Antidepressant efficacy of high-frequency transcranial magnetic stimulation over the left dorsolateral prefrontal cortex in double-blind sham-controlled designs: a meta-analysis

D. J. L. G. Schutter*

* Address for correspondence: D. J. L. G. Schutter, Ph.D., Experimental Psychology, Utrecht University, Heidelberglaan 2, 3584CS Utrecht, The Netherlands.
(Email: d.schutter@uu.nl)

Background. For more than a decade high-frequency repetitive transcranial magnetic stimulation (rTMS) has been applied to the left dorsolateral prefrontal cortex (DLPFC) in search of an alternative treatment for depression. The aim of this study was to provide an update on its clinical efficacy by performing a meta-analysis involving double-blind sham-controlled studies.

Method. A literature search was conducted in the databases PubMed and Web of Science in the period between January 1980 and November 2007 with the search terms 'depression' and 'transcranial magnetic stimulation'. Thirty double-blind sham-controlled parallel studies with 1164 patients comparing the percentage change in depression scores from baseline to endpoint of active versus sham treatment were included. A random effects meta-analysis was performed to investigate the clinical efficacy of fast-frequency rTMS over the left DLPFC in depression.

Results. The test for heterogeneity was not significant ($Q_T = 30.46, p = 0.39$). A significant overall weighted mean effect size, $d = 0.39$ [95% confidence interval (CI) 0.25–0.54], for active treatment was observed ($z = 6.52, p < 0.0001$). Medication resistance and intensity of rTMS did not play a role in the effect size.

Conclusions. These findings show that high-frequency rTMS over the left DLPFC is superior to sham in the treatment of depression. The effect size is robust and comparable to at least a subset of commercially available antidepressant drug agents. Current limitations and future prospects are discussed.

Received 16 November 2007; Revised 1 February 2008; Accepted 27 March 2008

Key words: Depression, dorsolateral prefrontal cortex, meta-analysis, transcranial magnetic stimulation, treatment.

Introduction

The World Health Organization has estimated that 121 million people worldwide currently suffer from depression and are in need of treatment. Conventional treatments of depression range from pharmacological agents and cognitive behavioural therapy to electroconvulsive shock therapy (ECT), but in the past decade transcranial magnetic stimulation (TMS) has been explored as an alternative application. TMS targets depression by modifying neuronal activity function with magnetically induced electrical currents in the brain. The technique, originally introduced in 1985, is non-invasive and safe, and can easily be applied to the scalp in a relatively painless manner. The main principle of TMS is based on Faraday’s law of electromagnetic induction, which states that a magnetic pulse situated near conductors will be transformed into an electric current. This electrical current will subsequently depolarize underlying cortical nerve cells tangentially oriented to the magnetic field (Bohning, 2000).

The theoretical background for applying fast-frequency repetitive TMS (rTMS) to the left prefrontal cortex in the treatment of depression may find its origin in earlier observations that depression following stroke was often associated with left prefrontal cortex damage, but not with damage to the right prefrontal cortex (Robinson & Szetela, 1981). Additional support for the involvement of the left prefrontal cortex in depression was provided by functional neuroimaging demonstrating left anterior hypoactivity in depressive patients (Baxter et al. 1989). This link may have been one of the reasons why researchers started to apply trains of fast-frequency rTMS over the left anterior part of the hemisphere in an attempt to locally enhance neural activity and alleviate depressive symptoms. Current views hold that restoring the balance
between left and right prefrontal cortex activity is in fact more important than establishing absolute increases in left-sided activity per se. In support, there is some evidence suggesting that inhibitory slow-frequency rTMS over the right dorsolateral prefrontal cortex (DLPFC) also has antidepressant properties (Klein et al. 1999). Nonetheless, ever since the first publication of an open-label study in 1993 that showed mood improvements in two depressed patients following fast-frequency TMS over the left DLPFC (Hoflich et al. 1993), the vast majority of researchers have adopted this strategy and explored the effects of fast-frequency rTMS over the left DLPFC in major depression (George et al. 1999).

Fast-frequency stimulation over the left DLPFC has several advantages over other biologically oriented treatments. TMS is associated with only mild physical discomfort, has no cognitive side-effects and may have neuroprotective properties (Post et al. 1999). The most commonly reported complaint is a headache, which usually responds promptly to a common analgesic. The main concern with rTMS is its potential to induce a seizure. Safety guidelines, including limits of stimulation intensity, monitoring of subjects, medical management of induced seizures and contra-indications to rTMS as described by the International Federation of Clinical Neurophysiology have helped to minimize seizure risk (Wassermann, 1998). Between 2001 and 2003 several quantitative reviews were published on the antidepressant properties of TMS (McNamara et al. 2001; Burt et al. 2002; Holtzheimer et al. 2002; Martin et al. 2002, 2003; Couturier, 2005). These studies report effect sizes ranging from no improvement whatsoever to clear beneficial effects following active treatment. The meta-analyses are, however, hampered by methodological issues, including small number of studies, using an endpoint instead of baseline-corrected depression scores, and heterogeneity of effect sizes. Overall, the meta-analyses nonetheless do suggest that depressive patients benefit more from active than from sham or no TMS treatment, but that clinical efficacy still needs to be proven. In recent years this has resulted in methodologically improved TMS studies (Fitzgerald et al. 2003), but also in a growing number of researchers and practitioners who are unsure whether TMS holds the promise of becoming a clinical treatment in biological psychiatry. In fact, rTMS is currently being reviewed by both the Food and Drug Administration (FDA) in the USA and the National Institute for Health and Clinical Excellence (NICE) in the UK for approval. The aim of this study was to provide an update on the status of fast-frequency rTMS over the left DLPFC in depression. To this end a meta-analysis was performed that included all available published clinical trials that have studied the antidepressant effects that applied at least five treatment sessions of high-frequency rTMS over the left DLPFC in double-blind sham-controlled designs exclusively.

Method

Study selection

Articles for inclusion were identified starting with conducting a literature search in the databases PubMed and Web of Science in the period between January 1980 and November 2007. The search criteria were 'depression' and 'transcranial magnetic stimulation' and yielded 577 hits in PubMed and 976 hits in Web of Science. Titles and abstract of the studies were screened for consideration. In addition, the reference lists of previous meta-analyses (McNamara et al. 2001; Burt et al. 2002; Holtzheimer et al. 2002; Martin et al. 2002, 2003; Couturier, 2005) and reviews (George et al. 1997, 2003; Gershon et al. 2003; Padberg & Möller, 2003; Loo & Mitchell, 2005; Herrmann & Ebmeier, 2006) were screened to minimize the risk of overlooking potentially suitable studies for inclusion. Candidate studies had to satisfy the following quality criteria based on the Cochrane Reviewers’ Handbook 4.1.4 and the Users’ Guide to the Medical Literature (Couturier, 2005):

1. Study validity: random allocation; patients and clinical raters were blind to treatment (double-blind); sham-controlled, parallel design, intent-to-treat analysis;
2. Adults with major depressive episode without psychotic features according to DSM-IV criteria;
3. High frequency (>5 Hz) rTMS over the left DLPFC, intensity >80% motor threshold (MT), at least five treatment sessions, sham condition; 45° and 90° from scalp or sham coil;
4. Primary outcome measure: baseline-corrected percentage change in scores on the Hamilton Depression Rating Scale (HAMD) or the Montgomery–Asberg Depression Rating Scale (MADRS).

Additional quality criteria were:

5. Participant’s treatment complete within 6 weeks after first session;
6. The article was published in a peer-reviewed English-language journal;
7. Study approved by a medical ethical committees or review board.

Thirty of the initially selected studies fulfilled the criteria for inclusion in the meta-analysis. Characteristics of the studies are listed in Table 1.
Data synthesis and analysis

Effect sizes were calculated for the difference in the absolute and percentage change in HAMD scores from baseline to outcome after the final session between ‘active’ and ‘sham’ rTMS. The effect size estimate used was Hedges’ $g$, which is an standardized mean difference that accounts for the fact that the sampling variance for ‘active’ and ‘sham’ groups are not always equal (Hedges & Olkin, 1985). When the absolute or percentage change was not reported or could not be calculated from the data, the corresponding author was contacted and asked to provide the necessary details for estimating the effect size. In one case, the reported $t$ values from paired sample comparisons and the net change in HAMD scores between baseline and final outcome in the ‘active’ and ‘sham’ conditions were used to estimate the pooled standard deviation for computing Hedges’ $g$. From these effect sizes the Hedges’ $d$ values were calculated to correct for a bias in effect size due to small group samples (Hedges & Olkin, 1985). Because of the small sample sizes in some of the treatment studies, non-parametric variances were chosen for the meta-analysis.

A common difficulty in TMS treatment trials is that the studies often do not have equal samples sizes and some sort of weighing is required. In addition to Hedges’ $d$, a weighted average was used to compute the cumulative effect size ($E$) for the present studies (Hedges & Olkin, 1985). The cumulative effect size represents the overall magnitude of the effect size and Stouffer’s $z$ statistic was used to test whether or not the cumulative effect size was different from chance. Additionally, the cumulative effect size was used in a random effects model to determine the total heterogeneity of the effect sizes, $Q_T$, and tested against the $\chi^2$ distribution with $29$ ($n-1$) degrees of freedom (Hedges, 1981). A significant $Q_T$ means that the variance of the effect sizes is greater than to be expected from sampling errors. This suggests that the observed variance can be explained by other variables besides treatment and should be further investigated.

A matter of concern in the interpretation of meta-analytical results is the possibility of an upward bias of the effect size due to the omission of unpublished studies with null effects. The failure of non-significant studies being published in the literature creating a publication bias is termed the ‘file drawer problem’ (Rosenthal, 1979). In addition to inspecting the funnel plot, one of the easiest methods to explore the robustness of the results to the possibility of publication bias is computing the fail-safe number. The fail-safe number of studies ($N_{fs}$) provides an estimation of how many non-significant or missing studies would be needed to render the observed meta-analytical results non-significant (Rosenthal’s method: $\alpha<0.05$) for active rTMS treatment.

All analyses were performed with MetaWin version 2 (Rosenberg et al. 2000).

Results

A total 1164 patients with major depression (mean ± S.D.: age $49.1 \pm 7.5$ years) were enrolled in the meta-analysis, of which 606 patients (age $49.5 \pm 7.8$ years) received real rTMS treatment and 558 patients (age $48.9 \pm 7.4$ years) received sham rTMS treatment. The majority of participants in the real ($n=451$) and sham rTMS treatment condition ($n=399$) were resistant to medication. Treatments were generally well tolerated and no deaths were reported. Moreover, no seizures were observed in the real rTMS treatment and only one patient reported having a seizure following a session of sham treatment (Mogg et al. 2008). The most commonly observed side-effects associated with rTMS were headaches, dizziness, nausea and painful local sensation. These side-effects are typically considered to be mild and respond promptly to analgesics. Considering the very low incidence of serious adverse events, rTMS when applied within the range of the International Federation of Clinical Neurophysiology (IFCN) safety guidelines can be considered a safe method.

The overall weighted mean effect size for treatment was 0.39 [95% confidence interval (CI) 0.25–0.54, $z=6.52$, $p<0.0001$]. An analysis of variance (ANOVA) did not provide evidence for a difference in effect size between medication-resistant ($n=17$) and non-medication resistant depression ($n=8$) [$F(1, 24)=0.03$, $p=0.87$]. An additional ANOVA comparing the difference of effect sizes between studies that applied <100% MT intensities ($n=14$) and studies that used 100–120% MT intensities ($n=16$) was not significant [$F(1, 29)=0.22$, $p=0.65$]. These results argue against the notion that medication resistance or intensity of rTMS play a major role in the antidepressant effect of rTMS. The mean effect size and 95% CI of the studies are plotted in Fig. 1.

The test for heterogeneity was not significant ($Q_T=30.46$, $p=0.39$), implying that the variance among the effect sizes was not greater than expected by sampling error. Moreover, visual inspection of the funnel plot, as depicted graphically in Fig. 2, showed that, given the typical symmetrical funnel, there is no reason to assume a bias in publishing positive results. The fail-safe number of studies was 269.6, indicating that at least 269 unpublished null-findings were needed to render the effect of active treatment statistically non-significant. It is unlikely that such a large
### Table 1. Study characteristics

<table>
<thead>
<tr>
<th>Study</th>
<th>Scale</th>
<th>TMS</th>
<th>n</th>
<th>Mean age ± S.D.</th>
<th>Parameters</th>
<th>Total pulses per session</th>
<th>No. of sessions</th>
<th>Medication resistant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. George et al. 1997</td>
<td>HAMD</td>
<td>Active</td>
<td>7</td>
<td>42.4 ± 15.5</td>
<td>10 Hz, 80% MT, 20 trains, 2 s on, 58 s off</td>
<td>800</td>
<td>10</td>
<td>No</td>
</tr>
<tr>
<td>2. Haag et al. 1997</td>
<td>HAMD</td>
<td>Active</td>
<td>6</td>
<td>51.2 ± 16.1</td>
<td>10 Hz, 90% MT, 5 trains, 5 s on, 55 s off</td>
<td>1250</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>3. Avery et al. 1999</td>
<td>HAMD</td>
<td>Active</td>
<td>4</td>
<td>44.3 ± 10.1</td>
<td>10 Hz, 80% MT, 20 trains, 5 s on, 55 s off</td>
<td>1000</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>4. Kimbrell et al. 1999</td>
<td>HAMD</td>
<td>Active</td>
<td>3</td>
<td>40.2 ± 15.1</td>
<td>20 Hz, 80% MT, 20 trains, 2 s on, 60 s off</td>
<td>800</td>
<td>10</td>
<td>N.A.</td>
</tr>
<tr>
<td>5. Loo et al. 1999</td>
<td>HAMD</td>
<td>Active</td>
<td>9</td>
<td>45.7 ± 14.7</td>
<td>10 Hz, 110% MT, 30 trains, 5 s on, 30 s off</td>
<td>1500</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>6. Padberg et al. 1999</td>
<td>HAMD</td>
<td>Active</td>
<td>6</td>
<td>63.5 ± 15.8</td>
<td>10 Hz, 90% MT, 5 trains, 5 s on, 30 s off</td>
<td>250</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>7. Berman et al. 2000</td>
<td>HAMD</td>
<td>Active</td>
<td>10</td>
<td>45.2 ± 9.5</td>
<td>20 Hz, 80% MT, 20 trains, 2 s on, 58 s off</td>
<td>800</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>8. Eschweiler et al. 2000</td>
<td>HAMD</td>
<td>Active</td>
<td>5</td>
<td>59 ± 5.1</td>
<td>10 Hz, 90% MT, 5 trains, 5 s on, 50 s off</td>
<td>250</td>
<td>10</td>
<td>N.A.</td>
</tr>
<tr>
<td>9. George et al. 2000</td>
<td>HAMD</td>
<td>Active</td>
<td>6</td>
<td>43.3 ± 11.6</td>
<td>10 Hz, 110% MT, 40 trains, 2 s on, 28 s off</td>
<td>1600</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>10. Garcia-Toro et al. 2001</td>
<td>HAMD</td>
<td>Active</td>
<td>10</td>
<td>60.7 ± 9.8</td>
<td>20 Hz, 80% MT, 20 trains, 2 s on, 58 s off</td>
<td>800</td>
<td>5</td>
<td>N.A.</td>
</tr>
<tr>
<td>11. Manes et al. 2001</td>
<td>HAMD</td>
<td>Active</td>
<td>10</td>
<td>42.2 ± 7.3</td>
<td>5 Hz, 110% MT, 40 trains, 8 s on, 22 s off</td>
<td>1600</td>
<td>10</td>
<td>N.A.</td>
</tr>
<tr>
<td>12. Nahas et al. 2003</td>
<td>HAMD</td>
<td>Active</td>
<td>10</td>
<td>43.3 ± 9.3</td>
<td>10 Hz, 100% MT, 20 trains, 5 s on, 25 s off</td>
<td>1000</td>
<td>10</td>
<td>N.A.</td>
</tr>
<tr>
<td>13. Szuba et al. 2001</td>
<td>HAMD</td>
<td>Active</td>
<td>10</td>
<td>39.7 ± 12.1</td>
<td>10 Hz, 100% MT, 20 trains, 5 s on, 25 s off</td>
<td>1000</td>
<td>10</td>
<td>N.A.</td>
</tr>
<tr>
<td>14. Boutros et al. 2002</td>
<td>HAMD</td>
<td>Active</td>
<td>10</td>
<td>33.4 ± 9.3</td>
<td>20 Hz, 80% MT, 20 trains, 2 s on, 58 s off</td>
<td>800</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>15. Padberg et al. 2002</td>
<td>HAMD</td>
<td>Active</td>
<td>20</td>
<td>61.2 ± 4.4</td>
<td>10 Hz, 90% MT (n=10), 100% MT (n=10), 15 trains, 10 s on, 30 s off</td>
<td>1500</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>16. Fitzgerald et al. 2003</td>
<td>MADRS</td>
<td>Active</td>
<td>20</td>
<td>42.2 ± 9.8</td>
<td>10 Hz, 100% MT, 20 trains, 5 s on, 25 s off</td>
<td>1000</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>17. Höppner et al. 2003</td>
<td>HAMD</td>
<td>Active</td>
<td>11</td>
<td>60.4 ± 6</td>
<td>20 Hz, 90% MT, 20 trains, 2 s on, 60 s off</td>
<td>800</td>
<td>10</td>
<td>No</td>
</tr>
<tr>
<td>18. Holtzheimer et al. 2004</td>
<td>HAMD</td>
<td>Active</td>
<td>10</td>
<td>40.4 ± 8.5</td>
<td>10 Hz, 110% MT, 32 trains, 5 s on, 30–60 s off</td>
<td>1600</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>19. Jorge et al. 2004</td>
<td>HAMD</td>
<td>Active</td>
<td>10</td>
<td>63.1 ± 8.1</td>
<td>10 Hz, 110% MT, 20 trains, 5 s on, 60 s off</td>
<td>1000</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>20. Koerselman et al. 2004</td>
<td>HAMD</td>
<td>Active</td>
<td>27</td>
<td>51 ± 15.4</td>
<td>20 Hz, 80% MT, 20 trains, 2 s on, 28 s off</td>
<td>800</td>
<td>10</td>
<td>No</td>
</tr>
<tr>
<td>21. Mosimann et al. 2004</td>
<td>HAMD</td>
<td>Active</td>
<td>15</td>
<td>60 ± 13.4</td>
<td>20 Hz, 100% MT, 40 trains, 2 s on, 28 s off</td>
<td>1600</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>22. Poulet et al. 2004</td>
<td>MADRS</td>
<td>Active</td>
<td>10</td>
<td>64.4 ± 13</td>
<td>10 Hz, 80% MT, 20 trains, 2 s on, 58 s off</td>
<td>400</td>
<td>10</td>
<td>No</td>
</tr>
</tbody>
</table>

D. J. L. G. Schutter
The aim of this meta-analysis was to investigate whether high-frequency rTMS applied over the left dorsolateral prefrontal cortex (DLPFC) is effective in treating depression. Table 2 presents the main outcomes of the meta-analysis.

Table 2: Main outcomes of the meta-analysis.

<table>
<thead>
<tr>
<th>Study</th>
<th>HAMD</th>
<th>Active/Sham</th>
<th>Sample size</th>
<th>Study effect size (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>23. Miniussi et al. 2005</td>
<td>HAMD</td>
<td>21-item</td>
<td>54 ± 12.7</td>
<td>17 Hz, 110% MT, 40 trains, 3 s on, 120 s off</td>
</tr>
<tr>
<td>24. Rossini et al. 2005</td>
<td>HAMD</td>
<td>21-item</td>
<td>53 ± 12.4</td>
<td>15 Hz, 100% MT, 30 trains, 2 s on, 28 s off</td>
</tr>
<tr>
<td>25. Rumi et al. 2005</td>
<td>HAMD</td>
<td>17-item</td>
<td>48.4 ± 13.7</td>
<td>5 Hz, 120% MT, 25 trains, 10 s on, 20 s off</td>
</tr>
<tr>
<td>26. Avery et al. 2006</td>
<td>HAMD</td>
<td>17-item</td>
<td>46.4 ± 12.1</td>
<td>17 Hz, 110% MT, 40 trains, 3 s on, 120 s off</td>
</tr>
<tr>
<td>27. Herwig et al. 2007</td>
<td>HAMD</td>
<td>21-item</td>
<td>39.5 ± 12.8</td>
<td>10 Hz, 110% MT, 100 trains, 2 s on, 8 s off</td>
</tr>
<tr>
<td>28. Loo et al. 2007</td>
<td>HAMD</td>
<td>21-item</td>
<td>38.9 ± 8.8</td>
<td>10 Hz, 110% MT, 30 trains, 5 s on, 25 s off</td>
</tr>
<tr>
<td>29. Mogg et al. 2008</td>
<td>HAMD</td>
<td>17-item</td>
<td>44.3 ± 10.3</td>
<td>10 Hz, 110% MT, 20 trains, 5 s on, 55 s off</td>
</tr>
<tr>
<td>30. O’Reardon et al. 2007</td>
<td>HAMD</td>
<td>21-item</td>
<td>44.2 ± 9.7</td>
<td>10 Hz, 120% MT, 75 trains, 4 s on, 26 s off</td>
</tr>
</tbody>
</table>

HAMD, Hamilton Depression Rating Scale; MADRS, Montgomery–Asberg Depression Rating Scale; TMS, transcranial magnetic stimulation; S.D., standard deviation; MT, motor threshold; N.A., not available.

Medication resistance is defined as the failure to respond to >2 trials of antidepressants or history of failed responses to electroconvulsive therapy.

Discussion

The aim of this meta-analysis was to investigate whether high-frequency rTMS applied over the left DLPFC is effective in treating depression. Table 2 presents the main outcomes of the meta-analysis.
DLPFC can be considered an effective treatment method in depression. The results show that rTMS treatment has significantly more antidepressant efficacy than sham treatment. The effect size, $d=0.39$, shows that there is little doubt that magnetically induced electrical currents in the brain improve depression.

However, an important point in studying the antidepressant effects of rTMS is the control condition. The vast majority of the studies (80% in this meta-analysis) use active stimulation with the coil oriented at a 45° or 90° angle. Even though the magnetic field intensity is oriented away from the target, it has been demonstrated that these forms of sham can be active (Lisanby et al. 2001). In addition, coil placements in the real and sham conditions can produce considerable variation in felt scalp sensations that may jeopardize the double-blind nature of the trial. In an attempt to overcome this limitation, several studies have made use of purpose-built sham coils that mimic the scalp sensations and sound click of real rTMS. Moreover, several groups are currently working on refining the quality of the control condition. There is some recent evidence that focal electrical stimulation of the scalp as a sham condition is capable of creating a true indistinguishable placebo condition (Arana et al. 2008). Evidently, at this point more work is needed but the initial results are promising.

Related to the previous point is the issue of successful blinding during treatment. In this meta-analysis only data points were included that were acquired during the blind phase of the study. One of the quality criteria for study inclusion was that patients and clinical raters were blind to the stimulation condition. Although the interaction between the physician who applied rTMS and the patients was kept to a minimum during treatment, the fact that the physician was not blind may nevertheless have influenced treatment outcome. Of the 30 studies, six studies statistically checked whether patients had remained blind during treatment (Berman et al. 2000; Fitzgerald et al. 2003; Jorge et al. 2004; Avery et al. 2006; Loo et al. 2007; Mogg et al. 2008). Five of the six studies reported that patients were unsuccessful in guessing their treatment condition. Only Mogg et al. (2008) reported that patients in the real condition were significantly better in guessing their treatment (70%). Notably, patients in the sham rTMS condition did not score above chance level (38%).

According to the authors, many patients in the real rTMS condition made their guess on the basis of mood improvements experienced during the actual treatment. In sum, even though blinding can be successful at this point, the nature of rTMS as well as the unavailability of an ideal sham condition makes it difficult for researchers to ascertain patients remain blind to the type of treatment. The HAMD and the MADRS are the most commonly used primary outcome measures of depression ratings. Importantly however, the HAMD emphasizes the somatic aspects of depression whereas the MADRS stresses the psychological symptoms of depression (Heo et al. 2007). Thus, different measurement instruments may yield different treatment outcomes. Removal of the TMS trials using the MADRS as the primary outcome measure did not affect heterogeneity ($Q_T=28.41, p=0.39$) or effect size ($d=0.39$), demonstrating that the rTMS trials using the MADRS did not bias the current results in any meaningful way. However, the method of meta-analysis has been criticized for combining dissimilar studies, publication bias and inclusion of poor-quality studies. In the present study these concerns were tackled by imposing stringent inclusion criteria, examining publication bias and heterogeneity. In fact, criticisms of meta-analyses are equally applicable to traditional, non-quantitative, narrative reviews of the literature (Rosenthal & DiMatteo, 2001). Furthermore, the fact that the tests for heterogeneity were not significant shows that the variance among the effect sizes of the different studies were not greater than expected by sampling error and the results obtained are reliable.

How do the current findings relate to other recent meta-analyses? In one of the latest meta-analytical studies, the effect sizes of 13 earlier rTMS studies (324 patients) from the meta-analyses of Martin et al. (2003) were compared to the effect sizes of five more recent rTMS studies (Gross et al. 2007). The results showed that the effect size of the more recent rTMS studies (246 patients) was estimated to be 0.76 (95% CI

### Table 2. Main results

<table>
<thead>
<tr>
<th>Comparison</th>
<th>No. of studies</th>
<th>Real</th>
<th>Sham</th>
<th>Total</th>
<th>Combined effect size</th>
<th>95% CI</th>
<th>$Z$</th>
<th>$p$</th>
<th>$Q_T$</th>
<th>$p(x^2)$</th>
<th>$N_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real versus sham</td>
<td>30</td>
<td>606</td>
<td>558</td>
<td>1164</td>
<td>0.39</td>
<td>0.25–0.54</td>
<td>6.52</td>
<td>&lt;0.0001</td>
<td>30.46</td>
<td>0.39</td>
<td>269</td>
</tr>
</tbody>
</table>

CI, Confidence interval; $Q_T$, total heterogeneity.
0.51–1.01) as compared to an effect size of 0.35 (95% CI 0.04–0.66) in the 13 earlier studies. These results suggest that more recent clinical trials are more effective in sorting antidepressant effects. In a recent meta-analysis a pooled effect size of 0.65 was reported and, according to the authors, indicated a clinical effect (Herrmann & Ebmeier, 2006). However, the test for heterogeneity was significant, which suggests that other variables besides rTMS played a mediating role in the observed treatment effect. The effect size found in the current meta-analysis seems to be more in line with meta-analyses reporting moderate effect sizes (e.g. Martin et al. 2003).

Although rTMS seems to show only moderate effects, there is meta-analytical evidence indicating that the moderate effect size presently observed is comparable to effect sizes seen in active placebo-controlled trials with pharmacological treatments (Joffe et al. 1996; Moncrieff et al. 1998, 2004). Moderate effect sizes have been observed for both tricyclic and tetracyclic agents. A recent meta-analysis examining the effects of antidepressant medication in 1320 patients with post-stroke depression showed that the pooled response rates in active and placebo arms were 65.2% and 44.4% respectively. The effect size was 0.23, suggesting a small to moderate, but significant, improvement in depression in the active group (Chen et al. 2006). In sum, these findings suggest that rTMS can be as effective as at least some of the commercially available antidepressant medications.

Even though rTMS may be as effective as antidepressant medications, questions still remain as to why rTMS treatment produces moderate effect sizes and how to optimize the antidepressant effects of rTMS. The current meta-analysis used baseline depression scores prior to entering treatment and these were compared to the depression scores directly after the final session. The number of studies that conduct follow-up measurements is small, and there is some evidence that TMS may suffer from a therapeutic onset delay, analogous to pharmacological medication. Two follow-up studies examining the antidepressant effects following 10 and 15 sessions of rTMS found improvements in the baseline-corrected percentage change in HAMD scores after 1 week ($d = 0.49$) and several weeks ($d = 0.44$) post-treatment respectively (Mosimann et al. 2004; Miniussi et al. 2005; Rossini et al. 2005; Rumi et al. 2005). Other studies have also published beneficial post-treatment effects from weeks to several months (Dannon et al. 2002; Avery et al. 2006), but zero findings have been reported as well (for a review see Martin et al. 2003). Logistic and experimental technical issues, as well as ethical concerns, make it difficult to conduct controlled clinical trials with sufficiently long follow-up assessments. These kinds of studies are, nevertheless, important to establish the temporal course of rTMS-related antidepressant effects and to elucidate the underlying physiological mechanisms.

Another issue concerns a selection bias in the patient population that may result in an underestimation of the antidepressant potential of rTMS. In the studies reported here, all patients suffered from major depression and many of the patients had failed to respond to at least two antidepressant drug treatments and/or ECT (see also Table 1). Prior studies that have investigated the antidepressant effects of ECT have found that medication-resistant patients often show small to moderate improvement and are more vulnerable to relapse (Prudic et al. 1996; Dannon et al. 2002). An additional variable that may stand in the way of large effect sizes is age (Sackheim, 1994). Mean age and standard deviation of the current patient population was 49.1 ± 7.5 years and some evidence exists suggesting that younger depressed patients respond better to antidepressant treatment (Lynes et al. 1996); but see Radziwon-Zaleska et al. (2006) for an exception. Age-related reductions in brain plasticity and increases in scalp to prefrontal cortex distance that result in smaller electrical currents reaching the target tissue are possible explanations (Grafman, 2000; Nahas et al. 2004).

Besides the population bias, it has been suggested that lengthening the duration of treatment further than the typical 10 sessions enhances antidepressant efficacy (Fitzgerald et al. 2003, 2006; Avery et al. 2006; Loo et al. 2007). A study by Fitzgerald et al. (2003), who applied 4 weeks of fast-frequency rTMS in major depressive patients, reported progressive clinical improvements on the MADRS, with $d > 0.80$. It should, however, be mentioned that only the initial 2 weeks were double-blind in this study. Nonetheless, the results provide some support for a positive relationship between treatment duration and clinical response.

The basic neural framework for applying fast-frequency rTMS comes from observations that depression is linked to left DLPFC hypovigilance (for a review see Davidson et al. 1999) and that increasing neuronal activity over time may have beneficiary effects. Homeostatic behavioural and brain function may, however, also require a balance between the left and the right prefrontal cortex (Schutter & van Honk, 2005b). In agreement with this, reducing neuronal activity of the right DLPFC with slow-frequency rTMS also has antidepressant effects. A double-blind sham-controlled study found proof for antidepressant effects of 10 daily sessions of slow-frequency (1 Hz) rTMS (120 pulses) over the right DLPFC, $d = 0.45$ (Klein et al. 1999). Moreover, a double-blind sham-controlled design of Fitzgerald et al. (2003) even found greater
reductions in baseline-corrected change in MADRS scores between 2 and 4 weeks of slow-frequency as compared to fast-frequency rTMS treatment, $d = 1.20$. Although some authors have reported no significant improvement after slow-frequency rTMS over the right DLPFC (Höppner et al. 2003; Kaufmann et al. 2004), this approach may nevertheless be an interesting alternative to fast-frequency TMS for other reasons as well. Slow-frequency rTMS is usually better tolerated by patients and minimizes the risk for adverse events (Wassermann, 1998; George et al. 1999; Post et al. 1999). The slow-frequency technique has even been successfully applied to treat intractable epilepsy (Joo et al. 2007). Of note, using an original combination of fast- and slow-frequency techniques (Loo et al. 2003), Fitzgerald et al. (2006) applied three trains of slow-frequency rTMS of 140-s duration over the right DLPFC, immediately followed by 15 trains of 5 s of fast-frequency rTMS over the left DLPFC in a double-blind design. Significant reductions in the endpoints of the MADRS scores were observed after 10 sessions of active as compared to sham treatment, $d = 0.5$.

In addition to methodological innovations, technical developments also play an important part in the search for clinically effective treatment protocols. The discovery of inducing long-lasting changes in neuronal excitability, wherein the cortex is stimulated with bursts of 50-Hz rTMS repeated every 0.2 s, is an exemplary technical innovation that will undoubtedly contribute to the fine-tuning of the stimulation parameters (Huang et al. 2005). Another exciting development is the construction of a specially designed coil that allows stimulation of deep brain structures (Roth et al. 2007) and may be able to directly reach the brain’s reward and motivation circuitry (Dunlop & Nemeroff, 2007). The DLPFC may not be the ideal target region for rTMS in depression and possible alternative regions have been identified (Chen et al. 2006). Electrophysiological scalp recordings and rTMS studies have presented evidence for parietal cortex involvement in depression (Keller et al. 2000; van Honk et al. 2003). Furthermore, neuroanatomical evidence and preliminary support for antidepressant properties of high-frequency rTMS over the medial cerebellum have been provided (Schutter et al. 2003; Schutter & van Honk, 2005a, b, 2006).

Finally, a potential setback of all studies is that the inclusion criteria are exclusively based on psychiatric evaluation and no information is available on possible depression-related brain disturbances. Clinically relevant response rates to rTMS may well depend on whether the depression is paralleled by identifiable neural disturbances. In agreement, it has been shown that resting-state metabolism in the anterior cingulate cortex predicts improvements in depression after 10 sessions of fast-frequency rTMS over the left DLPFC (Teneback et al. 1999). Thus, in addition to the psychiatric evaluation, information on neurobiological abnormalities can be helpful in establishing guidelines and clinical prognoses on whether rTMS will be effective or not.

In conclusion, the current meta-analysis included 30 double-blind sham-controlled treatment trials with 1164 patients in total. The results show that fast-frequency rTMS over the left DLPFC is superior to sham and may be as effective as at least a subset of commercially available antidepressant medications. In addition, TMS is a safe method and because of its few side-effects is well tolerated by patients. However, at this point caution should be exercised because the integrity of blinding and the lack of a proper control condition are considered limitations of rTMS trials. In addition, age bias, medication, suboptimal stimulation parameters, lack of biological information and follow-up assessments may stand in the way of exploiting the effects of rTMS. Nevertheless, ongoing methodological innovations and technological advancements in the field will without doubt further improve the quality and therapeutic efficacy of future rTMS trials. All in all, the present findings suggest that rTMS treatment may be an alternative for patients suffering from major (non-psychotic) depression, and especially for those patients who do not tolerate the side-effects associated with regular pharmacological treatment.

Acknowledgements

I thank Drs D. Avery, G. Eschweiler, R. Jorge, F. Koerselman, C. Miniussi, D. McLoughlin, A. Mogg, C. Loo and J. van Honk and three anonymous reviewers. This work was supported by an Innovational Research Grant (VENI 451-04-070) from the Netherlands Organization for Scientific Research (NWO).

Declaration of Interest

None.

References


Transcranial magnetic stimulation in depression


Dunlop BW, Nemeroff CB (2007). The role of dopamine in the pathophysiology of depression. *Archives of General Psychiatry* 64, 327–337.


