Meta-analysis of repetitive transcranial magnetic stimulation in the treatment of auditory verbal hallucinations: Update and effects after one month

C.W. Slotema a,⁎, A. Aleman b, Z.J. Daskalakis c, I.E. Sommer d

a Parnassia Bavo Psychiatric Institute, Lijnbaan 4, 2512 VA The Hague, the Netherlands
b Department of Cognitive Neuropsychiatry, University Medical Center Groningen, University of Groningen, Postbus 30001, 9700 RB Groningen, the Netherlands
c Centre for Addiction and Mental Health, University of Toronto, 250 College Street, 7th Floor, Toronto, Ontario, Canada M5T 1R8
d Department of Psychiatry, Neuroscience Division & Rudolf Magnus Institute for Neuroscience, University Medical Centre Utrecht, the Netherlands

Abstract

Objective: Several meta-analyses considering repetitive transcranial magnetic stimulation (rTMS) for auditory verbal hallucinations (AVH) have been performed with moderate to high mean weighted effect sizes. Since then several negative findings were reported in relatively large samples. The aim of this study was to provide an update of the literature on the efficacy of rTMS for AVH and to investigate the effect of rTMS one month after the end of treatment.

Data sources: A literature search was performed from 1966 through August 2012 using Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Embase Psychiatry, Ovid Medline, PsycINFO and PubMed. Randomized, double blind, sham-controlled studies of Controlled Trials, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Data sources:

Keywords:
Transcranial magnetic stimulation
Auditory verbal hallucinations
Meta-analysis

A R T I C L E   I N F O

Article history:
Received 10 April 2012
Received in revised form 20 August 2012
Accepted 28 August 2012
Available online 30 September 2012

Keywords:
Transcranial magnetic stimulation
Auditory verbal hallucinations
Meta-analysis

1. Introduction

Repetitive transcranial magnetic stimulation (rTMS) is a safe treatment method with only mild side effects such as transient headache and scalp discomfort during stimulation. Hoffman et al. (1999) were the first to investigate the effect of rTMS in the treatment of medication-resistant auditory verbal hallucinations (AVH). In their initial study, rTMS was directed at the left temporoparietal cortex at a frequency of 1 Hz. This site of stimulation is believed to overlay Brodmann area 40 (Homan et al., 1987) and was selected in light of results of a previous positron emission tomography (PET) study demonstrating activation in this brain area during AVH (Silbersweig et al., 1995). Furthermore, this region has a central role in speech perception (Ojemann, 1978; Fiez et al., 1996; Benson et al., 2001). Since that first study several randomized sham-controlled studies on this subject have been published, some with positive, others with negative findings. These publications have been summarized in four meta-analyses revealing moderate to high mean weighted effect sizes ranging from 0.54 to 1.0 (Aleman et al., 2007; Tranulis et al., 2008; Freitas et al., 2009; Slotema et al., 2010). Since the publication of these meta-analyses several negative studies have become available which included relatively large sample sizes. In addition, these previous meta-analyses focussed on the direct effects of rTMS at the end of the treatment trial, and did not account for the long-term effects after one month.
not explore effects of rTMS at follow-up after some weeks or months. The aim of this study is therefore twofold:

1. to conduct a new meta-analysis considering the effect of rTMS for AVH
2. to investigate the long term effects of rTMS for AVH.

2. Experimental/materials and methods

A search of the literature was performed using Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Embase Psychiatry 1997 through August 2012, Ovid Medline through August 2012, PubMed 1990 through August 2012 and PsycINFO 1990 through 2012. All articles were searched for cross-references.

The search terms auditory verbal hallucinations, auditory hallucinations, psychosis, psychotic features, transcranial magnetic stimulation, repetitive transcranial magnetic stimulation, TMS and repetitive TMS were used.

Criteria for inclusion were:

1. Treatment with rTMS using the frequency of 1 Hz.
2. The summed score of the Auditory Hallucination Rating Scale (AHRs) (Hoffman et al., 2003) was used as an outcome measure; if the summed score could not be obtained, the item ‘frequency’ of the AHRs or a Visual Analogue Scale such as the Hallucination Change Scale was used as a second or third choice respectively. In case the severity of AVH was not included as an outcome measure, the scores for a scale assessing positive symptoms (of schizophrenia) were obtained.
3. The study was performed in a double-blind, randomized controlled design using a sham condition. In a separate analysis cross-over studies were excluded as patients might not remain blinded during this design, which may influence the results.
4. Sufficient data were needed to compute Hedges’ g (i.e., sample size, means, standard deviations or exact t or p values for rTMS main effect for change scores).
5. The article was written in English.
6. When different publications had an overlap in patient samples, the article with the largest sample size was included.

The studies were screened conforming to the criteria for inclusion. The first author performed the screening.

2.1. Data extraction

The following data were acquired: number of treated patients per treatment condition, pre- and posttreatment (defined as immediately after the end of treatment and one month after cessation of treatment) means and standard deviations of the severity of AVH at baseline and at end of treatment, or exact F, t or p values. Furthermore, the study design and treatment parameters, such as frequency, percentage of the individual motor threshold, number of TMS pulses, number of sessions, focus of treatment and type of coil, were acquired. Finally, information considering dropouts and side effects was obtained.

Authors were contacted and invited to send additional data in case publications contained insufficient or incomplete results.

2.2. Effect size calculation

The mean weighted effect size, Hedges’ g, was computed with the aims of Comprehensive Meta-Analysis Version 2.0 (Biostat, Englewood, New Jersey) in a random effects model. First, the effect sizes were calculated for the mean change in symptom severity between pre- and posttreatment for the separate conditions and weighted according to sample size. In studies with three treatment conditions, the two actual treatments were compared separately with the sham condition. Then, meta-analytic methods were used to obtain a combined, weighted effect size. In order to investigate the effect of rTMS on the severity of psychosis, a separate analysis was conducted by using the changes assessed with the summed score of the positive items of the Positive And Negative Syndrome Scale. Separate analyses were confined to rTMS directed at the left temporoparietal area, rTMS targeted at all brain regions and interview-based clinician-rated versus self-report measures.

A homogeneity statistic, I², was computed to test whether the studies could be taken to share a common population effect size (Higgins et al., 2003). A percentage of 50% or higher indicates heterogeneity of the individual study effect sizes, which poses a limitation to a reliable interpretation of the results. Whenever significant heterogeneity was found, a moderator analysis was performed to investigate the potential moderating factors, such as localization of target area for stimulation, intensity of the individual motor threshold and number of TMS pulses. These parameters were correlated with Hedges’ g using Pearson’s correlations in Statistical Packages for the Social Sciences (SPSS, Chicago, Illinois) version 18.

The effect size can be overestimated in case of an omission of studies with negative results in the literature. Therefore, a fail-safe number was computed, which is an estimation of the number of missing studies that is needed to change the results of the meta-analysis to non-significant (Rosenthal, 1979). To visualize a putative decline of effect sizes over the last years, the mean weighted effect sizes were plotted against year of publication.

3. Results

Fifteen studies were included in the meta-analysis considering rTMS directed at the left temporoparietal area. Two additional studies were included for the meta-analysis for all rTMS foci. In Table 1 the number of studies that was excluded and the reason for exclusion are presented.

Three hundred and thirty-seven patients were included (Hoffman et al., 2000; McIntosh et al., 2004; Schonfeldt-Lecuona et al., 2004; Chibbaro et al., 2005; Fitzgerald et al., 2005; Hoffman et al., 2005; Lee et al., 2005; Poulet et al., 2005; Brunelin et al., 2006; Jandl et al., 2006; Rosa et al., 2007; Saha et al., 2008; Vercammen et al., 2009; Loo et al., 2010; Slotema et al., 2011; De Jesus et al., 2011; Blumberger et al., in press: Bais et al., in preparation); 209 and 197 patients received real rTMS targeted at the left temporoparietal area and sham-treatment respectively. Details of the treatment-paradigms of all studies are presented in Table 2. Ten out of fifteen studies included patients with therapy-resistant AVH and Hoffman et al. (2005) included 42 out of 50 patients with therapy-resistant AVH.

Table 1

<table>
<thead>
<tr>
<th>Number of excluded studies and reason for exclusion.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gateway</strong></td>
</tr>
<tr>
<td>ovid</td>
</tr>
<tr>
<td>No clinical trial 19</td>
</tr>
<tr>
<td>No RCT with placebo 17</td>
</tr>
<tr>
<td>No English 12</td>
</tr>
<tr>
<td>Severity of AVH 6</td>
</tr>
<tr>
<td>no outcome measure 3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
<tr>
<td>PubMed 76</td>
</tr>
<tr>
<td>No clinical trial 61</td>
</tr>
<tr>
<td>No RCT with placebo 83</td>
</tr>
<tr>
<td>No English 11</td>
</tr>
<tr>
<td>Severity of AVH 2</td>
</tr>
<tr>
<td>no outcome measure 104</td>
</tr>
</tbody>
</table>

RCT = randomized controlled trial.
AVH = auditory verbal hallucinations.
The mean weighted effect size was 0.44 (95% CI 0.19–0.68, $I^2 = 35.7$). This corresponds to a moderate effect. See Fig. 1 for the results of the meta-analysis. The fail-safe number was 996 studies. A moderator analysis did not reveal a correlation between specific treatment-paradigms (duration of each treatment, number of treatments, % of motor threshold for stimulation) and mean weighted effect sizes.

A separate analysis on 10 studies including only parallel designs with the temporoparietal cortex as a focus for treatment revealed a mean weighted effect size of 0.45 (95% CI 0.24–0.67, $I^2 = 35.6$) for 459 patients, corresponding to a moderate effect size. A meta-analysis of all rTMS-foci (n = 17 studies) revealed a mean weighted effect size of 0.33 (95% CI 0.17–0.50) for 459 patients, corresponding to a small but significant effect size. $I^2$ was 12.9, indicating low heterogeneity. A separate analysis of the interview-based clinician-rated measures resulted in a mean weighted effect size of 0.49 (95% CI 0.32–0.67, $I^2 = 12.5$). For self-report measures the mean weighted effect size was 0.42 (95% CI 0.07–0.77, $I^2 = 46.6$). The difference between self-report- and interview-based clinician-rated measures was not significant ($t=1.87$, df 19, $p=0.85$). No difference in effect could be revealed between therapy-resistant AVH and non-therapy-resistant AVH ($t=0.67$, df 16, $p=0.85$).

A confound analysis did not show a difference in effect size for studies measuring changes in hallucinations (Hallucination Change Scale) instead of level of hallucinations (Auditory Hallucination Rating Scale, Psychotic Symptoms Rating Scales, Positive And Negative Syndrome Scale) (Pearson’s correlation 0.289, $p=0.32$). Furthermore,
no difference in effect size could be revealed between the different questionnaires (Pearson's correlation 0.360, p=0.19).

Five studies with 61 patients in the real TMS condition and 57 in the sham condition were included for the meta-analysis on long-term effects of rTMS applied at the left temporoparietal cortex one month after the end of treatment. The mean weighted effect size was 0.40 (95% CI \(-0.23\) to 1.02), \(I^2 = 63.6\). The results of this meta-analysis are presented in Fig. 2. The fail-safe number was 10 studies. No correlation between treatment paradigms and mean weighted effect sizes was found.

Side effects and reasons for dropouts for parallel designs are described in Tables 3 and 4. These data could be obtained from eight out of eleven studies. No significant differences could be found in the number of dropouts between real and sham-treatment (Fisher's Exact Test p=0.11).

A funnel plot of the effect sizes and years of publication is presented in Fig. 3. The effect sizes of studies targeting the left temporoparietal area that have been published before 2007 were significantly higher than those of more recently published studies (Pearson's correlation \(-0.658\), p=0.002).

4. Discussion

Four previous meta-analyses of repetitive transcranial magnetic stimulation (rTMS) in the treatment of auditory verbal hallucinations (AVH), revealed moderate to large effect sizes (Aleman et al., 2007; Tranulis et al., 2008; Freitas et al., 2009; Slotema et al., 2010). A new meta-analysis was conducted to evaluate the effect of rTMS directed at the left temporoparietal area versus sham in the treatment of AVH. Furthermore, the long term effect of rTMS was investigated by exploring the results one month after the cessation of rTMS treatment.

In the current meta-analysis the mean weighted effect size of studies that targeted the left temporoparietal cortex was moderate (\(g = 0.44\)). In an additional analysis a significant, but small mean weighted effect size of 0.33 was found in favor of all rTMS-foci. The results did not change if separate analyses of (types of) scales were performed nor when cross-over studies were excluded from the meta-analyses. The influence of rTMS on the severity of psychosis as assessed using the PANSS positive subscore was small but significant (\(g = 0.28\)).
The effect of rTMS on AVH was no longer significant at one month follow-up. This short duration of the effect of rTMS is a matter of concern; an intensive daily treatment of 2 to 4 weeks with a small treatment effect combined with lack of persistence, may call into question its utility as a meaningful treatment for patients troubled with persistent symptoms.

The results of this meta-analysis show a lower mean weighted effect size of 0.33 than previous meta-analyses, which yielded effect sizes between 0.52 and 1.0 (Aleman et al., 2007; Tranulis et al., 2008; Freitas et al., 2009; Slotema et al., 2010), but are in line with a letter to the editor that has recently been published (Demeulemeester et al., in press #278). This decline over the years may be due to an initial positive-outcome or publication bias. When new treatment strategies are introduced, initial reports tend to include relatively small sample sizes and provide favorable results, while small sampled studies with negative findings do not become published (Emerson et al., 2010).

Such trends have led effect sizes to decrease with year of publication (Munafo and Flint, 2010). As rTMS is a relatively new technique, future studies may show less positive results. The scatter plot of the randomized, sham-controlled studies investigating the effect of rTMS for AVH per year of publication, presented in Fig. 3, does not show evidence for a publication bias as small studies with negative results can be found in the early literature of rTMS as well. However, the effect sizes of studies that have been published more recently are significantly lower. Thus, there seems to be a decrease of the effect of rTMS over the years.

A matter of concern is that all included studies except Slotema et al. (2011) and Hoffman et al. (2005) are underpowered; with a mean weighted effect size of 0.5, minimum correlation 0.5, power 0.8 and alpha 0.05, 34 participants are needed per condition (Maxwell and Delaney, 2004). Therefore, future studies should direct at including large samples in a multicentre setting as was performed by O’Reardon et al. (2007) for depression.

4.1. Limitations

The majority of studies included patients with therapy-resistant AVH. The literature proposed medication-resistance to be associated with smaller effect sizes, but this could not be confirmed in two meta-analyses considering rTMS for depression (Kirsch et al., 2008; Schutter, 2009), neither in this meta-analysis.

Furthermore, the results of the meta-analysis considering the effect of rTMS one month after treatment were not homogenous and the number of studies was low. This underlines the need for additional, larger trials, to identify moderator variables. Patient characteristics (e.g. less chronic populations) and localization of the TMS coil (e.g. targeting language regions with fMRI-based neuronavigation) might be very relevant in this regard.

5. Conclusion

This meta-analysis, which could include 17 randomized, sham-controlled trials, found a small, yet significant mean weighted effect

---

**Table 3**

<table>
<thead>
<tr>
<th>Side effects</th>
<th>Real TMS (n=219)</th>
<th>Sham treatment (n=145)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headache</td>
<td>28</td>
<td>5</td>
</tr>
<tr>
<td>Twitching facial musculature</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>Dizziness</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>58</td>
<td>10</td>
</tr>
</tbody>
</table>

**Table 4**

<table>
<thead>
<tr>
<th>Reasons for dropout</th>
<th>Real TMS (n=219)</th>
<th>Sham treatment (n=145)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worsening psychotic symptoms</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Deterioration of mental state</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Side effects</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Lack of effect and inability to attend appointments</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Unknown</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total n (%)</td>
<td>10 (4.6)</td>
<td>13 (9.1)</td>
</tr>
</tbody>
</table>

Fig. 3. Scatter plot of mean weighted effect sizes of rTMS directed at the left temporoparietal cortex versus year of publication.
size of 0.33 for repetitive TMS as compared to sham-treatment for treatment-resistant AVH. One month after treatment, the difference between real and sham rTMS appeared to be no longer significant.

The mean weighted effect size of rTMS is decreasing due to the inclusion of studies with larger patient samples and negative findings. This may imply that early studies, with typically small samples, overestimated the clinical efficacy of rTMS. Alternatively, only subgroups of patients may benefit from rTMS, and the more recent studies may have had small proportions of those patients due to chance. Clearly, more research is needed in order to optimize parameters and further evaluate the clinical relevance of this intervention.

Role of funding source
Neither funding nor approval by a Regional Committee for Medical Research Ethics Health Region was needed.

Contributors
IE Sommer and C.W. Slotema designed the study and wrote the study protocol. C.W. Slotema managed literature searches and analyses. C.W. Slotema undertook the statistical analysis. C.W. Slotema wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflict of interest
All authors declare that they have no conflicts of interest.

Acknowledgments
We have no acknowledgments.

References