European Psychiatry 27 (2012) 147–155

Available online at ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com/en

Review

Systematic review and meta-analysis of vagus nerve stimulation in the treatment of depression: Variable results based on study designs

J.L.R. Martin [a,*], E. Martín-Sánchez [b,1]

[a] School of Biomedical and Health Sciences, Universidad Europea de Madrid, Madrid, Spain
[b] Clinical Research Department, Health and Social Care Foundation of Castilla, La Mancha, Spain

ARTICLE INFO

Article history:
Received 18 February 2011
Received in revised form 18 July 2011
Accepted 22 July 2011
Available online 1 December 2011

Keywords:
Vagus nerve stimulation
VNS
Depression
Systematic review
Meta-analysis

ABSTRACT

Purpose: To determine the efficacy of vagus nerve stimulation (VNS) for treatment of depression.
Methods: We conducted a systematic review and meta-analysis of analytical studies. Efficacy was evaluated according to severity of illness and percentage of responders.
Results: We identified 687 references. Of these, 14 met the selection criteria and were included in the review. The meta-analysis of efficacy for uncontrolled studies showed a significant reduction in scores at the Hamilton Depression Rating Scale endpoint, and the percentage of responders was 31.8% ([23.2% to 41.8%], $P < 0.001$). However, the randomised control trial which covered a sample of 235 patients with depression, reported no statistically significant differences between the active intervention and placebo groups (OR = 1.61 [95%CI 0.72 to 3.62]; $P = 0.25$). To study the cause of this heterogeneity, a meta-regression was performed. The adjusted coefficient of determination (R^2_{Adj}) was 0.84, which implies that an 84% variation in effect size across the studies was explained by baseline severity of depression ($P < 0.0001$).
Conclusion: Currently, insufficient data are available to describe VNS as effective in the treatment of depression. In addition, it cannot be ruled out that the positive results observed in the uncontrolled studies might have been mainly due to a placebo effect.

1. Introduction

Vagus nerve stimulation (VNS) is an invasive technique consisting of inserting a pulse generator under the skin of the upper chest, which sends signals through the electrodes of the lead to the brain by way of the left vagus nerve [38,45]. In 1997, it was approved by the United States Food and Drug Administration (FDA) as an adjunctive therapy for reducing the frequency of seizures in adults and adolescents who were refractory to antiepileptic medications [46]. Clinical studies conducted on such patients indicated that the technique might possibly affect their mood [14]. This finding, together with the technique's apparent biological plausibility –inasmuch as the vagus nerve affords access to encephalic structures traditionally linked to neuropsychiatric disorders [1,40]– led to studies being undertaken on the technique's indication for use in the treatment of depression. In 2005, the FDA approved the use of VNS for treatment of major depressive disorder (MDD) in patients over the age of 18 years, refractory to other treatments, defined as patients who have not shown an satisfactory response to two or more adequate antidepressant treatments. To date, however, the only evidence of the technique's efficacy and safety when used on this type of patient comes from individual uncontrolled studies [2,6,36,41,43,44,48] of varying duration which have reported initially positive results. This is in contrast to the negative results reported by one, controlled, 10-week study that used a placebo [42]. The study conducted by George et al. in 2005 [17] compared two non-randomised groups of patients with refractory depression, and reported greater antidepressive efficacy for the group that received VNS along with treatment as usual than for the group that only received treatment as usual, after 12 months of intervention.

In view of the high prevalence of depression, coupled with the high rates of subjects refractory to pharmacological treatments [16,26], the consequences in terms of patients' health and quality of life [15,24], and the ensuing cost for health care systems [35,47], a coadjuvant non-pharmacological intervention for the treatment of such patients would appear to be a promising tool for clinical practice.

Accordingly, we felt that this technique's initial relevance in the treatment of refractory depression, the contradictory results

* Corresponding author. School of biomedical and Health Sciences, Universidad Europea de Madrid C/ Tajo, s/n 28670, Villaviciosa de Odón, Madrid, Spain.
E-mail address: joseluis.rodriguez@uem.es (J.L.R. Martin).
[1] Área de Investigación Clínica, Fundación Sociosanitaria de Castilla-La Mancha, C/ Huérfanos Cristinos, 5; 3 drcha, 45003 – Toledo, Castilla-La Mancha, Spain.

0924-9338/$ – see front matter © 2011 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.eurpsy.2011.07.006

observed to date and the lack of conclusive data on its clinical validity called for a systematic review and meta-analysis to ascertain the technique's efficacy and safety on the basis of the best evidence currently available.

2. Methods

2.1. Search strategy and selection criteria

We conducted a systematic review of the scientific literature available until December 2010, with a search of the following databases: Medline/PubMed; Embase; The Cochrane Controlled Trials Register (TRIALS CENTRAL); Pascal Biomed; and CINAL. This search was completed by using standard Internet search engines to comb web pages that were specific to the disorder targeted (depression) or hosted by medical technology firms that marketed the technique. In addition, we also consulted the web pages of registries of ongoing clinical trials, such as Current Controlled Trials and Clinical Trials.gov, and reviewed the references cited by all the pertinent studies located. For search purposes, we used medical subject headings and relevant search terms, such as "depressive disorder", "depression", "depressed patient", "vagal nerve stimulation", "vagus nerve stimulation" and "VNS". In cases where insufficient data were available, authors were contacted directly by e-mail. No restrictions were applied in terms of language or publication status.

Our selection criterion was defined as follows: any randomised controlled trial (RCT) in which the intervention studied was VNS, applied to any value of intensity, frequency and pulse width among patients in whom depressive symptomatology had been measured. Furthermore, since it was common knowledge that relatively few trials had been conducted with this technique, the reviewers also decided to locate all analytical studies in which VNS had been applied to and depressive symptomatology measured in the study patients pre- and postintervention (before-after designs).

After the abstracts of all the studies located had been examined, studies that failed to meet the inclusion criteria were ruled out and the text of all potentially eligible papers was read in full. The two reviewers then assessed the papers separately, and extracted and recorded the data on purpose-designed data-collection sheets. Points of disagreement were settled by discussion or, where necessary, by bringing in an external expert.

2.2. Data-extraction

The two reviewers extracted data on the following aspects: patients' sociodemographic features (age, distribution by gender, disease, baseline depressive symptomatology, study inclusion and exclusion criteria, and concomitant treatments); characteristics of the intervention (current intensity, pulse frequency, pulse width and stimulus on/off cycle); study design (type of study, sample size, duration of intervention, duration of follow-up and description of control group); and outcomes (depressive symptomatology postintervention, adverse events, and withdrawals).

2.3. Outcomes

The following were used as efficacy outcomes: level of depression as measured by depressive symptomatology scales; and percentage of responders, defined as subjects whose depressive symptomatology scores showed a $\geq 50\%$ change over baseline. These outcomes were analysed in the short (≤ 12 weeks), medium (> 12 and < 48 weeks) and long term (> 48 weeks for several time points depending of the global duration) by reference to subjects' baseline pathology, i.e., all depressive subjects/only subjects refractory to medication.

For safety analysis purposes, all adverse events reported by the respective studies were recorded and subsequently analysed by grouping them according to the Coding Symbols for Thesaurus of Adverse Reaction Terms (COSTART) classification [4].

2.4. Quality assessment

To assess clinical-trial quality, a quantitative scale [23] was used in tandem with a qualitative review, with special emphasis being laid on concealment of allocation, blinding methods, description and analysis of withdrawals. Analytical studies without a control group were subjected to a qualitative review.

2.5. Statistical analysis

In the case of RCT studies, initial and final means and standard deviations (SDs), as measured by different depressive symptomatology scales and quantified in the same direction, were recorded for each intervention group. Based on these values, change from baseline was then calculated. In the case of studies with before-after designs, pre- and post-test scores were recorded, and the difference between the standardised change score (d_{change}), and its corresponding variance was computed. For the purposes of this calculation, a Pearson correlation factor (r) of 0.5, between pre- and post-test scores was assumed and a bias-correction factor for small samples applied [3,12,34]. The effect of the treatment was quantified as the standardised mean difference (SMD), with a final estimator of the size of the global effect across all the studies being added. Weighting was performed according to level of study precision, using the inverse variance method [49].

Efficacy was also studied on the basis of the dichotomous variable, "number of responders", by calculating the percentage of responders in each intervention group and the response rate for studies with before-after designs.

Heterogeneity was statistically studied using the χ^2 test and quantified using the I^2 test [21], with values of $P < 0.05$ for the former and $> 50\%$ for the latter being deemed indicative of heterogeneity. To explain heterogeneity by reference to a continuous measure, a meta-regression was performed on the basis of the above assumptions, using the baseline severity of depression of all studies included in the meta-analysis as the explanatory variable and each study's effect size estimated by Hedges' g [13,31] as the dependent variable.

2.6. Ad-hoc analysis for severe adverse events

To ascertain, in the case of before-after studies without a control group, whether adverse events classed as severe were the result of the VNS intervention or, conversely, of the natural disease course, a meta-analysis was performed on the basis of the incidence density of such events (subjects treated X follow-up time). The results were then compared to the incidence density of the same events in active intervention arms of clinical trials conducted with selective serotonin reuptake inhibitors (SSRIs) in the two most recent systematic reviews published on the topic [7,8]. To prevent the incidence of these events being overestimated, all studies were included in the analysis, regardless of whether or not such events had occurred during subject follow-up.

All analyses were performed by fitting both fixed- and random-effects models. The RevMan 5.0 [39] and Comprehensive Meta-analysis statistical programmes were used [10].

3. Results

Our search located a total of 687 references, 546 of which were directly excluded after their respective Abstracts had been examined. The texts of the remaining 141 studies were read in full. Of these, 14 met the pre-established selection criteria and were included in the systematic review [2,6,14,18,19,22,27,30,36,41–44,48], with nine being ultimately meta-analysed [2,19,22,27,36,41,43,44,48]. The remaining five [6,14,18,30,42] could not be included in the meta-analysis for the following reasons: two failed to present the data available for analysis; one reported data in medians; and the other two were clinical trials that were individually described in the results section.

The main reasons for excluding the 127 studies reviewed were as follows: 66 failed to conform to an analytical design (case-report, review, letter to the editor); in 10, the intervention targeted for study was not VNS; one study was undertaken on animals; 21 did not measure depressive symptomatology; five contained duplicated data; 22 were conducted on patients drawn from other studies that had already been included; and, lastly, two studies in which depressive symptomatology had been measured were excluded because they had been undertaken on patients with Alzheimer's Disease, among whom self- or hetero-administered scales are not considered a good measure of level of depression [28].

Of the 14 studies included [2,6,14,18,19,22,27,30,36,41–44,48], eight were conducted on patients with depression [2,27,36,41–44,48] and the remaining six on patients with epilepsy [6,14,18,19,22,30], among whom depressive symptomatology was measured before and after the intervention. In terms of design, one of the studies that included patients with depression was an RCT [42], with the remainder being before-after studies without a control group. Of the studies conducted on patients with epilepsy, one was an RCT [14], and the remaining five had different analytical designs (Fig. 1).

3.1. Study population

The RCT conducted on patients with epilepsy [14] covered a sample of 11 subjects (5 women and 6 men), mean age 31.9 (SD 7.4) years. The RCT with depressive patients [42] covered a sample of 235 subjects. Of these, 139 women and 83 men, mean age 47.0 (SD 9.0) years in the VNS and 45.9 (SD 9.0) years in the placebo group, were included in the analysis. Prior duration of depressive symptomatology was 26.1 (SD = 11.0) and 24.9 (SD = 13.0) years in the active and control groups respectively. A total of 51.8% of patients in the active group and 53.6% in the sham group had received electroconvulsive therapy (ECT) before the study began.

The remaining studies [2,6,14,19,22,27,30,36,41,43,44,48] involved a total of 492 subjects, 304 of whom were women, mean age 33.9 to 50.0 years, with the sole exception of the 2005 Halböök study [18], which was not included in the analyses as it involved children, mean age 11.3 (SD 3.6) years. Of the above total, 149 subjects were diagnosed with epilepsy and 426 had a clinical diagnosis of depression (Table 1). Prior duration of depressive symptomatology as reported by the various studies was: 19.3 (SD = 13.1) years by Rush et al. [41]; 18.1 (SD = 10.9) years by Sackeim et al. [44]; 25.5 (SD = 11.9) years by Rush et al. [43]; 7.2 (SD = 1.9) years by Sperling et al. [48]; and 19.1 (SD = 10.5) years by Bajbouj et al. [2]. The remaining studies reported no data on the duration of previous history of depression. Rush et al. [41], Sackeim et al. [44], Rush et al. [43] and Bajbouj et al. [2] reported data on history of previous use of ECT, which was 63.0, 33.3, 52.7 and 50.0% respectively.

3.1.1. Quantitative results

3.1.1.1. Efficacy

3.1.1.1.1. Randomised controlled trials. The RCT covering a sample of 11 patients with epilepsy. This study with low statistical power reported no statistically significant differences between the two depression-therapy groups studied, i.e., high- versus low-stimulation ($P < 0.10$). No patient presented with previous history of depression or antidepressive treatments. VNS was applied for 24 weeks.

The only RCT that targeted patients with depression, applied VNS for 10 weeks and covered a total of 235 subjects, 222 of whom were included in the efficacy analysis. The principal outcome used was the response rate as measured by the Hamilton Depression Rating Scale (HDRS), which showed no statistically significant differences between the active intervention and placebo groups (OR = 1.61 [95%CI 0.72 to 3.62]; $P = 0.25$). Similarly, no statistically significant differences were observed for the continuous outcome measure, ''depressive symptomatology'', as measured by different depression scales, namely, $P = 0.78$ for the HDRS, $P = 0.23$ for the Montgomery-Asberg Depression Rating Scale (MADRS), and $P = 0.16$ for the 30-item Inventory of Depressive Symptomatology-Self-Report ($IDS\text{-}SR_{30}$). The only statistically significant difference in evidence was for outcome response rates measured by the $IDS\text{-}SR_{30}$ ($P = 0.03$).

3.1.1.1.2. Uncontrolled studies. The meta-analysis performed for the continuous measure, ''level of depression'', included nine before-after studies having a total of 447 subjects. The duration of the intervention and the characteristics of VNS application are described in Table 1. The results showed an SMD = 1.52 (95%CI 0.91 to 2.13), $P < 0.0001$ for a random effects model with a mean intervention duration of 33.87 weeks. The inter-study heterogeneity test yielded a statistical significance of $P < 0.0001$, with $I^2 = 91.49\%$ (Fig. 2A).

On stratifying by intervention duration, analysis of short-term efficacy (≤ 12 weeks) included five studies with a total of 194 subjects, yielded an SMD = 1.74 (1.00 to 2.47), $P < 0.0001$, and indicated high heterogeneity ($I^2 = 87.15\%$). In the medium term (> 12 weeks and < 48 weeks), two studies with 38 subjects were included and yielded an SMD = 0.35 (−0.40 to 1.11), a non-significant $P = 0.36$ and indicated heterogeneity ($I^2 = 55.76\%$). In the long term (48 weeks), the four studies involving 347 subjects were meta-analysed and showed an SMD = 2.27 (1.31 to 3.22), $P = 0.001$ and high heterogeneity ($I^2 = 92.29\%$). Finally, at 96 weeks, two studies involving 133 subjects showed an SMD = 2.52 (2.16 to 2.88), $P < 0.0001$;. $I^2 = 0.00\%$.

When the analysis was repeated and restricted to studies undertaken on patients with depression refractory to standard treatments, it covered seven studies having a total of 399 subjects, and yielded an SMD = 1.94 (1.36 to 2.52), $P < 0.0001$ and indicated a heterogeneity of $I^2 = 83.86\%$.

For the dichotomous variable, ''response rate'', six before-after studies reported satisfactory data for this outcome and were included, with a total of 408 subjects. The results showed a response rate of 31.8% (23.2% to 41.8%), $P = 0.001$ for a random effects model with a mean intervention duration of 20.13 weeks. The heterogeneity test yielded a value of $I^2 = 66.00\%$ (Fig. 3).

Analysis of subjects with depression refractory to standard treatments included five studies having a total of 380 subjects. This analysis yielded a response rate of 33.5% (23.9% to 44.8%), $P = 0.005$, and indicated high heterogeneity ($I^2 = 69.82\%$).

3.1.1.1.3. Explanation of heterogeneity: meta-regression. The results for the principal outcome, ''level of depression'', showed great disparity between the data furnished by the RCT, which found no statistically significant differences, and the size of the global effect

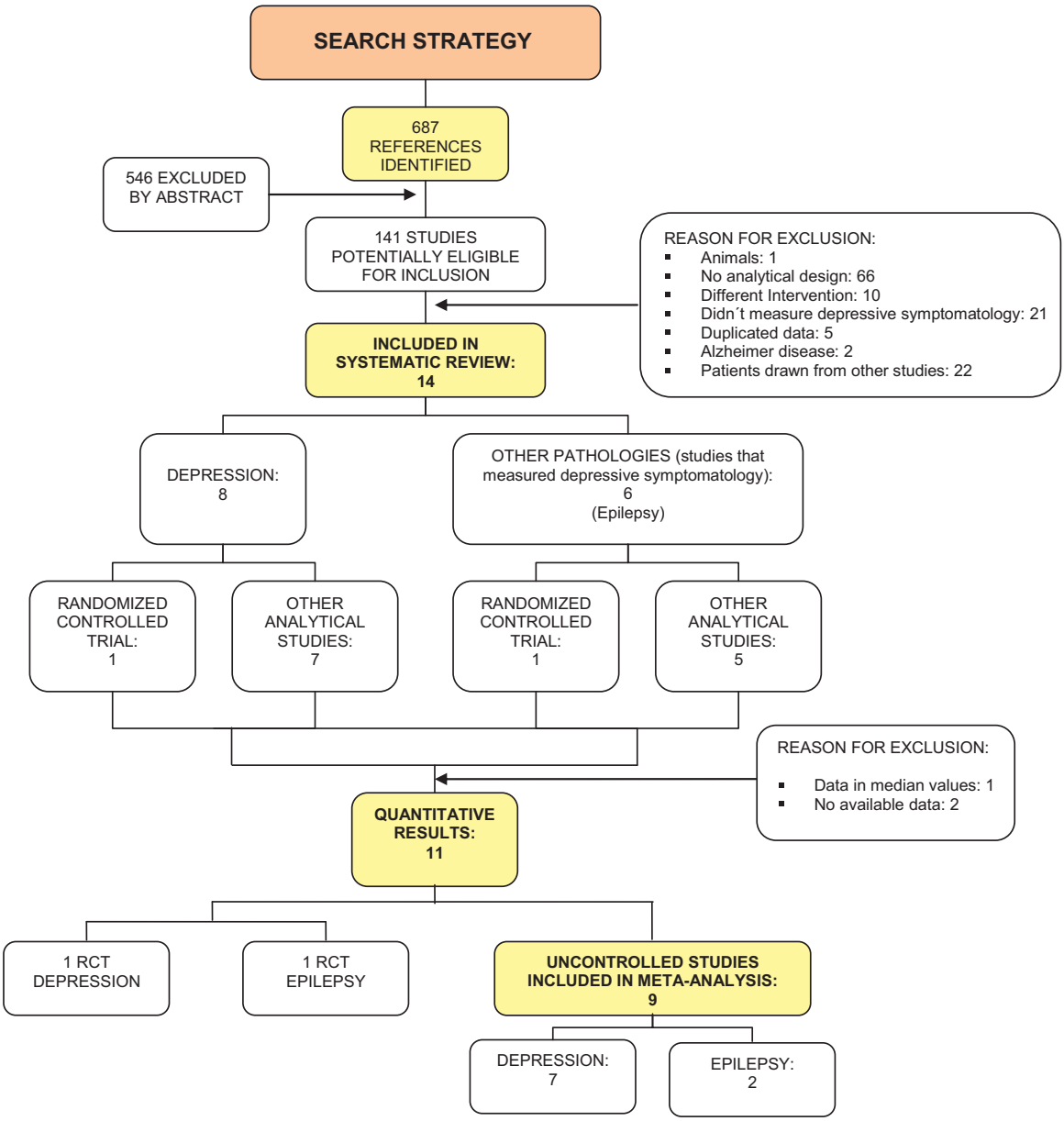

Fig. 1. Flow of studies through selection process. *RCT:* randomized controlled trial.

observed for the nine meta-analysed before-after studies. In addition, all these recent studies indicated high heterogeneity.

To study the cause of this heterogeneity, different moderator variables were studied, and subgroup analyses were performed for the following variables: subjects' pathology and baseline level of depression; year of publication; and study size. Most of the studies applied VNS within a range of current intensity values, depending on patients' tolerance. Similarly, they reported the parameters of application of the technique as a range of values, thereby rendering it impossible to perform the analysis using current intensity as a moderator variable. Finally, the studies were listed in order of subjects' baseline level of depression, and a possible relationship between this variable and the effect size of each of the studies included was graphically observable (Fig. 2B).

A meta-regression was thus performed, using subjects' mean baseline severity of depression in the studies included in the meta-analysis as the explanatory variable and each study's effect size estimated by Hedges' g [13] as the dependent variable (Fig. 4.). This regression model showed a slope of 0.07, meaning that for every one-point increase in baseline depression level, there was a 0.07-point increase in the magnitude of the observed effect. The adjusted coefficient of determination (R^2_{Adj}) was 0.84, which implies that an 84% variation in effect size across the studies was explained by baseline severity of depression ($P < 0.0001$). The meta-regression was repeated including only the 7 studies conducted on patients with depression, and yielded very similar results (slope = 0.11; $R^2_{Adj} = 0.66$; [$P < 0.0001$]).

3.1.1.2. Safety

3.1.1.2.1. Randomised controlled trials. The RCT conducted on patients with epilepsy, yielded no information on the technique's safety among the study patients.

The only RCT included which targeted patients with depression, reported three withdrawals, all in the VNS group and all due to severe adverse events (hoarseness, explantation due to infection, and suicide). With respect to the remaining adverse events, subjects who received VNS reported higher frequencies of

Table 1
Characteristics of included studies.

Study	Design		*N* (duration)	Pathology	VNS intervention				Baseline characteristics		
					Current intensity	Pulse frequency	Pulse width	Stimulation duration (on/off)	Gender (F/M)	Age[c,d]	Severity of depression[d] (scale)
Elger, 2000	RCT	VNS low stimulation condition	5 (24 w)	Epilepsy	Subjective limits of signal detection	No information	250 μs	60/600 s	5/6	31.9 (7.4)	10.8 (4.4) (MADRS)
		VNS high stimulation condition	6 (24 w)		1.75 mA (max)	No information	500 μs	30/300 s			10.8 (5.9) (MADRS)
Harden, 2000[a,b]	Experimental, non-randomized	VNS	20 (12 w)	Epilepsy	No information	No information	No information	No information	14/6	39.0 (9.1)	12.9 (8.2) (HRDS)
		Antieplileptic drugs	20 (12 w)						14/6	40.2 (13.3)	12.8 (8.8) (HDRS)
Hoppe, 2001[a]	Before and after		28 (28.8 w)	Epilepsy	No information	No information	No information	No information	10/18	35.4 (10.8)	9.7 (6.5) (BDI)
Chavel, 2003	Before and after		29 (96 w)	Epilepsy	No information	No information	No information	No information	12/17	33.9 (10.3)	(BDI)
Hallböök, 2005	Before and after		15 (36 w)	Epilepsy	0.25–1.5 mA	30 Hz	500 μs	30/300 s	5/10	11.3 (3.6)	10.3 (3.51)[e] (DSRS)
McGlone, 2008	Case control	VNS	16 (48 w)	Epilepsy	0.25–3.0 mA	30 Hz	500 μs	30/300 s	7/9	35.0 (8.0)	(GDS)
		Medical management	11 (48 w)						6/3	37.0 (6.7)	
		Cerebral resective surgery	10 (48 w)						6/4	36.0 (12.7)	
Rush, 2000[a]	Before and after		30 (10 w)	Depression	0.25–3.0 mA	20–30 Hz	500 μs	30/300 s	20/10	47.5 (7.5)	38.0 (5.5) (HDRS)
Sackeim, 2001[a]	Before and after		60 (10 w)	Depression	0.25–3.0 mA	20 Hz	500 μs	30/300 s	39/21	46.8 (8.7)	36.8 (5.8) (HDRS)
ÓKeane, 2005[a]	Before and after		11 (12 w)	Depression	1.5 mA (max)	No information	No information	30/300 s	9/2	43.09 (8.28)	35.5 (3.95) (HDRS)
Rush, 2005a	RCT	VNS	119 (10 w)	Depression	0.67 (0.25–3.5) mA	20 Hz	500 μs	30/300 s	66/46	47.0 (9.0)	28.8 (5.3) (HDRS)
		Sham	116 (10 w)						73/37	45.9 (9.0)	29.7 (5.2) (HDRS)
Rush, 2005b[a]	Before and after		205 (48 w)	Depression	1.0 mA	20 Hz	500 μs	30/300 s	131/74	46.3 (8.9)	28.0 (5.7) (HDRS)
Zobel, 2005[a]	Before and after		12 (4 w)	Depression	0.75–1.50 mA	20 Hz	500 μs	30/300 s	6/6	48.42 (15.28)	23.7 (5.6) (HDRS)
Corcoran, 2006[a]	Before and after		11 (24 w)	Depression	0.25 mA	20 Hz	500 μs	30/300 s	8/3	43.0 (8.72)	36.36 (3.44) (HDRS)
Neuhaus, 2007[a]	Before and after		13 (10 w)	Depression	0.94 ± 0.46 mA	20 Hz	250 μs	30/300 s	10/3	47.77 (11.77)	24.15 (3.75) (HDRS)
Marangell, 2008[a]	Before and after		10 (40 w)	Depression	0.25–2.0 mA	20 Hz	500 μs	30/300 s	7/2	46.89 (10.20)	20.9 (7.2) (HDRS)
Schlaepfer, 2008[a]	Before and after		74 (48 w)	Depression	0.25–2.0 mA	20 Hz	500 μs	30/300 s	50/24	47.4 (11.7)	34.0 (5.8) (HDRS)

VNS: vagus nerve stimulation; RCT: randomized controlled trial; w: weeks; max: maximum; μs: microseconds; s: seconds; mA: milliampère; Hz: Hertz; MADRS: Montgomery-Asberg Depression Rating Scale; HDRS: Hamilton Depression Rating Scale; BDI: Beck Depression Inventory; DSRS: Birleson Depression Self-Rating Scale; GDS: Geriatric Depression Scale.

[a] Studies included in the meta-analysis.

[b] Only VNS arms were used in the analysis.

[c] Years.

[d] Mean (SD).

[e] Average of medians.

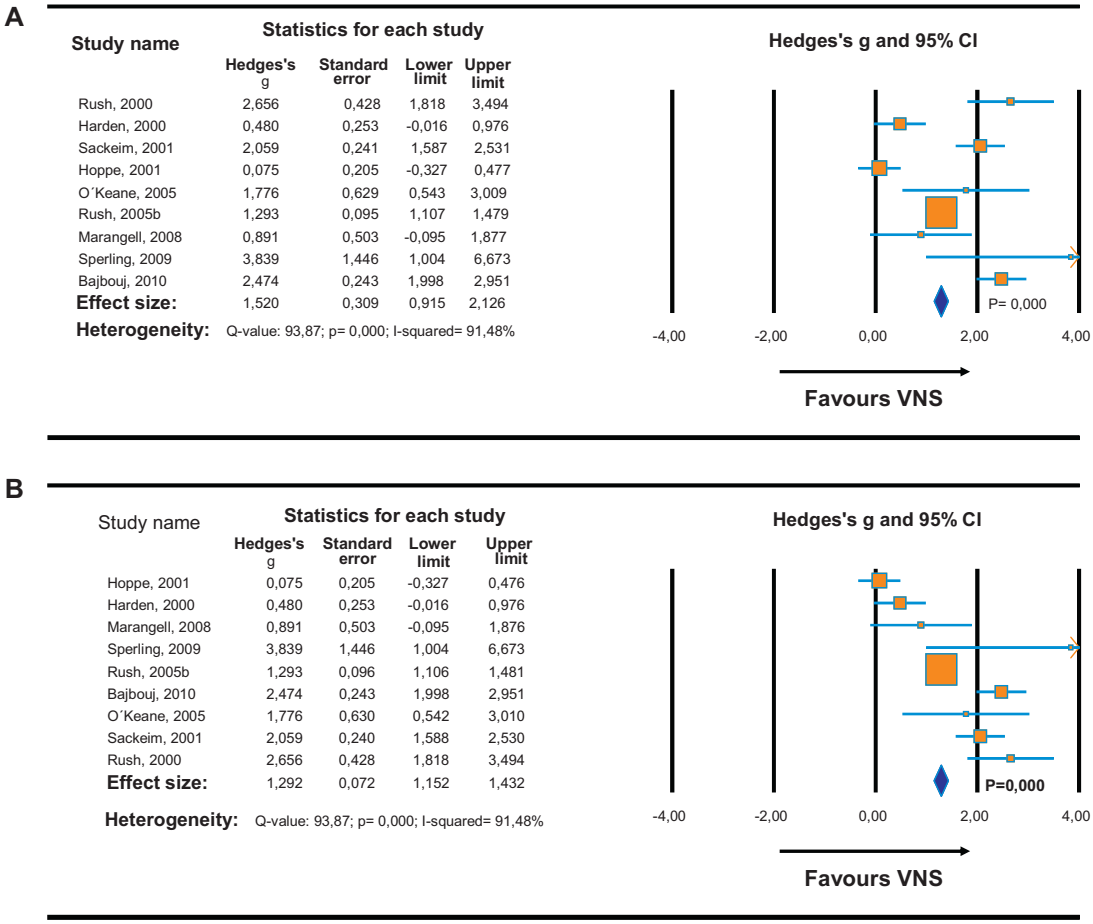

Fig. 2. A. Meta-analysis of efficacy: depressive symptomatology (studies listed by year of publication). B Meta-analysis of Fig. 2.A. listening the studies by patientś baseline severity of depression.

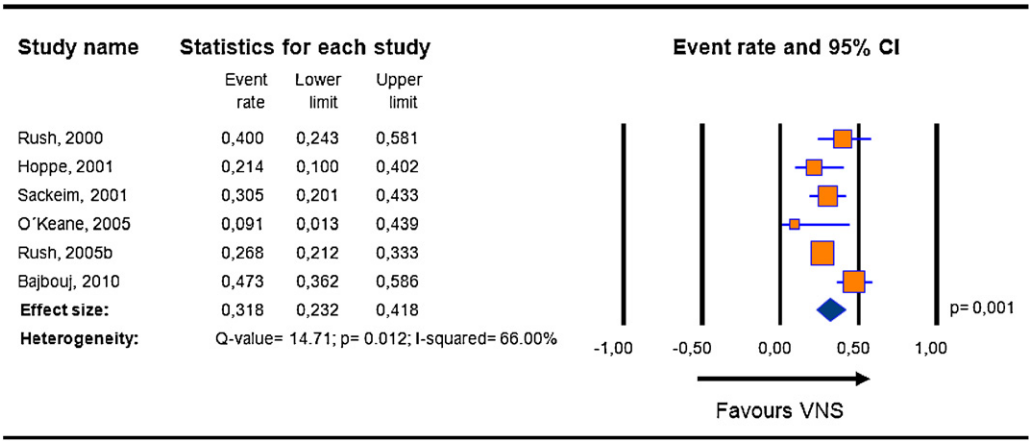

Fig. 3. Meta-analysis of efficacy: response rates.

appearance of all the events described, with rises of 30% for voice alteration, 20% for increased cough, 9% for dyspnea, 10% for dysphagia, 11% for neck pain, 6% for paresthesia, 6% for vomiting, 9% for laryngismus, 5% for dyspepsia, 6% for wound infection and 2% for palpitations.

3.1.1.2.2. Uncontrolled studies. Of the 12 studies covered, only six reported quantitative data on adverse events, with "suicide or attempted suicide", "mania or hypomania", and "worsening depression with need for hospitalisation" being included in the category of severe adverse events.

For suicide or attempted suicide, four before-after studies were meta-analysed, yielding a cumulative incidence of 4.6% for the 348 subjects included (2.8 to 7.3%), $P < 0.0001$, without any statistically appreciable heterogeneity ($I^2 = 0.00\%$). As these studies had no control groups, in order to study the intervention's influence on the occurrence of this event, without any possible confounding

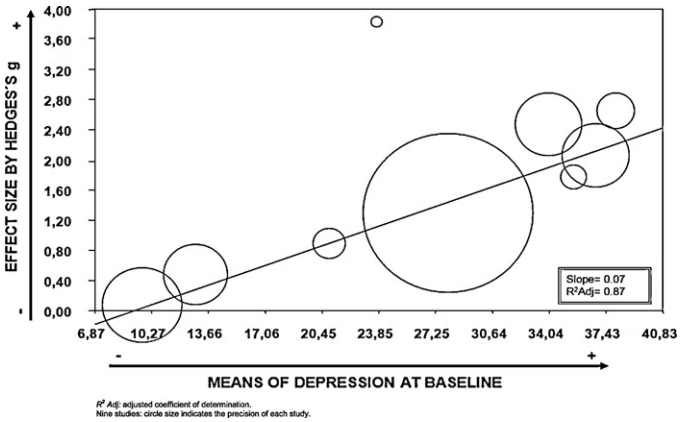

Fig. 4. Meta-regression: explanation of heterogeneity using baseline level of depression as moderator variable.

caused by natural disease course, incidence density was calculated for all the meta-analysed studies. This resulted in a figure of 0.090 suicides or attempted suicides per 100 subjects per week during the intervention period (0.056 to 0.145), and statistical homogeneity ($I^2 = 0.00\%$). In view of this being a meta-analysis of uncontrolled studies conducted on subjects with a baseline pathology entailing a risk of suicide, the incidence density found for this event was compared against that of a cohort comprising all subjects with the same condition and inclusion criteria, who had received treatment with SSRIs in the randomised clinical trials included in the two most recent reviews published on these drugs [7,8]. When the relevant analysis was performed, it showed an incidence density for the 1991 meta-analysed subjects of 0.093 suicides or attempted suicides per 100 subjects per week (0.065 to 0.133; $I^2 = 0.00\%$), thus highlighting the fact that the number of suicides or attempted suicides among depressive subjects was practically identical for both therapeutic interventions.

Mania or hypomania showed a cumulative incidence of 2.7% (1.4 to 5.1%; $I^2 = 0.00\%$) for the four studies and 368 subjects that were meta-analysed. Incidence density was 0.094 cases per 100 subjects per week (0.049 to 0.180; $I^2 = 0.00\%$).

For the same four studies, cumulative incidence of hospitalisations due to worsening depression was 12.1% (8.6 to 16.7%; $I^2 = 17.13\%$). Incidence density was 0.225 cases per 100 subjects/week (0.168 to 0.303).

Using the COSTART classification [4], the remaining adverse events reported were classified in accordance with an anatomo-physiological system and the specific function(s) that these affected. The respective short-, medium- and long-term rates were calculated for each event (Table 2).

4. Discussion

This systematic review found that there are insufficient data currently available to describe VNS as an effective technique for

Table 2
Adverse events: most frequently reported adverse events and serious adverse events.

Most frequently reported adverse events			
Body system Adverse events	Effect size		
	Short term: ≤ 12 w[a]	Medium term: $>12, <48$ w[a]	Long term: 48 w[a]
Body as a whole			
Incision site pain	19.3 (4.5 to 41.3)	No data	No data
Headache	11.9 (4.3 to 28.7)	3.9 (2.1 to 6.8)	3.7 (2.0 to 6.8)
Pain	15.8 (6.8 to 32.3)	6.9 (4.4 to 10.6)	6.2[b]
Chest pain	10.8 (4.2 to 24.9)	No data	No data
Neck pain	15.5 (12.2 to 19.5)	9.6 (6.3 to 14.5)	13.1 (9.5 to 17.7)
Infection	5.6 (2.4 to 12.8)	No data	No data
Respiratory System			
Voice Alteration	67.3 (50.7 to 80.5)	19.4 (0.6 to 90.8)[d]	22.9 (6.7 to 55.2)[d]
Pharyngitis	11.6 (5.8 to 21.8)	3.9 (2.2 to 6.9)	5.2 (3.1 to 8.6)
Dyspnea	15.2 (11.9 to 19.1)	12.7 (7.0 to 22.1)	15.0 (11.1 to 20.0)
Coughing	23.4 (15.6 to 33.6)	10.6 (4.6 to 22.7)	6.0 (3.7 to 9.5)
Digestive System			
Dysphagia	13.2 (10.2 to 17.0)	8.4[b]	4.1 (2.3 to 7.3)
Dyspepsia	7.1 (2.8 to 16.8)	3.3[b]	3.3[b]
Nausea	5.9 (3.8 to 9.0)	2.5 (1.2 to 5.1)	2.3 (1.0 to 5.0)
Nervous System			
Dizziness	7.0 (3.2 to 14.8)	No data	No data
Paresthesia	6.9 (2.9 to 15.7)	6.7[b]	4.3[b]
Hypertonia	10[b]	No data	3.3[b]
Twitching	4.4 (2.0 to 9.6)	No data	No data
Insomnia	4.5 (2.6 to 7.5)	2.2[b]	0.95[b]
Skin and appendages			
Rash-Pruritis	0.7 (3.9 to 12.2)	No data	No data

Serious adverse events			
Events	Number of studies (*N*)	Cumulative incidence[a]	Incidence density[c]
Suicide or suicide attempt	4 (348)	4,6 (2.8 to 7.3)	0,085 (0.052 to 0.14)
Worsening depression	4 (368)	12.1 (8.6 to 16.7)	0.225 (0.168 to 0.303)
Mania or hipomania	4 (368)	2.7 (1.4 to 5.1)	0.094 (0.049 to 0.180)

[a] Data in %.
[b] Only one study reported available data.
[c] Events per 100 subjects per week.
[d] No statistical significance; number of studies (without parentheses) and sample size (number of subject) showed in parentheses.

treatment of depression. In addition, it cannot be ruled out that the positive results observed in the uncontrolled studies might have been mainly due to a placebo effect [5,50]. These conclusions are based on the study of the two principal outcome measures provided by the studies included in the review, namely: responder rate, construed as a remission of at least 50% in depressive symptomatology; and the difference between subjects' depressive symptomatology at the beginning and end of the intervention, as measured by psychometric scales.

To date, the results of studies that have assessed the therapeutic potential of VNS in the treatment of depression have been essentially based on before-after designs without a control group, with HDRS figures showing a responder rate of up to 42% at two years and remissions in symptomatology of 22% [37], results similar to those observed in our meta-analysis. Nevertheless, the ability of before-after designs to show causality is very limited [20] and renders it impossible for the possible effect generated by the intervention to be quantified and distinguished from the placebo effect, regression toward the mean, spontaneous remission, or the Hawthorne effect due to patients feeling observed during the conduct of follow-up studies [29]. In 2008, Daban et al. published a systematic review [11] which included 18 uncontrolled studies and a single randomised clinical trial. They performed no meta-analyses and concluded that, despite the promising results yielded principally by open studies, the current evidence was not conclusive and more evidence for decision-making was therefore required.

This experimental weakness has highlighted the fact that any positive effect observed is dependent by as much as 84% on patients' baseline level of depression on entering the study. In other words, subjects with major or more severe depressive symptomatology at the start of studies obtained better results than did subjects with less severe symptomatology, and these results were linear for all the studies, since –as shown by the meta-regression– for every one-point increase in baseline depression levels, there was an increase of 0.07 points in the magnitude of the positive effect found. When the two non-randomised groups, namely, that administered VNS plus treatment as usual and that solely administered treatment as usual, were compared after 12 months of treatment, the VNS group was seen to have received a greater antidepressive benefit [17].

In contrast, in cases where a random control group was used, these results were not in evidence. In the only placebo-controlled trial included in our review, there were no significant differences between the two groups of patients (VNS versus sham) during the randomised blind phase of the study, and in a sample of 222 subjects, the number of responders was 17 in the VNS and 11 in the placebo group. Yet, when the stimulation was activated in both groups (VNS and sham), with the aim of conducting an open, long-term, follow-up study, a response was observed in 55 subjects (27%) at one year of treatment [43]. This response percentage was similar to that of the remaining follow-up studies and to those observed by our review based on the weighted sum –overall effect– of all the uncontrolled studies. Armed with these data, one cannot rule out a possible regression-toward-the-mean effect, whereby the first baseline measure taken in the most severe patients would tend towards the centre of the distribution in successive measures (i.e., becoming less extreme). In such cases, a simple placebo effect in all the subjects would logically mean that the more severe a subject was initially, the greater the perception of efficacy would then be.

Insofar as duration of the intervention is concerned, it should be noted that, despite the considerable inter-study heterogeneity found, the positive results of this technique were obtained in the short and long term, and no relevant effects were observed in the medium term (12 to 48 weeks). Whereas the short-term results may in great measure depend on the above-mentioned placebo effects, the long-term effects might be due to a long latency period before the therapeutic effect becomes evident. Nonetheless, in view of the research designs used to date, one cannot ignore the fact that the positive long-term effect could be due to natural disease course. Moreover, the meta-regression showed that the main variable for obtaining positive results was the severity of patients' condition at study commencement, regardless of the duration of the intervention. The response rates for subjects who were refractory proved to be similar to those observed for subjects who were not refractory to pharmacological medication.

With a population prevalence of 10% [25], MDD is likely to continue being one of the main causes of disability in years to come [32]. Such subjects, with a remission rate of only 30% and response to conventional antidepressants of 45 to 55%[33], should be a priority target in the search for interventions that serve to enhance their quality of life in the short, medium and long term. Accordingly, if this technique, with a 32% response among subjects refractory to medication, is regarded as a possible solution–even without its placebo effect being properly ascertained–consideration should also be given to its cost, which exceeds $ 25000 per implantation [9], its invasive nature, requiring two hours of surgery, and the risks associated with the intervention or its side-effects.

Among the adverse events observed, the most severe were mania or hypomania (2.7%), hospitalisation due to worsening of the depression (12.1%), and suicide or attempted suicide (4.6%). Since the study designs prevented us from establishing whether these percentages were due to the intervention or, conversely, to natural disease course, we obtained the incidence density of each of these events in the studies included. In the case of suicide, this incidence density was then compared to that of the cohort formed by the active arms of subjects having the same characteristics but undergoing SSRI treatment in the RCTs included in the systematic reviews [7,8]. The results failed to show significant differences between the two cohorts. Hence, based on current data, VNS would not appear to provoke a risk of suicidal conduct any higher than that observed for subjects treated exclusively with SSRIs. However, according to data from the RCT included in the review, note should also be taken of the fact that, among the possible side-effects, the following affected higher percentages of patients in the VNS versus the placebo group: voice alteration, dyspnea, dysphagia, paresthesia, laryngismus and wound infection, 6%.

The FDA approved the use of this technique in 2005 as a long-term adjuvant therapy among patients with chronic or recurrent depression who did not respond adequately to conventional antidepressants [47]. Yet, this decision was not based on results furnished by RCTs, as is mandatory for new pharmacological antidepressant treatments, so that the grade of evidence as to the benefit or causality of the effects observed with this technique has not been shown. For the selfsame reason, perhaps there may not be sufficient information available as to the technique's potential risks, which in this case could be more important than its beneficial effects. Before being approved in clinical practice, invasive interventions such as that assessed here should yield outcomes with a high degree of causality. Although it might seem that a non-pharmacological therapy approved as adjuvant in refractory patients could only have positive booster effects (even as a simple placebo effect) vis-à-vis the main treatment, the data provided by studies on adverse events or worsening depression with need for hospitalisation, render it essential for the positive and negative effects of this technique to be ascertained in exactly the same way as for any pharmacological treatment, i.e., exclusively using RCT designs.

To sum up, our findings show that, while the VNS technique has yet to show evidence of efficacy or effectiveness as a therapeutic

treatment for depression, it does nevertheless entail a substantial risk that its positive effects may be mediated by the placebo effect, regression toward the mean, spontaneous remission or the Hawthorne effect. These results are mainly based on uncontrolled studies, with small or medium sample sizes and intermediate quality levels. On the other hand, another limitation of our review is that the duration of the RCT included was 10 weeks. A solid evidence of the effect of this therapy should therefore be exclusively based on long-term clinical trials with a control group, aimed at monitoring the possible latency involved in the technique's producing its desired effect, with special attention being paid to any adverse events observed, since at present it is not at all clear whether the potential benefit of applying this technique in clinical practice is outweighed by the possible harm.

Disclosure of interest

The authors declare that they have no conflicts of interest concerning this article.

References

[1] Adrienne E, Dorr AE, Debonnel G. Effect of vagus nerve stimulation on serotonergic and noradrenergic transmission. J Pharmacol Exp Ther 2006; 318:890–8.
[2] Bajbouj M, Merkl A, Schlaepfer TE, Frick C, Zobel A, Maier W, et al. Two-year outcome of vagus nerve stimulation in treatment-resistant depression. J Clin Psychopharmacol 2010;30(3):273–81.
[3] Becker BJ. Synthesizing standardized mean-change measures. Br J Math Stat Psychol 1988;41:257–78.
[4] Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). Drug Saf 1999;20(2):109–17.
[5] Carpenter Jr WT. Placebo effect in depression. Am J Psychiatry 2009; 166(8):935.
[6] Chavel SM, Westerveld M, Spencer S. Long-term outcome of vagus nerve stimulation for refractory partial epilepsy. Epilepsy Behav 2003;4(3):302–9.
[7] Cipriani A, La FT, Furukawa TA, Signoretti A, Nakagawa A, Churchill R, et al. Sertraline versus other antidepressive agents for depression. Cochrane Database Syst Rev 2009;(2):CD006117.
[8] Cipriani A, Santilli C, Furukawa TA, Signoretti A, Nakagawa A, McGuire H, et al. Escitalopram versus other antidepressive agents for depression. Cochrane Database Syst Rev 2009;(2):CD006532.
[9] Cohen LJ, Allen Jr JC. Estimating the potential savings with vagus nerve stimulation for treatment-resistant depression: a payer perspective. Curr Med Res Opin 2008;24(8):2203–17.
[10] Comprehensive Meta-Analysis. Version 2.2.048 [computer program]. USA: 2008.
[11] Daban C, Martinez-Aran A, Cruz N, Vieta E. Safety and efficacy of Vagus Nerve Stimulation in treatment-resistant depression. A systematic review. J Affect Disord 2008;110(1–2):1–15.
[12] Dunlap WPCJMVJBBMJ. Meta-analysis of experiments with matched groups or repeated measures designs. Psychol Methods 1996;1:170–7.
[13] Egger M, Smith GD, Altman DG. Systematic reviews in health care. Meta-analysis in context, 2nd ed, Cornwall: Bodmin; 2001.
[14] Elger G, Hoppe C, Falkai P, Rush AJ, Elger CE. Vagus nerve stimulation is associated with mood improvements in epilepsy patients. Epilepsy Res 2000;42(2–3):203–10.
[15] Fekadu A, Wooderson SC, Markopoulo K, Donaldson C, Papadopoulos A, Cleare AJ. What happens to patients with treatment-resistant depression? A systematic review of medium to long-term outcome studies. J Affect Disord 2009;116(1–2):4–11.
[16] Gaynes BN. Identifying difficult-to-treat depression: differential diagnosis, subtypes, and comorbidities. J Clin Psychiatry 2009;70(Suppl. 6):10–5.
[17] George MS, Rush AJ, Marangell LB, Sackeim HA, Brannan SK, Davis SM, et al. A one-year comparison of vagus nerve stimulation with treatment as usual for treatment-resistant depression. Biol Psychiatry 2005;58(5):364–73.
[18] Hallbook T, Lundgren J, Stjernqvist K, Blennow G, Stromblad LG, Rosen I. Vagus nerve stimulation in 15 children with therapy resistant epilepsy; its impact on cognition, quality of life, behaviour and mood. Seizure 2005;14(7):504–13.
[19] Harden CL, Pulver MC, Ravdin LD, Nikolov B, Halper JP, Labar DR. A pilot study of mood in epilepsy patients treated with vagus nerve stimulation. Epilepsy Behav 2000;1(2):93–9.
[20] Harris AD, Bradham DD, Baumgarten M, Zuckerman IH, Fink JC, Perencevich EN. The use and interpretation of quasi-experimental studies in infectious diseases. Antimicrob Resist 2004;38:1586–91.
[21] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557–60.
[22] Hoppe C, Helmstaedter C, Scherrmann J, Elger CE. Self-reported mood changes following 6 months of vagus nerve stimulation in epilepsy patients. Epilepsy Behav 2001;2(4):335–42.
[23] Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996;17(1):1–12.
[24] Kennedy SH, Giacobbe P. Treatment resistant depression–advances in somatic therapies. Ann Clin Psychiatry 2007;19(4):279–87.
[25] Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003;289(23):3095–105.
[26] Little A. Treatment-resistant depression. Am Fam Physician 2009;80(2): 167–72.
[27] Marangell LB, Suppes T, Zboyan HA, Prashad SJ, Fischer G, Snow D, et al. A 1-year pilot study of vagus nerve stimulation in treatment-resistant rapid-cycling bipolar disorder. J Clin Psychiatry 2008;69(2):183–9.
[28] Mayer LS, Bay RC, Politis A, Steinberg M, Steele C, Baker AS, et al. Comparison of three rating scales as outcome measures for treatment trials of depression in Alzheimer disease: findings from DIADS. Int J Geriatr Psychiatry 2006;21(10):930–6.
[29] McCarney R, Warner J, Iliffe S, van HR, Griffin M, Fisher P. The Hawthorne Effect: a randomised, controlled trial. BMC Med Res Methodol 2007;7:30.
[30] McGlone J, Valdivia I, Penner M, Williams J, Sadler RM, Clarke DB. Quality of life and memory after vagus nerve stimulator implantation for epilepsy. Can J Neurol Sci 2008;35(3):287–96.
[31] McGrath RE, Meyer GJ. When effect sizes disagree: the case of r and d. Psychol Methods 2006;11(4):386–401.
[32] Michaud CM, Murray CJ, Bloom BR. Burden of disease–implications for future research. JAMA 2001;285(5):535–9.
[33] Milby AH, Halpern CH, Baltuch GH. Vagus nerve stimulation for epilepsy and depression. Neurotherapeutics 2008;5(1):75–85.
[34] Morris SBDRP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychol Methods 2002;7:105–25.
[35] Nemeroff CB. Prevalence and management of treatment-resistant depression. J Clin Psychiatry 2007;68(Suppl. 8):17–25.
[36] O'Keane V, Dinan TG, Scott L, Corcoran C. Changes in hypothalamic-pituitary-adrenal axis measures after vagus nerve stimulation therapy in chronic depression. Biol Psychiatry 2005;58(12):963–8.
[37] Rado J, Janicak PG. Vagus nerve stimulation for severe depression. J Psychosoc Nurs Ment Health Serv 2007;45(7):43–51.
[38] Ramani R. Vagus nerve stimulation therapy for seizures. J Neurosurg Anesthesiol 2008;20(1):29–35.
[39] Review Manager (RevMan) version 5.0 [computer program]. Copenhagen: The Cochrane Collaboration; 2008.
[40] Rush AJ, Siefert SE. Clinical issues in considering vagus nerve stimulation for treatment-resistant depression. Exp Neurol 2009;219(1):36–43.
[41] Rush AJ, George MS, Sackeim HA, Marangell LB, Husain MM, Giller C, et al. Vagus nerve stimulation (VNS) for treatment-resistant depressions: a multicenter study. Biol Psychiatry 2000;47(4):276–86.
[42] Rush AJ, Marangell LB, Sackeim HA, George MS, Brannan SK, Davis SM, et al. Vagus nerve stimulation for treatment-resistant depression: a randomized, controlled acute phase trial. Biol Psychiatry 2005;58(5):347–54.
[43] Rush AJ, Sackeim HA, Marangell LB, George MS, Brannan SK, Davis SM, et al. Effects of 12 months of vagus nerve stimulation in treatment-resistant depression: a naturalistic study. Biol Psychiatry 2005;58(5):355–63.
[44] Sackeim HA, Rush AJ, George MS, Marangell LB, Husain MM, Nahas Z, et al. Vagus nerve stimulation (VNS) for treatment-resistant depression: efficacy, side effects, and predictors of outcome. Neuropsychopharmacology 2001; 25(5):713–28.
[45] Sakas DE, Korfias S, Nicholson CL, Panourias IG, Georgakoulias N, Gatzonis S, et al. Vagus nerve stimulation for intractable epilepsy: outcome in two series combining 90 patients. Acta Neurochir Suppl 2007;97(Pt 2):287–91.
[46] Schachter SC. Vagus nerve stimulation therapy summary: five years after FDA approval. Neurology 2002;59(6 Suppl. 4):S15–20.
[47] Shelton RC, Osuntokun O, Heinloth AN, Corya SA. Therapeutic options for treatment-resistant depression. CNS Drugs 2010;24(2):131–61.
[48] Sperling W, Reulbach U, Kornhuber J. Clinical benefits and cost effectiveness of vagus nerve stimulation in a long-term treatment of patients with major depression. Pharmacopsychiatry 2009;42(3):85–8.
[49] Sutton AJ. AKJDSTSF. Methods for Meta- Analysis in Medical Research. Chichester: Wiley series in probability and statistics; 2000.
[50] Wallace M. Physiology of the placebo effect, and the evidence for changes in brain metabolism. N Z Med J 2010;123(1312):54–60.