or placebo, 5 capsules venlafaxine or placebo and 3 tablets quetiapine or placebo), using the double-dummy technique. Throughout the study, it was aimed to reach an adequate plasma level (for imipramine) or the maximum dose (for venlafaxine and quetiapine). To ensure blindness, blood was taken from each patient. Imipramine was initiated at 75 mg/day. After two days, this dose was doubled and after another 5 days, a first blood sample was drawn. Based on this plasma level
A randomized, double-blind study

the dose was adjusted at day 10 according to a dosing algorithm ranging from 75 to 450 mg/day to obtain a plasma level for imipramine plus desmethylinimipramine of 200-300 ng/ml (19). Further imipramine plasma levels were determined at every weekly visit and, if necessary, the dose was adjusted. Venlafaxine (extended release form) was started at 75 mg/day for the first 2 days, then 150 mg/day for days 3-5, 225 mg/day for days 6-9, 300 mg/day for days 10-16, and finally 375 mg/day. Quetiapine (immediate release form) was initiated at 100 mg/day for 2 days, then 200 mg/day for days 3-5, 400 mg/day for days 6-9, and finally 600 mg/day. All dose adjustments were done by an independent psychiatrist at the UMC Utrecht, based on information about plasma levels (imipramine) and on information (received by fax) from the treating clinician regarding tolerability. As concomitant psychotropic medication, only benzodiazepines at a maximum of 3 mg lorazepam equivalent per day were allowed. Treatment compliance was secured by intake under supervision and weekly checking the ward medication registration. Patients missing medication for ≥3 days were considered dropouts. To check whether blindness was preserved, the rater was asked to guess the actual given treatment at treatment week 5.

At baseline, diagnosis was confirmed with the Structured Clinical Interview for DSM-IV Axis I disorders. Baseline assessments also included a psychiatric history, assessment of adequacy of treatment(s) in the current episode using the Antidepressant Treatment History Form, a medical history and a physical examination including vital signs and routine laboratory assessments. All assessments were done by trained physicians.

Severity of depressive symptoms was assessed at baseline and then weekly using the HAM-D (17 items) and the Clinical Global Impressions (CGI) severity as well as the CGI-change scale. In addition, hallucinations or delusions including whether they were mood-congruent or mood-incongruent were documented at baseline and then weekly. HAM-D, CGI and the other data collection methodology was discussed at regular meetings with the researchers from all participating sites. Interrater reliability as indicated by the intraclass correlation coefficient and based on three patients and eight raters (from different study sites) was 0.93 (95%CI: 0.74; 1.00) for the HAM-D. Vital signs were assessed each week, as were adverse events by inquiring about any unpleasant feeling and, if present, rated mild, moderate or severe.

**Statistical Analysis**

The primary outcome measure was response to treatment, defined as a ≥50% decrease in HAM-D scores from baseline to study endpoint, and a final HAM-D score of ≤14. Secondary outcomes were response on the CGI-change, defined as a score of ‘much improved’ or ‘very much improved’, differences in mean change between
treatment arms in HAM-D and CGI-severity scores from baseline, and absence of psychotic features at endpoint. An additional secondary outcome was time to response. Although this was not predefined as an outcome measure in the protocol, we also analyzed remission rates (defined as HAM-D ≤7), because also remission of symptoms is considered an important outcome.24

As we were concerned with the possibility that the advantage of a better efficacy might be diminished by a poorer intolerance, we decided a priori that a difference in response rates of ≥25% between treatment groups would be clinically relevant. Therefore, the study was aimed at comparing three groups of 55 patients (α = 0.05 and β = 0.20).

All analyses were carried out according to the intention-to-treat (ITT) principle. Therefore, all subjects with baseline data available were included in the analysis. We compared the baseline characteristics of all randomized patients with those of patients who completed all follow-up visits. Differences in response rates, remission rates and proportions of absence of psychotic features at endpoint were expressed as risk differences (RD). In these analyses, we applied the ‘last observation carried forward’ (LOCF) method in case of drop out. Using logistic regression, we performed additional analyses to investigate whether the results changed when adjusting for inequalities in baseline indicators of response probability. For the primary outcome, a subgroup analysis was carried out among completers.

Differences between groups in change from baseline to endpoint of the HAM-D and the CGI, were first compared using the LOCF method. The unpaired t-test was used for these comparisons.

Next, we estimated differences in weekly change of the HAM-D and the CGI scores using linear mixed models for fixed and random effects.25 The mixed models analyses were also ITT analyses, using all available data. The dependency of the repeated assessments of the outcome variables was taken into account by including a random intercept and random slope for patients. Differences in mean change in HAM-D scores and CGI scores during follow-up between the treatment groups were quantified as the fixed effect of group by time interaction. To test for potential non-linearity in the change of HAM-D and CGI scores over time, we additionally included a time-squared variable. In all these analyses, time was entered as a continuous variable. Mixed models with time as a categorical variable were built to obtain estimated marginal mean HAM-D and CGI scores at each follow-up visit for graphical presentation. In addition, differences in time to response between treatment groups were evaluated on basis of ITT using survival analysis. To this end, Kaplan-Meier curves were constructed for the graphical comparison of the time-related cumulative proportion with response between treatment groups. Patients who dropped-out were censored. Differences were tested for statistical significance using the log rank test.
For testing observer blindness, agreement between actual treatment and treatment guess at treatment week 5 was assessed by calculating the Kappa statistic, with values less than 0.4 indicating poor agreement.26

Differences in the prevalence of adverse events between treatment groups were tested using Fisher’s exact test. The precision of the efficacy estimates was expressed as a 95% confidence interval (95% CI). The level of significance was set at 5% and p-values were two-sided. All analyses were conducted using SPSS 14.0 for Windows.

Results

Patient Sample

During the 5-year study period, 222 patients were assessed. Of these, 100 patients were not included (Figure 1).
Figure 1 Flowchart of participation in the study
In total 124 patients received a randomization number after the screening visit, but 2 of them withdrew informed consent prior to baseline and were excluded. Thus, the ITT sample consists of 122 patients, including 2 patients who never took study medication. Forty-two patients were randomized to imipramine (mean dose visit 5-9 (i.e. treatment period with adequate dosing): 254.4 mg/day, SD 101.1, mean plasma level 294 ng/ml, SD 75.2); 39 to venlafaxine (mean maximum dose: 372.3 mg/day, SD 14.2); and 41 to venlafaxine-quetiapine (mean maximum dose venlafaxine 373.4 mg/day, SD 11.2 and quetiapine 598.9 mg/day, SD 15.0). Mean daily benzodiazepine use (in mg equivalents of lorazepam) in the imipramine group was 1.3 mg (SD 1.1), in the venlafaxine group 1.2 mg (SD 1.2), and in the venlafaxine-quetiapine group 1.2 mg (SD 1.2).

Patients’ baseline characteristics are summarized in Table 1.
Table 1 Characteristics of the patients at baseline

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Imipramine (n = 42)</th>
<th>Venlafaxine (n = 39)</th>
<th>Venlafaxine + Quetiapine (n = 41)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, No. (%)</td>
<td>23 (54.8)</td>
<td>17 (43.6)</td>
<td>22 (53.7)</td>
<td>.55</td>
</tr>
<tr>
<td>Age, mean (SD), years</td>
<td>51.6 (9.6)</td>
<td>49.5 (12.0)</td>
<td>50.6 (11.2)</td>
<td>.71</td>
</tr>
<tr>
<td>Baseline HAM-D score, mean (SD)</td>
<td>32.0 (5.3)</td>
<td>31.6 (4.6)</td>
<td>31.6 (5.4)</td>
<td>.92</td>
</tr>
<tr>
<td>Baseline CGI-severity score, mean (SD)</td>
<td>5.6 (0.6)</td>
<td>5.7 (0.6)</td>
<td>5.5 (1.0)</td>
<td>.30</td>
</tr>
<tr>
<td>Hallucinations, No. (%)</td>
<td>9 (21.4)</td>
<td>11 (28.2)</td>
<td>9 (22.0)</td>
<td>.73</td>
</tr>
<tr>
<td>- Only mood-congruent, No. (%)</td>
<td>5 (11.9)</td>
<td>8 (20.5)</td>
<td>7 (17.1)</td>
<td></td>
</tr>
<tr>
<td>- Only mood-incongruent, No. (%)</td>
<td>2 (4.8)</td>
<td>3 (7.7)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>- Both mood-congruent and mood-incongruent, No. (%)</td>
<td>2 (4.8)</td>
<td>0 (0.0)</td>
<td>2 (4.9)</td>
<td></td>
</tr>
<tr>
<td>Delusions, No. (%)</td>
<td>36 (85.7)</td>
<td>38 (97.4)</td>
<td>38 (92.7)</td>
<td>.15</td>
</tr>
<tr>
<td>- Only mood-congruent, No. (%)</td>
<td>30 (71.4)</td>
<td>33 (84.6)</td>
<td>28 (68.3)</td>
<td></td>
</tr>
<tr>
<td>- Only mood-incongruent, No. (%)</td>
<td>1 (2.4)</td>
<td>0 (0.0)</td>
<td>1 (2.4)</td>
<td></td>
</tr>
<tr>
<td>- Both mood-congruent and mood-incongruent, No. (%)</td>
<td>5 (11.0)</td>
<td>5 (12.8)</td>
<td>9 (22.0)</td>
<td></td>
</tr>
<tr>
<td>Mood-incongruent, hallucinations and mood-incongruent delusions, No. (%)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>.</td>
</tr>
<tr>
<td>Previous episodes, No. (SD)</td>
<td>1.0 (1.2)</td>
<td>0.6 (0.8)</td>
<td>1.1 (1.4)</td>
<td>.28</td>
</tr>
<tr>
<td>Duration of current episode, mean (SD), weeks</td>
<td>25.6 (32.1)</td>
<td>42.7 (110.9)</td>
<td>40.4 (98.8)</td>
<td>.63</td>
</tr>
<tr>
<td>ATHF depression score, mean (SD)</td>
<td>1.45 (1.38)</td>
<td>1.36 (1.66)</td>
<td>1.51 (1.50)</td>
<td>.90</td>
</tr>
<tr>
<td>ATHF psychotic depression score, mean (SD)</td>
<td>0.90 (0.73)</td>
<td>0.64 (0.67)</td>
<td>0.78 (0.65)</td>
<td>.23</td>
</tr>
</tbody>
</table>

SD: Standard deviation; HAM-D: Hamilton Rating Scale of depression 17-item; CGI: Clinical Global Impression ATHF: Antidepressant Treatment History Form, current episode, score 0-5.
Except for gender and duration of current episode, there were no notable imbalances or significant differences between groups. Overall, 37% of patients had prior exposure to a SSRI, 15% to a TCA and 8% to venlafaxine (at inadequate dose).

Twenty-two patients (18.0%) dropped out: 6 patients because of serious adverse events (SAE, see Safety section below), 16 for other reasons (Figure 1). One hundred patients (82.0%) completed the study: imipramine 35/42 (83.3%) venlafaxine 31/39 (79.5%) and venlafaxine-quetiapine 34/41 (82.9%). Dropouts did not differ from completers in baseline characteristics.

One patient was treated during one year with 150 mg venlafaxine until 4 days before starting with venlafaxine as study medication and one patient for 3 weeks with venlafaxine 150 mg until 6 days before starting with venlafaxine-quetiapine. There were no other cases of pretreatment with study medication during the current episode.

**Efficacy Results**

Table 2 presents the efficacy results.
Table 2 Data on the efficacy outcome measures

<table>
<thead>
<tr>
<th></th>
<th>Imipramine (n = 42)</th>
<th>Venlafaxine (n = 39)</th>
<th>Venlafaxine + Quetiapine (n = 41)</th>
<th>Statistics</th>
<th>Im vs V</th>
<th>VQ vs Im</th>
<th>VQ vs V</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response on HAM-D, No. (%)</strong></td>
<td>22/42 (52.4)</td>
<td>13/39 (33.3)</td>
<td>27/41 (65.9)</td>
<td>RD = 19.1 (95%CI: 2.1; 40.2)</td>
<td>RD = 13.5 (95%CI: -7.5; 34.4)</td>
<td>RD = 32.5 (95%CI: 11.8; 53.2)</td>
<td></td>
</tr>
<tr>
<td><strong>Response on CGI, No. (%)</strong></td>
<td>20/42 (47.6)</td>
<td>11/39 (28.2)</td>
<td>24/41 (58.5)</td>
<td>RD = 19.4 (95%CI: 1.2; 40.1)</td>
<td>RD = 10.9 (95%CI: -10.4; 32.3)</td>
<td>RD = 30.3 (95%CI: 9.7; 51.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Remission on HAM-D, No. (%)</strong></td>
<td>9/42 (21.4)</td>
<td>11/39 (28.2)</td>
<td>17/41 (41.5)</td>
<td>RD = 6.8 (95%CI: 25.6; 12.0)</td>
<td>RD = 20.0 (95%CI: 0.5; 39.6)</td>
<td>RD = 13.3 (95%CI: 7.4; 33.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Change in HAM-D score from baseline to endpoint, mean (SD)</strong></td>
<td>17.1 (9.7)</td>
<td>13.9 (10.4)</td>
<td>18.4 (11.4)</td>
<td>3.2 (95%CI: 1.3; 7.7)</td>
<td>1.3 (95%CI: -3.3; 5.9)</td>
<td>4.5 (95%CI: -0.4; 9.4)</td>
<td></td>
</tr>
<tr>
<td><strong>Change in CGI score from baseline to endpoint, mean (SD)</strong></td>
<td>1.9 (1.8)</td>
<td>1.7 (1.6)</td>
<td>2.4 (1.8)</td>
<td>0.2 (95%CI: 0.6; 1.0)</td>
<td>0.5 (95%CI: -0.3; 1.3)</td>
<td>0.7 (95%CI: -0.1; 1.5)</td>
<td></td>
</tr>
<tr>
<td><strong>Mean difference in weekly decrease in HAM-D score</strong></td>
<td>0.41 (95%CI: 0.20; 1.02)</td>
<td>0.25 (95%CI: -0.34; 0.85)</td>
<td>0.66 (95%CI: 0.04; 1.28)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean difference in weekly decrease in CGI score</strong></td>
<td>0.09 (95%CI: 0.04; 0.15)</td>
<td>0.05 (95%CI: 0.05; 0.15)</td>
<td>0.11 (95%CI: 0.01; 0.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Absence of psychotic features at endpoint, No (%)</strong></td>
<td>26/42 (61.9)</td>
<td>23/39 (59.0)</td>
<td>29/41 (70.7)</td>
<td>RD = 2.9 (95%CI: 18.4; 24.2)</td>
<td>RD = 8.8 (95%CI: -11; 29.1)</td>
<td>RD = 11.7 (95%CI: -9.0; 32.5)</td>
<td></td>
</tr>
<tr>
<td><strong>Absence of psychotic features among HAM-D responders, No (%)</strong></td>
<td>21/22 (95.5)</td>
<td>12/13 (92.3)</td>
<td>26/27 (96.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Im = imipramine; V = venlafaxine; VQ = venlafaxine + quetiapine; HAM-D: Hamilton Rating Scale of depression 17-item; CGI: Clinical Global Impression; RD: Risk Difference; CI: Confidence Interval; SD: Standard deviation
* With the last observation carried forward. ** From linear mixed model. Bold: significant difference
On the primary endpoint (response on HAM-D at endpoint), venlafaxine-quetiapine was significantly more effective than venlafaxine, while there were no significant differences between venlafaxine-quetiapine and imipramine, or between imipramine and venlafaxine. The unadjusted ORs (95%CI) for imipramine versus venlafaxine, venlafaxine-quetiapine versus venlafaxine and venlafaxine-quetiapine versus imipramine were 2.20 (0.89; 5.41), 3.86 (1.53; 9.75), and 1.75 (0.72; 4.25), respectively. When we adjusted for group imbalances in gender and duration of the current episode, the ORs were similar, i.e. 2.43 (0.94; 6.28), 4.02 (1.56; 10.32) and 1.76 (0.72; 4.30), respectively. Similar, but non-significant differences were obtained for the secondary outcome measures. In the group of patients fulfilling responder criteria (LOCF), there were no patients with hallucinations. In each group there was one patient fulfilling responder criteria (LOCF) still with a delusion, which disappeared during the subsequent month while continuing the same medication.

Regarding remission, venlafaxine-quetiapine was significantly more effective than imipramine; however, in contrast to response, venlafaxine-quetiapine was not significantly more effective than venlafaxine alone, while there was also no significant difference between imipramine and venlafaxine.

In the linear mixed models analysis, a non-linear decrease over time was seen in each group for mean scores on HAM-D and, to a lesser extent, on CGI (Figure 2).
We adjusted for this leveling effect by retaining the significant ($p<.001$) time squared variable in all final models. From week 5 onwards, differences between treatment groups became more marked. The treatment by time interaction terms indicating treatment effects in the linear mixed models showed an advantage for venlafaxine-quetiapine. Patients receiving venlafaxine-quetiapine demonstrated a 0.66 point per week faster mean decrease in HAM-D scores than those receiving venlafaxine. Although not significant, the mean rate of decrease in HAM-D scores was also greater with imipramine compared to venlafaxine. The smallest and non-significant difference in HAM-D decline was between venlafaxine-quetiapine and imipramine. Essentially the same pattern was observed for the CGI: a substantial and significant difference in decrease per week was found for venlafaxine-quetiapine versus venlafaxine.

The Kaplan-Meier curves by treatment group are shown in Figure 3.

![Kaplan-Meier curves](image.png)

**Figure 3** Kaplan-Meier curves of the time-related cumulative proportion without response between treatment groups
Overall, the median time to response was 5 (95%CI: 4; 6) weeks. Differences between these curves were not significant (p = 0.23), possibly, because the differences between the curves only appeared after treatment week 5.

Dropouts did not influence effect estimates. When restricting to those who completed all visits (n=100) the results were similar: venlafaxine-quetiapine was significantly more effective than venlafaxine, while there were no significant differences between venlafaxine-quetiapine and imipramine, or between imipramine and venlafaxine.

Analysis of treatment guesses at treatment week 5 showed that agreement between guessed and actual medication was only slightly higher than expected on the basis of chance, with a very small Kappa statistic of 0.14.

**Safety Results**

There were 8 SAEs, and 6 of them dropped out from the study. Three dropouts switched to (hypo)mania: 1 mania and 1 hypomania in the imipramine group, and 1 hypomania in the venlafaxine-quetiapine group. The two hypomaniac patients had reached study response criterion at time of dropout. The three other patients dropped out because of extra-pyramidal symptoms (imipramine), liver dysfunction (venlafaxine) and urine retention (venlafaxine-quetiapine). All these six patients recovered after drug discontinuation.

Two other SAEs, both in the venlafaxine-quetiapine group, were suicide attempts: one strangulation attempt at treatment day 13, and one autointoxication with 200 mg oxazepam at treatment day 16. Both attempts had no medical consequences and both patients continued the study.

Other non-serious adverse events occurring in more than 10% of patients in either group are presented in Table 3.
Table 3 Adverse events* occurring in more than 10% of patients in any one group

<table>
<thead>
<tr>
<th>%</th>
<th>Imipramine</th>
<th>Venlafaxine</th>
<th>Venlafaxine + Quetiapine</th>
<th>P-value for difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Imi vs V</td>
</tr>
<tr>
<td>Blurred vision</td>
<td>16.7</td>
<td>10.3</td>
<td>2.4</td>
<td>.52</td>
</tr>
<tr>
<td>Cold hands</td>
<td>14.3</td>
<td>7.7</td>
<td>4.9</td>
<td>.49</td>
</tr>
<tr>
<td>Constipation</td>
<td>21.4</td>
<td>30.8</td>
<td>29.3</td>
<td>.45</td>
</tr>
<tr>
<td>Dizziness</td>
<td>38.1</td>
<td>23.1</td>
<td>41.5</td>
<td>.16</td>
</tr>
<tr>
<td>Dry mouth</td>
<td>57.1</td>
<td>30.8</td>
<td>41.5</td>
<td>.03</td>
</tr>
<tr>
<td>Headache</td>
<td>16.7</td>
<td>2.6</td>
<td>4.9</td>
<td>.32</td>
</tr>
<tr>
<td>Insomnia</td>
<td>16.7</td>
<td>2.6</td>
<td>4.9</td>
<td>.06</td>
</tr>
<tr>
<td>Memory</td>
<td>9.5</td>
<td>7.7</td>
<td>14.6</td>
<td>1.00</td>
</tr>
<tr>
<td>Nervousness</td>
<td>19.0</td>
<td>18.0</td>
<td>19.5</td>
<td>1.00</td>
</tr>
<tr>
<td>Other a.e.</td>
<td>16.7</td>
<td>23.1</td>
<td>17.1</td>
<td>.58</td>
</tr>
<tr>
<td>Restlessness</td>
<td>16.7</td>
<td>25.6</td>
<td>12.2</td>
<td>.42</td>
</tr>
<tr>
<td>Somnolence</td>
<td>11.9</td>
<td>2.6</td>
<td>26.8</td>
<td>.20</td>
</tr>
<tr>
<td>Strange/bad taste</td>
<td>7.1</td>
<td>0.0</td>
<td>12.2</td>
<td>.24</td>
</tr>
<tr>
<td>Tiredness</td>
<td>19.1</td>
<td>7.7</td>
<td>34.2</td>
<td>.20</td>
</tr>
<tr>
<td>Transpiration</td>
<td>16.7</td>
<td>12.8</td>
<td>12.2</td>
<td>.76</td>
</tr>
<tr>
<td>Tremor</td>
<td>14.3</td>
<td>10.3</td>
<td>7.3</td>
<td>.74</td>
</tr>
</tbody>
</table>

* Adverse event: an increase of at least 1 step on a 4-point scale (none - light - moderate - severe) at any visit after baseline compared to baseline and at least moderate
Im = imipramine; V = venlafaxine; VQ = venlafaxine + quetiapine
Bold: significant difference
Because of the increasing importance of metabolic syndrome with atypical antipsychotics (27), we also analyzed weight changes from baseline to endpoint (LOCF). The weight of the patients showed a non-significant increase with imipramine and venlafaxine (mean 0.4 kg; SD 3.0 and mean 0.5 kg; SD 4.0, respectively). With venlafaxine-quetiapine, however, the increase was substantial (mean 3.8 kg; SD 4.0) and significantly different from imipramine and venlafaxine (p <0.01). More than 7% weight increase (27) occurred in 3% of patients with imipramine, in 10% with venlafaxine and in 35% with venlafaxine-quetiapine. (p <0.001 and p = 0.009, respectively).

Discussion

This study demonstrated for the first time that a combination of a newer generation antidepressant (venlafaxine) plus an atypical antipsychotic (quetiapine) is more effective than monotherapy with a newer generation antidepressant (venlafaxine). On the primary (response on the HAM-D) and on several secondary outcomes, the difference reached significance and the 32.6% difference in response rates between venlafaxine-quetiapine and venlafaxine is clinically relevant. Spiker et al28 reported a similar outcome, but in their study the difference in remission rates between amitriptyline plus perphenazine compared with amitriptyline monotherapy was only significant among completers and not in the ITT analysis, which was not reported as such.

For the (post hoc) outcome remission, we found a significant difference between venlafaxine-quetiapine and imipramine, but not between venlafaxine-quetiapine and venlafaxine. We have no explanation for the large differences in remission rates among the responders (imipramine 45%, venlafaxine 100%, venlafaxine plus quetiapine 71%); it may be a chance finding. The remission results underscore the response results, suggesting that venlafaxine-quetiapine is not only better than venlafaxine but also better than imipramine, and that this difference with imipramine at response did not reach significance because of limited power. Imipramine at adequate plasma levels seems to be a relatively effective treatment for psychotic depression. The response rate of 52.4% in our study was almost the same as in previous studies by Bruijn et al., Van den Broek et al., and Birkenhager et al. (post hoc analysis of two studies: 22/40, 55.0%).29,30,31,32 It would be interesting to know how effective a combination of quetiapine with imipramine would have been.

Another question is what would have been the effect of monotherapy with quetiapine, or the combination with another antipsychotic. When we wrote the protocol, treatment with an antipsychotic alone seemed not appropriate, which was confirmed in our later meta-analysis.7 Now there are indications that, in addition
to its antipsychotic and antimanic effect, quetiapine also has an antidepressive effect. In recent studies, quetiapine monotherapy was found efficacious in patients with bipolar depression,\textsuperscript{33–34} as well as in unipolar MDD.\textsuperscript{35} It remains unclear to what extent patients with psychotic features were included in these studies. A possible reason for its antidepressive effect is that one of the metabolites of quetiapine (norquetiapine) is a noradrenalin reuptake inhibitor.\textsuperscript{36} Further studies with quetiapine as monotherapy for psychotic depression are warranted. With regard to other atypical antipsychotics, it is unclear whether they would have been as effective as quetiapine in this combination with venlafaxine. Other atypical antipsychotics have rarely been studied in psychotic depression. In one out of two RCTs olanzapine plus fluoxetine was found to be significantly more effective than olanzapine or placebo, while olanzapine alone was not more effective than placebo.\textsuperscript{37}

We used the older immediate release form of quetiapine. With the newer extended release form, it might have been possible to build up the dose to 600 mg quetiapine within one week. The immediate release form maybe was a slight disadvantage for venlafaxine plus quetiapine compared to the newer extended release form of quetiapine because with the extended release form treatment with the maximum dose of quetiapine would have been a week longer.

Our primary outcome measure (response on the HAM-D) did not include effect on psychosis. We assumed a priori that recovery of depression would also include disappearance of psychotic symptoms. We also did not apply a psychosis rating scale such as the PANSS or the BPRS: the PANSS is developed for the assessment of severity of primary psychotic disorders such as schizophrenia, and not of psychotic depression, and the BPRS does not specifically measure psychotic features and especially does not discriminate between mood-congruent and mood-incongruent psychotic features. Therefore we decided to assess psychotic features by asking the clinician to document the presence (or absence) of hallucinations or delusions including whether they were mood-congruent or mood-incongruent at baseline and then weekly.

Among the responders at the end of the trial three patients (one in each group) still had a mood-congruent delusion that disappeared in all these patients during follow-up.

In our study there were few dropouts and we regard this as one of its major strengths, particularly for the significance level and the internal validity of the findings. None of the SAEs had long-term consequences and, overall, tolerance of medication was high. We found a significantly higher rate of somnolence and tiredness in the venlafaxine-quetiapine group compared to venlafaxine. Weight increase was significantly higher with venlafaxine-quetiapine compared to both venlafaxine and to imipramine.

Differences between groups appeared after 5 weeks of treatment and reco-
very continued thereafter until the end of treatment. Therefore, a clinical implication could be that patients with a psychotic depression should be treated for at least 7 weeks (perhaps even longer) with venlafaxine-quetiapine before deciding whether to switch to another treatment. However, because this might be a chance finding it deserves follow-up in future studies.

The question remains as to what extent our findings are generalizable. In clinical practice, many patients with psychotic depression are treated as inpatients, as was the case in the present study. In most of our patients, psychotic symptoms were mood congruent; no patient had only mood incongruent hallucinations and/or mood incongruent delusions. Age is another consideration. We do not know if our results are applicable to elderly patients. The only (underpowered) RCT with patients aged over 50 years (mean age 73 years)38 showed no difference between adding perphenazine to nortriptyline (7/19) and adding placebo to nortriptyline (7/17). Whether our findings are applicable to other antidepressants or other antipsychotics is of course also questionable, and further RCTs are needed to address this question.

A methodological limitation of our study is the lack of a placebo group. A randomized study with a placebo arm would have been a better study design to answer the question whether the three treatments are efficacious. However, our question was whether the combination of an antidepressant plus an antipsychotic was more effective than an antidepressant alone, for which a placebo arm is not necessary. Moreover, we considered placebo treatment in these severely ill patients undesirable and we anticipated problems to get approval from our ethical review and with recruitment. Especially in the more severe patients with high HAM-D scores response to placebo is low.11 Several previous studies in psychotic depression found response rates of even less than 10 %.12-13 Two exceptions with a much higher placebo response are two recent studies.37,39 We think that differences in treatment setting (shorter duration of hospitalization), selection of less severely ill patients and/or study design between these two studies and our and previous studies may explain the differences in placebo response.

A second methodological limitation of our study may be that we broke the code for each patient two weeks after their participation in the study. Most importantly, this means that knowledge of the randomized condition of this patient could impossibly have affected the rating of psychopathology directly. However, it remains possible that an investigator, after deblinding, retrospectively linked certain side effects to the use of certain medications and that this knowledge, in subsequent patients, influenced the rating of psychopathology. This may have caused bias only if the link between side effects and medications was largely consistent, their recognition by the investigator was sufficiently precise, and resulted in systematic over- or underrating of psychopathology. It is our opinion that this mecha-
nism seems rather unlikely. Nevertheless, we checked for blindness during the study by asking the investigator to guess to which treatment arm the patient had been randomized. Agreement between guessed and actual medication was very low (Kappa 0.14), indicating that preservation of blindness of our study was high.

A third limitation can be some asymmetries in the design of the study. There is a slight difference in dosage of medication in the three groups. Based on plasma level at day 7 imipramine dosage was adjusted at day 10. Venlafaxine dosage was 300 mg at day 10 and 375 mg at day 17. Quetiapine was 600 mg at day 10. So, imipramine and quetiapine were at optimum and highest dose 7 days earlier than venlafaxine, leaving for venlafaxine 5 weeks and for imipramine and quetiapine 6 weeks treatment with the optimum or highest dose. This could have been a slight disadvantage for venlafaxine compared to imipramine. Another slight disadvantage for venlafaxine could have been that the imipramine dose was adjusted to plasma levels and venlafaxine was titrated upwards to maximum dose. However, in clinical reality there is no solution to this asymmetry because dose-level relationships have not been established for venlafaxine. Fifteen percent of the included patients had prior exposure to a TCA, 37% to a SSRI and 8% to venlafaxine. This could have been a small disadvantage for imipramine, but because venlafaxine 300 to 375 mg also is a serotonin and noradrenalin reuptake inhibitor this is unlikely.

A fourth limitation is that remission was a posthoc secondary outcome measure. We used remission as secondary outcome because it is considered more and more as an important outcome and because there is internationally agreement about the definition of remission, preventing us from a posthoc data driven interpretation of outcome measures.

An important limitation is that we included only 122 patients instead of the planned 155. Nevertheless, our study is among the larger studies on psychotic depression ever performed and the largest with this research question: is the combination of an antipsychotic and an antidepressant better than an antidepressant alone. Two other studies comparing the combination of an antipsychotic plus an antidepressant with antipsychotic monotherapy (instead of an antidepressant) in three treatment groups (including placebo) had a comparable number of patients (124 and 125, and respectively) but a much higher dropout rate (73/124; 59.9% and 66/125; 52.8%, respectively). Soon the largest study until now will be published with 259 subjects divided over two treatment arms, also comparing the combination of an antipsychotic plus an antidepressant with antipsychotic monotherapy (instead of an antidepressant). This study was like ours without a placebo group but also with a much higher dropout rate of 45%. In retrospect, our design was ambitious and it would have been better to have two rather than three treatment arms. The imipramine arm was not necessary to answer our main research question, namely, whether combination therapy with an antidepressant plus an antipsychotic is more
effective than monotherapy with an antidepressant. Post-hoc, the 122 patients we recruited would have provided enough power for such a two-armed study.

In conclusion: the clinical belief, expressed in many guidelines, that psychotic depression should be treated with a combination of an antidepressant and an antipsychotic and not with an antidepressant alone, can be considered evidence based with regard to venlafaxine-quetiapine combination therapy versus venlafaxine monotherapy. Whether this is also the case for imipramine monotherapy is likely, but cannot be concluded from the present data.

Acknowledgements

The authors thank the following for their support:

JP Selten, R van Ojen, I Klijn: independent psychiatrist for medication

G Hugenholtz, H Wassens, C Stuivenberg, T Gerbranda: central distribution of medication

F Ververs: laboratory (plasma levels imipramine)

EA Wijnia, K Nijssen, JW van de Maaskant, Julius Center

J Verploegh, J Leijtens: research nurses

F Kruisdijk, psychiatrist, Symfora, Amersfoort

MF de Vries, LJ Vos, I Sommer, C Schubart, UMC Utrecht

NPJT van Schayk, I Vosjan, I van Geel, I Kop, AstraZeneca

H Kornaat, M Kanters-van Buren, E Ides, M Griekspoor, Wyeth

AstraZeneca and Wyeth for financial support and providing the study medication
References


11. Khan A, Leventhal RM, Khan SR et al. Severity of depression and response to


