Placebo-Controlled, Antidepressant Clinical Trials Cannot Be Shortened to Less Than 4 Weeks' Duration: A Pooled Analysis of Randomized Clinical Trials Employing a Diagnostic Odds Ratio–Based Approach

Enrico Tedeschini, MD; Maurizio Fava, MD; and George I. Papakostas, MD

Objective: In double-blind, randomized, placebo-controlled clinical trials for major depressive disorder (MDD), the impact of study duration on outcome has not been adequately studied. Our aim was to examine whether placebo-controlled antidepressant trials in MDD could be shortened to less than 4 weeks. In order to accomplish this, we examined the relationship between a "positive" or "negative" finding early on (weeks 1–4), and outcome at end point.

Data Sources: MEDLINE/PubMed publication databases were searched for randomized, double-blind, placebo-controlled trials of antidepressants for adults with MDD published between January 1, 1980, and July 1, 2009 (inclusive).

Data Selection: One hundred seventy-five articles were found eligible. We obtained required measures during the required time points for 101 articles (57.7%). Final inclusion of articles was determined by consensus among the authors.

Data Synthesis: One hundred eighty-two drug-placebo comparisons from 104 clinical trials were pooled (29,213 patients). The strength of the relationship between early and end point outcome increased progressively. However, only at week 4 did the diagnostic odds ratio (27.44) indicate strong concordance between early and end point outcome. The specificity of early outcome as a predictor of end point outcome did not vary substantially from visit to visit (0.91–0.92), while the sensitivity increased proportionally with each visit (from 0.17 to 0.72).

Conclusions: The present analysis suggests that antidepressant clinical trials cannot be shortened to less than 4 weeks' duration, primarily due to the increased risk of erroneously concluding that an effective treatment is ineffective. Four weeks is the minimum adequate length of a trial in order to reliably detect drug versus placebo differences.

J Clin Psychiatry 2011;72(1):98–103

Submitted: *September 28, 2009; accepted January 14, 2010.*
Online ahead of print: *November 2, 2010 (doi:10.4088/JCP.09m05725blu).*
Corresponding author: *Enrico Tedeschini, MD, Depression Clinical and Research Program, Massachusetts General Hospital, 50 Staniford St, Boston, MA 02114 (etedeschini@partners.org).*

Major depressive disorder (MDD) is a highly prevalent illness that can often have a deleterious impact on the lives of those affected. Psychotherapy and antidepressant medications represent the mainstay of treatment for MDD. Double-blind, randomized, placebo-controlled clinical trials are considered the "gold standard" for the development of novel antidepressant therapies, and their successful design and conduct are critical to the advancement of the field. Unfortunately, however, even for compounds that have repeatedly been proven to be efficacious in treating MDD, differences in efficacy versus placebo are not always apparent throughout all clinical trials conducted. Such "failed" or "negative" trials may, in turn, lead to delays in bringing new treatments to the clinic as well as increase the costs of new treatment development.[1]

For this reason, over the last 2 decades, a number of researchers have investigated the relationship between various elements of clinical trial design and the likelihood of obtaining a "positive" result, including the severity of depression at baseline,[2,3] the choice of primary outcome measure,[4,5] the presence and duration of the placebo lead-in period,[6–8] and the effect of concomitant medications administered during the study.[9] Unfortunately, however, the impact of study duration on clinical trial outcome has not been adequately studied. In particular, evaluating whether the duration of clinical trials can be shortened while preserving their ability to detect a significant treatment effect, when one exists, is critical for several reasons. First, limiting the duration of trials could translate to reducing the overall cost of developing a new treatment, since fewer study visits would be required.[3,10] In addition, limiting the duration of the study would also reduce patient attrition rates, resulting in further cost savings per study. Decreasing study attrition would also result in more generalizable study results (since more patients from the original sample would complete the trial), and, from an analytic standpoint, also reduce reliance on complex statistical tests designed to address missing data derived from premature study termination.[11] Finally, from an ethical point of view, shorter clinical trials would allow researchers to minimize the duration of unnecessary exposure to either placebo or active (sometimes experimental) treatment.[1,12]

Walsh et al[12] and others[13] have noted that, in antidepressant, placebo-controlled clinical trials, there is often a statistical difference in mean depressive symptoms score reduction between antidepressant- and placebo-treated patients by the third week and almost always by the fourth week after randomization. In a previous study conducted by our group in which we pooled clinical trials of at least 4 weeks' duration, we did not find that treatment duration influenced the magnitude of the efficacy difference between antidepressants and placebo, leading us to conclude that extending trials

beyond 4 weeks does not significantly improve the likelihood of detecting a significant treatment effect.[3] However, since there is a paucity of placebo-controlled studies focusing on the use of antidepressants as monotherapy for MDD shorter than 4 weeks (we could find only 1), it was not possible for us to address whether those kinds of studies would have produced comparable results with studies of 4 weeks' duration. In the present analysis, we sought to examine whether placebo-controlled antidepressant trials in MDD could be shortened to less than 4 weeks. In order to accomplish this, we examined the relationship between a "positive" or "negative" finding early on (weeks 1–4), and outcome at end point in MDD placebo-controlled trials. The strength of the relationship between early and end point outcome was assessed by estimating the diagnostic odds ratio (DOR),[14] an overall measure of the statistical predictive ability of a test, in which outcome was considered to be the "test" and eventual outcome the "gold standard."

METHOD

Data Sources and Search Strategy

We sought to identify double-blind, randomized, placebo-controlled trials of antidepressants used as monotherapy for the treatment of MDD for possible inclusion in the meta-analysis. We defined antidepressants as pharmacologic agents that have or had received a letter of approval from US, Canadian, or European Union drug regulatory agencies for the treatment of MDD. According to this definition, the following pharmacologic agents met criteria to be considered antidepressants: amitriptyline, nortriptyline, imipramine, desipramine, clomipramine, trimipramine, protriptyline, dothiepin, doxepin, lofepramine, amoxapine, maprotiline, amineptine, nomifensine, bupropion, phenelzine, tranylcypromine, isocarboxazid, moclobemide, brofaromine, fluoxetine, sertraline, paroxetine, citalopram, escitalopram, fluvoxamine, zimelidine, tianeptine, trazodone, nefazodone, agomelatine, venlafaxine, desvenlafaxine, duloxetine, milnacipran, reboxetine, mirtazapine, and mianserin.

Eligible studies were first identified using searches of PubMed/MEDLINE, by cross-referencing the search term *placebo* with each of the above-mentioned agents. The PubMed/MEDLINE search was limited to articles that were published between January 1, 1980, and July 1, 2009 (inclusive). The cutoff year of 1980 was used in our search in order to decrease diagnostic variability, since the *DSM-III* was introduced in 1980. In order to expand our database, we then reviewed the reference list of all studies identified with PubMed/MEDLINE. Final inclusion of articles was determined by consensus among the authors.

Study Selection

We selected for randomized, double-blind, placebo-controlled trials of antidepressants used as monotherapy for the acute-phase treatment of MDD. We then selected studies that also met all of the following criteria:

1. Defined MDD according to *DSM-III*,[15] *DSM-III-R*,[16] *DSM-IV*,[17] Research Diagnostic Criteria,[18] or Feighner's Diagnostic Criteria[19];
2. Were at least 5 weeks in duration;
3. Focused on the use of antidepressants in their oral formulation;
4. Presented entirely original (not previously published) data;
5. Focused on the treatment of adult patients;
6. Did not exclusively focus on the treatment of patients with treatment-resistant depression or patients with other depressive disorders, including bipolar disorder, depression with psychotic features, dysthymic disorder, neurotic depression, or minor depression;
7. Did not exclusively focus on the treatment of MDD in patients with comorbid alcohol or substance use disorders or patients with a specific comorbid medical illness;
8. Involved the use of the Hamilton Depression Rating Scale (HDRS),[20] the Montgomery-Asberg Depression Rating Scale (MADRS),[21] or the Clinical Global Impressions-Improvement Scale (CGI-I)[22] as one of their outcome measures;
9. Reported the following outcomes of interest: an examination of whether there was a statistically significant difference in the change in depression scores (measured using 1 of the scales cited above) from baseline to a time point, between antidepressant- and placebo-treated patients (with the use of a 2-tailed test, and α set at .05); and
10. Reported the outcome of interest cited above for the following weeks: baseline, week 4, end point, and at least 1 intermediary visit between baseline and week 4.

Definitions

A study was considered positive at a given time point (weeks 1–4 or end point), if there was a greater reduction in depression scores from baseline to that given time point among antidepressant- than placebo-treated patients (2-tailed test, $\alpha = .05$). For consistency, the HDRS was chosen over the MADRS, and the MADRS over the CGI-I when response rates from multiple scales were reported. For each trial, it was recorded whether a trial was positive or negative at each time point (weeks 1–4) and at end point.

Quantitative Data Synthesis

The following measures were calculated utilizing data from early time points (weeks 1–4) versus end point.

- The number of either true positive results (TP) or true negative results (TN), which indicates at a given time point (weeks 1–4) how many times there is concordance between the result at a particular visit and the result at the end point (either positive or negative).

- The number of either false positive results (FP) or false negative results (FN), which indicates at a given time point (weeks 1–4) how many times there is discordance between the result at a particular visit and the result at the end point (either positive or negative).
- The sensitivity of a positive outcome at a particular visit with respect to the probability of obtaining a positive outcome at end point: this was defined as TP/(TP + FN)
- The specificity of a negative outcome at a particular visit with respect to the probability of obtaining a negative outcome at end point: this was defined as TN/(TN + FP)
- The reliability of the test was the degree to which further measurements in MDD randomized controlled trials (RCTs) show similar results: this was defined as TP/(TP + FP)
- The accuracy of the test was the degree of closeness of an early outcome to its eventual (end point) outcome in MDD RCTs: this was defined as (TP + TN)/(TP + FP + TN + FN)
- The DOR of the test of whether early outcome can predict eventual (end point) outcome in MDD RCTs: this was defined as (TP × TN)/(FP × FN)

RESULTS

Initially 7,257 abstracts were identified in PubMed/MEDLINE. Of these, 6,837 were excluded for a number of reasons (other topics, reviews). The remaining 420 abstracts described clinical trials of antidepressants used as monotherapy for depressive disorders. These 420 articles were obtained and reviewed thoroughly; 15 additional articles were identified after reviewing the reference lists of these 420 articles as well as of 2 large meta-analyses. Ninety-eight articles were excluded because they presented data published elsewhere; 25 manuscripts were excluded because they focused on children and/or adolescents with depression; and 40 were excluded because they focused on the treatment of depressive disorders other than MDD (bipolar disorder, MDD with psychotic features, dysthymic disorder, minor depression, or "neurotic depression"), because they focused on perinatal MDD, because the diagnosis of MDD was based on the *DSM-II*, or because they did not state which, if any, diagnostic criteria, were used to define MDD. One study was excluded because it focused on patients with treatment-resistant depression, 27 were excluded because they focused on the treatment of patients with depression and comorbid alcohol and/or drug use disorders, and 62 were excluded because they focused on the treatment of patients with depression and comorbid Axis III disorders. Three were excluded because they did not involve the use of an oral form of an antidepressant (selegiline), 3 because they were less than 4 weeks in duration, and 1 study because it did not involve the use of the HDRS, MADRS, or CGI-I.

Table 1. Antidepressants Compared to Placebo in 182 Comparisons From 104 Clinical Trials[a]

Antidepressant Class, No. (%)	Antidepressant	No. (%)
Tricyclic antidepressant, 37 (20.33)	Amitriptyline	9 (4.95)
	Desipramine	1 (0.55)
	Dothiepin	1 (0.55)
	Doxepin	2 (1.10)
	Imipramine	20 (11.00)
	Lofepramine	2 (1.10)
	Maprotiline	1 (0.55)
	Nortriptyline	1 (0.55)
Monoamine oxidase inhibitor, 3 (1.65)	Isocarboxazid	1 (0.55)
	Moclobemide	1 (0.55)
	Phenelzine	1 (0.55)
Selective serotonin reuptake inhibitor, 60 (32.97)	Citalopram	5 (2.75)
	Escitalopram	6 (3.30)
	Fluoxetine	14 (7.69)
	Fluvoxamine	8 (4.40)
	Paroxetine	18 (9.89)
	Sertraline	8 (4.40)
	Zimelidine	1 (0.55)
Serotonin-norepinephrine reuptake inhibitor, 47 (25.82)	Desvenlafaxine	13 (7.14)
	Duloxetine	13 (7.14)
	Venlafaxine	21 (11.54)
Other, 35 (19.23)	Agomelatine	5 (2.75)
	Bupropion	7 (3.85)
	Mianserin	3 (1.65)
	Mirtazapine	6 (3.30)
	Nefazodone	11 (6.04)
	Reboxetine	1 (0.55)
	Trazodone	2 (1.10)

[a]Percentages do not sum to 100 because of rounding.

Thus, a total of 175 articles were found eligible for inclusion in our pooled analysis (list available upon request). We were able to obtain required measures (statistical analysis testing whether or not the difference in the change in depression severity scores from baseline to that time point between antidepressant- and placebo-treated patients was statistically significant, using a 2-tailed test and $\alpha = .05$) during the required time points (baseline, end point, week 4, and at least 1 intermediary visit between baseline and week 4) for 101 (57.7%) of the 175 articles eligible for the pooled analysis. A total of 182 drug-placebo comparisons from 104 clinical trials were pooled, involving a total of 29,213 patients randomized to treatment with either an antidepressant (n = 18,446) or placebo (n = 10,767). Mean (SD) study duration was 7.5 (2.4) weeks, and the mean (SD) sample size per treatment arm was 102.1 (60.5) patients. The HDRS was used in 96 (92.3%) clinical trials, while the MADRS was used in the remaining 8 (7.7%) studies. Antidepressants tested against placebo in these 182 comparisons are listed in Table 1. One hundred thirty-five of 182 drug-placebo comparisons (74.2 %) were positive (indicated superior outcome for drug versus placebo) according to the a priori primary outcome measure used in each individual study as well as the definition of outcome employed in our present analysis.

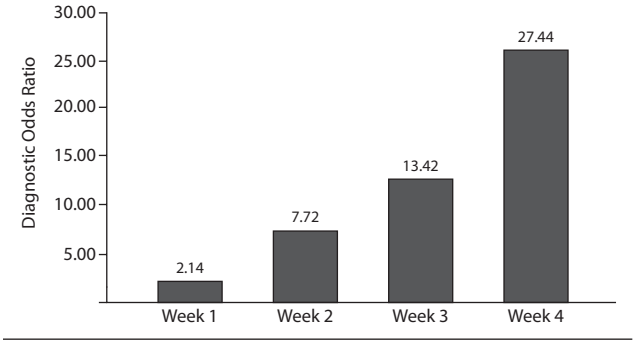
Analyses results are reported in Table 2. Diagnostic odds ratio values progressively increased from 2.14 to 7.72, 13.42, and 27.44, corresponding to weeks 1, 2, 3, and 4, respectively (Figure 1). The score at week 4 (27.44), according to

© Copyright 2010 Physicians Postgraduate Press, Inc.

Table 2. Results From Analyses of 182 Antidepressant-Placebo Comparisons From 104 Clinical Trials

Variable	Week 1	Week 2	Week 3	Week 4
Study, no. (%)	176 (96.7)	181 (99.5)	147 (80.8)	182 (100)
True positive, no.	22	56	61	97
True negative, no.	42	43	33	43
False positive, no.	4	4	3	4
False negative, no.	108	78	50	38
Reliability	0.85	0.93	0.95	0.96
Accuracy	0.36	0.55	0.64	0.77
Sensitivity	0.17	0.42	0.55	0.72
Specificity	0.91	0.91	0.92	0.91

Figure 1: Diagnostic Odds Ratio Values Corresponding to Each Study Week Compared to End Point (5 weeks or more)



the Jaeschke's guide,[23] corresponded to strong diagnostic evidence, while the remaining time points demonstrated weak (13.42, week 3; 7.72, week 2) to very weak (2.14, week 1) diagnostic evidence.

DISCUSSION

In a previous study conducted by our group utilizing data from trials of 4 weeks' duration or longer, we had found evidence suggesting that prolonging placebo-controlled, antidepressant clinical trials beyond 4 weeks did not lead to greater efficacy advantages for antidepressants versus placebo that were statistically significant.[3] However, due to the paucity of studies shorter than 4 weeks in duration (only 1 has been published, to date), it was not possible, using the data-analytic model employed in that meta-analysis, to ascertain whether, in fact, placebo-controlled antidepressant clinical trials could be shortened even further without compromising their ability to detect a significant treatment effect where one exists, or rule out the absence of such an effect where it did not exist.

The present analysis is, to our knowledge, the first ever published in the field of MDD that sought to examine whether the "optimal" duration of placebo-controlled antidepressant trials is less than 4 weeks. Using a DOR-based approach applied to a dataset of randomized, double-blind, clinical trials of antidepressants for MDD, we found that the strength of the relationship between early and end point outcome increased progressively from week 1 to week 4. According to Jaeschke's guide,[23] the fourth week value of the DOR (27.44) indicated a strong concordance between early and end point outcome, while for the 3 earlier time points, the DOR showed a weak (weeks 2 and 3) or a very weak (week 1) concordance. In accordance with this, while the reliability of the test did not vary dramatically from week to week (0.85–0.96), the accuracy of early outcome as a predictor of end point outcome progressively increased throughout the first 4 weeks of treatment (from 0.36 to 0.77). Taken together, these results suggest that placebo-controlled antidepressant trials cannot be shortened to less that 4 weeks' duration. In light of our previous findings,[3] the results of the present study suggest that 4 weeks is the optimal duration of a placebo-controlled antidepressant trial. This finding is also in line with previous observations by other groups[12,13] that had suggested that a statistical difference in mean depressive score reduction could often be detected within the fourth week of treatment.

One disadvantage of the DOR is that it informs the reader about the quality of "test performance" (in this case, early outcome as a predictor of end point outcome), without specifying whether it is a difference in the strength of the sensitivity or the specificity of a given test. Interestingly enough, in the present analysis, the specificity of early outcome as a predictor of end point outcome did not vary substantially from visit to visit (0.91–0.92), suggesting that a positive early outcome was very likely to yield a positive end point outcome, while the sensitivity of the test increased proportionally with each visit (from 0.17 to 0.72), suggesting that a negative result before week 4 might not have been followed by a negative end point result. Therefore, while trials of shorter duration would not be much more likely to report a significant treatment effect where one did not exist, they would be much more likely to miss a significant treatment effect where one was present.

Several limitations of the current work should be acknowledged when interpreting the findings. By far the most important limitation derives from our method of identifying studies to be included in pooled analyses. Specifically, only published studies were included in our dataset, since it is impossible to comprehensively obtain data from unpublished studies involving the use of antidepressants conducted since 1980 (only recently has the pharmaceutical industry been proactive about posting summary outcome data from a limited number of recently conducted, unpublished clinical trials on their Web sites). Since published studies may be more likely to have yielded positive results than unpublished ones (due to publication bias or the file-drawer phenomenon), it is quite possible that our dataset has, thus, been "enriched" with positive studies. Since negative studies at end point are more likely to be negative at earlier time points, and positive studies at end point more likely to be positive at earlier time points, focusing on published studies alone may have inflated true positive rates as well as false negative rates. In turn, inflating true positive rates may have led us to overestimating the sensitivity and underestimating the specificity of our test. Therefore, it would be interesting

to examine whether the inclusion of unpublished studies would strengthen or weaken our findings. However, while including unpublished studies might result in lower overall sensitivity and greater specificity, it is unlikely that the basic finding of our study (that sensitivity increases progressively from week 1 to week 4 while specificity remains relatively constant) would have been altered were we able to include all unpublished studies.

An additional limitation is that the present analysis was based on clinical trial–level data as opposed to individual patient–level data. Having individual patient–level data would have been much preferable, as it would have afforded us the opportunity to test whether individual patient characteristics influenced the relationship between early symptom improvement (or lack thereof) as a predictor of end point outcome, data that could then be applied to the design of studies of shorter duration. Finally, it is also important to keep in mind that the present dataset consists (predominantly), of clinical trials involving the use of agents with similar mechanisms of action (monoamine reuptake inhibitors and/or monoamine receptor blockers). Antidepressant drugs developed in the future may demonstrate much faster onset of action than monoamine reuptake inhibitors/receptor antagonists (ie, glutamatergic agents)[24] necessitating adaptations in trial design in order to "dynamically" capture more rapid improvements in mood (ie, shorter trial duration with more frequent assessments).

In conclusion, the results of the present analysis suggest that antidepressant clinical trials cannot be shortened to less than 4 weeks' duration, primarily due to the increased risk of erroneously concluding that an effective treatment is ineffective. Therefore, 4 weeks is the minimum adequate length of a trial in order to reliably detect drug versus placebo differences.

Drug names: bupropion (Wellbutrin, Aplenzin, and others), citalopram (Celexa and others), clomipramine (Anafranil and others), desipramine (Norpramin and others), desvenlafaxine (Pristiq), doxepin (Zonalon, Silenor, and others), duloxetine (Cymbalta), escitalopram (Lexapro and others), fluoxetine (Prozac and others), fluvoxamine (Luvox and others), imipramine (Tofranil and others), isocarboxazid (Marplan), milnacipran (Savella), mirtazapine (Remeron and others), nortriptyline (Pamelor, Aventyl, and others), paroxetine (Paxil, Pexeva, and others), phenelzine (Nardil), protriptyline (Vivactil and others), selegiline (Emsam, Zelapar, and others), sertraline (Zoloft and others), tranylcypromine (Parnate and others), trazodone (Oleptro and others), trimipramine (Surmontil and others), venlafaxine (Effexor and others).

Author affiliations: Depression Clinical and Research Program at Massachusetts General Hospital (MGH), Harvard Medical School, Boston (all authors); and the Department of Psychiatry, University of Modena and Reggio Emilia, Modena, Italy (Dr Tedeschini).

Potential conflicts of interests: **Dr Fava** has served as an advisor/consultant for Abbott, Amarin, Aspect Medical Systems, AstraZeneca, Auspex, Bayer, Best Practice Project Management, Biovail, BrainCells, Bristol-Myers Squibb, Cephalon, Clinical Trials Solutions, CNS Response, Compellis, Cypress, Dov, Eli Lilly, EPIX, Fabre-Kramer, Forest, GlaxoSmithKline, Grunenthal, Janssen, Jazz, Johnson & Johnson, Knoll, Labopharm, Lorex, Lundbeck, MedAvante, Merck, Methylation Sciences, Neuronetics, Novartis, Nutrition 21, Organon, Pamlab, Pfizer, PharmaStar, Pharmavite, Precision Human Biolaboratory, Roche, sanofi-aventis, Sepracor, Schering-Plough, Solvay, Somaxon, Somerset, Synthelabo, Takeda, Tetragenex, Transcept, Vanda, and Wyeth-Ayerst; has received research support from Abbott, Alkermes, Aspect Medical Systems, AstraZeneca, Bio Research, BrainCells, Bristol-Myers Squibb, Cephalon, Clinical Trial Solutions, Eli Lilly, Forest, Ganeden, GlaxoSmithKline, Johnson & Johnson, Lichtwer, Lorex, NARSAD, the National Center for Complementary and Alternative Medicine, the National Institute on Drug Abuse, the NIMH, Novartis, Organon Inc, Pamlab, Pfizer, Pharmavite, Roche, sanofi-aventis, Shire, Solvay, Synthelabo, and Wyeth-Ayerst; has served as a speaker for Advanced Meeting Partners, American Psychiatric Association, AstraZeneca, Belvoir, Boehringer-Ingelheim, Bristol-Myers Squibb, Cephalon, Eli Lilly, Forest, GlaxoSmithKline, Imedex, Novartis, Organon, Pfizer, PharmaStar, MGH Psychiatry Academy/Primedia, MGH Psychiatry Academy/Reed-Elsevier, UBC, Wyeth-Ayerst Laboratories; has equity Holdings in Compellis; and has patent applications for SPCD and for a combination of azapirones and bupropion in MDD and copyright royalties for the MGH Cognitive and Physical Functioning Questionnaire, the MGH Sexual Functioning Inventory, the MGH Antidepressant Treatment Response Questionnaire, the Discontinuation-Emergent Signs and Symptoms scale, and SAFER. **Dr Papakostas** has served as a consultant for AstraZeneca, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Evotec, Inflabloc, Jazz, Otsuka, Pamlab, Pfizer, Pierre Fabre, Shire, and Wyeth; has received honoraria from AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Evotec, GlaxoSmithKline, Inflabloc, Jazz, Lundbeck, Otsuka, Pamlab, Pfizer, Pierre Fabre, Shire, Titan, and Wyeth; has received research support from Bristol-Myers Squibb, Forest, and the National Institute of Mental Health; and has served on the speakers bureaus for Bristol-Myers Squibb and Pfizer. **Dr Tedeschini** reports no competing interests.
Funding/support: None reported.

REFERENCES

1. Fava M, Evins AE, Dorer DJ, et al. The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. *Psychother Psychosom*. 2003;72(3):115–127.
2. Stein DJ, Baldwin DS, Dolberg OT, et al. Which factors predict placebo response in anxiety disorders and major depression? an analysis of placebo-controlled studies of escitalopram. *J Clin Psychiatry*. 2006;67(11):1741–1746.
3. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? a meta-regression of double-blind, randomized clinical trials in MDD. *Eur Neuropsychopharmacol*. 2009;19(1):34–40.
4. Carmody TJ, Rush AJ, Bernstein I, et al. The Montgomery Asberg and the Hamilton ratings of depression: a comparison of measures. *Eur Neuropsychopharmacol*. 2006;16(8):601–611.
5. Faries D, Herrera J, Rayamajhi J, et al. The responsiveness of the Hamilton Depression Rating Scale. *J Psychiatr Res*. 2000;34(1):3–10.
6. Trivedi MH, Rush H. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*. 1994;11(1):33–43.
7. Faries DE, Heiligenstein JH, Tollefson GD, et al. The double-blind variable placebo lead-in period: results from two antidepressant clinical trials. *J Clin Psychopharmacol*. 2001;21(6):561–568.
8. Lee S, Walker JR, Jakul L, et al. Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? a meta-analytic evaluation. *Depress Anxiety*. 2004;19(1):10–19.
9. Wernicke JF, Sayler ME, Koke SC, et al. Fluoxetine and concomitant centrally acting medication use during clinical trials of depression: the absence of an effect related to agitation and suicidal behavior. *Depress Anxiety*. 1997;6(1):31–39.
10. Gelenberg AJ, Thase ME, Meyer RE, et al. The history and current state of antidepressant clinical trial design: a call to action for proof-of-concept studies. *J Clin Psychiatry*. 2008;69(10):1513–1528.
11. Leon AC, Mallinckrodt CH, Chuang-Stein C, et al. Attrition in randomized controlled clinical trials: methodological issues in psychopharmacology. *Biol Psychiatry*. 2006;59(11):1001–1005.
12. Walsh BT, Seidman SN, Sysko R, et al. Placebo response in studies of major depression: variable, substantial, and growing. *JAMA*. 2002;287(14):1840–1847.
13. Stahl SM, Nierenberg AA, Gorman JM. Evidence of early onset of antidepressant effect in randomized controlled trials. *J Clin Psychiatry*. 2001;62(suppl 4):17–23.
14. Egger M, Smith GD, Altman DG. *Systematic Reviews in Health Care: Meta-Analysis in Context*. London, England: BMJ Books; 2001.

15. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition. Washington, DC: American Psychiatric Association; 1980.
16. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders,* Third Edition, Revised. Washington, DC: American Psychiatric Association; 1987.
17. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition*:* Washington, DC: American Psychiatric Association; 1994.
18. Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. *Arch Gen Psychiatry*. 1978;35(6):773–782.
19. Feighner JP, Robins E, Guze SB, et al. Diagnostic criteria for use in psychiatric research. *Arch Gen Psychiatry*. 1972;26(1):57–63.
20. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23(1):56–62.
21. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134(4):382–389.
22. Guy W. *ECDEU Assessment Manual for Psychopharmacology. US Dept Health, Education, and Welfare publication (ADM) 76-338*. Rockville, MD: National Institute of Mental Health; 1976:218–222.
23. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature, III: how to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA*. 1994;271(9):703–707.
24. Zarate CA Jr, Singh JB, Carlson PJ, et al. A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry*. 2006;63(8):856–864.

Editor's Note: We encourage authors to submit papers for consideration as a part of our Early Career Psychiatrists section. Please contact Marlene P. Freeman, MD, at mfreeman@psychiatrist.com.

© COPYRIGHT 2010 PHYSICIANS POSTGRADUATE PRESS, INC.