Clinical trial design in non-invasive brain stimulation psychiatric research

ANDRÉ RUSSOWSKY BRUNONI[1] & FELIPE FREGNI[2,3]

1 Hospital Universitário, University of São Paulo, São Paulo, Brazil
2 Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA
3 Laboratory of Neuromodulation, Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital and Massachusetts General Hospital, Harvard Medical School, Boston, MA

Abstract

Major depressive disorder (MDD) trials – investigating either non-pharmacological or pharmacological interventions – have shown mixed results. Many reasons explain this heterogeneity, but one that stands out is the trial design due to specific challenges in the field. We aimed therefore to review the methodology of non-invasive brain stimulation (NIBS) trials and provide a framework to improve clinical trial design. We performed a systematic review for randomized, controlled MDD trials whose intervention was transcranial magnetic stimulation (rTMS) or transcranial direct current stimulation (tDCS) in MEDLINE and other databases from April 2002 to April 2008. We created an unstructured checklist based on CONSORT guidelines to extract items such as power analysis, sham method, blinding assessment, allocation concealment, operational criteria used for MDD, definition of refractory depression and primary study hypotheses. Thirty-one studies were included. We found that the main methodological issues can be divided in to three groups: (1) issues related to phase II/small trials, (2) issues related to MDD trials and, (3) specific issues of NIBS studies. Taken together, they can threaten study validity and lead to inconclusive results. Feasible solutions include: estimating the sample size a priori; measuring the degree of refractoriness of the subjects; specifying the primary hypothesis and statistical tests; controlling predictor variables through stratification randomization methods or using strict eligibility criteria; adjusting the study design to the target population; using adaptive designs and exploring NIBS efficacy employing biological markers. In conclusion, our study summarizes the main methodological issues of NIBS trials and proposes a number of alternatives to manage them. *Copyright © 2011 John Wiley & Sons, Ltd.*

Introduction

Despite the significant neuropsychopharmacology advancement in the past decades, it is still needed to develop new somatic interventions for psychiatric disorders

(Baghai *et al.*, 2006; Marder, 2006). Non-invasive brain stimulation (NIBS) is a novel therapeutic intervention facing rapid development, with preliminary studies showing significant clinical gains in several neurologic and psychiatric conditions such as major depressive

disorder (MDD) (Gross *et al.*, 2007), schizophrenia (Freitas *et al.*, 2009), stroke (Fregni and Pascual-Leone, 2007), and Parkinson's disease (Elahi and Chen, 2009). We use the term "NIBS" here to refer to devices applied to the scalp in order to induce electrical currents in specific brain areas. The two main examples are transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS).

Besides the initial encouraging results from NIBS clinical studies, it is unclear whether these effects are related to specific properties of the technique as two MDD meta-analyses studies (Herrmann and Ebmeier, 2006; Schutter, 2008) failed to identify predictors of response, meaning that such effects might not be reproducible in different contexts. Such mixed results might also be explained by study design issues, such as small sample sizes, lack of generalizability and insufficient blinding. This might also be true for several antidepressant studies that may fail due to methodological problems (Gelenberg *et al.*, 2008).

Therefore, a critical overview of the methodology of NIBS trials might improve its clinical development. We aimed to review their methodology as to discuss potential pitfalls and implications for designing future trials. We focused our review on repetitive transcranial magnetic stimulation (rTMS) (a device that generates a strong magnetic field that induces an electric current to modulate a cortical focal brain area) and tDCS (a device that polarizes cortical tissue and thereby modifies neuronal excitability) trials for MDD which has a greater number of studies – although most issues can be applied for other neuropsychiatric disorders.

Methods

We performed a systematic review on all tDCS and rTMS trials published from April 2002 to April 2008. Although vagal nerve stimulation (VNS) is a neuromodulatory device used for MDD treatment, it was not included because we reckon this technique as minimally invasive (as it envolves surgery) and also because the study design methodology of VNS has been already reviewed in the past years when validating VNS intervention to epilepsy (Groves and Brown, 2005). We chose to look for articles after 2002 as Martin *et al.*'s meta-analysis (2003) reviewed previous trials. The detailed literature search and selection criteria are described elsewhere (Brunoni *et al.*, 2009); in summary, we searched for the keywords "depression", "transcranial magnetic stimulation" and "sham" on MEDLINE, Web of Science, Cochrane and Scielo as to include only randomized, double-blinded,

sham-controlled, parallel trials – 29 rTMS studies were found (Anderson *et al.*, 2007; Avery *et al.*, 2006; Bortolomasi *et al.*, 2007; Boutros *et al.*, 2002; Bretlau *et al.*, 2008; Fitzgerald *et al.*, 2006; Fitzgerald *et al.*, 2003; Garcia-Toro *et al.*, 2006; Hausmann *et al.*, 2004; Herwig *et al.*, 2007; Herwig *et al.*, 2003; Holtzheimer *et al.*, 2004; Hoppner *et al.*, 2003; Januel *et al.*, 2006; Jorge *et al.*, 2008; Jorge *et al.*, 2004; Koerselman *et al.*, 2004; Loo *et al.*, 2003; Loo *et al.*, 2007; Mogg *et al.*, 2008; Mosimann *et al.*, 2004; O'Reardon *et al.*, 2007; Poulet *et al.*, 2004; Rossini *et al.*, 2005a, 2005b; Rumi *et al.*, 2005; Stern *et al.*, 2007; Su *et al.*, 2005). The same criteria were used for tDCS, identifying two additional studies (Boggio *et al.*, 2008; Fregni *et al.*, 2006a), for a total of 31.

We then examined each trial for potential methodological issues threatening study validity. We constructed an unstructured checklist based on the CONSORT statement and other methodological reviews analyzing pitfalls in clinical trials (Boutron *et al.*, 2007; Glasser and Howard, 2006; Hopewell *et al.*, 2008; Leucht *et al.*, 2008; Zlowodzki *et al.*, 2006). Four main aspects were addressed:

(i) Statistical validity – we checked whether the studies were vulnerable to type I and type II errors by looking for their power analysis as well as for the adequacy of statistical tests and the primary hypothesis (if it was clearly defined and if it was based on response or score changes). We also performed a power analysis (using Stata 10® software) in negative studies.

(ii) Internal validity – we checked whether the studies were robust enough to demonstrate causality by looking for evidence of allocation concealment, potential blinding violation, sham method utilized, methods for handling missing data [intention-to-treat (ITT) and last-observation-carried-forward (LOCF) approaches] and evidence of carryover effects.

(iii) Construct validity – we examined the appropriateness of the operational criteria used for major depression, depression severity and treatment resistant depression (TRD). We also checked whether the studies verified measures of effectiveness as to assess not only statistical but also clinically meaningful differences.

(iv) External generalizability – we classified each study in five groups to check whether their conclusions can be generalized: (1) efficacy in general conditions; (2) long-term efficacy; (3) efficacy in specific conditions or subgroups; (4) new clinical or stereotaxic approaches and; (5) "add-on" strategies.

Regarding the nomenclature for combination strategies, it should be underscored that the term "add-on" is used, for some authors, to refer to subjects on antidepressant drugs (Fitzgerald *et al.*, 2006); while others use the term "augmentation" to refer to "add-on" strategies (Herwig *et al.*, 2007). Here, we chose to use the terms applied in pharmacological trials (Altshuler *et al.*, 2003) – therefore, studies whose strategy was to simultaneously combine non-pharmacological and pharmacological interventions at the beginning of the trial are referred as "add-on" or "accelerating" studies.

Results

General characteristics

The 31 studies included 1505 patients, with a mean age of 50.8 years, 819 (54%) of whom were female. Most of the studies had small sample sizes as 28 studies remained below the 75th percentile, with sample sizes ranging from 10 to 68 subjects, whereas three studies had sample sizes of 99 or more (Herwig *et al.*, 2007; O'Reardon *et al.*, 2007; Rossini *et al.*, 2005b).

Statistical validity

Only 10 (33%) of the studies performed power analysis calculations to estimate sample size, two of them showing negative results. Seven of the 21 studies not estimating sample size showed negative findings – a power analysis of them was performed (Table 1).

Regarding statistical tests, we observed most studies performed more than five tests for efficacy, as depression was measured by at least two rating scales and outcomes were expressed as continuous (score change) as well as categorical (response/remission rates). The tests most used were: ANCOVA – generally covariating for

baseline depression scores; repeated-measures ANOVA – commonly to explore for the effects of time, treatment and their interaction and the chi square test for addressing differences in responders for categorical variables. Logistic regressions and multiple linear regressions were the tests used for identifying predictors of response.

Finally, 16 (51%) studies did not define a priori the primary depression rating scale and 18 (58%) did not define if the primary outcome assessment was either continuous (score change) or categorical (number of responders).

Internal validity

Table 2 summarizes the internal validity assessment of the studies. All studies used a "single-blinded with external blinded rater evaluation" approach to preserve blinding and eight (25%) studies assessed blinding violation. Nineteen (61%) studies used an ITT approach – these studies handled with missing data from drop-out patients by using the LOCF method.

Regarding sham method, nine (30%) studies used a form of sham coil (i.e. a coil that did not generate a magnetic field and was only used for blinding purposes), 17 (55%) studies used an angled coil for blinding and five studies (16%) used other methods: a shielded coil (Jorge *et al.*, 2008), a coil generating a small magnetic field (O'Reardon *et al.*, 2007), and a turned-off tDCS device (Boggio *et al.*, 2008; Fregni *et al.*, 2006a).

Finally, although all the studies referred that patients were randomized, only 12 (39%) described the methodology used.

Construct validity

Table 3 shows the operational criteria used for construct validity. All studies based their criteira on the Diagnostic

Table 1 Power analyses of the studies showing negative results, active and sham stands for the endpoint depression scores in the active and sham rTMS groups, respectively

Author	Active	Sham	Cohen's *d*	Sample size	Estimated power (%)
Boutros *et al.*, 2002	27.09 ± 13.17	26.42 ± 13.35	0.05	21	5
Hausmann *et al.*, 2004	15.8 ± 9.5	20.2 ± 10.9	0.44	38	23
Holtzheimer *et al.*, 2004	14.6 ± 3.2	15.3 ± 3	0.23	15	7
Hoppner *et al.*, 2003	30.3 ± 4	29 ± 11	0.18	30	6
Loo *et al.*, 2003	32 ± 5	27 ± 5	1	19	58
Mosimann *et al.*, 2004	23.3 ± 7.2	20.4 ± 6.6	0.44	24	10
Poulet *et al.*, 2004	11.22 ± 6.1	18.12 ± 9.6	0.88	19	45

Note: The measure of the effect size is expressed in Cohen's *d*. The power analyses were performed using Stata 10® software.

and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV): 11 (35%) studies used a structured checklist for diagnosis, while in 20 (65%) it was made by

Table 2 Internal validity of the studies

Internal validity of the studies	Number of trials (%)
Blinding	
Single-blinded with external raters	31 (100%)
Blinding assessment	8 (25%)
Attrition	
Intention-to-treat approach	19 (61%)
Allocation concealment	
Not reported or unclear	19 (61%)
Sealed envelopes	4 (13%)
Random number list	8 (25%)
Sham method	
Angled coil	17 (55%)
Standard coil with deceiving	9 (30%)
Other methods	5 (16%)
Wash-out period	
Unclear	6 (19%)
Less than a week	6 (19%)
One to three weeks	6 (19%)
More than three weeks	7 (22%)

Table 3 Construct validity of the studies

Construct validity of the studies	Number of studies (%)
Definition of treatment resistant depression	
Not defined	9 (30%)
At least one AD failed	17 (55%)
At least two AD failed	3 (10%)
One or more AD failed	2 (5%)
Diagnostic of depression	
Unstructured	20 (65%)
MINI	4 (13%)
SCID	6 (19%)
SCAN	1 (3%)
Depression rating scale	
HDRS	29 (93%)
MADRS	10 (33%)
BDI	17 (55%)

AD, antidepressant drug; HDRS, Hamilton Depression Rating Scale; MADRS, Montgomery–Asberg Depression Rating Scale; BDI, Beck Depression Inventory; MINI, Mini International Neuropsychiatric Interview; SCID, Structured Clinical Interview for DSM disorders; SCAN, Schedule for Clinical Assessment in Neuropsychiatry.

clinical interview. Nine (33%) studies presented no definition of TRD, while 17 (55%) defined it correcly.

All studies used more than one depression rating scale, but only 14 (36%) defined the primary scale: 10 (33%) used the HDRS (Hamilton Depression Rating Scale) and four (13%), the MADRS (Montgomery–Asberg Depression Rating Scale).

Finally, the studies that performed power analyses and those that defined a priori whether the outcome assessment would be either continuous or categorical verified the clinical effectiveness of intervention; e.g. O'Reardon et al. (2007) defined a meaningful outcome as 25% reduction in depression rating scale as compared to sham group.

External generalizability

Table 4 shows that 11 (33%) studies assessed the general efficacy of NIBS and three (9%), the long-term efficacy. Six (19%) studies assessed new NIBS approaches, such as rTMS stimulation three times a week (Anderson et al., 2007) or twice daily (Loo et al., 2007), a bilateral simultaneuous stimulation (Loo et al., 2003), a stereotaxic guided stimulation (Herwig et al., 2003), an alternated high-frequency/low-frequency stimulation (Garcia-Toro et al., 2006) and a right/left sided alternated stimulation (Fitzgerald et al., 2006). Five (17%) studies assessed NIBS in specific populations, such as elderly patients (Mosimann et al., 2004) or vascular depression (Jorge et al., 2008). Finally, six studies (25%) addressed the efficacy of rTMS as an "add-on" strategy and 20 (65%) studies addressed the efficacy in TRD patients.

Discussion

We identified a number of methodological issues across the 31 NIBS trials reviewed, which can be divided in three groups: (1) issues inherent to small and exploratory studies, e.g. underpowered analysis; (2) issues common in MDD

Table 4 External validity of the studies

External validity of the studies	Number of studies (%)
General topics	
Refractoriness	20 (65%)
Patients in use of antidepressant drug	22 (71%)
"Accelerating" studies	6 (25%)
Augmentation studies	13 (42%)
Study hypotheses	
General efficacy of intervention	11 (33%)
Long-term efficacy	3 (10%)
Specific conditions/subgroups	5 (17%)
New clinical approaches	6 (20%)
"Add-on" studies	6 (25%)

trials, for instance, diagnostic validity; (3) issues specific to NIBS trials such as imperfect blinding and generalizability of the results. We first explore the concepts of internal validity versus generalizability and thereafter discuss these issues as to propose a potential alternatives to overcome them – which are summarized in Table 5.

Internal validity versus generalizability in NIBS trials

There is no definite answer regarding whether a clinical trial should focus on either internal validity (i.e. to prove causality between the intervention and the outcome) or generalizability (i.e. to aim to generalize the observed results outside the experimental situation) (Glasser and Howard, 2006). Usually, small studies focus on internal

validity because of strict eligibility criteria and lack of power to detect effects in a heterogeneous population, thereby grouping homogeneous groups which will only differ by the treatment applied. Such approach increases the chance of positive results if treatment is effective – in fact, small studies also perform multiple comparisons for this reason. However, these studies have low power and are prone to fail if internal validity is harmed or the treatment has a modest effect size. Besides, even underpowered trials are useful because: (1) they provide data on the best outcomes to be used in subsequent trials; (2) they provide information on parameters of stimulation and between-subjects variability; (3) they can be pooled together in meta-analyses. However, NIBS studies focusing on generalizability are necessary to validate

Table 5 Summary of the main issues identified and possible alternatives to handle them

Topic	Issues	Alternatives
General	Nomenclature – indistinctly use of the terms "add-on", "augmentantion", "acceleration".	Use the same nomenclature of pharmacological trials: *augmentantion* when associating NIBS to a patient already on drugs; *accelerating/potentiation/add-on* when NIBS and drug therapy start simultaneously; *switch* when drugs are stopped and NIBS is initiated.
Statistics	Increased probability of type I error (false positive).	Determine a priori the statistical test and only one primary hypothesis; refer other comparisons as exploratory.
Statistics	Increased probability of type II error (false negative).	Calculate sample size for an $ES \approx 0.5$; use adaptive designs; force as few variables as possible in a GLM.
Statistics	Violation of assumptions of normality.	Increase the sample size; use non-parametric tests; normalize the data, limit the number of variables imputed.
Construct validity	Diagnostic of MDD, TRD and use of rating scales.	Use structured interview of diagnosis; define refractoriness dichotomically or use a scale of degree of resistance.
Design	Randomization.	Describe methods of randomization and allocation.
Design	Blinding of a non-pharmacological intervention.	Exclude patients non-naïve to rTMS; avoid contact of subjects; use external raters; avoid crossover design; assess blinding.
Design	Carryover effects of NIBS are unknown.	Avoid crossover design; perform drug washout; use statistical methods to handle with carryover effects.
Generalizability	Patients with TRD have a smaller response.	Measure the degree of resistance; stratify during randomization; analyze as a covariate; perform a subgroup analysis.
Generalizability	Accelerating trials have two active groups, requiring a large sample size.	Use patients with severe MDD to decrease variability; add a third sham/placebo group; use a factorial design.
Generalizability	Augmentation trials have a wide variability of the sample.	Control TRD, gender and other predictors; use patients taking the same drug and dosage for more than 12 weeks.
Generalizability	Switch trials might not reflect clinical practice.	Enroll patients that do not want or are unable to take oral pills, e.g. pregnant women, patients with several medical conditions.

GLM, general linear model; ES, effect size; MDD, major depressive disorder; TRD, treatment-resistant depression; NIBS, non-invasive brain stimulation; rTMS, repetitive transcranial magnetic stimulation.

the intervention as an effective treatment. The study of O'Reardon *et al.* (2007) is an example of an effectiveness study.

Issues related to exploratory studies

Hypothesis testing

Most trials reviewed are vulnerable to both type I and type II errors. Type I error increases when multiple statistical comparisons are performed – in fact, several studies carried out five or more statistical tests for efficacy. However, the majority of negative studies were underpowered, suggesting that these studies had false negative conclusions – in fact, this finding is common in clinical trials, as a meta-analysis of negative findings in orthopedic trials showed that type II error was 90% (Lochner *et al.*, 2001). Possible solutions here are: (1) to determine a priori the statistical method that will be used for the primary outcome (Hopewell *et al.*, 2008); (2) to claim other analyses as secondary or exploratory; (3) to perform sample size estimations before the trial starts (however the primary hypothesis needs to be defined for the sample size calculation) – given the results based on NIBS meta-analyses, the estimated effect size for sample size calculations might range from a Cohen's d of 0.4 to 0.75 (Gross *et al.*, 2007; Schutter, 2008); (4) to use adaptive designs, as we discuss later.

Statistics

Most trials reviewed used more complex statistical approaches, such as mixed analysis of variances (ANOVAs) or multivariate ANOVAs. The advantage of using these models is to compare several dependent and independent variables simultaneously without increasing the probability of type I or II errors. However, possible caveats are that: (1) usually several variables are imputed in these general linear models, which might violate the recommended assumptions for using such models (for instance, maximal number of events or entries according to the number of independent variables) (Ottenbacher *et al.*, 2004) and; (2) the use of parametric tests is controversial because: the sample size of most studies were not large enough to fit the Central Limit Theorem assumptions; psychometric scales are more ordinal than continuous scales (Bech, 1988), and; tests for normality were often not reported. To address these issues, we suggest: (1) to determine a priori the statistical tests; (2) to perform normality tests, such as the Kolmogorov–Smirnov test, especially when the sample size is small; (3) to limit the number of variables imputed in the covariate adjustments,

usually using only baseline depression scores and, perhaps, degree of resistance; (4) to handle with known predictors of response not only through statistics, but also in the study design, i.e. at the enrollment and randomization phases (defining adequate inclusion/exclusion criteria and using stratified randomization approach, for instance), as such approach would not undermine statistical power.

Randomization

Most studies did not report the randomization methods (a process that ensure that each subject in a trial has the same probability of being assigned for any group) and allocation (the technique used to direct a subject to his/her assigned group preserving randomization and blinding) methods, therefore, we could not evaluate to which extent this issue jeopardizes the internal validity of NIBS trials.

Issues related to MDD trials

The main threat identified here refers to construct validity (i.e. how the diagnosis performed relates to the "real" diagnosis – which concerns to generalizability) and reliability (i.e. the interrater agreement when performing a diagnosis) – in fact, this is also an issue in antidepressant drug trials (Gelenberg *et al.*, 2008). In most studies, MDD diagnosis was made through unstructured interviews, which is a threat to validity – as differential diagnoses, such as bipolar disorders and personality disorders, might have less sensitivity to diagnose these conditions (Zimmerman and Mattia, 1999) – and also to reliability, as unstructured interviews have lesser inter-realibility agreeement than structured and semistructured interviews (Williams *et al.*, 2002). Another threat is that the trials reviewed used four different diagnostic criteria to define TRD – in fact, TRD refers to a specific subgroup (15–40%) of patients failing to achieve significant clinical improvement after at least two antidepressant trials (Berlim and Turecki, 2007). Since this subgroup of patients presents lower response rates to antidepressant treatments (Rush *et al.*, 2006) and most NIBS trials enroll TRD patients – as treating with drugs first might be more cost-effective – future trials need to consider assessing TRD by: (1) using refractoriness criteria as defined by two failed antidepressant trials of different antidepressant classes as commonly referred in the literature (Berlim and Turecki, 2007) or; (2) grading (not only identifying) depression resistance as it is known that the degree of resistance is associated with antidepressant response (Lisanby *et al.*, 2009; Rush *et al.*, 2006).

Finally, it is not clear whether the primary outcome should be based on score change or depression remission: although a complete depression treatment should aim remission of symptoms (Moller, 2008), the use of score change can be useful to assess improvement in TRD patients that present low remission rates (Trivedi, 2006) as well as to increase study power.

Issues related to NIBS trials

Blinding

Blinding is a set of techniques used to keep subjects and/or study staff unaware of the intervention administered. Insufficient blinding might bias results by increasing the effectiveness of the active group (by increased staff care and/or accomplishment of patient's expectancies) and decreasing in the control group (Hrobjartsson and Gotzsche, 2003; Noseworthy et al., 1994). Also, the optimal NIBS blinding technique has not been established yet: some researchers use an angled coil in control group, therefore resulting in staff unblinding; while others prefer to use a sham coil as placebo intervention – which might be a better method of blinding, however the sensation associated with it is still not the same as that induced by the real coil. Along these lines, tDCS devices might provide a more reliable blinding method, since, accordingly to the set of parameters specified, the skin sensation associated with the use of active stimulation might be null after a couple of seconds of stimulation (Boggio et al., 2008; Gandiga et al., 2006) — yet, at high intensities, patients might notice tingling or other symptoms associated with stimulation, compromising blinding. Therefore, alternative blinding methods – especially for rTMS – should be used, such as avoiding contact between subjects of different groups, excluding patients non-naïve to rTMS, avoiding crossover trials and blinding raters to treatment applied (Boutron et al., 2007).

Blinding assessment

Only eight studies reported methods to address blinding, mainly by asking subjects or staff personnel to guess where patients were allocated. However, although blinding assessment can be useful to identify whether the blinding was broken (Boutron et al., 2007), this method is controversial (Fergusson et al., 2004) – e.g. correct guessing might be related to side effects and amelioration of symptoms and not to the blinding process. Therefore, blinding realiability might be obtained through other methods, such as familiarizing staff to study procedures, performing team meetings to discuss potential blinding

breaking and testing the blinding technique before applying it in a efficacy trial.

Placebo effect of NIBS devices

Non-pharmacological interventions might have a greater placebo response than pharmacological ones (Kaptchuk et al., 2000) for two theoretical reasons: one is that subjects might have greater expectancies regarding a non-pharmacological intervention and the other is that sham interventions might have an active effect (Dincer and Linde, 2003). However, such interventions are more difficult to blind, patients might be skeptical (as with a completely different approach) and therefore have less expectancy, and NIBS devices are often used in TRD patients – these factors that might decrease placebo response. Along these lines, a recent meta-analysis showed that the placebo sham response is similar to a placebo pill response (Brunoni et al., 2009). More studies are needed to explore the possibility of NIBS devices having a greater sham response.

Carryover effects

The carryover effect (i.e. the effect of one intervention being carried over to another) threats internal validity of NIBS trials because in both "add-on" and "switch" strategies, non-pharmacological and pharmacological effects might be mixed together; thereby, whether or not a difference is observed, it would not be possible to solely attribute it to the NIBS. To acknowledge this threat, the wash-out period of antidepressant drugs should be reported, being of at least four half-lives of the drug (Leucht et al., 2008). The issue is even more controversial in NIBS crossover designs, in which the effect duration of non-pharmacological interventions is not yet known; therefore, such trials might need a specific analysis of potential carryover effects (comparing baselines before each intervention or including treatment effect in the model).

Treatment-resistant depression

It is still unclear whether NIBS studies should either include or exclude refractory patients – although refractoriness can theoretically decrease NIBS efficacy as remission rates decay as refractoriness increases (Fregni et al., 2006b; Lisanby et al., 2009; Rush et al., 2006); placebo response might also decrease with refractoriness (Brunoni et al., 2009) and, therefore, NIBS effects could actually be enhanced by a lower placebo response in control group – in fact, an open tDCS study showed promising results in TRD patients (Ferrucci et al., 2009). Useful approaches

might be measuring the degree of resistance of each patient with severity rating scales and perform covariate adjustments or, alternatively, classify the subject dichotomically and stratify accordingly during randomization.

Along these lines, other predictors of response, such as gender (Huang *et al.*, 2008) and age (Fregni *et al.*, 2006b) should also be taken into account when designing NIBS trials – these variables might be handled through stratification and/or explored through covariate and multivariate adjustments – although such analyses might increase type I error as well as require a greater sample size to avoid type II error.

Study design

Although most studies enrolled patients who were previously using antidepressant drugs, pivotal studies such as the study of O'Reardon *et al.* (2007) enrolled drug-free patients and Herwig *et al.* (2007) enrolled patients in an accelerating study. Each approach has its strenghts and weaknesses (Figure 1).

Accelerating studies combine a NIBS and a pharmacological intervention, theoretically fastening response (Padberg and Moller, 2003), which might be useful for patients with severe and/or refractory depression. On the contrary, such design suffers from the lack of a true placebo group as at least one active intervention is conducted for each patient. Therefore, this approach might increase response rates in both groups, thereby demanding a large sample size to detect a significant difference.

Drug-free patients designs are attractive because there is no confounding effect between NIBS and antidepressant drugs and such studies might be useful for patients who might not be taking antidepressants, as we discuss later. However, it is not known yet whether maintenance treatment

with NIBS is effective because there are few follow-up studies (Demirtas-Tatlidede *et al.*, 2008). In addition, this adds complexity to the study by requesting a wash-out period.

Finally, augmentation studies do not evolve wash-out periods and keep patients on drugs after the trial is over. Also, such approach might also have the greatest external generalizability as it would enroll patients who are still symptomatic after the pharmacological treatment – a putative target population for NIBS approaches. The drawback is that exploratory studies would enroll patients with various grades of refractoriness and different therapeutic regimens; while pivotal NIBS augmentation trials would be the most complex, needing a previous run-in antidepressant drug trial to select eligible patients (e.g.those who did not achieve clinical response).

Improving future NIBS studies

To conclude, we discuss a few areas of improvement for future NIBS studies.

Use of adaptive designs

Adaptive designs allow modification in on-going trials without compromising power (Chow and Chang, 2008). Such designs allow to reduce sample size and trial duration (Sagkriotis and Scholpp, 2008), drop weaker treatments (Chow and Chang, 2008) and change primary outcomes during the trial (Banerjee and Tsiatis, 2006); being increasingly used for phase II/III trials. NIBS studies could take advantage of such designs – for instance, by testing either different stimulation "doses" or stimulation sites in the scalp, dropping in advance weaker treatments; or in sample size re-estimation to achieve statistical power.
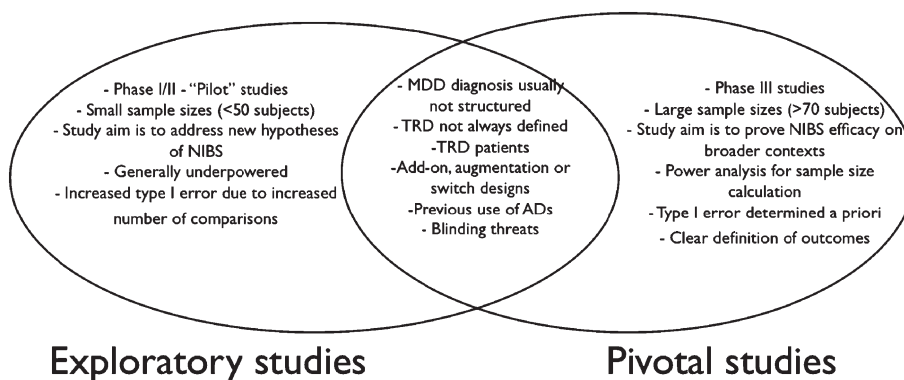

Figure 1 Summary of the non-invasive brain stimulation (NIBS) studies reviewed, comparing similarities and differences of initial exploratory trials to pivotal trials.

Use of translational research

Translational research utilizes knowledge from basic sciences to improve clinical research (Woolf, 2008). Therefore, NIBS efficacy can also be explored through biological findings emerging from MDD translational studies, such as: (1) biological serum markers, such as brain-derived neurotrophic factor (BDNF), which is related to neuroplasticity in key brain areas related to depression (Brunoni et al., 2008a) with two meta-analyses showing that the serum levels are related to depression severity and response (Brunoni et al., 2008b; Sen et al., 2008); (2) neuroimaging studies, such as magnetic resonance imaging (MRI) spectroscopy, which measures the cerebral bioenergetic metabolism that is usually low in MDD (Iosifescu et al., 2008), in fact, a recent spectroscopty study showed that brain metabolism is altered after NIBS (Rango et al., 2008); (3) quantitative electroencephalograph (EEG) recordings, which measures brain activity and might be a predictor of antidepressant treatment (Hunter et al., 2007) – also EEG could be used to index brain activity during NIBS, as this technique can be used to index cortical activity (that is the target for NIBS treatment). This would be useful to adjust tDCS/rTMS dosage individually.

Use of NIBS in selected subgroups of patients

There are some patient groups where depression treatment is challenging and where NIBS could be an alternative treatment strategy. Pregnant women, for instance, are often hindered from drug treatment because of safety concerns (Campagne, 2007). Therefore, NIBS could be an adequate acute and maintenance treatment during pregnancy (Klirova et al., 2008). Along these lines, geriatric patients could also benefit of NIBS approaches, since they usually present co-morbidities and polypharmacy (Rabheru, 2004) hindering optimal antidepressant treatment.

Conclusion

Conducting NIBS trials is challenging as specific methodological issues of brain stimulation are summed with the common design problems faced in depression and in pilot trials. Taken together, these pitfalls can seriously burden study validity and lead to inconclusive results. In our review, we aimed to discuss such pitfalls and, consequently, to provide further directions for future clinical trials on NIBS for MDD.

Acknowledgments

References

Altshuler L.L., Frye M.A., Gitlin M.J. (2003) Acceleration and augmentation strategies for treating bipolar depression. *Biological Psychiatry*, **53**(8), 691–700.

Anderson I.M., Delvai N.A., Ashim B., Ashim S., Lewin C., Singh V., Sturman D., Strickland P.L. (2007) Adjunctive fast repetitive transcranial magnetic stimulation in depression. *British Journal of Psychiatry*, **190**, 533–534.

Avery D.H., Holtzheimer P.E., 3rd, Fawaz W., Russo J., Neumaier J., Dunner D.L., Haynor D. R., Claypoole K.H., Wajdik C., Roy-Byrne P. (2006) A controlled study of repetitive transcranial magnetic stimulation in medication-resistant major depression. *Biological Psychiatry*, **59**(2), 187–194.

Baghai T.C., Volz H.P., Moller H.J. (2006) Drug treatment of depression in the 2000s: An overview of achievements in the last 10 years and future possibilities. *World Journal of Biological Psychiatry*, **7**(4), 198–222.

Banerjee A., Tsiatis A.A. (2006) Adaptive two-stage designs in phase II clinical trials. *Statistics in Medicine*, **25**(19), 3382–3395.

Bech P. (1988) Rating scales for mood disorders: Applicability, consistency and construct validity. *Acta Psychiatrica Scandinavica, Supplement*, **345**, 45–55.

Berlim M.T., Turecki G. (2007) What is the meaning of treatment resistant/refractory major depression (TRD)? A systematic review of current randomized trials. *European Neuropsychopharmacology*, **17**(11), 696–707.

Boggio P.S., Rigonatti S.P., Ribeiro R.B., Myczkowski M.L., Nitsche M.A., Pascual-Leone A., Fregni F. (2008) A randomized, double-blind clinical trial on the efficacy of cortical direct current stimulation for the treatment of major depression. *International Journal of Neuropsychopharmacology*, **11**(2), 249–254.

Bortolomasi M., Minelli A., Fuggetta G., Perini M., Comencini S., Fiaschi A., Manganotti P. (2007) Long-lasting effects of high frequency repetitive transcranial magnetic stimulation in major depressed patients. *Psychiatry Research*, **150**(2), 181–186.

Boutron I., Guittet L., Estellat C., Moher D., Hrobjartsson A., Ravaud P. (2007) Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Medicine*, **4**(2), e61.

Boutros N.N., Gueorguieva R., Hoffman R.E., Oren D.A., Feingold A., Berman R.M. (2002) Lack of a therapeutic effect of a 2-week subthreshold transcranial magnetic stimulation course for treatment-resistant depression. *Psychiatry Research*, **113**(3), 245–254.

Bretlau L.G., Lunde M., Lindberg L., Unden M., Dissing S., Bech P. (2008) Repetitive transcranial magnetic stimulation (rTMS) in combination with escitalopram in patients with treatment-resistant major depression: A double-blind, randomised, sham-controlled trial. *Pharmacopsychiatry*, **41**(2), 41–47.

Brunoni A.R., Boggio P.S., Fregni F. (2008a) Can the 'yin and yang' BDNF hypothesis be used to

predict the effects of rTMS treatment in neuropsychiatry? *Medical Hypotheses*, **71**(2), 279–282.

Brunoni A.R., Lopes M., Fregni F. (2008b) A systematic review and meta-analysis of clinical studies on major depression and BDNF levels: implications for the role of neuroplasticity in depression. *The International Journal of Neuropsychopharmacology/Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)*, **11**(8), 1169–1180.

Brunoni A.R., Lopes M., Kaptchuk T.J., Fregni F. (2009) Placebo response of non-pharmacological and pharmacological trials in major depression: A systematic review and meta-analysis. *PLoS ONE* **4**(3), e4824.

Campagne D.M. (2007) Fact: Antidepressants and anxiolytics are not safe during pregnancy. *European Journal of Obstetrics & Gynecology amd Reproductive Biology*, **135**(2), 145–148.

Chow S.C., Chang M. (2008) Adaptive design methods in clinical trials – a review. *Orphanet Journal of Rare Diseases*, **3**, 11.

Demirtas-Tatlidede A., Mechanic-Hamilton D., Press D.Z., Pearlman C., Stern W.M., Thall M., Pascual-Leone A. (2008) An open-label, prospective study of repetitive transcranial magnetic stimulation (rTMS) in the long-term treatment of refractory depression: Reproducibility and duration of the antidepressant effect in medication-free patients. *Journal of Clinical Psychiatry*, **69**(6), 930–934.

Dincer F., Linde K. (2003) Sham interventions in randomized clinical trials of acupuncture – a review. *Complementary Therapies in Medicine*, **11**(4), 235–242.

Elahi B., Chen R. (2009) Effect of transcranial magnetic stimulation on Parkinson motor function – systematic review of controlled clinical trials. *Movement Disorders*, **24**(3), 357–363.

Fergusson D., Glass K.C., Waring D., Shapiro S. (2004) Turning a blind eye: The success of blinding reported in a random sample of randomised, placebo controlled trials. *British Medical Journal*, **328**(7437), 432.

Ferrucci R., Bortolomasi M., Vergari M., Tadini L., Salvoro B., Giacopuzzi M., Barbieri S., Priori A. (2009) Transcranial direct current stimulation in severe, drug-resistant major depression. *Journal of Affective Disorders*, **118**(1–3), 215–219.

Fitzgerald P.B., Benitez J., de Castella A., Daskalakis Z.J., Brown T.L., Kulkarni J. (2006) A randomized, controlled trial of sequential bilateral repetitive transcranial magnetic stimulation for treatment-resistant depression. *The American Journal of Psychiatry*, **163**(1), 88–94.

Fitzgerald P.B., Brown T.L., Marston N.A., Daskalakis Z.J., De Castella A., Kulkarni J. (2003) Transcranial magnetic stimulation in the treatment of depression: A double-blind, placebo-controlled trial. *Archives of General Psychiatry*, **60**(10), 1002–1008.

Fregni F., Boggio P.S., Nitsche M.A., Marcolin M. A., Rigonatti S.P., Pascual-Leone A. (2006a) Treatment of major depression with transcranial direct current stimulation. *Bipolar Disorder*, **8**(2), 203–204.

Fregni F., Marcolin M.A., Myczkowski M., Amiaz R., Hasey G., Rumi D.O., Rosa M., Rigonatti S. P., Camprodon J., Walpoth M., Heaslip J., Grunhaus L., Hausmann A., Pascual-Leone A. (2006b) Predictors of antidepressant response in clinical trials of transcranial magnetic stimulation. *The International Journal of Neuropsychopharmacology/Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)*, **9**(6), 641–654.

Fregni F., Pascual-Leone A. (2007) Technology insight: Noninvasive brain stimulation in neurology-perspectives on the therapeutic potential of rTMS and tDCS. *Nature Clinical Practice Neurology*, **3**(7), 383–393.

Freitas C., Fregni F., Pascual-Leone A. (2009) Meta-analysis of the effects of repetitive transcranial magnetic stimulation (rTMS) on negative and positive symptoms in schizophrenia. *Schizophrenia Research*, **108**(1-3), 11–24.

Gandiga P.C., Hummel F.C., Cohen L.G. (2006) Transcranial DC stimulation (tDCS): A tool for double-blind sham-controlled clinical studies in brain stimulation. *Clinical Neurophysiology*, **117**(4), 845–850.

Garcia-Toro M., Salva J., Daumal J., Andres J., Romera M., Lafau O., Echevarria M., Mestre M., Bosch C., Collado C., Ibarra O., Aguirre I. (2006) High (20-Hz) and low (1-Hz) frequency transcranial magnetic stimulation as adjuvant treatment in medication-resistant depression. *Psychiatry Research*, **146**(1), 53–57.

Gelenberg A.J., Thase M.E., Meyer R.E., Goodwin F.K., Katz M.M., Kraemer H.C., Potter W.Z., Shelton R.C., Fava M., Khan A., Trivedi M.H., Ninan P.T., Mann J.J., Bergeson S., Endicott J., Kocsis J.H., Leon A.C., Manji H.K., Rosenbaum J.F. (2008) The history and current state of antidepressant clinical trial design: A call to action for proof-of-concept studies. *Journal of Clinical Psychiatry*, **69**(10), 1513–1528.

Glasser S.P., Howard G. (2006) Clinical trial design issues: At least 10 things you should look for in clinical trials. *Journal of Clinical Pharmacology*, **46**(10), 1106–1115.

Gross M., Nakamura L., Pascual-Leone A., Fregni F. (2007) Has repetitive transcranial magnetic stimulation (rTMS) treatment for depression improved? A systematic review and meta-analysis comparing the recent vs. the earlier rTMS studies. *Acta Psychiatrica Scandinavica*, **116**(3), 165–173.

Groves D.A., Brown V.J. (2005) Vagal nerve stimulation: A review of its applications and potential mechanisms that mediate its clinical effects. *Neuroscience & Biobehavioral Reviews*, **29**(3), 493–500.

Hausmann A., Kemmler G., Walpoth M., Mechtcheriakov S., Kramer-Reinstadler K., Lechner T., Walch T., Deisenhammer E.A., Kofler M., Rupp C.I., Hinterhuber H., Conca A. (2004) No benefit derived from repetitive transcranial magnetic stimulation in depression: A prospective, single centre, randomised, double blind, sham controlled "add on" trial. *Journal of Neurology, Neurosurgery and Psychiatry*, **75**(2), 320–322.

Herrmann L.L., Ebmeier K.P. (2006) Factors modifying the efficacy of transcranial magnetic stimulation in the treatment of depression: A review. *Journal of Clinical Psychiatry*, **67**(12), 1870–1876.

Herwig U., Fallgatter A.J., Hoppner J., Eschweiler G.W., Kron M., Hajak G., Padberg F., Naderi-Heiden A., Abler B., Eichhammer P., Grossheinrich N., Hay B., Kammer T., Langguth B., Laske C., Plewnia C., Richter M.M., Schulz M., Unterecker S., Zinke A., Spitzer M., Schonfeldt-Lecuona C. (2007) Antidepressant effects of augmentative transcranial magnetic stimulation: Randomised multicentre trial. *British Journal of Psychiatry*, **191**, 441–448.

Herwig U., Lampe Y., Juengling F.D., Wunderlich A., Walter H., Spitzer M., Schonfeldt-Lecuona C. (2003) Add-on rTMS for treatment of depression: A pilot study using stereotaxic coil-navigation according to PET data. *Journal of Psychiatric Research*, **37**(4), 267–275.

Holtzheimer P.E., 3rd, Russo J., Claypoole K.H., Roy-Byrne P., Avery D.H. (2004) Shorter duration of depressive episode may predict response to repetitive transcranial magnetic stimulation. *Depression and Anxiety*, **19**(1), 24–30.

Hopewell S., Clarke M., Moher D., Wager E., Middleton P., Altman D.G., Schulz K.F. (2008) CONSORT for reporting randomized controlled

trials in journal and conference abstracts: Explanation and elaboration. *PLoS Medicine*, **5**(1), e20.

Hoppner J., Schulz M., Irmisch G., Mau R., Schlafke D., Richter J. (2003) Antidepressant efficacy of two different rTMS procedures. High frequency over left versus low frequency over right prefrontal cortex compared with sham stimulation. *European Archives of Psychiatry and Clinical Neuroscience*, **253**(2), 103–109.

Hrobjartsson A., Gotzsche P.C. (2003) Placebo treatment versus no treatment. *Cochrane Database of Systematic Reviews*, (1), CD003974.

Huang C.C., Wei I.H., Chou Y.H., Su T.P. (2008) Effect of age, gender, menopausal status, and ovarian hormonal level on rTMS in treatment-resistant depression. *Psychoneuroendocrinology*, **33**(6), 821–831.

Hunter A.M., Cook I.A., Leuchter A.F. (2007) The promise of the quantitative electroencephalogram as a predictor of antidepressant treatment outcomes in major depressive disorder. *Psychiatric Clinics of North America*, **30**(1), 105–124.

Iosifescu D.V., Bolo N.R., Nierenberg A.A., Jensen J.E., Fava M., Renshaw P.F. (2008) Brain bioenergetics and response to triiodothyronine augmentation in major depressive disorder. *Biological Psychiatry*, **63**(12), 1127–1134.

Januel D., Dumortier G., Verdon C.M., Stamatiadis L., Saba G., Cabaret W., Benadhira R., Rocamora J.F., Braha S., Kalalou K., Vicaut P.E., Fermanian J. (2006) A double-blind sham controlled study of right prefrontal repetitive transcranial magnetic stimulation (rTMS): Therapeutic and cognitive effect in medication free unipolar depression during 4 weeks. *Progress in Neuropsychopharmacology & Biological Psychiatry*, **30**(1), 126–130.

Jorge R.E., Moser D.J., Acion L., Robinson R.G. (2008) Treatment of vascular depression using repetitive transcranial magnetic stimulation. *Archives of General Psychiatry*, **65**(3), 268–276.

Jorge R.E., Robinson R.G., Tateno A., Narushima K., Acion L., Moser D., Arndt S., Chemerinski E. (2004) Repetitive transcranial magnetic stimulation as treatment of poststroke depression: A preliminary study. *Biological Psychiatry*, **55**(4), 398–405.

Kaptchuk T.J., Goldman P., Stone D.A., Stason W.B. (2000) Do medical devices have enhanced placebo effects? *Journal of Clinical Epidemiology*, **53**(8), 786–792.

Klirova M., Novak T., Kopecek M., Mohr P., Strunzova V. (2008) Repetitive transcranial magnetic stimulation (rTMS) in major depressive episode during pregnancy. *Neuroendocrinology Letters*, **29**(1), 69–70.

Koerselman F., Laman D.M., van Duijn H., van Duijn M.A., Willems M.A. (2004) A 3-month, follow-up, randomized, placebo-controlled study of repetitive transcranial magnetic stimulation in depression. *Journal of Clinical Psychiatry*, **65**(10), 1323–1328.

Leucht S., Heres S., Hamann J., Kane J.M. (2008) Methodological issues in current antipsychotic drug trials. *Schizophrenia Bulletin*, **34**(2), 275–285.

Lisanby S.H., Husain M.M., Rosenquist P.B., Maixner D., Gutierrez R., Krystal A., Gilmer W., Marangell L.B., Aaronson S., Daskalakis Z.J., Canterbury R., Richelson E., Sackeim H.A., George M.S. (2009) Daily left prefrontal repetitive transcranial magnetic stimulation in the acute treatment of major depression: Clinical predictors of outcome in a multisite, randomized controlled clinical trial. *Neuropsychopharmacology*, **34**(2), 522–534.

Lochner H.V., Bhandari M., Tornetta P., 3rd. (2001) Type-II error rates (beta errors) of randomized trials in orthopaedic trauma. *Journal of Bone and Joint Surgery (American)*, **83–A**(11), 1650–1655.

Loo C.K., Mitchell P.B., Croker V.M., Malhi G.S., Wen W., Gandevia S.C., Sachdev P.S. (2003) Double-blind controlled investigation of bilateral prefrontal transcranial magnetic stimulation for the treatment of resistant major depression. *Psychological Medicine*, **33**(1), 33–40.

Loo C.K., Mitchell P.B., McFarquhar T.F., Malhi G.S., Sachdev P.S. (2007) A sham-controlled trial of the efficacy and safety of twice-daily rTMS in major depression. *Psychological Medicine*, **37**(3), 341–349.

Marder S.R. (2006) Initiatives to promote the discovery of drugs to improve cognitive function in severe mental illness. *Journal of Clinical Psychiatry*, **67**(7), e03.

Martin J.L., Barbanoj M.J., Schlaepfer T.E., Thompson E., Perez V., Kulisevsky J. (2003) Repetitive transcranial magnetic stimulation for the treatment of depression. Systematic review and meta-analysis. *British Journal of Psychiatry*, **182**, 480–491.

Mogg A., Pluck G., Eranti S.V., Landau S., Purvis R., Brown R.G., Curtis V., Howard R., Philpot M., McLoughlin D.M. (2008) A randomized controlled trial with 4-month follow-up of adjunctive repetitive transcranial magnetic stimulation of the left prefrontal cortex for depression. *Psychological Medicine*, **38**(3), 323–333.

Moller H.J. (2008) Outcomes in major depressive disorder: The evolving concept of remission and its implications for treatment. *World Journal of Biological Psychiatry*, **9**(2), 102–114.

Mosimann U.P., Schmitt W., Greenberg B.D., Kosel M., Muri R.M., Berkhoff M., Hess C.W., Fisch H.U., Schlaepfer T.E. (2004) Repetitive transcranial magnetic stimulation: A putative add-on treatment for major depression in elderly patients. *Psychiatry Research*, **126**(2), 123–133.

Noseworthy J.H., Ebers G.C., Vandervoort M.K., Farquhar R.E., Yetisir E., Roberts R. (1994) The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology*, **44**(1), 16–20.

O'Reardon J.P., Cristancho P., Pilania P., Bapatla K.B., Chuai S., Peshek A.D. (2007) Patients with a major depressive episode responding to treatment with repetitive transcranial magnetic stimulation (rTMS) are resistant to the effects of rapid tryptophan depletion. *Depression and Anxiety*, **24**(8), 537–544.

Ottenbacher K.J., Ottenbacher H.R., Tooth L., Ostir G.V. (2004) A review of two journals found that articles using multivariable logistic regression frequently did not report commonly recommended assumptions. *Journal of Clinical Epidemiology*, **57**(11), 1147–1152.

Padberg F., Moller H.J. (2003) Repetitive transcranial magnetic stimulation: Does it have potential in the treatment of depression? *CNS Drugs*, **17**(6), 383–403.

Poulet E., Brunelin J., Boeuve C., Lerond J., D'Amato T., Dalery J., Saoud M. (2004) Repetitive transcranial magnetic stimulation does not potentiate antidepressant treatment. *European Psychiatry*, **19**(6), 382–383.

Rabheru K. (2004) Special issues in the management of depression in older patients. *Canadian Journal of Psychiatry*, **49**(3 Suppl 1), 41S–50S.

Rango M., Cogiamanian F., Marceglia S., Barberis B., Arighi A., Biondetti P., Priori A. (2008) Myoinositol content in the human brain is modified by transcranial direct current stimulation in a matter of minutes: A 1H-MRS study. *Magnetic Resonance in Medicine*, **60**(4), 782–789.

Rossini D., Lucca A., Zanardi R., Magri L., Smeraldi E. (2005a) Transcranial magnetic stimulation in treatment-resistant depressed

patients: A double-blind, placebo-controlled trial. *Psychiatry Research*, **137**(1–2), 1–10.

Rossini D., Magri L., Lucca A., Giordani S., Smeraldi E., Zanardi R. (2005b) Does rTMS hasten the response to escitalopram, sertraline, or venlafaxine in patients with major depressive disorder? A double-blind, randomized, sham-controlled trial. *Journal of Clinical Psychiatry*, **66**(12), 1569–1575.

Rumi D.O., Gattaz W.F., Rigonatti S.P., Rosa M.A., Fregni F., Rosa M.O., Mansur C., Myczkowski M.L., Moreno R.A., Marcolin M.A. (2005) Transcranial magnetic stimulation accelerates the antidepressant effect of amitriptyline in severe depression: A double-blind placebo-controlled study. *Biological Psychiatry*, **57**(2), 162–166.

Rush A.J., Trivedi M.H., Wisniewski S.R., Nierenberg A.A., Stewart J.W., Warden D., Niederehe G., Thase M.E., Lavori P.W., Lebowitz B.D., McGrath P.J., Rosenbaum J.F., Sackeim H.A., Kupfer D.J., Luther J., Fava M. (2006) Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *The American Journal of Psychiatry*, **163**(11), 1905–1917.

Sagkriotis A., Scholpp J. (2008) Combining proof-of-concept with dose-finding: Utilization of adaptive designs in migraine clinical trials. *Cephalalgia*, **28**(8), 805–812.

Schutter D.J. (2008) Antidepressant efficacy of high-frequency transcranial magnetic stimulation over the left dorsolateral prefrontal cortex in double-blind sham-controlled designs: A meta-analysis. *Psychological Medicine*, 1–11.

Sen S., Duman R., Sanacora G. (2008) Serum brain-derived neurotrophic factor, depression, and antidepressant medications: Meta-analyses and implications. *Biological Psychiatry*, **64**(6), 527–532.

Stern W.M., Tormos J.M., Press D.Z., Pearlman C., Pascual-Leone A. (2007) Antidepressant effects of high and low frequency repetitive transcranial magnetic stimulation to the dorsolateral prefrontal cortex: A double-blind, randomized, placebo-controlled trial. *Journal of Neuropsychiatry and Clinical Neurosciences*, **19**(2), 179–186.

Su T.P., Huang C.C., Wei I.H. (2005) Add-on rTMS for medication-resistant depression: A randomized, double-blind, sham-controlled trial in Chinese patients. *Journal of Clinical Psychiatry*, **66**(7), 930–937.

Trivedi M.H. (2006) Major depressive disorder: Remission of associated symptoms. *Journal of Clinical Psychiatry*, **67**(Suppl. 6), 27–32.

Williams J.W., Jr, Noel P.H., Cordes J.A., Ramirez G., Pignone M. (2002) Is this patient clinically depressed? *Journal of the American Medical Association*, **287**(9), 1160–1170.

Woolf S.H. (2008) The meaning of translational research and why it matters. *Journal of the American Medical Association*, **299**(2), 211–213.

Zimmerman M., Mattia J.I. (1999) Psychiatric diagnosis in clinical practice: Is comorbidity being missed? *Comprehensive Psychiatry*, **40**(3), 182–191.

Zlowodzki M., Jonsson A., Bhandari M. (2006) Common pitfalls in the conduct of clinical research. *Medical Principles and Practice*, **15**(1), 1–8.