Durability of antidepressant response to repetitive transcranial magnetic stimulation: Systematic review and meta-analysis

Suhan Senova a, b, c, 1, Gonçalo Cotovio a, d, e, h, 1, Alvaro Pascual-Leone f, g, Albino J. Oliveira-Maia a, d, e, h, *

a Champalimaud Clinical Centre, Champalimaud Centre for the Unknown, Lisboa, Portugal
b Neurosurgery and Pneumology Departments, Assistance Publique-Hôpitaux de Paris (APHP), Groupe Henri-Mondor Albert-Chenevier, Créteil, France
c Equipe 14, U955 INSERM, Institut Mondor de Recherche Biomédicale and Faculté de Médecine, Université Paris Est, Créteil, France
d Department of Psychiatry and Mental Health, Centro Hospitalar de Lisboa Ocidental, Lisboa, Portugal
e NOVA Medical School | Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisboa, Portugal
f Berenson-Allen Center for Noninvasive Brain Stimulation, Division of Cognitive Neurology, Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA
g Instituto Guttman de Neurorehabilitación, Universitat Autonoma de Barcelona, Badalona, Barcelona, Spain
h Champalimaud Research, Champalimaud Centre for the Unknown, Lisboa, Portugal

A R T I C L E   I N F O

Article history:
Received 5 April 2018
Received in revised form 21 September 2018
Accepted 1 October 2018
Available online xxx

Keywords:
Depression
Repetitive transcranial magnetic stimulation
Durability
Relapse
Meta-analysis

A B S T R A C T

Background: The therapeutic options for treatment-resistant depression (TRD) encompass a range of neuromodulatory techniques, including repetitive transcranial magnetic stimulation (rTMS). While rTMS is safe and has documented short-term efficacy, durability of antidepressant effects is poorly established.

Objective: Assess existing evidence regarding durability of rTMS-induced antidepressant response.

Methods: We performed a systematic review of studies reporting antidepressant outcome measures collected three or more months after the end of an induction course of rTMS for depression. Among responders to the induction course, we used a meta-analytic approach to assess response rates at 3 (m3), 6 (m6) or 12 (m12) months after induction, and studied predictors of responder rates using meta-regression.

Results: Nineteen studies published between 2002 and 2018 were included. Eighteen were eligible for analysis at m3 (732 patients) and m6 (695 patients) and 9 at m12 (247 patients). Among initial responders, 66.5% sustained response at m3 (95% CI = 57.1–74.8%, I² = 27.6%), 52.9% at m6 (95% CI = 40.3–65%, I² = 0%), and 46.3% at m12 (95% CI = 32.6–60.7%, I² = 0%), in the absence of any major bias. Random-effects meta-regressions further demonstrated that a higher proportion of women, as well as receipt of maintenance treatment, predicted higher responder rates at specific time-points.

Conclusions: rTMS is a durable treatment for depression, with sustained responder rates of 50% up to 1 year after a successful induction course of treatment. Maintenance treatment may enhance the durability of the antidepressant effects of rTMS, and should be considered in clinical practice, as well as systematically explored in future clinical trials.

© 2018 Elsevier Inc. All rights reserved.

Introduction

Antidepressant medication and psychotherapy are first line treatments for patients suffering from major depression [1]. In case of insufficient benefit from, or intolerance to, medication, patients may be offered a number of neuromodulatory options, namely electroconvulsive therapy (ECT), repetitive transcranial magnetic stimulation (rTMS) and vagal nerve stimulation (VNS). Transcranial direct current stimulation (tDCS) and deep brain stimulation (DBS) are still investigational [2–6]. Choosing among these options requires careful weighing of their characteristics, namely durability of antidepressant effects, of particular importance since these patients typically suffer from multiple depressive episodes across their lifetime [7,8]. However, durability of antidepressant effects of rTMS has been insufficiently examined to date.

Efficacy of dorsolateral prefrontal cortex (DLPFC) rTMS for major depression has been documented by a large number of studies, including multicenter trials, and analyzed in several meta-analyses (e.g. Refs. [4,9–12]). Efficacy has been demonstrated for both
medication-free and medicated patients [13–17], and the United States Food and Drug Administration (FDA) has formally cleared four different devices for rTMS treatment of patients with prior treatment-failure for depression. The protocol leading to initial FDA clearance was based on the application of daily treatment sessions, 5 days a week, for up to six weeks, with stimulation delivered over the left hemisphere at high frequency (typically 10 Hz or higher) and high intensity (120% of the resting motor threshold) [18,19]. There is less evidence to support rTMS protocols using lower stimulation frequency or stimulation intensities [20].

Durability of antidepressant effects after rTMS induction has been assessed on short timescales, typically no longer than 3 months after treatment [14,21–23]. A previous meta-analysis including randomized studies with follow-up phases ranging from 2 to 16 weeks, none of which used rTMS maintenance treatment, concluded that rTMS has a small antidepressant effect during follow-up [22]. Meta-analyses taking into consideration longer follow-up phases, and considering options for maintenance treatment, have not been reported, probably because there are very few randomized controlled trials (RCT) with long-term follow-up. Indeed, in RCT’s for major depressive disorder (MDD), sustained blinding for long periods after the end of treatment raises difficult logistic and ethical questions. However, considering the high relapse rate for major depressive disorder [7], understanding treatment effects for only three months is not sufficiently informative. This is particularly important when considering the emergence of chronic neurostimulation treatment options for treatment resistant depression (TRD), such as VNS, DBS [5,24,25] and epidural prefrontal cortical stimulation [26], that offer the potential for prolonged efficacy.

Nevertheless, several open-label rTMS studies have retained patients for very long follow-up phases, of up to 6 years [27–29]. A meta-analysis including such studies, similar to what has been performed for electroconvulsive therapy [30], could thus begin to provide quantitative responses for the question of long-term durability of the antidepressant effects of rTMS [31], as well as explore the best alternatives for relapse prevention after successful rTMS [8,32]. Thus, the primary objective of this meta-analysis was to provide a systematic overview of the existing evidence regarding post-rTMS durability of response for patients followed for at least 3 months after acute treatment. The secondary goal was to evaluate the impact of maintenance rTMS treatment on durability of antidepressant effects, since this has been proposed as a potential alternative for relapse prevention, but its efficacy is not clearly established [33]. Considering both the previously mentioned meta-analysis for short-term follow-up, and trials including a longer follow-up period, we hypothesized that the efficacy of acute rTMS decreases over time, but is prolonged by rTMS maintenance treatment.

Material & methods

This systematic review and meta-analysis was performed according to PRISMA Guidelines [34].

Search strategy

An electronic literature search in PubMed and Web of Science was performed, including papers published up to June 2018. Mesh terms used are described in supplementary table 1. Two authors (SS and GC) independently eliminated ineligible reports in sequential steps, judging first by title and abstract review, followed by full-text screening. Hand-searches of the reference sections of prior reviews, as well as of previously identified eligible studies, were carried out to identify additional eligible studies. Inconsistencies in eligibility assessments were resolved by consensus, when necessary with participation of a third author. Quality assessment for each study was performed with the Newcastle-Ottawa Quality Assessment Scale for Cohort Studies [35].

Study eligibility criteria

Inclusion criteria

1. prospective or retrospective study reported in a peer-reviewed publication;
2. use of rTMS induction for treatment of a depressive episode (unipolar or bipolar) diagnosed by clinical judgment or formal diagnostic criteria (e.g., DSM-IV);
3. treatment response defined consistently in each study through clinical judgment, use of formal diagnostic criteria and/or clinician-rated depression severity rating scales (Hamilton Depression Rating Scale — HAM-D; Cognitive Global Impression Scale — CGI);
4. patients considered to be rTMS responders or remitters after rTMS induction were monitored at one or more time points, at least 3 months after the end of rTMS induction;
5. responder rates at each time-point were defined by the investigators of the original study and reported as the percentage of the initial responder or remitter sample, or could be directly extracted from the data in the manuscript.

Exclusion criteria

1. case studies or case series with less than 4 patients;
2. inclusion of non-affective psychosis, dementia, neurological disease, alcoholism or unstable medical conditions among the patient sample;
3. inclusion of patients younger than 18 years;
4. review or meta-analysis not reporting original data.

Data extraction

Data were extracted from each selected manuscript, using the text, tables, and/or figures, independently by three authors of this meta-analysis. Discrepancies were resolved through consensus, prior to data analysis. The primary outcome was the responder rate and the endpoints were three months after rTMS induction treatment: m3 (three to four months), m6 (five to six months) and m12 (ten to twelve months). Response was defined in each study according to specific criterion chosen by the authors, and expressed at each endpoint relatively to the number of patients who responded to induction rTMS and entered the follow-up phase. An intention-to-treat analysis approach was chosen, such that patients lost to follow-up were considered as non-responders.

Additional data were extracted to assess potential predictors of responder rates. Age and gender distributions were extracted as potential demographic predictors of response. Clinical variables of interest were depression severity prior to rTMS induction, duration of the affective disorder, number of previous episodes, duration of the current episode and percentage of patients with bipolar depression. Inclusion of patients with TRD was also extracted as a clinical variable. Since TRD is defined variably according to different criteria [36], for the purposes of this metanalysis we specified TRD according to stage II from the Thase & Rush criteria [37]. Parameters of the rTMS induction treatment were also extracted: intensity, rTMS frequency, number of pulses per treatment and total number of rTMS treatments. Other treatment-related variables were the study-defined criterion of response to rTMS induction, percentage of patients receiving left DLPFC rTMS, and delivery of maintenance treatment.

Please cite this article in press as: Senova S, et al., Durability of antidepressant response to repetitive transcranial magnetic stimulation: Systematic review and meta-analysis, Brain Stimulation (2018), https://doi.org/10.1016/j.brs.2018.10.001
rTMS treatment. For this study, maintenance rTMS was defined as any rTMS session delivered after the induction cycle.

Meta-analyses

Statistical analyses were performed using Comprehensive Meta Analysis Version 3. Due to differences in study designs and patient populations, mean responder rates with 95% confidence intervals (CIs) were calculated with a random-effects model. I² statistic was performed to assess heterogeneity. Univariate meta-regression analyses were performed to identify predictors of responder rates. Subgroup analyses were performed at m3, m6 and m12, to compute mean responder rates with 95% CIs according to maintenance treatment, namely among studies exclusively with, versus exclusively without maintenance treatment.

Publication bias analysis

Publication bias was assessed by visual inspection of funnel plots of standard error vs. logit of responder rates, whenever more than 10 studies were available. Duval and Tweedie’s Trim-and-Fill analysis was used to test if the funnel plot was symmetrical around the overall mean weighted responder rates of all studies. To test if studies with lower effect sizes differed systematically from studies with higher effect sizes we used the Begg and Mazumdar Rank Order Correlation (Kendall’s tau b) between the standardized effect sizes and the standard error of the mean (SEM) in each study, as well as the Egger’s regression of 1/SEM (predictor) on the standardized effect sizes.

Results

Search results

In initial literature searches, 207 articles were found and screened for eligibility, 30 of which were chosen for full-text inspection after title and abstract review. Among the references quoted in these articles, 32 more were selected for full-text inspection. Finally, 23 articles published between 2002 and 2018 were included in the systematic review (Fig. 1): 3 RCT [33,38,39], 14 open-label prospective studies [28,40–50,52,53] and 2 retrospective studies [54,55]. For one RCT [39] only the open label extension could be included.

Four studies [51,56–58] were not considered for quantitative review since response and relapse rates for responders to initial course of rTMS could not be extracted. Among the remaining 19 studies included in quantitative review, quality was fair to good (Supplementary Table 2). Four studies included only MDD patients with TRD, defined as non-responders to two or more courses of adequate pharmacotherapy [39,40,49,53], while 9 studies included MDD patients both with and without TRD [28,33,41,43,44,46,47,50,54]. The 6 remaining studies included patients with bipolar depression [38,42,45,48,52,55], two of which including only patients with TRD [38,52]. The criterion to define responders was also variable, as is further detailed in Table 1 and Supplementary Table 3.

One of the central objectives of the quantitative synthesis was regarding the effects of maintenance treatment. Thus, one of the open-label prospective studies [52] and one RCT [38], reporting data separately for maintenance and non-maintenance groups, were split accordingly, and considered as two distinct studies for meta-analyses. Among the studies delivering rTMS maintenance treatments, the design of the maintenance protocols was highly variable. In 7 studies, rTMS maintenance was administered according to planned treatment schemes, ranging from weekly sessions [48] to clusters of 5 sessions administered in two and one-half days every month [45], as well as extended taper [38,40,46,52,55] (Supplementary Table 4). In three studies, rTMS maintenance was administered as ‘rescue’ treatment, after evidence of symptom worsening [28,43,47]. Finally, in one RCT [33], while all patients were offered symptomatic re-introduction of rTMS, a subgroup of patients was randomized to receive additional planned maintenance, with a once-monthly session. Since this frequency of rTMS maintenance is well below the standard in other studies (Supplementary Table 4), both groups were considered together and as receiving symptomatic re-introduction of rTMS.

Responder rates at m3, m6 and m12

After a successful index rTMS induction course, among 732 patients from 18 studies, 66.5% (95% CI = 57.1–74.8%, I² = 27.6%) were still responders at m3 (Fig. 2a). At m6, among 695 patients from 18 studies, 52.9% (95% CI = 40.3–65%, I² = 0%) were still responders (Fig. 2b). At m12, among 247 patients from 9 studies, 46.3% (95% CI = 32.6–60.7%, I² = 0%) were still responders (Fig. 2c). While, at m3 and m6, cumulative analyses suggested a trend for responder rates to be higher for older studies (Supplementary Figures 1a and 2a), one-study removed analyses showed the final results were not dependent on any particular study at any time-point (Supplementary Figure 1b, 2b and 3b). Heterogeneity of the studies was low to moderate for all time-points.

Several sensitivity analyses were conducted to verify the nature of these findings when excluding studies according to several different criteria. One study [33] was identified as an outlier with a higher risk of bias due to presenting the highest loss to follow-up at all time points (>33.3% at m3, >50% at m6 and >66.6% at m12; Supplementary Table 5), as well as the lowest response criterion to induction rTMS (25%; Table 1). Another study was considered as potentially problematic since responders were defined according to clinical judgment only, without referring to any clinical score of depression [49]. We thus repeated the meta-analysis after excluding these trials (Supplementary Figure 6b). Response criteria used by the authors to include responders in the follow-up phase could also influence the results. In order to address this potential bias, we performed an additional sensitivity analysis using only studies defining response as at least 50% reduction in the HAMD score (Supplementary Figure 6c). Furthermore, and despite the fact that the percentage of patients with bipolar disorder in each study was not found to be a predictor of response rate at any time point (see section on predictors of responder rates), we conducted sensitivity analyses considering studies that included only patients with major depressive episode (Supplementary Figure 6d) or studies that included only patients with bipolar disorder (Supplementary Figure 6e). Finally, sensitivity analyses were conducted considering only patients with (Supplementary Figure 6f) or without (Supplementary Figure 6g) chronic depression, defined according to the duration of the current depressive episode. When comparing with the full dataset (Supplementary Figure 6a) these several sensitivity analyses rendered similar results.

Estimation of bias

While at m3 and m6 we gathered information from more than 10 studies, allowing for assessments of bias, at m12 only 9 studies were found and assessment of bias was not possible. For both m3 and m6 the funnel plot was symmetrical at visual inspection (Fig. 3a and b), the Begg and Mazumdar’s tests and the Egger’s tests were not significant (p > 0.4 for both tests), and there was no imputed study after the Duval and Tweedie’s Trim and Fill analysis. Thus, overall, our meta-analysis did not suffer from obvious publication bias.
Predictors of responder rates at m3 and m6

Univariate meta-regressions with random-effects analyses were performed to assess potential predictors of responder rates at m3 and m6 (Table 2, Supplementary Figures 4 and 5). Such predictors were not assessed at m12, due to the low number of studies at this time-point. The percentage of women included in each study was a significant positive predictor of being a responder at m3 ($p = 0.009$, effect size: $0.04$, 95% CI $0.01$–$0.07$, Z-value: 2.6) and m6 ($p = 0.02$, effect size: $0.04$, 95% CI $0.007$–$0.07$, Z-value: 2.37). There was a trend for lower response criterion to induction rTMS to be a predictor of worse response rate both at m3 ($p = 0.07$) and m6 (0.05), and for higher quality of studies to be a predictor of worse response rate at m3 ($p = 0.05$). Finally, maintenance treatment was a significant positive predictor of responder rates at m3 ($p = 0.03$, effect size: $0.91$, 95% CI $0.088$–$1.74$, Z-value: 2.17) and m6 ($p = 0.02$, effect size: $0.95$, 95% CI $0.12$–$1.79$, Z-value: 2.24). The remaining variables, including rTMS protocol parameters in the acute treatment such as stimulation location (Left DL-PFC), % of motor threshold (MT), pulses/day, frequency and induction days, were either not significant predictors of responder rates at m3 and m6, or there was not enough data to assess their effects using meta-regression (Table 2).

Subgroup analyses

Since the question of maintenance treatment is of crucial clinical relevance, we performed subgroup analyses to compute mean responder rates separately for patients receiving or not receiving maintenance rTMS at m3 and m6. At m3, responder rate was 35.8% higher for the 383 patients from 8 studies receiving maintenance (76.2%, 95% CI $63$–$85.8$, $I^2 = 0$%) than for the 349 patients from 8 studies that did not receive maintenance (56.1%, 95% CI $43.6$–$67.8$, $I^2 = 38.7$%), while at m6, responder rate was 58.7% higher for the 417 patients receiving maintenance in 10 studies (61.1%, 95% CI $49.8$–$71.3$, $I^2 = 0$%) than for the 278 patients from 8 studies that did not receive maintenance (38.5%, 95% CI $21.9$–$58.3$, $I^2 = 0$%). Analyses comparing planned and symptomatic re-introduction of rTMS schemes did not demonstrate any significant differences between the two approaches (Supplementary Table 4).

Discussion

In this systematic review and meta-analysis, we describe relevant findings with regards to durability of the antidepressant effects of rTMS treatment, among patients who respond to an acute course of rTMS. Indeed, 66.5% of these patients sustained response...
# Table 1

Demographic, clinical and methodological data extracted from each study included in the quantitative review.

<table>
<thead>
<tr>
<th>Study</th>
<th>Demographic</th>
<th>Diagnosis</th>
<th>Disease Characteristics (Baseline)</th>
<th>rTMS Parameters</th>
<th>Response Criteria</th>
<th>Maintenance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First Author</td>
<td>Year</td>
<td>Type</td>
<td>n</td>
<td>Age (y)</td>
<td>Fem. (%)</td>
</tr>
<tr>
<td>Dannon [41]</td>
<td>2002</td>
<td>Pros.</td>
<td>23</td>
<td>56.8</td>
<td>66.7</td>
<td>0</td>
</tr>
<tr>
<td>Boutros [39]</td>
<td>2002</td>
<td>RCT</td>
<td>5</td>
<td>48.8</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>Li [48]</td>
<td>2004</td>
<td>Pros.</td>
<td>7</td>
<td>44.9</td>
<td>71.4</td>
<td>100</td>
</tr>
<tr>
<td>Benadhira [40]</td>
<td>2005</td>
<td>Pros.</td>
<td>8</td>
<td>0</td>
<td>Yes</td>
<td>100</td>
</tr>
<tr>
<td>O'Reardon [28]</td>
<td>2005</td>
<td>Pros.</td>
<td>10</td>
<td>49.6</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>Eranti [44]</td>
<td>2007</td>
<td>Pros.</td>
<td>4</td>
<td>0</td>
<td>No</td>
<td>22</td>
</tr>
<tr>
<td>Cohen [54]</td>
<td>2009</td>
<td>Retr.</td>
<td>204</td>
<td>43.3</td>
<td>51</td>
<td>0</td>
</tr>
<tr>
<td>López-Ibor [49]</td>
<td>2010</td>
<td>Pros.</td>
<td>26</td>
<td>0</td>
<td>Yes</td>
<td>29.0</td>
</tr>
<tr>
<td>Janicak [47]</td>
<td>2011</td>
<td>Pros.</td>
<td>8</td>
<td>46.6</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>Benadhira (M) [38]</td>
<td>2017</td>
<td>RCT</td>
<td>10</td>
<td>48.6</td>
<td>61.2</td>
<td>0</td>
</tr>
<tr>
<td>Benadhira (NM) [38]</td>
<td>2017</td>
<td>RCT</td>
<td>7</td>
<td>56</td>
<td>83.3</td>
<td>n.c.</td>
</tr>
</tbody>
</table>

BD: bipolar disorder; CGI: Clinician-Reported Clinical Global Impressions.; CGI-S: Clinician-Reported Clinical Global Impressions-Severity of Illness scale; Dis. Dur.: disease duration; Eps. Dur.: episode duration; Fem. (%): % of women; HAMD: Hamilton depression score; M: maintenance; n: number of included patients; n.c.: not computable; n.: not determined; NM: non maintenance; Pl.: planned maintenance; Prior Eps.: prior episodes; Pros.: prospective studies; RCT: randomized controlled studies; Red.: reduction; Retr.: retrospective studies; Sy: symptoms; Sym.: symptomatic re-introduction of rTMS; TRD: treatment resistant depression; y: year. Study type classifications were performed according to previous definitions [59,60].

a Despite the absence of predefined response criterion for the inclusion in the follow-up phase, all the patients had a >50% reduction from the baseline HAMD score when entering the maintenance phase according to the table of results displayed in the article.

b One of the randomized studies comparing two groups of patients was [33]. However, in this study, patients of both groups received rTMS maintenance as needed (based on clinical observation), while scheduled rTMS maintenance was the randomized intervention. For this reason, both groups were recipients of rTMS maintenance, as defined for this manuscript, and we thus considered the data from the two groups jointly, and classified [33] as a prospective study. While [36] also describes results from a RCT of acute rTMS, the data included in this meta-analysis are the results from an open-label extension of the active arm of the RCT.

c Bipolar patients were included but their proportion could not be determined.

Please cite this article in press as: Senova S, et al., Durability of antidepressant response to repetitive transcranial magnetic stimulation: a systematic review and meta-analysis, Brain Stimulation (2018), https://doi.org/10.1016/j.brs.2018.10.001.
after 3 months, with responder rates decreasing progressively across time, and 46.3% maintaining response 1 year after induction treatment. These results expand the conclusions drawn from a previous meta-analysis [22] with shorter follow-up periods, of only up to 16 weeks. Here we show that the benefit of rTMS might last up to 1 year in close to half of the patients who respond to acute treatment. Furthermore, we found that studies including more women had higher responder rates at 3 and 6 months, in confirmation of the findings described by Kedzior and colleagues [10]. Nevertheless, since results were not reported separately according to gender, we cannot conclude that women have a better outcome than men, and suggest that future studies report outcomes by gender so this finding could be better explored. Importantly, rTMS maintenance treatment, when applied to those responding to acute rTMS treatment, was associated with higher sustained responder rates after 3 months, and especially after 6 months, which were respectively 35.8% and 58.7% greater than those reported in the absence of rTMS maintenance. This could reflect a more important

Fig. 2. Meta-analyses. Meta-analyses of Responder rates at 3 months follow-up (m3, panel a), 6 months follow-up (m6, panel b) and 12 months follow-up (m12, panel c), among patients benefitting from an acute course of rTMS. The analyses show that there are sustained responder rates to rTMS, of up to 67%, from 3 months to 1 year after an initial treatment course. CI: confidence interval; M: study arm including only patients with maintenance rTMS; NM: study arm including only patients without maintenance rTMS.
decline of responder rates in patients who do not receive rTMS maintenance treatment, for whom response rate was 31.4% lower at m6 relative to m3, than in those who did receive rTMS maintenance, where response rate was only 19.8% lower at m6 relative to m3. These findings, which have not been explored in previous meta-analyses, suggest that rTMS maintenance should be systematically considered for patients who respond to an initial treatment course of rTMS for depression, to enhance the chances to maintain the benefits of rTMS for a longer period.

While this meta-analysis suggests that there is value in maintenance rTMS, it provides limited insight regarding the protocol that should be used, specifically the duration and frequency of treatments, as well as optimal stimulation parameters. Among the 11 studies in which rTMS maintenance was offered, 4 described protocols including symptomatic re-introduction of rTMS schedules [28,33,43,47], whereas the remaining 7 were performed with planned maintenance schedules [38,40,45,46,48,52,55]. The available data did not suggest superiority for any of the two types of approach. Furthermore, the design of planned strategies varied significantly across studies, regarding both frequency, ranging from twice-weekly [57] to monthly [55], as well as duration, ranging from 3 months [40] to up to 2 years [45]. The methods for

Table 2
Effect sizes with 95% CI and p-values for moderators of response rates at 3 months (m3) and 6 months (m6) follow-up, calculated using univariate meta-regression.

<table>
<thead>
<tr>
<th>All Studies</th>
<th>M3</th>
<th>95% CI</th>
<th>p-value</th>
<th>M6</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintenance rTMS</td>
<td>0.91</td>
<td>0.088–1.74</td>
<td>0.03</td>
<td>0.96</td>
<td>0.12–1.79</td>
<td>0.03</td>
</tr>
<tr>
<td>% of Patients with Bipolar Depression</td>
<td>0.75</td>
<td>–1.75–3.24</td>
<td>0.56</td>
<td>0.60</td>
<td>–1.28–2.49</td>
<td>0.53</td>
</tr>
<tr>
<td>Response Criterion</td>
<td>0.04</td>
<td>–0.004–0.09</td>
<td>0.07</td>
<td>0.04</td>
<td>0.0006–0.07</td>
<td>0.05</td>
</tr>
<tr>
<td>Age</td>
<td>–0.03</td>
<td>–0.15–0.09</td>
<td>0.61</td>
<td>0.03</td>
<td>–0.10–0.15</td>
<td>0.7</td>
</tr>
<tr>
<td>% of Women</td>
<td>0.04</td>
<td>0.01–0.07</td>
<td>0.009</td>
<td>0.04</td>
<td>0.007–0.08</td>
<td>0.02</td>
</tr>
<tr>
<td>Baseline Clinical Score</td>
<td>–0.05</td>
<td>–0.18</td>
<td>0.46</td>
<td>–0.004</td>
<td>–0.14–0.13</td>
<td>0.95</td>
</tr>
<tr>
<td>Episode Duration</td>
<td>n.e.d.</td>
<td>n.e.d.</td>
<td>n.e.d.</td>
<td>n.e.d.</td>
<td>n.e.d.</td>
<td>n.e.d.</td>
</tr>
<tr>
<td>% of Patients with Left DLPFC Stimulation</td>
<td>0.0009</td>
<td>–0.01–0.02</td>
<td>0.9</td>
<td>0.004</td>
<td>–0.01–0.01</td>
<td>0.59</td>
</tr>
<tr>
<td>Motor Threshold</td>
<td>–0.0009</td>
<td>–0.04–0.02</td>
<td>0.61</td>
<td>–0.007</td>
<td>–0.05–0.04</td>
<td>0.76</td>
</tr>
<tr>
<td>Number of Pulses/day</td>
<td>–0.0001</td>
<td>–0.001–0.0009</td>
<td>0.85</td>
<td>–0.0001</td>
<td>–0.0007–0.0005</td>
<td>0.77</td>
</tr>
<tr>
<td>Frequency of Stimulation</td>
<td>–0.05</td>
<td>–0.17–0.08</td>
<td>0.44</td>
<td>–0.60</td>
<td>–0.17–0.04</td>
<td>0.25</td>
</tr>
<tr>
<td>Number of Induction Days</td>
<td>–0.04</td>
<td>–0.15–0.07</td>
<td>0.46</td>
<td>–0.05</td>
<td>–0.15–0.05</td>
<td>0.34</td>
</tr>
<tr>
<td>Quality of the Study</td>
<td>–0.27</td>
<td>–0.54–0.002</td>
<td>0.05</td>
<td>0.027</td>
<td>–0.59–0.05</td>
<td>0.1</td>
</tr>
</tbody>
</table>

DL-PFC: dorsolateral prefrontal cortex; rTMS: repeated transcranial magnetic stimulation; y: year.
n.e.d.: not enough data were gathered to run metaregressions on episode duration, disease duration and on number of previous episodes of depression.

Please cite this article in press as: Senova S, et al., Durability of antidepressant response to repetitive transcranial magnetic stimulation: Systematic review and meta-analysis, Brain Stimulation (2018), https://doi.org/10.1016/j.brs.2018.10.001
symptomatic re-introduction of rTMS were also not fully described, namely with regards to the exact criteria for re-introduction of rTMS. Ultimately, there was too much variability, and insufficient replication, to allow for clear meta-analytic comparisons. Thus, a definitive maintenance strategy or treatment schedule cannot be fully defined at this moment. Nevertheless, we believe these results highlight the need for controlled studies of symptomatic re-introduction and/or planned maintenance strategies for maintenance rTMS, that could provide further support not only for rTMS as a viable long-term treatment regimen for depressive disorders, but also help define which strategies should be used.

Our results are in line with what has been reported for durability of other neurostimulation approaches to treat MDD or TRD [61]. A meta-analysis assessing the relapse rate after a successful course of ECT in MMD patients showed that close to 40% of patients, both with and without maintenance ECT, relapsed after 6 months [30]. For tDCS 2 RCT’s have assessed the long-term effectiveness of tDCS and both reported a relapse rate of approximately 50% at m6 follow-ups [2]. Regarding invasive neuromodulation, where stimulation is typically delivered chronically, VNS is associated with a 52% response rate after 6 months [62], while DBS of the ventral capsule/ventral striatum is associated with response rates of 43.7% at m3 and 40% at m6 follow-up [63]. Consistently, in a 2014 meta-analysis, Berlim and colleagues found that the 12-month response rate following DBS treatment was 39.9%. While these numbers are similar to those reported here, it is important to underline that patient populations in the several studies are not necessarily equivalent. Importantly, the patients included in our meta-analysis were not exclusively TRD patients, while all patients offered VNS or DBS suffer from TRD. Moreover for VNS or DBS studies response rates were computed for all patients included at the beginning of the studies, whereas in our and other meta-analyses [30] response rates during follow-up were computed for responders to an acute course of treatment.

In the assessment of the included publications, we did not find evidence of relevant publication bias, supporting the validity of our results. However, the data available for this study suffers from other potential limitations. First and foremost, while our inclusion and exclusion criteria were not overly restrictive, the number of studies with long enough follow-up phases was low. Furthermore, some did not report responder or relapse rates, and thus could not be considered for meta-analyses. Importantly, most studies reported data for only one or two time-points, such that the data reported for each time-point of the meta-analysis does not represent the same group of studies. Heterogeneity of study methodology or study population is also of concern. However, sensitivity analyses removing outlier studies or studies with atypical antidepressant response criteria (i.e., anything other than HAMD ≥ 50%), as well as comparing studies including specific patient subgroups, essentially confirmed the findings in the full sample (supplementary Fig 6).

Importantly, we could not draw any strong conclusion regarding the durability of response to rTMS for MDD versus bipolar patients (supplementary Fig. 6d and 6e) nor for patients with versus without chronic depression (supplementary Fig. 6f and 6g). Future studies systematically and directly comparing these subgroups of patients are required to address these questions. Nevertheless, and while this study does not provide level 1 evidence given limitations in the available literature, it provides clear directions for additional well designed rTMS studies, for well-defined groups of patients, with at least 3 months’ post-treatment follow-up, and sound reporting of outcome measures accordingly to common reported criteria, at standard time-points.

Another important limitation results from the fact that our analyses included almost exclusively open label and naturalistic studies. We believe this reflects important ethical and logistic challenges in the conduction of long-term randomized and sham-controlled studies among depressed patients. In fact, one of the studies used in the meta-analyses described a randomized protocol, testing the efficacy of a planned rTMS maintenance protocol to prevent relapse [33], offered in addition to symptomatic re-introduction of rTMS that was available for all patients, and in the absence of a sham stimulation intervention for the control arm of the study. This limited the use of this study design to address the relevance of maintenance treatment, given that both study arms effectively received maintenance. Thus, we remained limited in the ability of extracting causal relationships from the data described here, since there is insufficient control for potential confounders. For example, it is possible that, during follow-up/maintenance phases, new antidepressants could be initiated, which could contribute towards the overall duration of the antidepressant response to induction rTMS and, albeit to a lesser degree, towards the advantages associated with maintenance rTMS. For these reasons, we strongly believe that studies with a RCT design are necessary to confirm the findings reported here, most importantly, to confirm the utility and define the optimal strategy for rTMS maintenance treatment. We expect that the findings reported here could help refine the design of future clinical trials in this field.

In conclusion, a systematic review and meta-analysis of the available literature suggests that rTMS is a durable treatment for depression, with sustained responder rates of 46–67%, from 3 months to 1 year after a successful initial treatment course. Of critical importance for current rTMS practice in depression treatment, maintenance rTMS delivered to those that respond to induction rTMS was found to enhance the rates of sustained responders across time-points, mainly 6 months after rTMS induction, when responder rates were almost 60% higher than among patients that did not receive maintenance rTMS.

Funding

SS was funded by Assistance Publique des Hôpitaux de Paris. APL was supported in part by the Berenson-Allen Foundation, the Sidney R. Baer, Jr. Foundation, the National Institutes of Health (R21 MH099196), and Harvard Catalyst | The Harvard Clinical and Translational Science Center (NCRR and the NCATS NIH, UL1 RR025758). Fundação para a Ciência e a Tecnologia funded GC through a Doctoral Scholarship (SFRH/BD/130210/2017), and AJO-M through a Junior Research and Career Development Award from the Harvard Medical School - Portugal Program (HMSP-IC/0020/2011). The content of this study is solely the responsibility of the authors and does not necessarily represent the official views of the Fundação para a Ciência e Tecnologia, Harvard Catalyst, Harvard University and its affiliated academic health care centers, the National Institutes of Health, or the Sidney R. Baer Jr. Foundation.

Declarations of interest

APL serves on the scientific advisory boards for Nexstim, Neuronix, Starlab Neuroscience, Neuroelectrics, Axilum Robotics, and Neosync; and is listed as an inventor on several issued and pending patents on the real-time integration of transcranial magnetic stimulation with electroencephalography and magnetic resonance imaging. SS, GC and AJO-M report no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brs.2018.10.001.


