

Antidepressant Treatment of Depression: A Metaanalysis

Russell Joffe, MD¹, Stephen Sokolov, MD², David Streiner, PhD³

Objective: To carry out a metaanalysis of antidepressant studies to calculate the effect sizes for antidepressant effect in depressive disorder.

Method: A metaanalysis of all antidepressant studies that included an active comparison drug as well as placebo was used to calculate the effect size. Articles were selected from a MEDLINE search for the period January 1966 to June 1995. Forty-nine studies were included in the metaanalysis.

Results: The effect sizes for antidepressant treatment are moderately larger than for placebo. A larger effect size was observed in studies where objective diagnostic criteria for depression were used.

Conclusions: We conclude that the superior efficacy of antidepressants over placebo can be demonstrated.

(Can J Psychiatry 1996;41:613–616)

Key Words: depression, metaanalysis, antidepressant

Antidepressants are the most commonly used treatment for patients with major depressive disorders. All currently available compounds within the various classes of antidepressant have been shown to be more effective than placebo in the treatment of major depression. Moreover, several metaanalyses report equal efficacy between the 2 major classes of antidepressants, the tricyclic antidepressants (TCAs) and selective serotonin reuptake inhibitors (SSRIs) (1–3). In a recent metaanalysis, however, Greenberg and colleagues (4) question the robustness of the therapeutic effect of antidepressants, suggesting that, if outcomes were assessed according to stricter criteria, less subject to clinical bias, then effect sizes are much more modest than usually reported. They argue that the active comparison drug, rather than the new drug under study, is the most appropriate

treatment to compare with placebo in the metaanalysis because positive expectations of success are fewer for the active comparison than for the new drug under study and are therefore not as subject to observer bias. Using this design, they reported very modest effect sizes in their metaanalysis (4). The findings of Greenberg and others (4) need to be carefully considered. Several methodological concerns about their study, however, limit their conclusions. First, only 22 studies were included in the metaanalysis, which raises the possibility that these studies may not be representative of antidepressant studies in general. Second, 13 of the 22 studies concerned a single antidepressant, trazodone, further limiting the generalizability of their findings. Last, studies with unspecified diagnostic criteria for depression and varying durations of acute treatment were included, which may bias against finding a superior efficacy of active drug over placebo. We decided, therefore, to carry out a metaanalysis of antidepressant studies that included an active comparison or reference drug as well as a placebo to calculate the effect sizes for antidepressant effect in depressive disorder.

Method

MEDLINE was searched for the period 1966 to June 1995 for references containing the following MeSH terms in common: “depressive disorder,” “antidepressive agents,” and “English language.” MEDLINE was searched for references

Manuscript received May 1996, revised September 1996.

¹Professor, Department of Psychiatry, McMaster University, Hamilton, Ontario.

²Assistant Professor, Department of Psychiatry, McMaster University, Hamilton, Ontario.

³Professor, Departments of Clinical Epidemiology and Biostatistics and of Psychiatry, McMaster University, Hamilton, Ontario.

Address reprint requests to: Dr R Joffe, McMaster University Medical Centre, Room 3G56, 1200 Main Street West, Hamilton, ON L8N 3Z5

Drug	Number of Studies
Imipramine	32
Amitriptyline	12
Fluvoxamine	7
Trazodone	6
Moclobemide	6
Alprazolam	4
Fluoxetine	3
Nefazadone	3
Paroxetine	3
Zimelidine	3

containing each of the MeSH terms and the title word "placebo." A list of references was generated, and each reference was reviewed for its appropriateness to the study on the basis of the following criteria, developed by all the investigators: 1) human subjects, 2) unipolar depression, 3) ages 18 to 70 years, 4) "medically well," 5) randomized, double-blind, placebo-controlled trial, 6) acute treatment (that is, not a maintenance or continuation trial), 7) 3 parallel treatment cells (investigational drug, reference drug, and placebo), and 8) a trial of at least 4 weeks' duration. All 3 investigators reviewed studies for eligibility for inclusion, and interobserver reliability was good ($\kappa = 0.80$).

All of these articles were obtained and reviewed and were supplemented by articles contained in the original meta-analysis by Greenberg and others (4).

Articles were included in the final analysis only if they had response measures from which accurate data could be extracted. Observer measures included the Hamilton Rating Scale for Depression (HAM-D), both 17- and 21-item forms, the Montgomery-Asberg Depression Rating Scale (MADRS), and the Clinical Global Impression Scale (CGI). Although self-rating scales had been used in many of the studies, the data could not be extracted in the vast majority of studies in a form suitable for analysis.

Ninety-one studies met the inclusion criteria for analysis. Of these, however, 42 did not provide response measures that could be extracted in a form suitable for analysis. Therefore, 49 studies were included in the final analysis. (These studies are listed in the Bibliography.)

Results

Study Characteristics

Ninety-one studies met eligibility criteria, but response data could be extracted from only 49 studies. Of these, 5 (10.2%) took place in an inpatient setting, 40 (81.6%) in an outpatient setting, and one (2.0%) used both types of patients;

the setting was not specified in 3 studies (6.1%). The most frequently used set of criteria was DSM-III ($n = 23$; 46.9%), followed by DSM-III-R and Research Diagnostic Criteria ($n = 6$ each; 12.2%), Feighner ($n = 3$; 6.1%), and ICD-9 ($n = 1$; 2.0%). No criteria were specified in 10 cases (20.4%). The duration of follow-up was positively skewed, ranging from 4 to 106 weeks, with a mode of 6 weeks (mean = 7.78, SD = 14.39). The drugs used in the studies are listed in Table I.

The 17-item HAM-D was used in 33 studies (67.3%), the 21-item form in 13 studies (26.5%), the MADRS in one study (2.0%), and the CGI in 2 others (5.1%).

Effect Sizes

The effect size (ES) for the change score was calculated as the difference between the baseline and final score divided by the standard deviation (SD) of the baseline score. Since most studies did not report this latter figure, a conservative estimate for each scale was taken from the literature and used for each study that employed that scale. The mean ES for the drugs was 1.57 (SD = 0.53) and for the placebo groups 1.02 (SD = 0.36).

The difference in change scores between the active and placebo conditions was calculated by subtracting the raw change score (baseline to follow-up) for the placebo from the raw change score for the drug and dividing by the SD of the baseline score. This could not be done for one study, since none of the patients in the placebo arm completed the study. Following the procedure of Hedges and Olkin (5), the mean ES for each condition was calculated according to the equation:

$$\bar{d} = \frac{\sum d_j / \sigma_{(d_j)}^2}{\sum 1 / \sigma_{(d_j)}^2}$$

where d_j is the difference score for the j th study. The formula weights each difference ES by the sample size using the formula:

$$\sigma_{(d_j)}^2 = \frac{n_p + n_a}{n_p n_a} + \frac{d_j^2}{2(n_p + n_a)}$$

where n_p is the final sample size of the placebo group and n_a the sample size of the active group. The final sample size was the number of patients entering the trial minus the dropouts, when known. The effect of this weighting is that larger studies have more influence on the average value than do smaller studies. The consistency of the ESs across studies was calculated by:

$$Q = \sum \frac{(d_j - \bar{d})^2}{\sigma_{(d_j)}^2}$$

The mean ES for the active drugs relative to the placebo was 0.497, indicating that 69% of the patients on medication did better than the average person in the control condition. There was significant heterogeneity of ESs ($Q[95] = 175.51$, $P < 0.0001$), meaning that it was legitimate to search for variables that might explain these differences.

A regression equation was calculated, with 3 predictor variables: duration of the trial, type of drug (TCA versus SSRI), and the use of objective diagnostic criteria. Duration was not significant, but the use of criteria was ($\beta = 0.422$, $t = 3.027$, $P = 0.004$). Studies that used objective criteria had a mean ES of 0.520, and those which did not had a mean ES of 0.382. There was a tendency for TCAs to have a larger ES than SSRIs (0.511 versus 0.464), but the coefficient for this term was not significant ($\beta = 0.249$, $t = 1.784$, $P = 0.081$).

Discussion

Our metaanalysis suggests that the ESs for antidepressant treatment are moderately large when compared with placebo. Furthermore, the ES for the TCAs was nonsignificantly larger than for SSRIs. In the majority of cases, the TCAs were the active comparison or reference treatment rather than the new antidepressant under study. This suggests that even if this group is used, the superior therapeutic effect of antidepressants over placebo can be demonstrated. The ESs we observed are comparable to those reported in previous metaanalyses (1–3), with the exception of that of Greenberg and collaborators (4). Of interest, we found that the use of standardized criteria for the diagnosis of depression affected the ES. Specifically, studies that used objective criteria had a significantly higher mean ES than those which did not. This may be one of the factors contributing to the small ESs reported by Greenberg and others (4), considering that many of the studies included in their metaanalysis were performed prior to the introduction of the DSM-III criteria for psychiatric diagnosis.

Our metaanalysis addressed some of the limitations of the one performed by Greenberg and others (4). First, it included a larger number of studies and many that were performed after the introduction of standardized diagnostic criteria. Second, it evaluated the impact of various factors on treatment response, including the duration of the antidepressant trial. Third, a much larger number of studies were included in which a wider range of antidepressants were compared (see Table I). By contrast, Greenberg and others (4) included only 22 studies, which were focused largely on 2 antidepressants, trazodone and amoxapine. Our metaanalysis, however, did not include an evaluation of self-rated measures of depression. We found in our careful evaluation of the 49 studies that only a handful yielded sufficient data to be subject to such an analysis. The possibility that patient- rather than observer-rated data may be less subject to the bias of expectation of the clinician is an interesting possibility which should be considered further in future studies. Nonetheless, our findings,

together with other metaanalyses (1–3) and an extensive literature on a range of antidepressants, suggest that the therapeutic effect of antidepressants is easily demonstrated and is significantly better than placebo.

Clinical Implications

- Antidepressants are superior to placebo.
- The therapeutic benefits of antidepressants can be demonstrated.
- The data do support a specific therapeutic efficacy of antidepressants.

Limitations

- The studies included were limited to those in MEDLINE.
- The studies were limited to the English language.
- Only published studies were included.

References

1. Anderson IM, Tomenson BM. The efficacy of selective serotonin re-uptake inhibitors in depression: a meta-analysis of studies against tricyclic antidepressants. *J Psychopharmacol* 1994;8:238–42.
2. Kasper S, Fuger J, Moller HJ. Comparative efficacy of antidepressants. *Drugs* 1992;43:11–23.
3. Song S, Freemantle N, Sheldon TA, House A, Watson P, Long A, and others. Selective serotonin re-uptake inhibitors: meta-analysis of efficacy and acceptability. *BMJ* 1993;306:683–7.
4. Greenberg RP, Bornstein RF, Greenberg MD, Fisher S. A meta-analysis of antidepressant outcome under “blinder” conditions. *J Consult Clin Psychol* 1992; 60:664–9.
5. Hedges LV, Olkin I. *Statistical methods for meta-analysis*. Orlando: Academic Press; 1995.

Bibliography

1. Amsterdam JD, Case WG, Csanalosi E, Singer M, Rickels K. A double-blind comparative trial of zimelidine, amitriptyline and placebo in patients with mixed anxiety and depression. *Pharmacopsychiatry* 1986;19:115–9.
2. Bakish D, Wiens A, Ellis J, Alda M, Lapierre Y. A double-blind, placebo-control comparison of moclobemide and amitriptyline in the treatment of depression. *Can J Psychiatry* 1992;37(1 Suppl):12S–17S.
3. Cassano GB, Conti L, Massimetti G, Mengali F, Waekelin JS, Levine J. Use of a standardized documentation system (BLIPS/BDP) in the conduct of a multicentre international trial comparing fluvoxamine, imipramine and placebo. *Psychopharmacol Bull* 1986;22:52–8.
4. Chouinard G, Saxena BM, Nair NPV, Kutcher SP, Bakish D, Bradwejn J, and others. A Canadian multicentre placebo-control study of a fixed dose of brofaromine, a reversible selective MAO-A inhibitor, in the treatment of major depression. *J Affect Disord* 1994;32:105–14.
5. Claghorn J, Gershon S, Goldstein BJ, Behmetz S, Bush DF, Huitfeldt B. A double-blind evaluation of zimelidine in comparison to placebo and amitriptyline in patients with major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 1983;7:367–92.
6. Cohn JB, Wilcox C. A comparison of fluvoxamine, imipramine and placebo in patients with major depressive disorder. *J Clin Psychiatry* 1985;46:26–31.
7. Cunningham LA, Borison RL, Carman JS, Chouinard G, Crowder JE, Diamond BI, and others. A comparison of venlafaxine, trazodone, and placebo in major depression. *J Clin Psychopharmacol* 1994;14:99–106.
8. Dominguez R, Jacobson AF, Weiss BL, Goldstein BJ, Steinbook RM. A placebo-control comparison of amoxapine and imipramine in the treatment of depressed outpatients. *Current Therapeutic Research* 1981;29:714–27.
9. Dominguez RA, Goldstein BJ, Jacobson AF, Steinbook RM. A double-blind placebo-control study of fluvoxamine and imipramine in depression. *J Clin Psychiatry* 1985;46:84–7.

10. Doogan DP, Langdon CJ. A double-blind, placebo-controlled comparison of sertraline and dothiepin in the treatment of major depression in general practice. *Int Clin Psychopharmacol* 1994;9:95-100.
11. Dunbar GC, Cohn JB, Fabre LF, Feighner JP, Fieve RR, Mendels J, and others. A comparison of paroxetine, imipramine and placebo in depressed outpatients. *Br J Psychiatry* 1991;159:394-8.
12. Escobar JI, Gomez J, Constain C, Rey J, Santacruz H. Controlled clinical trial with trazodone, a novel antidepressant: a self American experience. *J Clin Pharmacol* 1980;20:124-30.
13. Fabre LF, McLendon DM, Gainey A. Trazodone efficacy in depression: a double-blind comparison with imipramine and placebo in day hospital type patients. *Current Therapeutic Research* 1979;25:826-34.
14. Feighner JP. Trazodone, a triazolopyridine derivative, in primary depressive disorder. *J Clin Psychiatry* 1980;41:250-5.
15. Feighner JP, Aden GC, Fabre LF, Rickels K, Smith WT. Comparison of alprazolam, imipramine and placebo in the treatment of depression. *JAMA* 1983;249:3057-64.
16. Feighner JP, Boyer WF, Meredith CH, Hendrickson GG. A double-blind comparison of fluvoxamine, imipramine and placebo in outpatients with major depression. *Int Clin Psychopharmacol* 1989;94:127-34.
17. Feighner JP, Boyer WF, Meredith CH, Hendrickson GG. A placebo-controlled inpatient comparison of fluvoxamine maleate and imipramine in major depression. *Int Clin Psychopharmacol* 1989;4:239-44.
18. Feighner JB, Pambakain R, Fowler RC, Boyer WF, D'Amico MF. A comparison of nefazodone, imipramine, and placebo in patients of moderate to severe depression. *Psychopharmacol Bull* 1989;25:219-21.
19. Ferguson JM, Mendels J, Manowitz NR. Dothiepin versus doxepin in major depression: results of a multicentre, placebo-control trial. *J Clin Psychiatry* 1994;55:258-63.
20. Fontaine R, Ontivros A, Elie R. A double-blind comparison of nefazodone, imipramine and placebo in major depression. *J Clin Psychiatry* 1994;55:234-41.
21. Gelenberg AJ, Wojcik JD, Falk WE, Spring B, Brokman AW, Glavin-Nadeau M. Clovoxamine in the treatment of depressed outpatients: a double-blind parallel-group comparison against amitriptyline and placebo. *Compr Psychiatry* 1990;31:307-14.
22. Gershon SM, Mann JJ, Newton R, Gunther BJ. Evaluation of trazodone in the treatment of endogenous depression: results of a multicentre double-blind study. *J Clin Psychopharmacol* 1981;(1 Suppl):39S-44S.
23. Hicks F, Robins E, Murphy GE. Comparison of adinazolam, amitriptyline and placebo in the treatment of melancholic depression. *Psychiatry Res* 1988;23:221-7.
24. Hormazabal L, Omer LMO, Ismail S. Cyanopramine and amitriptyline in the treatment of depressed patients—a placebo-control study. *Psychopharmacology* 1985;86:205-8.
25. Lapiere YD, Browne M, Horn E, Oyewumi LK, Sarantidis D, Roberts N, and others. Treatment of major affective disorder with fluvoxamine. *J Clin Psychiatry* 1987;48:65-8.
26. Larsen JK, Holm P, Hoyer E, Mejlhede A, Mikkelsen FL, Olesen A, and others. Moclobemide and clomipramine in reactive depression: a placebo-controlled randomized trial. *Acta Psychiatr Scand* 1989;79:530-6.
27. Lydiard RB, Laird LK, Morton WA, Steele TE, Kellner C, Laraia MT, and others. Fluvoxamine, imipramine and placebo in the treatment of depressed outpatients: effects on depression. *Psychopharmacol Bull* 1989;25:68-71.
28. Mann JJ, Georgotas A, Newton R, Gershon S. A controlled study of trazodone, imipramine and placebo in outpatients with endogenous depression. *J Clin Psychopharmacol* 1981;1:75-80.
29. March JS, Kobak KA, Jefferson JW, Mazza J, Greist JH. A double-blind, placebo-control trial of fluvoxamine versus imipramine in outpatients with major depression. *J Clin Psychiatry* 1990;51:200-2.
30. Merideth CH, Feighner JP. A double-blind controlled evaluation of zimeldine, imipramine and placebo in patients with primary affective disorders. *Acta Psychiatr Scand* 1983;68(308 Suppl):70S-79S.
31. Muijen M, Roy B, Silverstone T, Mehmet A, Christie M. A comparative clinical trial of fluoxetine, mianserin and placebo in depressed outpatients. *Acta Psychiatr Scand* 1988;78:384-90.
32. Norton KRW, Sireling LI, Bhat AV, Rao B, Paykel ES. A double-blind comparison of fluvoxamine, imipramine and placebo in depressed patients. *J Affect Disord* 1984;7:297-308.
33. Paykel ES, Rowan PR, Parker RR, Bhat AV. Response to phenelzine, amitriptyline in subtypes of outpatient depression. *Arch Gen Psychiatry* 1982;39:1041-9.
34. Rickels SK, Case WG, Werbolowsky JA, Csanalosi I, Schless A, Weise CC. Amoxapine and imipramine in the treatment of depressed outpatients: a controlled study. *Am J Psychiatry* 1981;138:20-4.
35. Rickels K, Weise CC, Zal HM, Csanalosi I, Werblowsky J. Lofepamine and imipramine in unipolar depressed outpatients. *Acta Psychiatr Scand* 1982;66:104-20.
36. Rickels K, Feighner JP, Smith WT. Alprazolam, amitriptyline, doxepin and placebo in the treatment of depression. *Arch Gen Psychiatry* 1985;42:134-41.
37. Rickels K, Chung HR, Csanalosi IB, Hurowitz AM, London J, Wiseman K, and others. Alprazolam, diazepam, imipramine and placebo in outpatients with major depression. *Arch Gen Psychiatry* 1987;44:862-6.
38. Rickels K, London J, Fox I, Hassman H, Csanalosi I, Weise C. Adinazolam, diazepam, imipramine and placebo in major depressive disorder: a controlled study. *Pharmacopsychiatry* 1991;24:127-31.
39. Rickels K, Schweizer E, Clary C, Fox I, Weise C. Nefazodone and imipramine in major depression: a placebo-control trial. *Br J Psychiatry* 1994;154:802-5.
40. Spring B, Gelenberg AJ, Garvin R, Thompson S. Amitriptyline, clovoxamine and cognitive function. A placebo-control comparison in depressed outpatients. *Psychopharmacology* 1992;108:327-32.
41. Shrivastava RK, Shrivastava SHP, Overweg N, Blumhart CL. A double-blind comparison of paroxetine, imipramine and placebo in major depression. *J Clin Psychiatry* 1992;53(2 Suppl):48S-51S.
42. Smith WT, Glaudin V, Panagides J, Gilvary E. Mirtazapine versus amitriptyline versus placebo in the treatment of major depressive disorder. *Psychopharmacol Bull* 1990;26:191-6.
43. Thompson J, Rankin H, Ashcorft W, Yates CM, McQueen JK, Cummings SW. The treatment of depression in general practice: a comparison of L-tryptophan, amitriptyline, and a combination of L-tryptophan in amitriptyline with placebo. *Psychol Med* 1982;12:741-51.
44. Trapp GA, Handorf CR, Larach V. Trazodone in the treatment of depressed inpatients. *Psychopharmacol Bull* 1979;15:25-7.
45. Udabe RU, Marquez GA, Traballi CA, Portes N. Double-blind comparison of moclobemide, imipramine and placebo in depressed patients. *Acta Psychiatr Scand* 1990;360:54-6.
46. UK Moclobemide Study Group. A multi-centre comparative trial of moclobemide, imipramine and placebo in major depressive disorder. *Int Clin Psychopharmacol* 1994;9:109-13.
47. Versiani M, Oggero U, Alterwain P, Capponi R, Dajas F, Heinze-Martin G, and others. A double-blind comparative trial of moclobemide versus imipramine and placebo in major depressive episodes. *Br J Psychiatry* 1989;155(6 Suppl):72S-77S.
48. Versiani M, Nardi AE, Mundim FD, Alves AB. Moclobemide, imipramine and placebo in the treatment of major depression. *Acta Psychiatr Scand* 1990;360:57-8.
49. Wilcox CS, Cohn JB, Katz BB. A double-blind, placebo-controlled study comparing mianserin and amitriptyline in moderately depressed outpatients. *Int Clin Psychopharmacol* 1994;9:271-9.

Résumé

Objectif : Réaliser une méta-analyse des études sur les antidépresseurs pour calculer l'ampleur de l'effet antidépresseur dans le cadre du trouble dépressif.

Méthode : Une méta-analyse de toutes les études sur les antidépresseurs portant sur un médicament de comparaison actif, de même qu'un placebo, a servi à calculer l'ampleur de l'effet. Les articles étaient choisis à partir d'une recherche de MEDLINE pour la période de 1966 à juin 1995. Quarante-neuf études ont fait l'objet de la méta-analyse.

Résultats : L'ampleur de l'effet du traitement antidépresseur est passablement plus importante que celle du placebo. Dans les études faisant appel à des critères diagnostiques objectifs de dépression, on a observé une ampleur plus importante de l'effet.

Conclusions : Nous concluons qu'il est possible de démontrer que l'efficacité des antidépresseurs est supérieure à celle du placebo.