Journal of Psychiatric Research 46 (2012) 574–581

Contents lists available at SciVerse ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/psychires

Facing depression with botulinum toxin: A randomized controlled trial

M. Axel Wollmer [a,*], Claas de Boer [b], Nadeem Kalak [a], Johannes Beck [a], Thomas Götz [a], Tina Schmidt [b], Muris Hodzic [c], Ursula Bayer [a], Thilo Kollmann [a], Katja Kollewe [d], Daniela Sönmez [b], Katja Duntsch [b], Martin D. Haug [e], Manfred Schedlowski [f], Martin Hatzinger [g], Dirk Dressler [d], Serge Brand [a], Edith Holsboer-Trachsler [a], Tillmann H.C. Kruger [b]

[a] Psychiatric Hospital of the University of Basel, Wilhelm-Klein-Str. 27, 4012 Basel, Switzerland
[b] Department of Psychiatry, Social Psychiatry and Psychotherapy, Medical School Hannover, Germany
[c] Praxis am Riehenring, Basel, Switzerland
[d] Department of Neurology, Medical School Hannover, Germany
[e] Department of Plastic, Reconstructive and Aesthetic Surgery, University Hospital Basel, Switzerland
[f] Institute of Medical Psychology and Behavioral Immunobiology, University Clinic Essen, Germany
[g] Psychiatric Hospital, Solothurn, Switzerland

ARTICLE INFO

Article history:
Received 30 November 2011
Received in revised form
24 January 2012
Accepted 30 January 2012

Keywords:
Facial feedback
Emotional contagion
Major depression
Botulinum neurotoxin
Randomized controlled study

ABSTRACT

Positive effects on mood have been observed in subjects who underwent treatment of glabellar frown lines with botulinum toxin and, in an open case series, depression remitted or improved after such treatment. Using a randomized double-blind placebo-controlled trial design we assessed botulinum toxin injection to the glabellar region as an adjunctive treatment of major depression.

Thirty patients were randomly assigned to a verum (onabotulinumtoxinA, $n = 15$) or placebo (saline, $n = 15$) group. The primary end point was change in the 17-item version of the Hamilton Depression Rating Scale six weeks after treatment compared to baseline.

The verum and the placebo groups did not differ significantly in any of the collected baseline characteristics. Throughout the sixteen-week follow-up period there was a significant improvement in depressive symptoms in the verum group compared to the placebo group as measured by the Hamilton Depression Rating Scale ($F_{(6,168)} = 5.76$, $p < 0.001$, $\eta^2 = 0.17$). Six weeks after a single treatment scores of onabotulinumtoxinA recipients were reduced on average by 47.1% and by 9.2% in placebo-treated participants ($F_{(1,28)} = 12.30$, $p = 0.002$, $\eta^2 = 0.31$, $d = 1.28$). The effect size was even larger at the end of the study ($d = 1.80$). Treatment-dependent clinical improvement was also reflected in the Beck Depression Inventory, and in the Clinical Global Impressions Scale.

This study shows that a single treatment of the glabellar region with botulinum toxin may shortly accomplish a strong and sustained alleviation of depression in patients, who did not improve sufficiently on previous medication. It supports the concept, that the facial musculature not only expresses, but also regulates mood states.

1. Introduction

Affecting 121 million people, depression is one of the leading causes of disability in the world (WHO, http://www.who.int/mental_health/management/depression/definition/en/). Although there are various effective treatments, therapy response is unsatisfactory and depression becomes a chronic condition in a considerable proportion of patients (Gilmer et al., 2005). Thus, there is a need to develop further therapeutic techniques to improve the course and the prognosis of depressive disorders.

Negative emotions, like anger, fear, and sadness, that are prevalent in depression are associated with activation of the corrugator and procerus muscles in the glabellar region of the face (Ekman and Friesen, 1978). Accordingly, in patients with depressive disorders facial electromyography reveals a relative overactivity of the corrugator muscles during different affective imagery paradigms (Schwartz et al., 1976). Activation of the corrugator muscles is also correlated with psychomotor agitation in depression and contributes to facial features like the 'omega melancholicum' and Veraguth's folds (Greden et al., 1985). Charles Darwin (1872)

* Corresponding author. Tel.: +41 61 325 5319/5351; fax: +41 61 325 5585.
E-mail address: axel.wollmer@unibas.ch (M.A. Wollmer).

0022-3956
doi:10.1016/j.jpsychires.2012.01.027

perceived these features as a very specific expression of sadness and attributed them to the activity of so-called 'grief muscles' in the glabellar region (Darwin, 1872).

"Refuse to express a passion, and it dies", this aphorism by William James (1890) refers to the facial feedback hypothesis, which implies mutual interaction between emotions and facial muscle activity (Darwin, 1872; James, 1890). In fact there is experimental evidence that voluntary contraction of facial muscles can channel emotions, which are conversely expressed by activation of these muscles (Strack et al., 1988; Larsen et al., 1992).

Injection of botulinum toxin to the glabellar region inhibits the activity of the corrugator and procerus muscles. This effect is used in the cosmetic treatment of frown lines (Carruthers and Carruthers, 1992). Facial treatment with botulinum toxin has become the most frequent intervention in esthetic medicine with estimated applications of several million per year (American Society of Plastic Surgeons Report of the 2010 Plastic Surgery Statistics, http://www.plasticsurgery.org/Documents/news-resources/statistics/2010-statisticss/Top-Level/2010-US-cosmetic-reconstructive-plastic-surgery-minimally-invasive-statistics2.pdf).

Treatment of the glabellar region with botulinum toxin produces a relative change in facial expression from angry, sad, and fearful to happy and can impact on emotional experience (Heckmann et al., 2003; Davis et al., 2010). Recipients of this treatment reported an increase in emotional wellbeing beyond the cosmetic benefit (Sommer et al., 2003). Specifically, reduced levels of fear and sadness were observed (Lewis and Bowler, 2009). The treatment also attenuated the activation of limbic brain regions during voluntary contraction of the corrugator and procerus muscles, indicating that feedback from the facial musculature may modulate the processing of emotions (Hennenlotter et al., 2009). Hence, the processing time for sentences with negative affective connotation was prolonged in women after glabellar botulinum toxin treatment and the treatment interfered with the ability to decode the facial expression of other people (Havas et al., 2010; Neal and Chartrand, 2011). The capacity of this intervention to counteract negative emotions may also be of clinical use. Accordingly, preliminary data from an open case series with ten female patients indicate that it may reduce the symptoms of depression (Finzi and Wasserman, 2006). We hypothesized that facial psychomotor features associated with depression are not just epiphenomena but integral components of the disorder and may be targeted in its therapy. To explore, if attenuation of these features may produce alleviation in the affective symptoms, we conducted a randomized controlled trial of botulinum toxin injection to the glabellar region as an adjunctive treatment of major depression.

2. Materials and methods

2.1. Trial design

At two centers, the Psychiatric University Hospital of the University of Basel, Switzerland and the Medical School Hannover, Germany we conducted a randomized, placebo-controlled, double-blind trial from August 2009 through October 2010. The trial has been registered with ClinicalTrials.gov, number, NCT00934687. The study was investigator-initiated and was carried out independently of any commercial entity. It was funded solely by the Gottfried & Julia Bangerter-Rhyner-Stiftung, Bern, Switzerland, a private foundation that supports medical research (http://www.bangerter-stiftung.ch/). The study protocol was approved by the institutional review boards, the local ethic committees and the Swiss and German regulatory authorities. It comprised seven visits (baseline, two weeks, four weeks, six weeks, eight weeks, twelve weeks, and sixteen weeks after baseline).

A power analysis based on the observations from the open case series (Finzi and Wasserman, 2006), allowing for an equal placebo control group with a theoretical 50% improvement, indicated that a sample size of <30 participants would be sufficient to detect comparable effects with a power of $>80\%$ at a significance level of $p < 0.05$.

2.2. Participants

Participants were recruited from local psychiatric outpatient units, psychiatrists in private practice, or through advertisements placed in the local press. The method under investigation, i.e. botulinum toxin treatment, was not explicitly mentioned in the advertisement, in order to avoid attracting candidates who were primarily motivated by receiving this treatment for cosmetic reasons. Both men and women were included. Inclusion criteria were: age 25–65 years, on-going major depressive disorder (DSM-IV 296.xx) diagnosed according to the Structured Clinical Interview for Axis I DSM-IV disorders (SCID I; $\geq$15 points on the Hamilton Depression Rating Scale at screening) with or without a history of dysthymic disorder (DSM-IV 300.4), and a moderate to severe vertical glabellar line during maximum voluntary frowning according to a four-point clinical severity score (Honeck et al., 2003) as well as qualitatively and quantitatively stable treatment with one or, at most, two antidepressants for at least four weeks. For ethical reasons we did not include untreated patients unless they had not responded to at least one treatment trial with an antidepressant during on-going index episode and were reluctant to undergo another one.

Exclusion criteria were: psychotic symptoms, suicidal tendency, clinical severity requiring immediate intervention, further DSM-IV axis I diagnoses, clinically manifest personality disorder, severe premenstrual syndrome or premenstrual dysphoric syndrome, regular occurrence of migraine or other forms of cephalalgia, psychological strain associated with glabellar frown lines, contra-indications of botulinum toxin treatment, previous treatment with botulinum toxin, further psychopharmacological treatment other than a demand medication with limited amounts of sedatives or hypnotics (lorazepam up to 4 mg/week, oxazepam, up to 60 mg/week, zolpidem up to 20 mg/week, or zopiclone up to 15 mg/week), and on-going disorder-specific psychotherapy or any other specific therapy of depression.

From all participants we collected psychopathological findings and a personal, medical, and psychiatric history. Response to antidepressant treatment during on-going index episode was assessed according to the Massachusetts General Hospital Antidepressant Treatment Response Questionnaire (MGH-ATRQ; Fava and Davidson, 1996). Treatment resistance was rated on a graduated scale (Thase and Rush, 1997). Participants were subjected to a general physical and neurological examination. Complementary somatic diagnostic procedures were disposed if organic reasons for depression had not been sufficiently excluded previously. All patients provided written informed consent after complete description of the study and before inclusion. Participants promised to leave their antidepressant medication unchanged during the first six weeks of the trial. Thereafter, they were allowed to change their treatment in consultation with their physician.

2.3. Interventions

For the verum condition onabotulinumtoxinA (Vistabel®, Botox® Cosmetic, Allergan) was dissolved in 0.9% NaCl solution (B. Braun Medical) at a concentration of 100U/2.5 ml. Injections were

made using insulin syringes with 30G needles at five specific points in the glabellar region (Supplemental Fig. 1). Women received 29 U of onabotulinumtoxinA in total. We injected 7 U to the procerus muscle, 6 U bilaterally to the medial part of the corrugator muscles, and 5 U bilaterally to the lateral part of the corrugator muscles. The same injection scheme was applied in the open case series (Finzi and Wasserman, 2006). To account for their higher muscle mass, men received two more units at each injection site, i.e. 39 U in total. For the placebo condition we injected identical volumes of 0.9% NaCl solution according to the described injection scheme. The intervention took place once at the end of the baseline visit.

2.4. Outcomes

At each study visit participants were assessed using the following instruments and measures: The validated German translation of the Structured Interview Guide for the Hamilton Depression Rating Scale with Atypical Depression Supplement (SIGH-ADS) (Williams and Terman, 2003), the Beck Depression Inventory (BDI) self-rating questionnaire, the Clinical Global Impressions Scale (CGI), the four-point Clinical Severity Score for Glabellar Frown Lines (CSS-GFL) (Honeck et al., 2003), a standardized photograph of the face during maximum frowning, and an inquiry regarding side effects and changes in concomitant treatment. SIGH-ADS and CGI raters were experienced psychiatrists and clinical psychologists. A rater training held in preparation of the study yielded a high inter-rater reliability for the SIGH-ADS (intraclass correlation coefficient, ICC(3,1) = 0.97). If possible, participants were examined by the same raters throughout the entire study.

The primary end point of the trial was change in the 17-item version of the Hamilton Depression Rating Scale (HAM-D_{17}) score at the visit after six weeks versus baseline. The clinical response measured by the HAM-D_{17} score was classified as nonresponse ($<$25% reduction), partial response (25%–49% reduction), response ($\geq$50% reduction), or remission (HAM-D_{17} score $\leq$7).

2.5. Randomization and blinding

Participants were randomized using a computer-generated (Software R, R Development Core Team) randomization list comprising one block with 30 positions and a pre-defined allocation ratio of 1:1. It was provided by the Clinical Trial Unit of the University of Basel and was administered by a physician who was not otherwise involved in the study. When a new participant was included he randomized the participant according to the running number of the randomization list and instructed the study nurse of the respective center to prepare verum or placebo injections. The drug account was administered by the study nurse.

Syringes prepared for verum or placebo injection were optically indistinguishable from each other. Participants were instructed that the extent and the kinetics of the paralytic onabotulinumtoxinA effect were inter-individually variable and that the injection of saline may induce a temporary reduction in frown lines, owing to swelling of the treated area. To conceal cosmetic changes from psychometric raters, participants wore an opaque surgical cap, which covered glabella and forehead during the examinations. Assessment and photo documentation of glabellar frown lines, as well as assessment of local side effects, were conducted by other study personnel. Participants did not receive a feedback on cosmetic changes. At the end of the trial we asked participants and psychometric raters to guess group allocation in a forced decision and on a graduated scale (Bang et al., 2004). Because these factors may have an influence on placebo effects, we asked participants to rate treatment expectancy and rationale credibility according to a credibility/expectancy questionnaire (Devilly and Borkovec, 2000). Moreover, participants were asked for an appraisal of their cosmetic change on a five point Likert scale (0 = very negative, 1 = somewhat negative, 2 = neutral, 3 = somewhat positive, 4 = very positive).

2.6. Statistical analyses

Data were analyzed using two-way analyses of variance (ANOVA) for repeated measures with the factors group (verum vs. placebo) and time (seven time points throughout the sixteen-week follow-up period as well as two time points, baseline and six weeks after baseline, for the primary end point). In the case of deviation from sphericity, ANOVAs were performed using Greenhouse–Geisser (ε) corrected degrees of freedom. Inner subject effects of time by group interactions are reported. Where appropriate, post-hoc tests with a Bonferroni–Holm adjustment were used, applying Student's *t*-tests or, in the case of variance inhomogeneity, the more robust Welch-test. These tests were also applied for continuous baseline variables. For categorical variables Fisher's exact tests were employed. All tests were two-sided. Variance is reported as standard deviation (SD) unless otherwise noted. Effect sizes are reported as partial η^2, Cohen's d, or odds ratio (OR). To assess linear relationship between scales, Pearson's correlation coefficients (r) were calculated. Test results with an alpha level $\leq$0.05 are reported as significant. Statistical analyses were conducted using SPSS 19.0 for Windows.

3. Results

3.1. Participant flow and baseline characteristics

We assessed 263 subjects for eligibility in order to include the pre-defined number of 30 participants. Only ten patients were referred to the study by their physicians. Among these patients the yield rate was 60%. In detail, one participant was recruited by a local psychiatric outpatient unit and five were recruited via psychiatrists in private practice. The remaining 24 participants were included through advertisements in the local press. As expected, this strategy was associated with a much lower yield rate (9.5%). The major causes for exclusion were psychiatric comorbidity and medication, and declining to follow the study protocol. Six subjects specifically refused to undergo onabotulinumtoxinA treatment and three subjects could not be included only because they were not able to produce an at least moderate glabellar line at maximum frowning. Once randomized, all participants received the allocated intervention and completed the trial up to the final visit. The investigators consistently reminded and motivated the participants to follow the study protocol and come to the visits. Thus, there were no dropouts or protocol violations and no participant had to be excluded from the analysis (CONSORT diagram, Supplemental Fig. 2).

Verum and placebo groups did not differ significantly in any of the demographic or clinical baseline variables (Table 1).

3.2. Efficacy outcomes

Over the seven time points throughout the sixteen-week duration of the trial the following observations were made: Glabellar frown lines at maximum frowning were attenuated by almost one point on the four-point Clinical Severity Score for Glabellar Frown Lines, CSS-GFL, in the verum group. This effect occurred within two weeks, remained constant until the visit after eight weeks and started to wear off slightly thereafter. In contrast CSS-GFL remained unchanged in the placebo group (ANOVA,

Table 1
Baseline characteristics of study participants, verum (onabotulinumtoxinA) vs. placebo.

Variable	Total ($N = 30$)		Verum ($N = 15$)		Placebo ($N = 15$)		Test		Significance
Demographic data									
	N	%	*N*	%	*N*	%			
Women	23	76.7	12	80.0	11	73.3	Fisher's exact		$p = 1.00$
Living alone	14	46.7	5	33.3	9	60.0	Fisher's exact		$p = 0.27$
Centre Basel	13	43.3	6	40.0	7	46.7	Fisher's exact		$p = 1.00$
	Mean	SD	Mean	SD	Mean	SD		df	
Age, years	50.57	8.91	52.2	7.54	48.93	10.10	$t = 1.00$	28	$p = 0.32$
Education, years	13.50	2.35	14.07	2.40	12.93	2.22	$t = 1.34$	28	$p = 0.19$
BMI	26.76	7.72	26.83	7.62	26.68	8.08	$t = 0.05$	28	$p = 0.96$
Specification of depression									
	N	%	*N*	%	*N*	%			
Recurrent F33.xx	24	80.0	11	73.3	13	86.7	Fisher's exact		$p = 0.65$
Moderate F3x.1x	20	66.7	11	73.3	9	60.0	Fisher's exact		$p = 0.70$
Somatic syndrome F3x.x1	16	53.3	9	60.0	7	46.7	Fisher's exact		$p = 0.72$
History of dysthymic disorder	8	26.67	3	20.0	5	33.3	Fisher's exact		$p = 0.68$
Chronic IE	13	43.3	6	40.0	7	46.7	Fisher's exact		$p = 1.00$
	Mean	SD	Mean	SD	Mean	SD		df	
Duration of MDD, years	16.03	12.67	16.40	11.97	15.67	13.74	$t = 0.16$	28	$p = 0.88$
Duration of IE, months	28.53	28.01	26.40	28.60	30.67	28.23	$t = 0.41$	28	$p = 0.68$
Number of episodes	7.13	9.12	9.07	12.27	5.20	3.71	$t = 1.17$	28	$p = 0.25$
HAM-D_{17}	20.07	4.37	21.40	4.31	18.73	4.15	$t = 1.73$	28	$p = 0.10$
HAM-D_{21}	22.23	4.29	23.27	3.83	21.20	4.60	$t = 1.34$	28	$p = 0.19$
ADS	7.53	4.18	8.80	4.16	6.27	3.92	$t = 1.72$	28	$p = 0.10$
SIGH-ADS	27.60	7.42	30.20	7.58	25.00	6.49	$t = 2.02$	28	$p > 0.05$
BDI	25.33	8.14	27.00	8.91	23.67	7.21	$t = 1.13$	28	$p = 0.27$
CGI	4.50	0.73	4.67	0.62	4.33	0.82	$t = 1.26$	28	$p = 0.22$
Antidepressant medication									
	N	%	*N*	%	*N*	%			
SSRI	10	33.3	6	40.0	4	26.7	Fisher's exact		$p = 0.70$
SNRI	9	30.0	4	26.7	5	33.3	Fisher's exact		$p = 1.00$
Mirtazapine	6	20.0	1	6.7	5	33.3	Fisher's exact		$p = 0.17$
Tricyclic compounds	6	20.0	4	26.7	2	13.3	Fisher's exact		$p = 0.65$
Other compounds	6	20.0	2	13.3	4	26.7	Fisher's exact		$p = 0.65$
Combination	8	26.7	3	20.0	5	33.3	Fisher's exact		$p = 0.68$
	Mean	SD	Mean	SD	Mean	SD		df	
Treatment trials during IE	1.27	0.69	1.40	0.83	1.13	0.52	$t = 1.06$	28	$p = 0.30$
Treatment resistance stage	0.83	0.65	1.00	0.66	0.67	0.62	$t = 0.93$	28	$p = 0.36$
Current antidepressants	1.23	0.50	1.13	0.52	1.33	0.49	$t = 1.09$	28	$p = 0.29$
Adequate dose	1.03	0.56	1.00	0.54	1.07	0.59	$t = 0.32$	28	$p = 0.75$
Stable treatment, months	19.35	20.68	18.30	15.82	20.46	25.48	$t = 0.28$	28	$p = 0.78$
Further characteristics									
CSS-GFL	2.33	0.48	2.20	0.41	2.47	0.52	$w = 1.56$	28	$p = 0.13$
SCID II	4.04	2.33	3.71	2.40	4.36	2.31	$t = 0.73$	26	$p = 0.48$

Besides those for major depressive disorder (DSM-IV 296.xx) all participants also fulfilled the ICD-10 criteria for mild to moderate depressive episode. Depression was specified according to ICD-10 (F32.00-F33.11). Chronic duration of on-going index episode (IE) was defined as ≥24 months. ADS is the score for atypical depression symptoms, SIGH-ADS is the sum score of HAM-D_{17} + ADS. SSRI (selective serotonin reuptake inhibitors), SNRI (serotonin and noradrenalin reuptake inhibitors), mirtazapine, tricyclic compounds, and other compounds (bupropion, moclobemide, trazodone, or St. John's wort extract) refer to the number of participants treated with the respective medication. Combination indicates the number of participants treated with two antidepressants. Treatment resistance was staged according to Thase and Rush, 1997. Adequate dose is the number of current antidepressants given in at least the smallest recommended dose for the treatment of major depression. SCID II is the number of traits in the Structured Clinical Interview for Axis II DSM-IV disorders (SCID II) questionnaire with above threshold scores. Clinical evaluation of these findings did not result in the diagnosis of a manifest personality disorder in any of the study participants. Elevated scores were rather owing to the in most cases longstanding depressive disorder, which made it difficult for the patients to differentiate between depression-related states and personality traits. For Welch-test uncorrected degrees of freedom (df) are reported. All statistical tests were two-sided. There was a trend towards higher SIGH-ADS scores in the verum group ($p = 0.053$). Introduction of the SIGH-ADS score at baseline as a covariate did not in principle change the results of the ANOVAs (data not shown).

$n = 30$; $F_{(6,168)} = 13.59$, $\varepsilon = 0.50$, $\eta^2 = 0.33$, $p < 0.001$, two-sided; Fig. 1A). There was a significant clinical improvement in depression over time in the verum group, but not in the placebo one. Apart from a fluctuation at the visit after eight weeks this effect increased until the final visit. It was measured by the Hamilton Depression Rating Scale, HAM-D_{17}, score (ANOVA, $n = 30$; $F_{(6,168)} = 5.76$, $\varepsilon = 0.74$, $\eta^2 = 0.17$, $p < 0.001$, two-sided; Fig. 1B) and by the Beck Depression Inventory, BDI, score (ANOVA, $n = 30$; $F_{(6,168)} = 3.79$, $\varepsilon = 0.51$, $\eta^2 = 0.12$, $p = 0.01$, two-sided; Fig. 1C) and was also reflected in an improvement in the Clinical Global Impressions Scale, CGI, (ANOVA, $n = 30$; $F_{(6,168)} = 7.91$, $\varepsilon = 0.66$, $\eta^2 = 0.22$, $p < 0.001$, two-sided; Fig. 1D). The external rating (HAM-D_{17}) and the self-assessment (BDI) of depression symptoms were correlated (Pearson's product moment correlational analysis, $n = 30$; $r = 0.75$, $p < 0.001$).

As for the primary end point (HAM-D_{17} at the visit after six weeks versus baseline), there was a significant improvement in the verum group compared to the placebo group with a large effect size (−10.07 ± 8.16 points, −47.1% vs. −1.73 ± 4.25 points, −9.2%; ANOVA, Student's *t*-test, $n = 30$; $F_{(1,28)} = 12.30$, $\eta^2 = 0.31$,

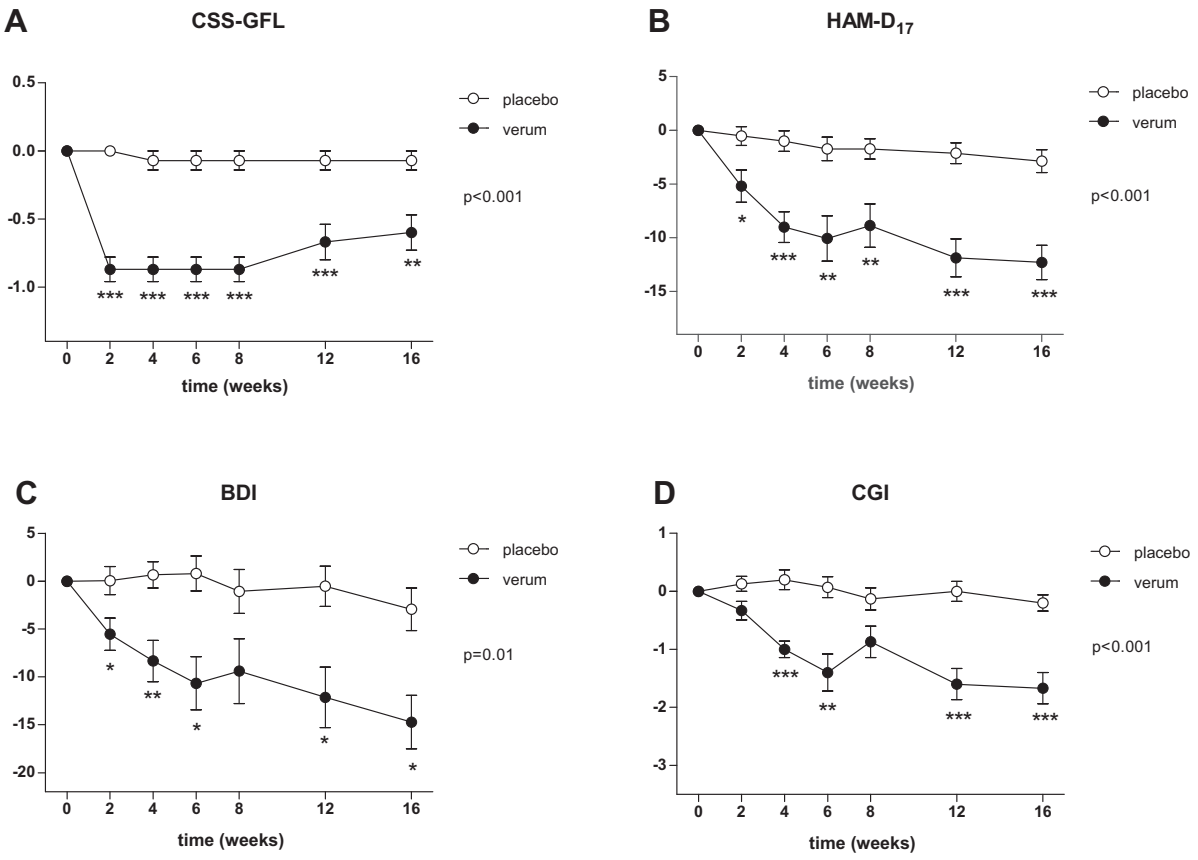

Fig. 1. Efficacy outcomes are presented as score changes (mean ± standard error, SEM) in the CSS-GFL (A), HAM-D_{17} (B), the BDI (C), and the CGI (D) In the onabotulinumtoxinA group ($n = 15$) and the placebo group ($n = 15$) at the visits after two, four, six, eight, twelve, and sixteen weeks versus baseline (set to zero). *P*-values refer to inner subject effects of time by group interactions from ANOVAs with Greenhouse–Geisser-corrected degrees of freedom. Asterisks indicate time points with significant group differences in post-hoc tests after Bonferroni-Holm adjustment of *p*-values for initially six comparisons ($^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$). All participants were included and all statistical tests were two-sided. All values and detailed statistics corresponding to the graphs are available in the supplemental data section (Supplemental table 1). Between the visit after eight weeks and the visit after sixteen weeks eleven participants underwent some change in their treatment. Nine of these changes were counteractive to the group difference, i.e. intensifications of therapy in the placebo group or cutback of therapy in the verum group. Exclusion of the participants with treatment changes did not in principle affect the group differences observed in the last two visits or influence the run of the HAM-D_{17}, BDI, and CGI curves (data not shown).

$t_{(28)} = 3.51$, $d = 1.28$, $p = 0.002$, two-sided; Table 2). Accordingly, the partial response rate (25%–49% HAM-D_{17} score reduction; 86.7% vs. 26.7%; Fisher's exact test, $n = 30$; OR = 17.9, 95% CI = 2.7–116.9, $p = 0.003$, two-sided) and the response rate (≥50% HAM-D_{17} score reduction; 60.0% vs. 13.3%; Fisher's exact test, $n = 30$; OR = 9.8, 95% CI = 1.6–59.7, $p = 0.02$, two-sided) were significantly higher in the verum group than in the placebo one (Fig. 2). Of the collected baseline characteristics only female sex was significantly

Table 2
Efficacy outcome measures six weeks after baseline.

Variable	Verum ($N = 15$)		Placebo ($N = 15$)				Test	Effect size	Significance
	Mean	SD	Mean	SD					
CSS-GFL	1.33	0.62	2.40	0.63			$F = 50.40$	$\eta^2 = 0.64$	$p < 0.001$
HAM-D_{17}	**11.33**	**7.22**	**17.00**	**6.52**			***F* = 12.30**	***η*² = 0.31**	***p* = 0.002**
HAM-D_{21}	12.47	7.76	18.53	6.84			$F = 9.87$	$\eta^2 = 0.26$	$p = 0.004$
ADS	6.00	5.18	6.27	1.94			$F = 5.30$	$\eta^2 = 0.16$	$p = 0.029$
SIGH-ADS	17.33	11.27	23.27	7.41			$F = 12.05$	$\eta^2 = 0.30$	$p = 0.002$
BDI	16.33	11.55	24.47	11.41			$F = 11.88$	$\eta^2 = 0.30$	$p = 0.002$
CGI	3.27	1.33	4.40	0.99			$F = 15.83$	$\eta^2 = 0.36$	$p < 0.001$
					Difference	CI			
Δ CSS-GFL	−0.87	0.35	−0.07	0.26	−0.8	−1.02 to −0.58	$t = 7.10$	$d = 2.59$	$p < 0.001$
Δ HAM-D_{17}	**−10.07**	**8.16**	**−1.73**	**4.25**	**−8.34**	**−13.00 to −3.68**	***t* = 3.51**	***d* = 1.28**	***p* = 0.002**
Δ HAM-D_{21}	−10.80	8.87	−2.67	4.69	−8.13	−13.21 to −3.05	$t = 3.14$	$d = 1.14$	$p = 0.004$
Δ ADS	−2.80	2.96	0.00	3.66	−2.80	−5.18 to −0.42	$t = 2.30$	$d = 0.84$	$p = 0.029$
Δ SIGH-ADS	−12.87	10.32	−1.73	6.92	−11.14	−17.43 to −4.85	$t = 3.47$	$d = 1.27$	$p = 0.002$
Δ BDI	−10.67	10.77	0.80	7.06	−11.47	−17.99 to −4.95	$t = 3.45$	$d = 1.26$	$p = 0.002$
Δ CGI	−1.40	1.24	0.07	0.70	−1.47	−2.19 to −0.75	$w = 3.98$	$d = 1.46$	$p < 0.001$

The table reports the scores at the visit after six weeks and the Δ values between this visit and the baseline. ANOVAs were calculated with two time points, baseline and six weeks after baseline (df = 1,28). Effect sizes are reported as partial η^2 and Cohen's d, referring to ANOVA inner subject effects of time by group interactions and Student's *t*-test or Welch-tests (df = 28) of Δ values, respectively. All statistical tests were two-sided. The primary end point is highlighted in bold type.

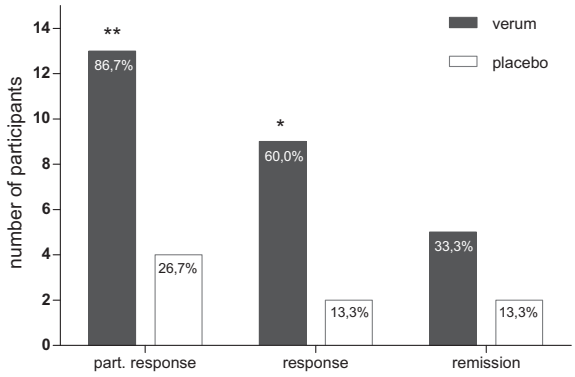

Fig. 2. Partial response, response, and remission rates are presented as number and percentage of participants in the onabotulinumtoxinA group and the placebo group at the visit after six weeks. Significant group differences were observed for partial response (** Fisher's exact, p=0.003, OR=17.9, 95% CI=2.7–116.9) and response (* Fisher's exact, p=0.02, OR=9.8, 95% CI=1.6–59.7). Originally, response rate defined as a 30% reduction in the HAM-D score was registered as the primary end point (ClinicalTrials.gov, NCT00934687), which is obviously also positive. For reasons of comparability with other studies it was changed to the primary end point described in the text.

associated with response (Fisher's exact test, $n = 30$; $p = 0.03$, two-sided; Supplemental Table 2). Hence, differences between the verum and the placebo group were slightly more pronounced, when the primary end point was analyzed for female participants alone (-11.67 ± 7.48 points, -53.9% vs. -2.36 ± 4.37 points, -13.0%; ANOVA, Student's t-test, $n = 23$; $F_{(1,21)} = 12.95$, $\eta^2 = 0.38$, $t_{(21)} = 3.60$, $d = 1.52$, $p = 0.002$, two-sided). Even larger effect sizes were observed at the end of the sixteen-week follow-up period (HAM-D_{17}, whole sample $d = 1.80$; Supplemental Table 3; female participants $d = 2.32$).

3.3. Side effects and blinding quality

The study treatment was well tolerated. Apart from local irritation immediately after injection, short episodes of headache during the first weeks were the only relevant and possibly treatment-related adverse events during the trial. They occurred both in the verum and in the placebo group (40.0% vs. 26.7%, Fisher's exact, $p = 0.7$).

At the end of the trial 90% of the participants guessed their group allocation correctly with firm conviction. They named presence or absence of cosmetic change as the major unblinding factor. Psychometric raters guessed 60% of group allocations correctly. They were less convinced about their guesses and improvement in depressive symptoms or lack of it was the major clue. Treatment expectancy and rationale credibility ratings did not differ between groups at baseline and did not predict clinical outcome (Supplemental Table 2). Treatment response was not associated with the appraisal of the cosmetic change (Supplemental Table 2).

4. Discussion

In the present study, we show, for the first time using a randomized controlled trial design, that a single treatment of the glabellar region with onabotulinumtoxinA may reduce the symptoms of major depression. This effect developed within few weeks and persisted until the end of the sixteen-week follow-up period. The effect sizes in our study were large and the response and remission rates were high for a trial with patients suffering from partly chronic and treatment resistant depression (Rush et al., 2006).

Botulinum toxin treatment offers some favorable characteristics: The long-running effect of a single dose may circumvent problems associated with poor therapy adherence, which is a critical issue in the treatment of depressed patients (Serna et al., 2010). Because of the long treatment intervals it may also be an economic treatment option (Beer, 2010). Moreover, the safety and tolerability record of botulinum toxin injections to the glabellar region is excellent (Brin et al., 2009).

It is unknown how botulinum toxin ameliorates depression. Several mechanisms may be involved: We assume that reduced proprioceptive feedback from the paralyzed facial muscles is a relevant mechanism of mood improvement. This mechanism is supported by other studies that show effects of botulinum toxin treatment on emotional perception (Hennenlotter et al., 2009; Havas et al., 2010; Neal and Chartrand, 2011), as well as by experiments in which facial muscle activity was influenced by other means (Strack et al., 1988; Larsen et al., 1992). Thus, botulinum toxin treatment may act comparable to a passive and uninterrupted relaxation exercise. It is possible that a more positive facial expression and improved feedback both from one's own face in the mirror and from social interaction partners may have contributed to mood enhancement (Heckmann et al., 2003). We largely exclude a merely esthetic benefit as the major cause of mood improvement, because we did not include patients who were cosmetically concerned about their frown lines. Notably, treatment response did not depend on a positive appraisal of the cosmetic change. One participant actually attained remission but disliked the change in facial appearance. Animal studies have shown that retrograde axonal transport of botulinum toxin is possible (Antonucci et al., 2008). Thus, we cannot exclude involvement of central pharmacological botulinum toxin effects including possible pharmacodynamic or pharmacokinetic interactions with the concomitant antidepressant therapy.

Placebo effects play a relevant part in the clinical improvement in depression under therapy and we assume that they also contribute to the clinical change observed in the present study (Brunoni et al., 2009; Rief et al., 2009). The method-immanent cosmetic effect of onabotulinumtoxinA treatment entailed a high unblinding rate among participants. This may have caused a bias in placebo effects in favor of the verum group, leading to underestimation of the placebo proportion of clinical improvement. However, high unblinding rates are common in antidepressant trials and a recent study shows that open-label placebo treatment can still induce clinical improvement (Perlis et al., 2010; Even et al., 2000; Kaptchuk et al., 2010). This is also in line with the substantial improvement observed in the placebo arm of the PREEMPT study, which recently led to registration of onabotulinumtoxinA injections to the head-neck area, including the glabellar region as treatment for chronic migraine (Dodick et al., 2010). Notably, one participant in the verum group of our study whose structurally fixed frown line persisted after treatment was convinced to be on placebo, but still attained remission of depression. Unblinding of psychopathological raters was below the range reported in other antidepressant trials (Even et al., 2000). It arose from treatment outcome rather than predicted it. Thus, it is unlikely that observer-expectancy effects may have accounted for a large proportion of the group difference. The improvement we observed in the placebo group was only marginal and fell short of our expectations. This may be ascribed to the high proportion of participants with a long duration of depression and poor response to previous treatment (Brown et al., 1992). Although the two groups of our study did not differ significantly in any of the collected baseline variables, we cannot fully exclude that hidden selection biases may have contributed to the particularly small improvement in the placebo group and a consequential inflation of the outcome differences between the groups.

This is a pilot study and its results will have to hold their ground in larger trials. It has some limitations that should be addressed: The proportion of male participants was too low to make a statement on efficacy in men; however, there may be a gender effect in favor of women. Generalizability is also limited by the selection of patients with glabellar frown lines, although only three subjects were excluded owing to lack of these lines. It is unclear, whether the treatment can also alleviate depression in hypomimic patients without this psychomotor phenotype. In our study onabotulinumtoxinA was applied essentially as an adjunctive intervention in patients with stable pharmacological therapy. Therefore, it is possible that botulinum toxin is not an antidepressant per se but may rather augment the antidepressant effect of the concomitant medication. However, one participant in the verum group was devoid of any antidepressant medication and still attained remission. Remission of participants without previous or present antidepressant medication also occurred in the open case series (Finzi and Wasserman, 2006). In spite of these observations, it remains to be established, whether botulinum toxin is also effective as a primary intervention in treatment-naïve patients. Moreover, it is uncertain if patients with severe or bipolar depression will benefit from botulinum toxin treatment. One theoretical hazard of the method is that modification of facial expression may conceal psychological strain in patients who do not respond or have not yet responded psychopathologically.

Sad facial expression is not only conferred by muscles of the glabellar region but also by muscles in lower sections of the face. It is possible that treatment of the depressor angulis oris and the mentalis muscles, for example, may also have mood-elevating effects and may enhance the clinical effect of the glabellar injection of botulinum toxin. Modulation of mood states with botulinum toxin may also be effective in the treatment of other clinical conditions involving negative emotions, like anxiety disorders.

In summary our study provides new evidence that botulinum toxin injection to the glabellar region may be an effective, safe, and sustainable intervention in the treatment of depression. It provides clinical support for the concept that the facial musculature not only expresses, but also regulates, mood states. At this juncture, the indication-compatible treatment of dynamic glabellar rhytides with botulinum toxin may be considered for depressed patients with the objective of inducing mood-lifting side effects.

Role of the funding source

The study was investigator-initiated and was carried out independently of any commercial entity. It was funded solely by the Gottfried & Julia Bangerter-Rhyner-Stiftung, Bern, Switzerland, a private foundation that supports medical research (http://www.bangerter-stiftung.ch/).

Contributors

M.A.W. was the principal investigator in Basel, Switzerland. He initiated the study here, wrote the grant application, the study protocol and the manuscript, recruited, characterized, treated, and rated participants, and revised the statistical analyses. C.d.B. recruited and rated participants, kept records of study data, and contributed to the discussion of the manuscript. N.K. was in charge of the logistics for participant recruitment, kept records of study data, and assisted in the statistical analyses and the preparation of tables. J.B. was involved in the baseline characterization of participants, kept records of study data, and contributed to the design of the study and to the discussion of the manuscript. T.G. was involved in the baseline characterization and rating of participants and kept records of study data. T.S. recruited and rated participants and kept records of study data. Mu.Ho. recruited and characterized participants. U.B. assisted in the preparation of study documentation and examined cosmetic outcomes. T.K. assisted in writing the grant application and examined cosmetic outcomes. K.K. treated participants and was involved in their rating. K.D. treated participants and was involved in their rating. M.D.H. gave advice regarding botulinum toxin dosage and supervised the treatment of participants. M.S. contributed to the interpretation of data and to the discussion of the manuscript. Ma.Ha. was involved in the study design and grant application. D.D. provided infrastructure and contributed to the discussion of the manuscript. S.B. did the statistical analyses and contributed to the discussion of the manuscript. E.H.T. contributed to the study design, the grant application, and the discussion of the manuscript. T.H.C.K. was the principal investigator in Hannover, Germany, he initiated the study here, contributed to the grant application, the study protocol and the manuscript, recruited, characterized, treated, and rated participants, and prepared all figures of the manuscript.

Conflict of interest

M.A.W. received honoraria for talks from Merz and Novartis. J.B received honoraria for talks from Ely Lilly and Brystol-Myers Squibb. T.G. has been working for Novartis from March to December 2011. K.K received honoraria from Allergan, Ipsen and Merz. D.D. received compensation for consulting from Allergan, Ipsen, Merz and Solstice/Eisai. E.H.-T received research grants from Actelion, Cephalon, Eli Lilly, Essex/MSD, Servier, and Vifor, recieved compensation for CME activity from AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Essex/MSD, Permamed, Pfizer, Servier, and Vifor, and is a member of the advisory boards of Eli Lilly, Lundbeck, and Pfizer. T.H.C.K. received honoraria for talks from Boehringer, Servier, Pfizer, and Ferring. These activities were all unrelated to the study. C.d.B., N.K., T.S., Mu.Ho., U.B., T.K., D.S., K.D., M.D.H., M.S., Ma.Ha., and S.B. declare no conflict of interest.

Acknowledgments

We thank the Gottfried & Julia Bangerter-Rhyner-Stiftung, Bern, Switzerland for funding this study. We should also like to thank our study nurses Simone Lüscher, Heike Gorzolla, Gabriele Dierks, and Linda Becker for their technical assistance, and Beatrice Pfeiffer and Claudette Lang for the photo documentation. Our thanks also go to Lars Walz and Thomas Zumbrunn from the Clinical Trial Unit, University Hospital Basel, for their advice on the study design and documentation, as well as for providing the randomization list, to Oliver Prince for administering the randomization list, to Berthold Rzany for providing an atlas with standard photos of glabellar frown lines, and to Janet Williams for allowing us to use the SIGH-ADS.

Appendix. Supplementary material

Supplementary material related to this article can be found online at doi:10.1016/j.jpsychires.2012.01.027.

References

Antonucci F, Rossi C, Gianfranceschi L, Rossetto O, Caleo M. Long-distance retrograde effects of botulinum neurotoxin A. Journal of Neuroscience 2008;28:3689–96.

Bang H, Ni L, Davis CE. Assessment of blinding in clinical trials. Controlled Clinical Trials 2004;25:143–56.

Beer K. Cost effectiveness of botulinum toxins for the treatment of depression: preliminary observations. Journal of Drugs in Dermatology 2010;9:27–30.

Brin MF, Boodhoo TI, Pogoda JM, James LM, Demos G, Terashima Y, et al. Safety and tolerability of onabotulinumtoxinA in the treatment of facial lines: a meta-

analysis of individual patient data from global clinical registration studies in 1678 participants. Journal of the American Academy of Dermatology 2009;61: 961–970.e1-11.
Brown WA, Johnson MF, Chen MG. Clinical features of depressed patients who do and do not improve with placebo. Psychiatry Research 1992;41:203–14.
Brunoni AR, Lopes M, Kaptchuk TJ, Fregni F. Placebo response of non-pharmacological and pharmacological trials in major depression: a systematic review and meta-analysis. PLoS One 2009;4:e4824.
Carruthers JD, Carruthers JA. Treatment of glabellar frown lines with C. botulinum-A exotoxin. Journal of Dermatologic Surgery & Oncology 1992;18:17–21.
Darwin CR. The expression of emotion in man and animals. London: Murray,; 1872.
Davis JI, Senghas A, Brandt F, Ochsner KN. The effects of BOTOX injections on emotional experience. Emotion 2010;10:433–40.
Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. Journal of Behavior Therapy and Experimental Psychiatry 2000; 31:73–86.
Dodick DW, Turkel CC, DeGryse RE, Aurora SK, Silberstein SD, Lipton RB, et al. OnabotulinumtoxinA for treatment of chronic migraine: pooled results from the double-blind, randomized, placebo-controlled phases of the PREEMPT clinical program. Headache 2010;50:921–36.
Ekman P, Friesen WV. Facial action coding system: a technique for the measurement of facial movement. Palo Alto: Consulting Psychologists Press; 1978.
Even C, Siobud-Dorocant E, Dardennes RM. Critical approach to antidepressant trials. Blindness protection is necessary, feasible and measurable. British Journal of Psychiatry 2000;177:47–51.
Fava M, Davidson KG. Definition and epidemiology of treatment-resistant depression. Psychiatric Clinics of North America 1996;19:179–200.
Finzi E, Wasserman E. Treatment of depression with botulinum toxin A: a case series. Dermatologic Surgery 2006;32:645–9.
Gilmer WS, Trivedi MH, Rush AJ, Wisniewski SR, Luther J, Howland RH, et al. Factors associated with chronic depressive episodes: a preliminary report from the STAR-D project. Acta Psychiatrica Scandinavica 2005;112:425–33.
Greden JF, Genero N, Price HL. Agitation-increased electromyogram activity in the corrugator muscle region: a possible explanation of the “Omega sign”? American Journal of Psychiatry 1985;142:348–51.
Havas DA, Glenberg AM, Gutowski KA, Lucarelli MJ, Davidson RJ. Cosmetic use of botulinum toxin-a affects processing of emotional language. Psychological Science 2010;21:895–900.
Heckmann M, Teichmann B, Schröder U, Sprengelmeyer R, Ceballos-Baumann AO. Pharmacologic denervation of frown muscles enhances baseline expression of happiness and decreases baseline expression of anger, sadness, and fear. Journal of the American Academy of Dermatology 2003;49:213–6.
Hennenlotter A, Dresel C, Castrop F, Ceballos-Baumann AO, Wohlschläger AM, Haslinger B. The link between facial feedback and neural activity within central circuitries of emotion–new insights from botulinum toxin-induced denervation of frown muscles. Cerebral Cortex 2009;19:537–42.
Honeck P, Weiss C, Sterry W, Rzany B, Gladys study group. Reproducibility of a four-point clinical severity score for glabellar frown lines. British Journal of Dermatology 2003;149:306–10.
James W. The principles of psychology. New York: Holt; 1890.
Kaptchuk TJ, Friedlander E, Kelley JM, Sanchez MN, Kokkotou E, Singer JP, et al. Placebos without deception: a randomized controlled trial in irritable bowel syndrome. PLoS One 2010;5:e15591.
Larsen RJ, Kasimatis M, Frey K. Facilitating the furrowed brow: an unobtrusive test of the facial feedback hypothesis applied to unpleasant affect. Cognition Emotion 1992;6:321–38.
Lewis MB, Bowler PJ. Botulinum toxin cosmetic therapy correlates with a more positive mood. Journal of Cosmetic Dermatology 2009;8:24–6.
Neal DT, Chartrand TL. Embodied emotion perception: amplifying and dampening facial feedback modulates emotion perception accuracy. Social Psychological and Personality Science; 2011. 1948550611406138.
Perlis RH, Ostacher M, Fava M, Nierenberg AA, Sachs GS, Rosenbaum JF. Assuring that double-blind is blind. American Journal of Psychiatry 2010;167:250–2.
Rief W, Nestoriuc Y, Weiss S, Welzel E, Barsky AJ, Hofmann SG. Meta-analysis of the placebo response in antidepressant trials. Journal of Affective Disorders 2009; 118:1–8.
Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. American Journal of Psychiatry 2006; 163:1905–17.
Schwartz GE, Fair PL, Salt P, Mandel MR, Klerman GL. Facial muscle patterning to affective imagery in depressed and nondepressed subjects. Science 1976;192: 489–91.
Serna MC, Cruz I, Real J, Gascó E, Galván L. Duration and adherence of antidepressant treatment (2003–2007) based on prescription database. European Psychiatry 2010;25:206–13.
Sommer B, Zschocke I, Bergfeld D, Sattler G, Augustin M. Satisfaction of patients after treatment with botulinum toxin for dynamic facial lines. Dermatologic Surgery 2003;29:456–60.
Strack R, Martin LL, Stepper S. Inhibiting and facilitating conditions of facial expressions: a nonobstrusive test of the facial feedback hypothesis. Journal of Personality and Social Psychology 1988;54:768–77.
Thase ME, Rush AJ. When at first you don’t succeed: sequential strategies for antidepressant nonresponders. Journal of Clinical Psychiatry 1997;58(Suppl. 13):23–9.
Williams JBW, Terman M. Structured interview guide for the Hamilton depression rating scale with atypical depression supplement (SIGH-ADS). New York, USA: New YorkState Psychiatric Institute; 2003.