Transcranial Magnetic Stimulation in the Management of Mood Disorders

Charlotte L. Allan a  Lucie L. Herrmann b  Klaus P. Ebmeier a

a Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, and b Institute of Psychiatry, King’s College, London, UK

Key Words
Transcranial magnetic stimulation · Depression · Affective disorders · Meta-analysis

Abstract

Background: Many trials of transcranial magnetic stimulation (TMS) have used small samples and, therefore, lack power. Here we present an up-to-date meta-analysis of TMS in the treatment of depression. Methods: We searched Medline and Embase from 1996 until 2008 for randomized sham-controlled trials, with patients and investigators blinded to treatment, and outcome measured using a version of the Hamilton Depression Rating Scale (or similar). We identified 1,789 studies. Thirty-one were suitable for inclusion, with a cumulative sample of 815 active and 716 sham TMS courses. Results: We found a moderately sized effect in favour of TMS [Random Effects Model Hedges’ g = 0.64, 95% confidence interval (95% CI) = 0.50–0.79]. The corresponding Pooled Peto Odds Ratio for treatment response (≥ 50% reduction in depression scores) was 4.1 (95% CI = 2.9–5.9). There was significant variability between study effect sizes. Meta-regressions with relevant study variables did not reveal any predictors of treatment efficacy. Nine studies included follow-up data with an average follow-up time of 4.3 weeks; there was no mean change in depression severity between the end of treatment and follow-up (Hedges’ g = −0.02, 95% CI = −0.22 to +0.18) and no heterogeneity in outcome. Discussion: TMS appears to be an effective treatment; however, at 4 weeks’ follow-up after TMS, there had been no further change in depression severity. Problems with finding a suitably blind and ineffective placebo condition may have confounded the published effect sizes. If the TMS effect is specific, only further large double-blind randomized controlled designs with systematic exploration of treatment and patient parameters will help to define optimum treatment indications and regimens.

Background

Affective disorders are common, and are associated with considerable morbidity and mortality [1, 2]. Although pharmacological and psychological therapies for mood disorders have improved management and outcome, some patients show poor or minimal response to such treatments. Therefore, the prospect of using repetitive transcranial magnetic stimulation (TMS), which appears to be safe with minimal side effects [3], has aroused great interest, especially for patients with treatment-resistant depression [4]. In particular, TMS has been proposed as an alternative treatment to electroconvulsive therapy since it does not have cognitive side effects, or risks associated with anaesthesia [5]. TMS has now been approved for use in treatment-resistant depression in Canada, Aus-
tralia, New Zealand, the European Union and Israel [4]. More recently, the US Food and Drug Administration have approved TMS for use in treatment-resistant, major depressive disorder [6].

In addition to treatment for depression, TMS has been used as a research tool, and has shown potential for treating a range of disorders including: tinnitus [7], chronic pain [8], mania, obsessive-compulsive disorder, post-traumatic stress disorder, panic disorder and schizophrenia [9]. However, there is little evidence that TMS is effective in conditions other than depression [9].

Studies to date have tended to involve relatively small numbers of subjects, and vary in terms of the site and frequency of stimuli used. To review the available evidence in an objective fashion, we present a systematic review with meta-analysis of randomized controlled trials using TMS in the treatment of depression. Although we focus on the use of TMS in the treatment of depression, treatment of mania and of patients with bipolar disorder will be briefly discussed.

**What Is TMS?**

Repetitive TMS is non-invasive, and its action is based on the induction of an electromagnetic field. When an alternating current is passed through a coil applied over the patient’s scalp, a magnetic field is generated in a specific area of the cerebral cortex, depending on coil position. If sufficient intensities of coil current are used, the associated magnetic field induces electrical currents in the brain that lead to neuronal depolarization [10].

The site most commonly used for the treatment of depression is the left prefrontal cortex, which has been highlighted for its involvement in the regulation of positive and negative emotion [11, 12]. However, right prefrontal and bilateral stimulation have also been used. Different intensities can be used for stimulation, typically ranging from 80 to 120% of motor threshold. Stimulation frequencies range from ‘high frequency’ (5–20 Hz) to ‘low frequency’ (≤1 Hz). Based on hypotheses suggesting relative underactivity of the left and overactivity of the right dorsolateral prefrontal cortex (DLPFC) in depression, higher (stimulating) frequencies are applied to the left, low (‘quenching’) frequencies to right DLPFC [13]. TMS is generally well tolerated, with mild side effects limited to transient scalp discomfort or pain [14]. It does not appear to have cognitive side effects, and risk of seizures is very low [3, 15].

When evaluating a novel technique such as TMS, double-blind, randomized controlled trials are the gold standard. However, in the context of TMS, the development of an effective sham control has been difficult [9]. There have been technological problems with creating identical scalp sensation, pain and noise compared to the active treatment, without providing any therapeutic benefit [16–18]. There is also the issue that it is difficult to blind the operator to whether the treatment is active or sham [17]. Much work has been undertaken to design an acceptable sham treatment and this will hopefully help to improve blinding and avoid inadvertent induction of cortical activity [18–21].

**TMS in the Treatment of Depression: A Meta-Analysis**

**Method**

Medline and Embase databases were searched from 1996 until the end of 2008, using the search terms 'Depression OR Depressed' and 'Transcranial OR TMS'. A total of 1,789 studies were identified. Titles and abstracts of the articles were examined to determine whether or not they could be included. If data were published repeatedly as a whole or in parts, the most recent or most inclusive publication was used. Studies were included in the meta-analysis if they met the following inclusion criteria: (a) randomized parallel or cross-over design, with sham control; (b) both patients and investigators (i.e. baseline and outcome assessors as opposed to the treatment team) unaware of the treatment conditions; (c) samples consisted of ≥10 patients in each group; (d) participants in each treatment group had a diagnosis of major depressive episode, and (e) outcome was measured using a version of the Hamilton Depression Rating Scale (HDRS) or Montgomery-Åsberg Depression Rating Scale (MADRS), with baseline and follow-up scores that were available or could be derived from published tables or figures. Altogether, 1,757 studies did not satisfy the above inclusion criteria. The remaining 31 studies were considered suitable for our meta-analysis and included a total of 815 depressed patients, randomized to active TMS and 716 randomized to sham TMS.

Data were extracted and recorded in a structured fashion: (a) study design; (b) sample characteristics (e.g. age, whether patients had treatment-resistant depression); (c) treatment parameters (stimulation frequency and intensity, number of treatment sessions, site of stimulation); (d) mean and standard deviation of the outcome measure before and after treatment, and if available at follow-up, and (e) number of responders (as defined by a reduction of 50% or more in depression rating scores).

In cross-over trials, only data from the initial stage of the trial were used to avoid carryover effects between trial stages [22]. In trials with more than one experimental group compared with the same controls, only one of the comparisons was included. When the experimental groups did not differ significantly in terms of efficacy, they were pooled. Where outcome data were reported with several different rating scales, data were extracted in a hierarchical fashion (17-item HDRS first, if this was not available, the
21-item HDRS, then MADRS). Outcome data were collected immediately after completion of the TMS course and at follow-up 1–12 weeks after the end of the TMS course. Data from studies in which participants received any kind of post-randomization open TMS during the follow-up period were excluded.

**Data Analysis**

Efficacy was first investigated by calculating random model effect sizes based on changes in depression scales (Hedges’ g; Comprehensive Meta-Analysis Software) [23]. The second stage of analysis involved computing Peto’s Odds Ratios for treatment response (≥50% reduction in depression scores) in a subset of the studies that provided such data. Fail-Safe N, Kendall’s τ, Egger’s Regression Intercept and Duval and Tweedie’s Trim and Fill were used to test for bias [23].

To test the influence of confounding variables, regression analysis was used to assess whether there were any variables that predicted the magnitude of effect sizes in individual studies. These included: site (left DLPFC or right DLPFC), total number of stimuli, stimulus intensity, stimulus frequency, year of publication, diagnosis (patients with only unipolar depression versus at least some patients with bipolar disorder), mean age of participants, treatment resistance and medication effects. A medication effect was considered most likely if patients were started on a new antidepressant or discontinued their previous medication within a period of 7 days before to 7 days after the first TMS treatment. Finally, for those studies with a follow-up period beyond the end of treatment, further change in depression scores to follow-up was analysed using Hedges’ g.

**Results**

**Efficacy Analysis**

Thirty-one studies were suitable for inclusion [13–15, 24–51]. This yielded 32 comparisons since 1 study had two participant groups which were analysed separately with their own controls [50], and led to an overall sample size of 1,317. Out of 699 patients receiving active TMS, 250 (35.8%) were classified as responders; of 618 patients who received sham TMS, 93 (15.0%) met the criteria for clinical response. The pooled Peto Odds Ratio was 4.1 (95% CI = 2.85–5.90). Heterogeneity between studies exceeded that expected by chance (Q = 85.2). Duval and Tweedie’s Trim and Fill procedure trimmed the 10 studies with the highest Peto’s Odds Ratios and added them on the other side of 1, resulting in a more conservative odds ratio of 2.61 (95% CI = 1.78–3.82; Q = 85.2).

**Predictor Analysis**

There were no significant predictors of TMS efficacy, after examining the following using meta-regression: site, stimulus intensity, stimulus frequency, year of publication, diagnosis mean age of participants, treatment resistance and medication effects. However, both Hedges’ g (p = 0.02) and Peto’s Odds Ratio (p = 0.04) were negatively correlated with total stimulus number. After removing the variance attributable to numbers of stimuli, there were still outcome differences between studies that were greater than expected by chance. There was greater efficacy (measured by Peto’s Odds Ratio) of TMS, when patients entering the studies were previously treatment resistant (p = 0.04).

**Follow-Up Data**

A total of 9 studies reported results up to 12 weeks after completion of TMS (weighted average: 4.33 weeks) [27, 35, 36, 38, 39, 42, 44, 48, 51]. Effect size was calculated, when comparing outcome immediately after completion of TMS, with outcome at the end of the follow-up period. There was no mean change in depression severity between the end of treatment and follow-up (Hedges’ g = 0.02, 95% CI= −0.22 to +0.18) and no heterogeneity in outcome.
TMS in Mania

With the premise that location of TMS stimulation can induce either happiness or sadness depending on prefrontal side of stimulation and stimulation frequency [52, 53], Grisaru et al. [54] undertook a blinded, controlled trial to investigate the use of TMS in mania. They found significant improvements in mood (p < 0.01) for patients treated with right, rather than left prefrontal TMS at 20 Hz. Although an open-label study using right high-frequency prefrontal TMS in the treatment of bipolar mania found that TMS was associated with a reduction in manic symptoms [55], Kaptsan et al. [56] could not replicate its result in a sham-controlled study.

Animal modelling has been used in an attempt to provide more evidence regarding the use of TMS in mania. Shaldivin et al. [57] used an amphetamine-induced hyperactivity model of mania in rats, and found that once-daily TMS treatment for up to 7 days significantly reduced amphetamine-induced hyperactivity whereas twice-daily TMS over the same time period enhanced hyperactivity. Taken together, evidence for the use of TMS as a treatment for mania is disappointing.

TMS in Bipolar Depression

TMS has been used in patients with depression in the context of bipolar affective disorder. Studies in our meta-analysis included patients with both unipolar and bipolar depression, and the underlying diagnosis was not a confounding variable. Importantly, the use of TMS for the treatment of depression in patients with bipolar disorder does not appear to induce mania [33, 58, 59].

Discussion

Our meta-analysis of changes in depression scales, and clinical improvement of more than 50% on such scales, confirms a moderate statistically significant effect of active therapy. This concurs with a recent meta-analysis in treatment-resistant depression [60]. Although we aimed to limit variability by strict inclusion criteria, the meta-analysis using Hedges’ g and Peto’s Odds Ratio showed significant heterogeneity between studies, greater than would be expected by chance. Significant variability between studies may exist due to sampling differences between patient groups, differences in treatment protocol and the measurements used. The consequence of significant variability in study outcomes means that we cannot put forward one ‘successful’ study protocol that would be predicted to replicate the pooled effect size.

Bias in meta-analyses, and for that purpose in all reviews, occurs if small positive studies are more likely to be published than small negative studies. We excluded studies with fewer than 10 subjects per group to reduce the influence of publication bias. Further, a variety of methods exist (and are implemented in Comprehensive Meta-Analysis [23]) that can estimate the potential bias due to non-publication of small negative studies: while Fail-Safe N in meta-analyses provides some reassurance, both Egger’s regression and Kendall’s τ showed a significant degree of publication bias. Duval and Tweedie’s Trim and Fill procedures helped arrive at a more realistic estimate of the effect size. Even after such corrections, a significant treatment effect remains.

Significant variability of outcomes raises the possibility that measurable differences in protocol, participants and methods between studies may predict the outcome and thus be responsible for the heterogeneity of results. Unfortunately, none of the examined variables were plausible, independent predictors of efficacy. The inverse relationship of Hedges’ g and Peto’s Odds Ratio with total number of stimuli was determined entirely by the low treatment effect size in the large study by O’Reardon et al. [14]. As it implies an inverse ‘dose effect’ that would certainly run counter to the hypothesis commonly proposed, and the hypothesis underpinning the study itself [14]. Previous studies have found that age [61] and treatment resistance are negative predictors of improvement in depressive symptoms [61, 62]. On the other hand, treatment resistance has also been reported as a positive predictor of treatment outcome [63]. It is possible that this counterintuitive effect, which we were able to confirm, is due to a number of confounding factors. Firstly, there is the issue that investigators may be more likely to persuade ethics committees that depressed patients who have not responded to previous treatment should receive sham treatment, and secondly, there is a preference to use add-on designs for studies of patients who are not treatment resistant. Other positive predictors we could not replicate in the meta-regression include: shorter duration of current episode, absence of comorbid anxiety and presence of sleep disturbance [62, 63].

In conclusion, despite great interest and enthusiasm for TMS as a treatment for affective disorders, particularly for depression, there is no evidence for lasting treatment effects beyond 12 weeks. As TMS involves a great number of variables relating to the treatment itself (fre-
quency, intensity, number and arrangement of stimuli, and coil location) as well as to the protocol (number, duration and time course of sessions), it is possible that the optimum treatment protocol is yet to be discovered. Such variability in design may well account for the heterogeneity of outcome. Meta-regression may not be powerful enough to tease apart the causes of variability so that large studies that try to account for the contribution of protocol variations will be necessary. Although the data published so far are significantly biased, widely acceptable corrections for such bias suggest that only marginally smaller effect sizes remain after accounting for this effect. Doubts remain about the validity of the sham conditions used in most TMS studies: patients may recognize the active treatment, which will enhance the effect of active treatment, but certain sham treatments may affect the cerebral cortex, making it possible that sham TMS is an ‘active placebo’ that may reduce the observed treatment effect. It is unclear which of these effects will prevail. Nonetheless, for many patients, TMS seems to be an acceptable and attractive treatment that is associated with minimal side effects; it is therefore likely to stay for the foreseeable future.

References


