

Effectiveness of paroxetine in the treatment of acute major depression in adults: a systematic re-examination of published and unpublished data from randomized trials

Corrado Barbui MD, Toshiaki A. Furukawa MD, Andrea Cipriani MD

ABSTRACT

Background: Concern has been raised about the efficacy of antidepressant therapy for major depression in adults. We undertook a systematic review of published and unpublished clinical trial data to determine the effectiveness and acceptability of paroxetine.

Methods: We searched the Cochrane Collaboration Depression, Anxiety and Neurosis Controlled Trials Register, the Cochrane Central Register of Controlled Trials, the Glaxo-SmithKline Clinical Trial Register, MEDLINE and EMBASE up to December 2006. Published and unpublished randomized trials comparing paroxetine with placebo in adults with major depression were eligible for inclusion. We selected the proportion of patients who left a study early for any reason as the primary outcome measure because it represents a hard measure of treatment effectiveness and acceptability.

Results: We included in our review 29 published and 11 unpublished clinical trials, with a total of 3704 patients who received paroxetine and 2687 who received with placebo. There was no difference between paroxetine and placebo in terms of the proportion of patients who left the study early for any reason (random effect relative risk [RR] 0.99, 99% confidence interval [CI] 0.88–1.11). Paroxetine was more effective than placebo, with fewer patients who did not experience improvement in symptoms of at least 50% (random effect RR 0.83, 99% CI 0.77–0.90). Significantly more patients in the paroxetine group than in the placebo group left their respective studies because of side effects (random effect RR 1.77, 95% CI 1.44–2.18) or experienced suicidal tendencies (odds ratio 2.55, 95% CI 1.17–5.54).

Interpretation: Among adults with moderate to severe major depression in the clinical trials we reviewed, paroxetine was not superior to placebo in terms of overall treatment effectiveness and acceptability. These results were not biased by selective inclusion of published studies.

Une version française de ce résumé est disponible à l'adresse www.cmaj.ca/cgi/content/full/178/3/296/DC1

CMAJ 2008;178(3):296-305

Although most treatment guidelines recommend one of the selective serotonin reuptake inhibitors in the pharmacologic treatment of moderate to severe depression in adults,^{1,2} concerns have been raised in recent years about the efficacy of these drugs in alleviating symptoms of depression.^{3,4}

First, in trials of antidepressants, the choice of the outcome of interest is often problematic. Whereas in other fields of medicine the definition of outcome measures may be relatively straightforward, efficacy in the treatment of depression may be an elusive concept, typically measured by rating scales. The use of rating scales as outcome measures has often been questioned by physicians, who seldom use them to define patients' improvement under clinical circumstances.⁵ In addition, improvement in symptoms is typically documented as the difference between baseline and post-treatment scores. Although from a practical viewpoint this approach seems reasonable, in that it allows physicians to make a reasoned judgment in terms of proportions of patients (and not in terms of means and standard deviations), it may systematically magnify the effect of new drugs relative to placebo.³ As a consequence of this criticism, the field of mental health has seen a renewal of interest in "hard measures" of treatment effectiveness.^{6–8} Hard measures include suicide attempts, treatment switching, hospital admissions, job loss or dropping out of the trial itself. Hard outcomes may also have a role in re-analyses of clinical trial data, where they may offer a "down-to-earth" evaluation of effectiveness and acceptability.

A second concern is that the modest differences between active antidepressants and placebo, as calculated in systematic reviews of clinical trial data,^{9–12} might be explained in part by the selective inclusion of specific subsets of studies, such as those submitted to regulatory authorities by drug companies, those that were eventually published or those supported by the manufacturers of selective serotonin reuptake inhibitors.¹³

In the present systematic review, we used a hard measure to determine the effectiveness and acceptability of paroxetine,

From the Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, World Health Organization Collaborating Centre for Research and Training in Mental Health (Barbui, Cipriani), University of Verona, Verona, Italy; and the Department of Psychiatry and Cognitive-Behavioral Medicine (Furukawa), Nagoya City University Graduate School of Medical Sciences, Mizuho-cho, Mizuho-ku, Nagoya, Japan

a commonly prescribed antidepressant belonging to the group of selective serotonin reuptake inhibitors. Paroxetine was chosen for 3 reasons: first, it is one of the most frequently prescribed antidepressant drugs both in primary and secondary care (as indicated by a search of the US National Center for Health Statistics database, www2.cdc.gov/drugs/); second, GlaxoSmithKline, the company that launched paroxetine, has recently adopted and implemented a disclosure policy to provide easy access to all published and unpublished data from clinical trials that it has sponsored; and third, the effectiveness and acceptability of paroxetine are particularly relevant in view of recent concerns about its safety profile, especially in terms of suicidal tendencies among pediatric and adult patients with major depression.^{14–17}

Methods

Studies, participants and interventions

We included only randomized placebo-controlled trials in our systematic review. For trials with a crossover design, only results from the first randomization period were considered. Included and excluded studies were collected according to the flow diagram of the Quality of Reporting of Meta-analyses statement.¹⁸ Considerable care was taken to exclude duplicate publications.

Study participants were adults (≥ 18 years of age) of either sex with a diagnosis of major depression. We included studies using any criteria to define depression. Concurrent diagnosis of another medical disorder was not considered an exclusion criterion.

We included trials if they compared paroxetine with placebo in the treatment of acute major depression. Included trials used any flexible-dose regimen of paroxetine, any fixed-dose regimen of 20 mg/d or more, and any pharmaceutical form of the drug.

Primary outcome measure

For the purposes of this review, we selected the proportion of patients who left a study early for any reason, a proxy measure of treatment discontinuation, as the primary outcome measure because it represents a hard measure of treatment effectiveness and acceptability.

Secondary outcome measures

We also analyzed 2 secondary outcomes: the proportion of patients who had a reduction of at least 50% in the Hamilton Depression Rating Scale or the Montgomery–Asberg Depression Rating Scale or who scored “much improved” or “very much improved” on the Clinical Global Impression scale; and the group mean change from baseline to end point (or the group mean scores at the end of the trial) on the Hamilton Depression Rating Scale or the Montgomery–Asberg Depression Rating Scale or any other depression scale.

Tolerability was evaluated with the following measures: proportion of patients who left the study because of side effects; proportion of patients with any adverse events; proportion of patients with any serious adverse events, according to the definition used by the authors of the study in question; proportion of patients who completed suicide; and propor-

tion of patients with suicidal tendencies, including those who completed or attempted suicide or experienced worsening of suicidal thoughts or emotional lability (a term used in antidepressant trials to indicate attempted suicide, self-harm, suicidal thoughts, crying and mood fluctuations).

Search strategy for identification of studies

To identify relevant studies, we searched the Cochrane Collaboration Depression, Anxiety and Neurosis Controlled Trials Register and the Cochrane Central Register of Controlled Trials up to Dec. 31, 2006. We used the following terms: keyword = “major depression” or “depression” and free-text = “paroxetin*.” Details of the search strategy used for the Cochrane Collaboration Depression, Anxiety and Neurosis Controlled Trials Register can be obtained at www.iop.kcl.ac.uk/iop/ccdan/index.htm. In addition, we searched MEDLINE (1966–2006) and EMBASE (1974–2006) using the search terms “paroxetine” and “randomized controlled trial” or “random allocation” or “double-blind method.” Articles published in languages other than English were included. The reference lists of relevant papers and previous systematic reviews were searched for other relevant studies.

We searched the GlaxoSmithKline Clinical Trial Register (<http://ctr.gsk.co.uk/medicinelist.asp>) using the term “paroxetine.” In addition, the Clinical Study Results Database (www.clinicalstudyresults.org), a web-based repository for clinical trials that was developed to render clinical study results for US-marketed pharmaceuticals more transparent, was searched for other relevant published and unpublished studies. The websites of the US Food and Drug Administration (www.fda.gov/cder/index.html) and the UK Medicines and Healthcare Products Regulatory Agency (www.mhra.gov.uk) were also searched for additional information. Interrater agreement for article selection was checked with the kappa coefficient.

Data extraction

Using a standardized form, 2 reviewers (C.B. and A.C.) independently extracted data on participant characteristics, intervention details and outcome measures. Disagreements were resolved by discussion and consensus with the third author (T.A.F.).

Suicide-related events were coded according to the US Food and Drug Administration coding system:¹⁹ code 1 = completed suicide, code 2 = suicide attempt, code 3 = preparatory acts toward imminent suicidal behaviour, code 4 = suicidal ideation, code 5 = self-injurious behaviour with unknown intent, code 6 = not enough information (fatal) and code 9 = not enough information (nonfatal).

For continuous outcomes, we extracted the mean change from baseline to the particular end point or, if the mean change was unavailable, the mean scores at end point, along with the standard deviation (SD) or standard error of this value and the number of patients included in the analysis.²⁰ The extracted scores for depression were based on the Hamilton Depression Rating Scale or Montgomery–Asberg Depression Rating Scale or any other depression scale. When only the standard error was reported, it was converted to the SD according to Altman and Bland.²¹ When neither SD nor stan-

standard error was reported for a particular study, we used the mean of known SDs for the group of included studies (calculated according to Furukawa and colleagues²²).

Study quality

Two reviewers (C.B. and A.C.) independently assessed trial quality in accordance with the *Cochrane Handbook for Systematic Reviews of Interventions*.²³ Studies were given a quality rating, ranging from C (poorest quality) to A (best quality), where C = inadequate concealment (e.g., alternation or an open random number table), B = inadequate details about how the randomization procedure was carried out, and A = adequate measures to conceal allocation (e.g., serially numbered, opaque, sealed envelopes; numbered or coded bottles or containers). Interrater agreement for quality assessment was checked by calculating the kappa score.

Statistical analysis

Trials comparing different doses or forms of paroxetine with placebo were converted into 2-arm trials by summing samples and averaging doses.²⁴ All analyses were defined a priori. We determined the main outcome by calculating the proportion of patients (relative to the total number of patients who underwent randomization) who did not complete the study. The number of patients who experienced a response to treatment was calculated on an intention-to-treat basis: dropouts were always included in such analyses. For studies in which data for dropouts had been carried forward and included in the efficacy evaluation (last observation carried forward), the data were analyzed according to the primary studies; when dropouts had been excluded from any assessment in the primary studies, we considered them to be drug failures.

For dichotomous outcomes, with rare exceptions, we calculated the relative risk (RR) on the basis of the random-effects model, because this type of model takes into account differences between studies even if there is no statistically significant heterogeneity.²³ For certain rare outcomes, we calculated Peto odds ratios (Peto OR) for the following patients: those with any serious adverse events, those who completed suicide and those who experienced suicidal tendencies. RR or OR values greater than 1.0 indicate that the results favour placebo over paroxetine. We checked the data to see if analysis using a fixed-effects model would make any substantive difference for outcomes that were not statistically significantly heterogeneous. When the overall results were significant, we calculated the number needed to treat as the inverse of the risk difference.

We analyzed continuous scores from different outcome scales using standardized mean differences. A random-effects model was used.

We calculated a 99% confidence interval (CI) for all efficacy estimates, as recommended by Cipriani and associates.²⁴ We adopted this approach, instead of an approach using the 95% CI, to generate the widest estimate of likely true effect. We set the level of significance at 0.01, because we were making multiple comparisons and we reasoned that only robust differences between treatments should inform clinical practice. In other words, we gave priority to avoiding a type I

rather than a type II error.²⁴ Conversely, we calculated 95% CIs for all tolerability estimates.

Visual inspection of graphs was used to investigate the possibility of statistical heterogeneity. This was supplemented with, primarily, the I^2 statistic. We interpreted I^2 estimates of at least 50% as indicating high levels of heterogeneity.²⁵

Subgroup analyses

We planned the following subgroup analyses in advance: trials using a paroxetine fixed-dose regimen of 20 mg/d versus trials using flexible-dose regimens or fixed-dose regimens of more than 20 mg/d; trials involving patients with moderate to severe major depression (defined as a group mean baseline score of 18 or more on the Hamilton Depression Rating Scale or Montgomery–Asberg Depression Rating Scale) versus trials involving patients with mild to moderate major depression; and published versus unpublished trials.

Results

Characteristics of included studies

The original searches yielded 72 studies. Of these, 21 were excluded because of inappropriate design (not a randomized or placebo-controlled study) (see online Appendix 1 for references, available at www.cmaj.ca/cgi/content/full/178/3/296/DC2). Of the remaining 51 studies retrieved for more detailed evaluation, 42 met our inclusion criteria and 40 provided data suitable for re-analysis (Figure 1) (see online Appendix 2 for

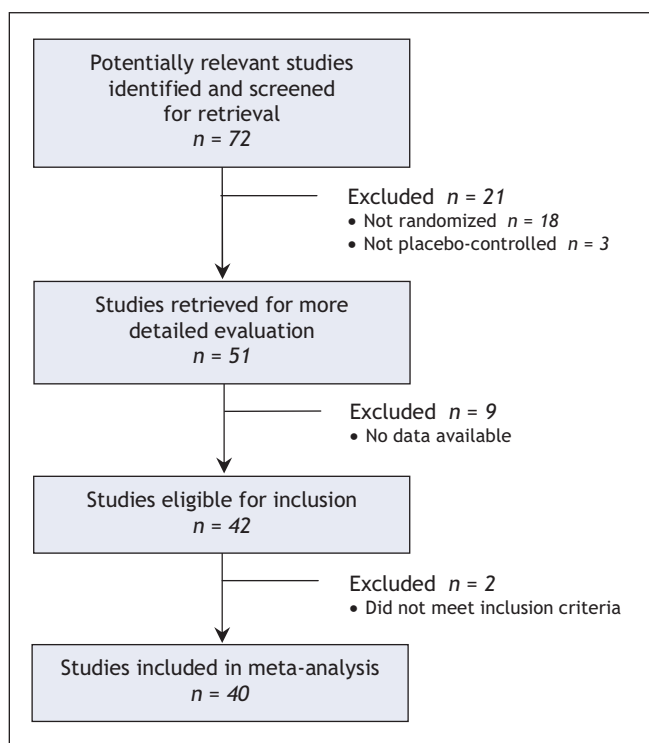


Figure 1: Selection of randomized placebo-controlled trials for systematic review of effectiveness of paroxetine in the treatment of major depression in adults.

references, available at www.cmaj.ca/cgi/content/full/178/3/296/DC2). There was substantial interrater agreement for article selection ($\kappa = 0.77$). Half of the studies ($n = 20$) had recruited fewer than 100 participants, and almost all ($n = 38$) were double-blind (see online Appendix 3, available at www.cmaj.ca/cgi/content/full/178/3/296/DC2). The mean length of follow-up was 7.5 weeks (SD 2.3).

The majority of trials ($n = 47$) enrolled outpatients, and in well over half of the studies ($n = 31$) the diagnosis of major depression was based on criteria of the *Diagnostic and Statistical Manual of Disorders* (third edition, third edition revised or fourth edition) or the International Statistical Classification of Diseases and Related Health Problems, 10th revision. Elderly subjects (over 65 years of age) were included in 10 of the studies (see online Appendix 3, available at www.cmaj.ca/cgi/content/full/178/3/296/DC2). In 36 studies, people with moderate to severe depression were enrolled, whereas in the other 4 studies, people with mild to moderate symptoms were enrolled. Patients with medical comorbidities were included in 5 studies. The description of concealment of allocation was rated as B in all but one study (rated A). Interrater agreement was good for description of concealment of allocation ($\kappa = 1.00$) and for blinding ($\kappa = 0.79$).

Primary outcome

Data from the 40 trials (involving a total of 3704 patients who received paroxetine and 2687 who received placebo) showed an absence of a positive effect of paroxetine in terms of the proportion of patients who discontinued treatment for any reason (random effect RR 0.99, 99% CI 0.88–1.11) (Figure 2). There was no between-study heterogeneity ($I^2 = 21.0\%$).

Depression scores

Data extracted from 22 trials (involving a total of 3032 patients who received paroxetine and 2080 who received placebo) showed a statistically significant positive effect of paroxetine in terms of the proportion of patients who did not show an improvement of at least 50% on depression measures (random effect RR 0.83, 99% CI 0.77–0.90; number needed to treat to avoid 1 additional failure 9, 99% CI 7–14) (Figure 3). In other words, a greater proportion of patients who received paroxetine than of those given placebo had an improvement of 50% or more. There was no statistically significant between-study heterogeneity ($I^2 = 19.2\%$). In terms of continuous outcomes, data extracted from 34 trials (involving a total of 3325 patients who received paroxetine and 2439 who received placebo) showed a statistically significant positive effect of paroxetine in terms of mean difference (standardized mean difference -0.31 , 99% CI -0.40 to -0.22) (Figure 4). There was no statistically significant between-study heterogeneity ($I^2 = 31.5\%$).

Tolerability

Data from 38 trials (involving a total of 3593 patients who received paroxetine and 2572 who received placebo) showed that significantly more patients assigned to receive paroxetine left the study because of side effects (random effect RR 1.77, 95% CI 1.44–2.18; number needed to harm 17, 95% CI 14–

25). There was no statistically significant between-study heterogeneity ($I^2 = 2.0\%$). In terms of patients reporting adverse events, data from 35 trials (involving a total of 3327 patients who received paroxetine and 2382 who received placebo) showed that significantly more patients assigned to receive paroxetine than of those given placebo reported any adverse event (random effect RR 1.15, 95% CI 1.11–1.19; number needed to harm 9, 95% CI 7–11) (Figure 5). There was no statistically significant between-study heterogeneity ($I^2 = 17.6\%$). Data from 34 trials (involving a total of 3327 patients who received paroxetine and 2382 who received placebo) showed no statistically significant difference between paroxetine and placebo in terms of patients reporting any serious adverse event (Peto OR 1.27, 95% CI 0.88–1.83). There was no statistically significant between-study heterogeneity ($I^2 = 0\%$).

Suicidal tendencies

Two patients, one who was receiving paroxetine and one who was receiving placebo, committed suicide. Data from 15 trials (involving a total of 2353 patients who received paroxetine and 1386 who received placebo) showed that significantly more patients assigned to receive paroxetine than of those given placebo experienced suicidal tendencies (US Food and Drug Administration codes 1 to 9) (Peto OR 2.55, 95% CI 1.17–5.54; number needed to harm 142, 95% CI 7–3333). There was no statistically significant between-study heterogeneity ($I^2 = 0\%$) (Figure 6).

Subgroup analyses

Overall treatment estimates from trials that used a 20-mg dose of paroxetine did not differ substantially from estimates from trials that used doses larger than 20 mg in terms of the proportion of patients discontinuing treatment for any reason. Similarly, overall treatment estimates for trials involving patients with moderate to severe depression did not differ substantially from estimates from trials involving patients with mild to moderate depression; however, only 4 studies involved patients with mild to moderate depression, and these yielded estimates with wide confidence intervals. Finally, there was no difference between drug and placebo groups on the basis of whether the study was published or not.

Interpretation

In this systematic review of published and unpublished studies comparing paroxetine with placebo in adults with major depression, we found that the drug was not superior to placebo in terms of the proportion of patients who discontinued treatment for any reason. However, when we examined the results using specific measures of depression, we found that paroxetine was significantly superior to placebo.

For continuous measures, the effect of paroxetine was small. However, it could be argued that concluding that paroxetine is not effective on the basis of a small mean difference may be erroneous, given that a small difference may have an important impact on a substantial number of patients.²⁶ For dichotomous measures, we found that, for every 100 adults with depression who received treatment with

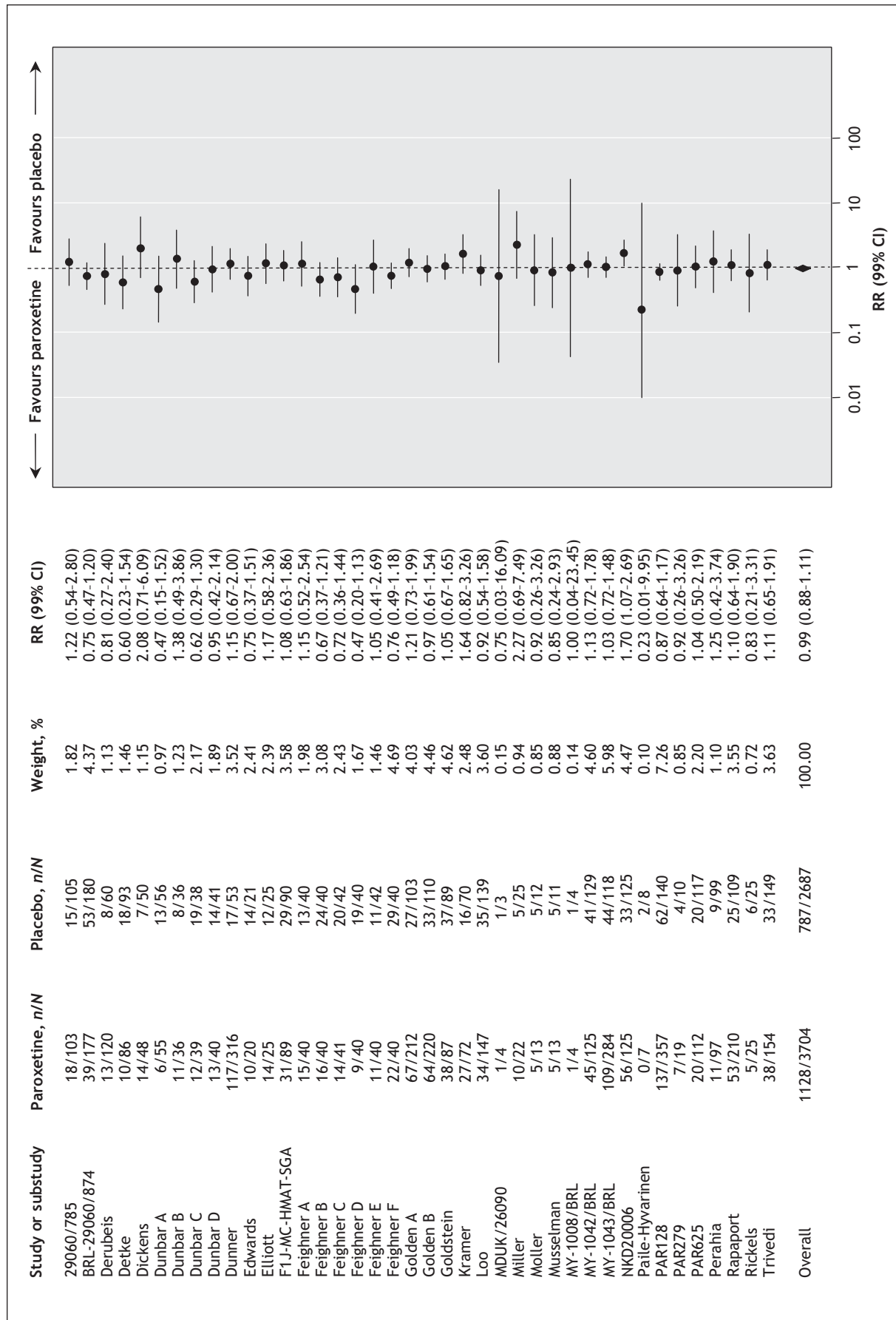


Figure 2: Random effects meta-analysis of the effect of paroxetine versus placebo on the proportion of patients who discontinued treatment for any reason. *n* = number of patients leaving the study for any reason, *N* = total number of patients, RR = relative risk, CI = confidence interval. Study designations are defined in online Appendix 2 (www.cmaj.ca/cgi/content/full/178/3/296/DC2).

paroxetine, 53 experienced a favourable response; of these, however, 42 would have had a favourable response to placebo, such that the response was attributable to paroxetine in only 11 cases. This means that physicians would need to expose 100 patients to paroxetine to provide benefit to 11. This modest effect was offset by tolerability problems, as indicated by patients who left studies because of adverse effects and by patients who reported adverse effects (even though they remained in the study).

A limitation of the present analysis is that the original studies were not designed or powered to determine the number of patients discontinuing treatment as a primary outcome. However, the nature of this hard outcome is such that the design should not have affected the relative comparison, and pooling all available studies addressed the issue of lack of power, as evidenced by the narrow confidence intervals. It might also be argued that, although patients in antidepressant trials typically discontinue treatment because of inefficacy or side effects, some individuals may have dropped out once they felt better. Although we acknowledge this possibility, in the field of antidepressant therapy patients are strongly encouraged to continue with the medication once they feel better. If patients did not continue with prescribed therapy once they felt better, as required in a double-blind clinical trial, then this seems an indication that treatment did not work or was problematic. From this perspective, treatment discontinuation seems to offer a practical evaluation of both effectiveness and tolerability of antidepressant therapy.

According to Moncrieff and Kirsch³ a very small difference in symptom score can translate

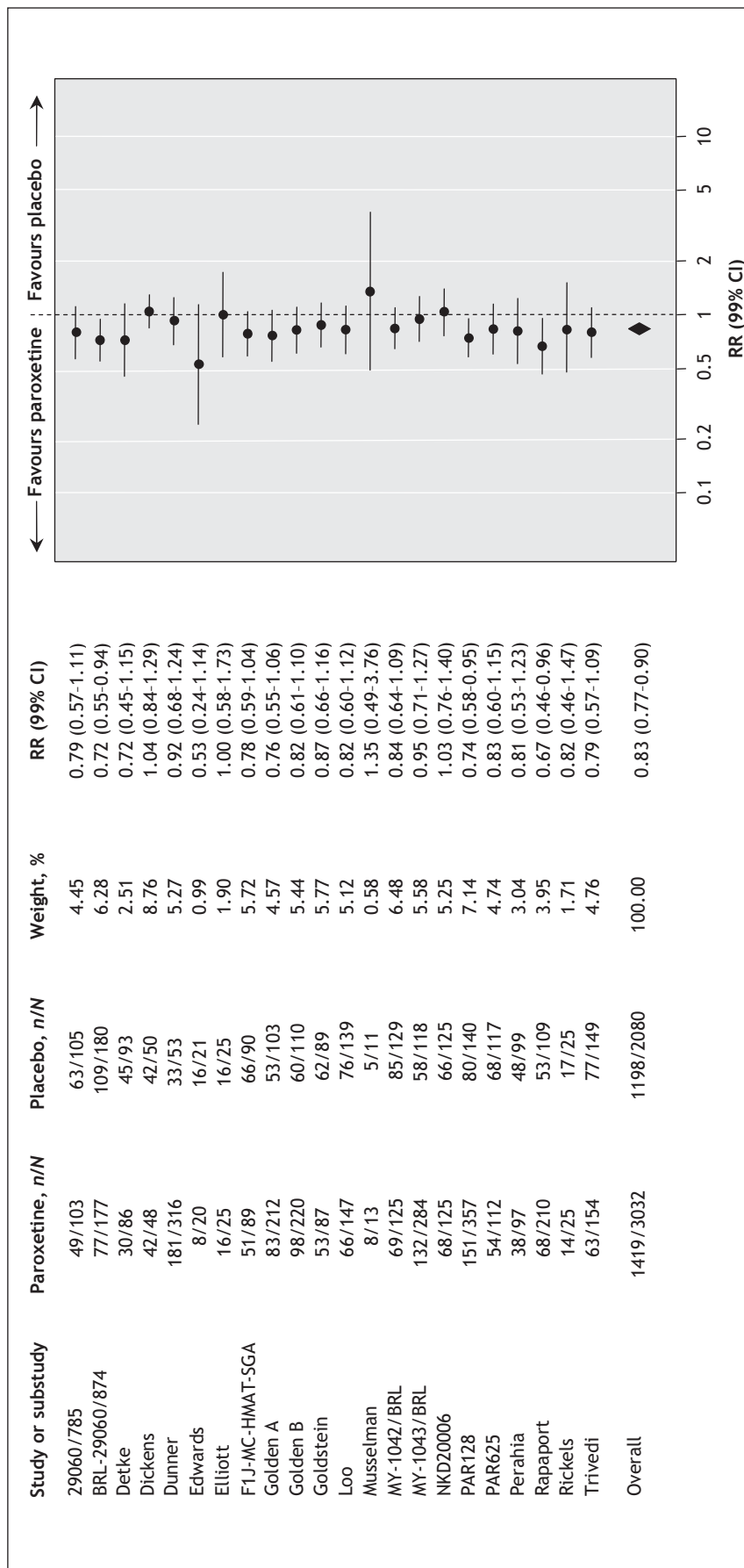


Figure 3: Random effects meta-analysis of the effect of paroxetine versus placebo on the proportion of patients who showed less than 50% improvement in measures of depression. n = number of patients who did not experience an improvement of at least 50% on depression measure, N = total number of patients, RR = relative risk, CI = confidence interval. Study designations are defined in online Appendix 2 (www.cmaaj.ca/cgi/content/full/178/3/296/DC2).

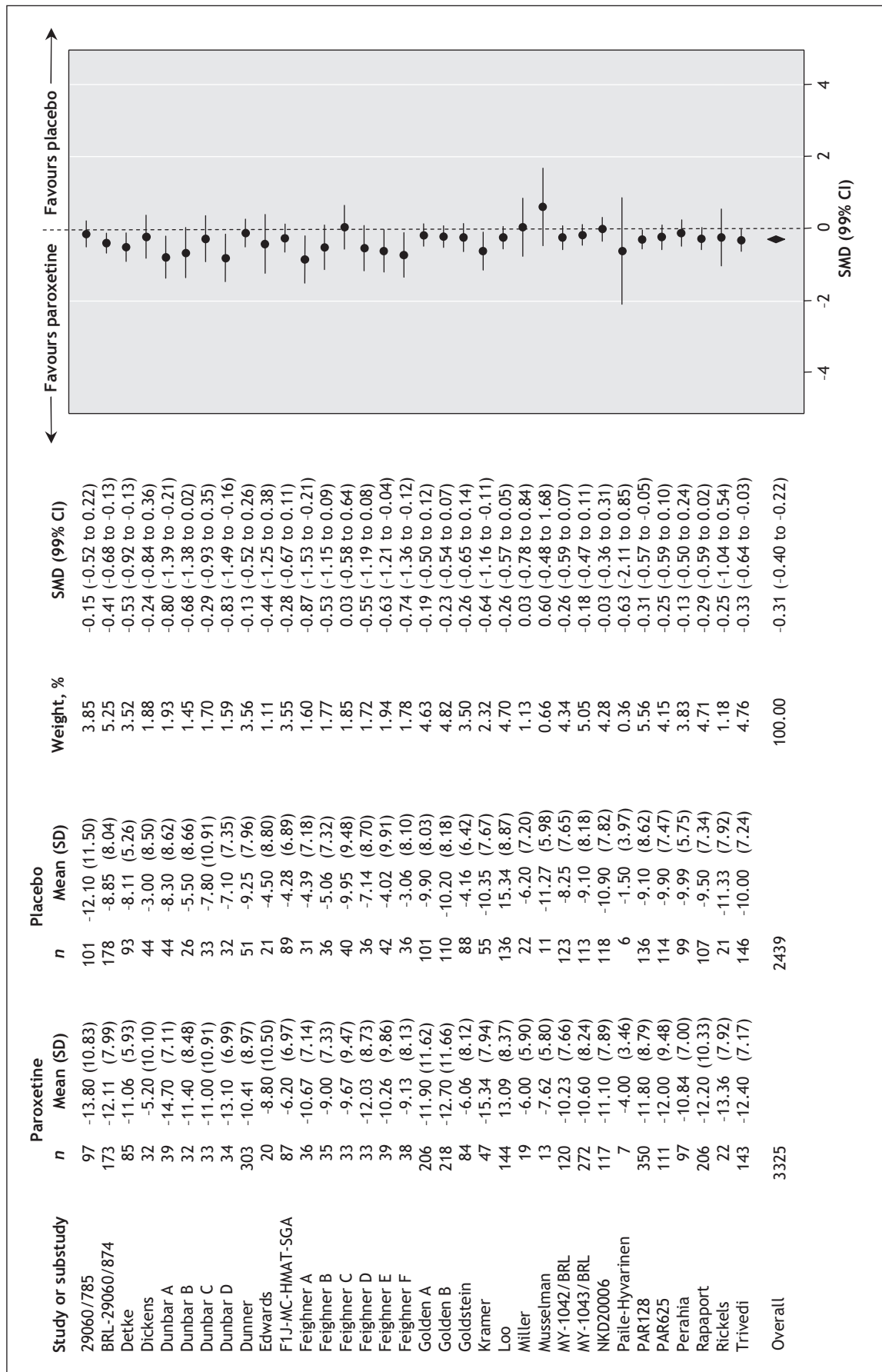


Figure 4: Random effects meta-analysis of the effect of paroxetine versus placebo on standardized depression outcomes. SD = standard deviation, SMD = standardized mean difference, CI = confidence interval. Study designations are defined in online Appendix 2 (www.cmaj.ca/cgi/content/full/178/3/296/DC2).

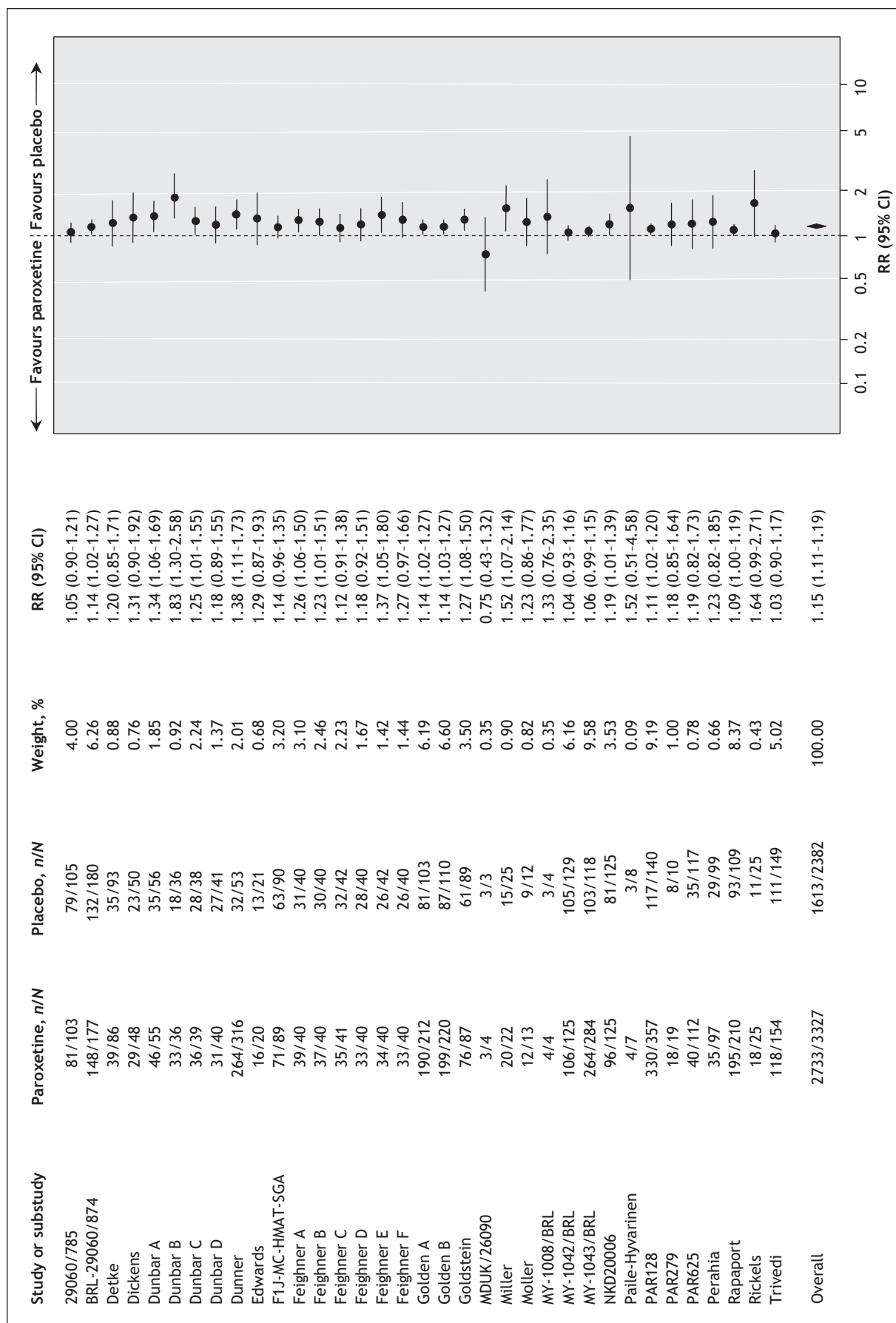


Figure 5: Random effects meta-analysis of the effect of paroxetine versus placebo on the proportion of patients who reported any adverse effects. *n* = number of patients who reported any adverse effects, *N* = total number of patients, RR = relative risk, CI = confidence interval. Study designations are defined in online Appendix 2 (www.cmaj.ca/cgi/content/full/178/3/296/DC2).

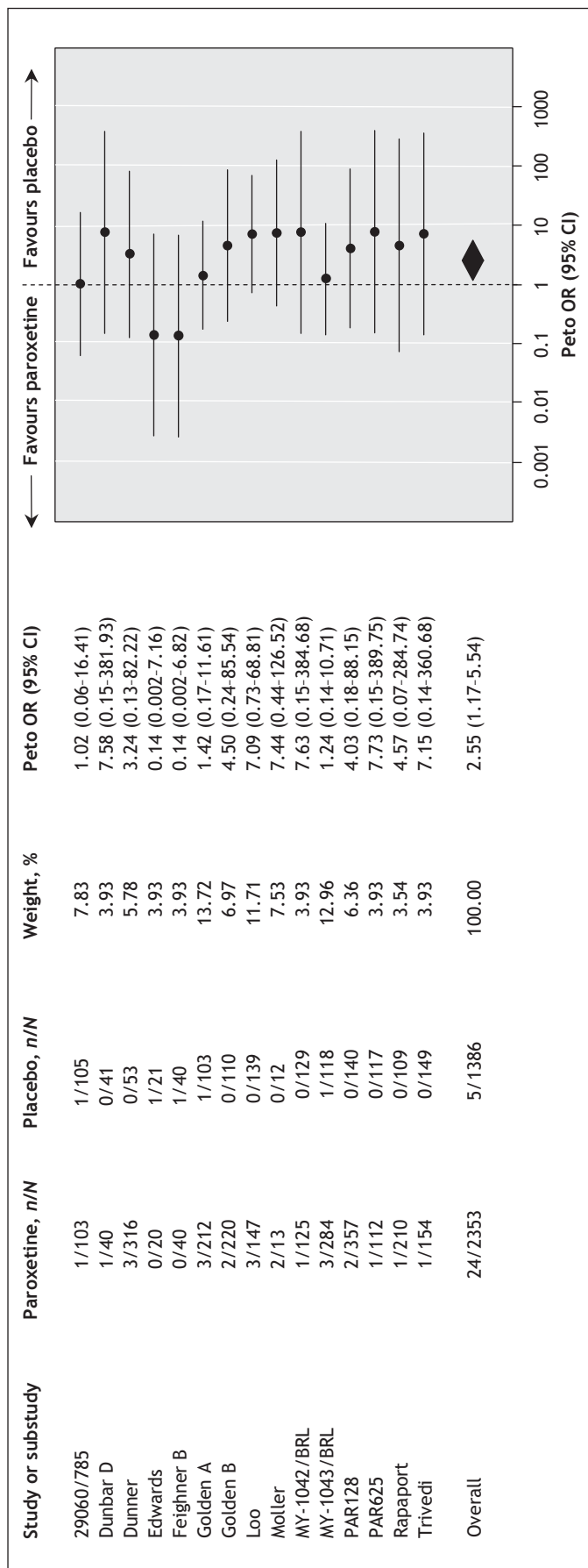


Figure 6: Random effects meta-analysis of the effect of paroxetine versus placebo on the proportion of patients who reported suicidal tendencies (as indicated by the coding system of the US Food and Drug Administration¹⁵). n = number of patients who reported suicidal tendencies, N = total number of patients, RR = relative risk, CI = confidence interval. Study designations are defined in online Appendix 2 (www.cmaj.ca/cgi/content/full/178/3/296/DC2).

into a very large difference in the proportion of patients who “improved.” For a dichotomous outcome we found an absolute risk difference of 11%, and for a continuous outcome we found a standardized mean difference of -0.31 . These results indicate that, at least in this data set, a small mean difference translated into a modest risk difference.

There remains uncertainty about the safety of paroxetine and other selective serotonin reuptake inhibitors, which may cause worsening of suicidal ideas in vulnerable people. The present analysis, which suggests that paroxetine is associated with a statistically significant increase in the risk of suicidal tendencies, expands the results of previous re-analyses of GlaxoSmithKline data.^{16,17} In particular, in the analysis carried out by GlaxoSmithKline of suicide attempts by adults with major depression, the frequency was higher among patients who received paroxetine than among those who received placebo (11/3455 v. 1/1978; OR 6.7, 95% CI 1.1-149.4).²⁷ The recently released re-analysis by the US Food and Drug Administration of 372 placebo-controlled antidepressant trials involving almost 100 000 patients with any psychiatric disorders confirmed these figures by showing that, among the selective serotonin reuptake inhibitors and newer antidepressants, only paroxetine was significantly associated with an excess risk of suicidal behaviour (US Food and Drug Administration codes 1, 2 and 3) (OR 2.76, 95% CI 1.16-6.60).¹⁹

Two considerations should be highlighted. First, the absolute numbers of patients with suicidal tendencies were very low, which leaves the possibility that reporting or not reporting a few cases would completely change the overall outcome.²⁸ In addition, we found disclosure of suicidal tendencies in the GlaxoSmithKline Clinical Trial Registry somewhat problematic. In some cases, for example, patients were reported to have “emotional liability,” with a note explaining that this term was used to indicate “completed or attempted suicide, suicidal tendencies or mood fluctuations.” Clearly, details on each case would have made the analysis more accurate. Second, we still do not know if suicidal tendencies are proxy measures of suicide attempts.

In conclusion, we found that paroxetine was not superior to placebo in terms of the proportion of study participants discontinuing treatment for any reason; in terms of depression measures, paroxetine exerted a modest antidepressant effect relative to placebo. Further research is clearly needed to identify and characterize patients who will experience a response to antidepressant therapy before such treatment is provided. Current research in pharmacogeno-

mics²⁹⁻³¹ may provide useful insights for everyday clinical practice, as may systematic reviews of individual-patient data,³² which should allow study and identification of clinically useful moderators of treatment effect.

This article has been peer reviewed.

Competing interests: None declared for Corrado Barbui and Andrea Cipriani. Toshiaki Furukawa has received research funds and speaker fees from Asahi Kasei, Astellas, Dai-Nippon, Eisai, Eli Lilly, GlaxoSmithKline, Janssen, Kyowa Hakko, Meiji, Organon, Pfizer, Tsumura, Yoshitomi and Zelia. His research has also been funded by the Japanese Ministry of Education, Science and Technology and by the Japanese Ministry of Health Labor and Welfare.

Contributors: All of the authors contributed to the design of the systematic review, coordination of the study, collection and management of the data, the statistical analysis, the interpretation of the results and the writing of the manuscript. All of the authors saw and approved the final version of the manuscript.

Acknowledgement: This study was funded in part by the Japanese Ministry of Education, Science and Technology and by the Japanese Ministry of Health Labor and Welfare.

REFERENCES

- National Collaborating Centre for Mental Health. *Depression (amended). Management of depression in primary and secondary care*. London (UK): National Institute for Health and Clinical Excellence; 2007.
- Bauer M, Whybrow PC, Angst J, et al. World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of unipolar depressive disorders. Part 2: Maintenance treatment of major depressive disorder and treatment of chronic depressive disorders and subthreshold depressions. *World J Biol Psychiatry* 2002;3:69-86.
- Moncrieff J, Kirsch I. Efficacy of antidepressants in adults. *BMJ* 2005;331:155-7.
- Moncrieff J. The antidepressant debate. *Br J Psychiatry* 2002;180:193-4.
- Keller MB, Lavori PW. The adequacy of treating depression. *J Nerv Ment Dis* 1988;176:471-4.
- Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967;20:637-48.
- March JS, Silva SG, Compton S, et al. The case for practical clinical trials in psychiatry. *Am J Psychiatry* 2005;162:836-46.
- Barbui C, Cipriani A. Evidence-based psychopharmacology: an agenda for the future. *Evid Based Ment Health* 2007;10:4-6.
- Joffe R, Sokolov S, Streiner D. Antidepressant treatment of depression: a meta-analysis. *Can J Psychiatry* 1996;41:613-6.
- Williams JW Jr, Mulrow CD, Chiquette E, et al. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. *Ann Intern Med* 2000;132:743-56.
- Khan A, Warner HA, Brown WA. Symptom reduction and suicide risk in patients treated with placebo in antidepressant clinical trials: an analysis of the Food and Drug Administration database. *Arch Gen Psychiatry* 2000;57:311-7.
- Taylor MJ, Freemantle N, Geddes JR, et al. Early onset of selective serotonin reuptake inhibitor antidepressant action: systematic review and meta-analysis. *Arch Gen Psychiatry* 2006;63:1217-23.
- Barbui C, Cipriani A. Publication bias in meta-analyses [letter]. *Arch Gen Psychiatry* 2007;64:868.
- Whittington CJ, Kendall T, Fonagy P, et al. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 2004;363:1341-5.
- Hammad TA, Laughren T, Racoosin J. Suicidality in pediatric patients treated with antidepressant drugs. *Arch Gen Psychiatry* 2006;63:332-9.
- Aursnes I, Tveit IF, Gaasemyr J, et al. Suicide attempts in clinical trials with paroxetine randomised against placebo. *BMC Med* 2005;3:14.
- Aursnes I, Tveit IF, Gaasemyr J, et al. Even more suicide attempts in clinical trials with paroxetine randomised against placebo. *BMC Psychiatry* 2006;6:55.
- Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;354:1896-900.
- FDA briefing document for Dec. 13, 2006, meeting of Psychopharmacologic Drugs Advisory Committee. Rockville (MD): US Food and Drug Administration, Center for Drug Evaluation and Research; 2006, p. 1-65. Available: www.fda.gov/ohrms/dockets/ac/06/briefing/2006-4272b1-index.htm (accessed 2007 Aug 29).
- Norman GR. Issues in the use of change scores in randomized trials. *J Clin Epidemiol* 1989;42:1097-105.
- Altman DG, Bland JM. Detecting skewness from summary information. *BMJ* 1996;313:1200.
- Furukawa TA, Barbui C, Cipriani A, et al. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 2006;59:7-10.
- Higgins J, Green S, editors. *Cochrane handbook for systematic reviews of interventions* 4.2.6. Chichester (UK): John Wiley & Sons; 2006 [updated 2006 Sept].
- Cipriani A, Brambilla P, Furukawa T, et al. Fluoxetine versus other types of pharmacotherapy for depression. *Cochrane Database Syst Rev* 2005;(4):CD004185.
- Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557-60.
- Guyatt GH, Juniper EF, Walter SD, et al. Interpreting treatment effects in randomised trials. *BMJ* 1998;316:690-3.
- Update April 5, 2006. Briefing document: Paroxetine adult suicidality analysis: major depressive disorder and non-major depressive disorder. GlaxoSmithKline; 2006. Available: http://us.gsk.com/docs-pdf/media-news/briefing_doc.pdf (accessed 2007 Dec 14).
- Cipriani A, Barbui C, Geddes JR. Suicide, depression, and antidepressants. *BMJ* 2005;330:373-4.
- Murphy GM Jr, Hollander SB, Rodrigues HE, et al. Effects of the serotonin transporter gene promoter polymorphism on mirtazapine and paroxetine efficacy and adverse events in geriatric major depression. *Arch Gen Psychiatry* 2004;61:1163-9.
- Malhotra AK, Murphy GM Jr, Kennedy JL. Pharmacogenetics of psychotropic drug response. *Am J Psychiatry* 2004;161:780-96.
- Murphy GM Jr, Kremer C, Rodrigues HE, et al. Pharmacogenetics of antidepressant medication intolerance. *Am J Psychiatry* 2003;160:1830-5.
- Clarke MJ, Stewart LA. Meta-analyses using individual patient data. *J Eval Clin Pract* 1997;3:207-12.

Correspondence to: Dr. Corrado Barbui, Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, University of Verona, Piazzale Scuro 10, 37134 Verona, Italy; fax +39-045-585871 (for the World Health Organization Collaborating Centre for Research and Training in Mental Health); corrado.barbui@univr.it