

## Evidence b(i)ased medicine—selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications

Hans Melander, Jane Ahlqvist-Rastad, Gertie Meijer, Björn Beermann

### Abstract

**Objectives** To investigate the relative impact on publication bias caused by multiple publication, selective publication, and selective reporting in studies sponsored by pharmaceutical companies.

**Design** 42 placebo controlled studies of five selective serotonin reuptake inhibitors submitted to the Swedish drug regulatory authority as a basis for marketing approval for treating major depression were compared with the studies actually published (between 1983 and 1999).

**Results** Multiple publication: 21 studies contributed to at least two publications each, and three studies contributed to five publications. Selective publication: studies showing significant effects of drug were published as stand alone publications more often than studies with non-significant results. Selective reporting: many publications ignored the results of intention to treat analyses and reported the more favourable per protocol analyses only.

**Conclusions** The degree of multiple publication, selective publication, and selective reporting differed between products. Thus, any attempt to recommend a specific selective serotonin reuptake inhibitor from the publicly available data only is likely to be based on biased evidence.

### Introduction

Drug treatment should rely on solid evidence, and it is now generally recognised that the standard basis for treatment guidelines is systematic literature reviews or meta-analyses of all randomised controlled trials. However, as meta-analyses are usually limited to publicly available data, several factors can give rise to biased conclusions. These include selection of studies submitted or accepted for publication,<sup>1 2</sup> inclusion of undetected duplicate publications,<sup>3 4</sup> and selective reporting (such as failure to report intention to treat results). Several actors (editors, investigators, and sponsors) affect whether and how scientific results reach the public domain. In clinical trials of drugs the role of the sponsor is especially important. The sponsor usually has access to all data on a specific product and has an obvious conflict of interest.<sup>5</sup>

Several authors have provided direct evidence of publication bias by investigating the publication status of protocols submitted to ethics committees or research organisations.<sup>6–11</sup> These investigators did not, however, examine whether there was a biased selection of results reported in the studies that were eventually published. The objective of our study was to investigate the relative impact on bias caused by multiple publication, selective publication, and selective reporting in studies sponsored by the pharmaceutical industry.

### Material and methods

#### Studies submitted to drug regulatory authority

Five selective serotonin reuptake inhibitors were approved in Sweden between 1989 and 1994 for treating major depression. Forty two short term (4–8 weeks) placebo controlled clinical trials with the approved doses were submitted to the Swedish drug regulatory authority and formed the basis for the approvals. When applying for marketing authorisation, the applicants are obliged to submit full reports of all studies performed by the applicants as well as all available information on any study not performed by the applicants. Thus, it is reasonable to assume that the submitted studies have not been subject to selection bias.

#### Studies published

We identified published versions of the submitted studies through a computer aided search in Medline (PubMed), Embase, and PsycINFO (Psychological Abstracts); scrutiny of reference lists with special focus on review articles and meta-analyses; and inquiries to the sponsoring companies. For each submitted study, we investigated the publication status and the degree of multiple publication. We classified a published article reporting results from a single submitted study as a stand alone publication, whereas we classified articles based on data from two or more submitted studies as pooled publications.

#### Comparison of studies

We chose the percentage of patients responding to treatment as the criterion for comparing results from submitted reports with those from published articles. Response was defined as at least a 50% reduction of the initial score on the Hamilton depression rating

Medical Products Agency, Box 23, S-751 03 Uppsala, Sweden

Hans Melander  
*senior biostatistician*

Jane Ahlqvist-Rastad  
*senior medical officer*

Gertie Meijer  
*documentalist*

Björn Beermann  
*professor*

Correspondence to:  
H Melander  
hans.melander@mpa.se

bmj.com 2003;326:1171

scale (HDRS) in most studies. In four studies response rates were based on the Montgomery Åsberg depression rating scale or the clinical global impression of change. In the pooled analyses of response rates we combined the estimate from the individual studies, with the inverse of the variance of the estimates as weights.<sup>12</sup>

## Results

We identified 38 publications presenting data from 38 of the 42 studies submitted to the drug regulatory authority.<sup>13–50</sup> They were published between 1983 and 1999 and included duplicate publications and pooled analyses. The sponsoring companies confirmed the completeness of our search.

### Multiple publication

Figure 1 shows the degree of multiple publication: this varied from no duplicate publication (drug 3) to extensive multiple publication (drug 1) with three stand alone publications appearing twice and two subsets of studies published as pooled publications three times each.

For drug 1, there were no cross references between the pooled analyses of the same subsets of studies. For each of the subsets, the first author was different in two of the pooled analyses, and the third publication had a single author. Many of the studies had appeared previously as stand alone publications, but reference to these in the pooled publications was given in two cases only, once for each subset. Some of the analyses were presented as a pooled analysis of stand alone centres and some as a multicentre study. For both subsets of

studies, the pooled results differed slightly between the publications.

For drug 2, eight studies resulted in three pooled publications based on different combinations of studies. The pooled analyses based on two and eight studies appeared simultaneously (as “a double blind comparison” and “a large multicentre study” respectively) with one author in common but without cross reference. Later, the five study analysis was presented as an intention to treat reanalysis of the per protocol analysis in the eight study publication without revealing that three studies were omitted. Nor was it said that two of the included studies had been published earlier as stand alone publications.

The pooled publication of studies of drug 4 was denoted as a review of multicentre controlled studies without identification of the included studies. Two of the studies later appeared as stand alone publications without acknowledgement of their earlier inclusion in a pooled publication. There was no author name in common in the pooled and stand alone publications.

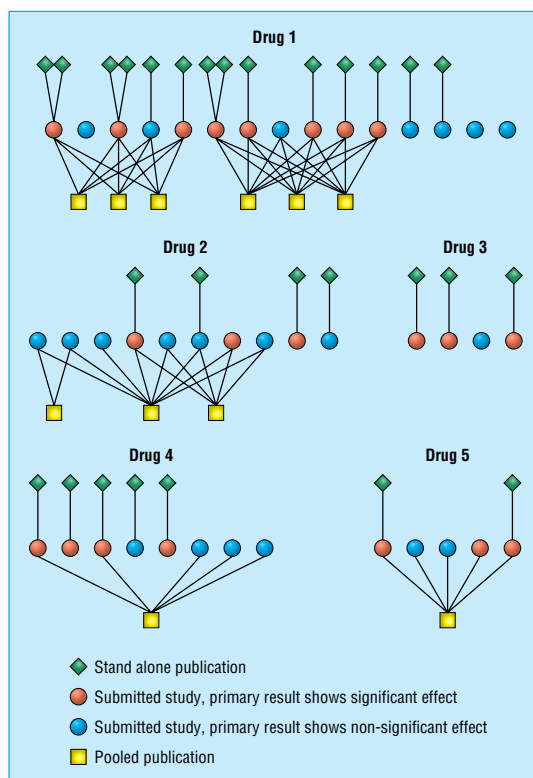
For drug 5, the pooled analysis was presented as a meta-analysis of the five available placebo controlled studies, clearly identified by the name of the principal investigator. Reference was given to one previous stand alone publication. The other stand alone publication appeared seven years later without reference to the pooled publication.

### Selective publication

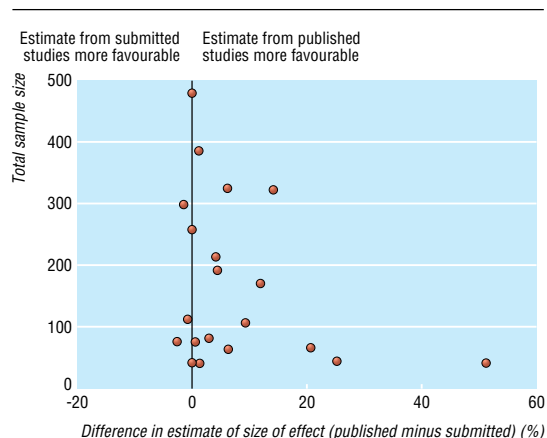
Of the 42 submitted studies, 21 found the test drug to be significantly more effective than placebo in the primary variable (fig 1). Nineteen of these studies appeared as stand alone publications. Only six of the 21 studies not showing significant results were published as stand alone publications. Of the four studies that never reached the public domain, all showed non-significant results with respect to the primary variable.

### Selective reporting

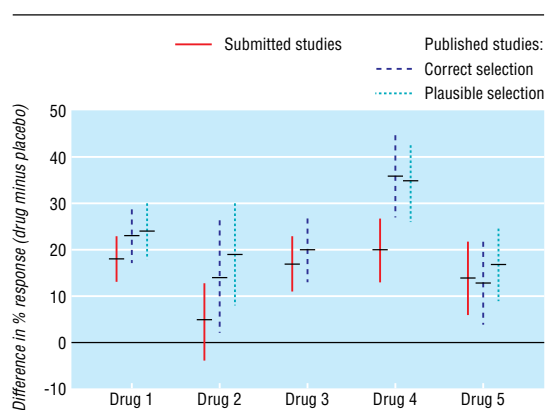
All but one of the study reports submitted to the regulatory agency presented results from two or more alternative analyses (intention to treat and per protocol). Only two of the stand alone publications presented an intention to treat analysis as well as a per protocol analysis. The remaining stand alone publications presented only one analysis, which tended to be the more favourable per protocol analysis. In the 15 stand alone and five pooled publications reporting differences in percentage response, patients who withdrew or who could not be evaluated were usually ignored in the calculations of the response rates. As figure 2 shows, this could result in large overestimates compared with the intention to treat analysis based on the submitted reports, where patients who withdrew or could not be evaluated were considered to be non-responders. In one extreme case the published difference in the percentage of patients responding to treatment was 51%, whereas no difference was seen in the intention to treat analysis. In five other cases the size of the overestimation was 10–25%. The degree of overestimation tended to be higher in smaller studies.



**Fig 1** Publication pattern for studies of the five selective serotonin reuptake inhibitors approved in Sweden between 1989 and 1994 for treating major depression



**Fig 2** Difference in estimated size of treatment effect (% response to drug minus response to placebo) from published studies and estimate from intention to treat analysis of submitted studies plotted against total sample size.



**Fig 3** Differences (95% confidence intervals) in response rate (% response to drug minus response to placebo). Results from pooled analyses of all submitted studies, correct selection of published studies (duplicates excluded), and plausible selection of published studies (including probably undetectable duplicates)

### Comparison of pooled results from submitted and published studies

In 41 of the 42 submitted studies data on response rate were provided or could easily be calculated on an intention to treat basis. In total, 15 stand alone publications and five pooled publications reported response rates based on data from 32 studies. For each drug, we compared a pooled analysis of all studies submitted to the regulatory agency with a pooled analysis of a correct selection of published studies in which all duplicates were excluded. We also made a pooled analysis of published studies including those duplicates that probably could not be identified as such without access to information about all studies. In this second selection we excluded duplicates with at least one author in common and only minor differences with respect to patient numbers and efficacy results but included any duplicates unidentifiable because of lack of cross reference between pooled publications and stand alone publications.

The pooled analyses of published data generally gave larger differences in response rate (drug minus placebo) than did the estimates from all submitted studies (fig 3). The result of the comparison was

conspicuous for two products. The estimate based on evidence from all available studies on drug 2 indicated a marginal effect, whereas the pooled analysis of published data gave an estimated effect size similar to that for most of the other drugs. Similarly, the analyses based on published data gave the impression that drug 4 was substantially more effective than the other drugs, whereas the analysis based on submitted studies did not. Since the estimates based on the published studies for these two drugs included data from all the submitted studies, the overestimations are due to selective reporting rather than selective publication. Overall, there were only minor differences in response rates between the correct selections of published studies and the plausible selections. Thus, in this material there is no indication of any major bias due to multiple publication.

### Discussion

In a cohort of studies submitted to the Swedish regulatory agency to secure marketing approval for five selective serotonin reuptake inhibitors for the treatment of major depression we have found evidence of duplicate publication, selective publication, and selective reporting. There was a high frequency of duplication due to the inclusion of different subsets of studies in several pooled publications. Studies showing significant differences between efficacy of drug and placebo were three times more likely to appear as stand alone publications than were studies with non-significant results. Although both intention to treat analyses and per protocol analyses were available in the submissions to the regulatory agency, only 24% of the stand alone publications reported the usually less favourable intention to treat results. In our material this selective reporting was the major cause for bias in overall estimates based on published data.

### Strengths and limitations of study

To our knowledge, access to full reports and study protocols for all studies, published as well as unpublished, is unique to our investigation. This has enabled us to study the impact of different sources of publication bias. It also allowed us to elucidate the sometimes complex pattern of publications. Our investigation is restricted to one class of antidepressant drugs, but there is no reason to believe that drug manufacturers have different policies for reporting and publishing studies of different drugs. Indeed, in a review of an antiemetic drug a similar pattern of duplicate publication has been reported.<sup>4</sup> Thus, our results are likely to be valid for other classes of drugs with a similar structure of the efficacy documentation—that is, several studies with small to medium sample size.

The percentage of submitted studies with full stand alone publications in our investigation (60%) is similar to what has been reported by others. In a review of five similar investigations the percentage of full publications ranged from 48% to 80% (median 62%).<sup>51</sup> The ratio of stand alone publications with significant results to those with non-significant studies was 3.2 in our investigation, which is somewhat higher than the overall corresponding ratio of 2.3 reported for the above investigations.<sup>51</sup> This difference might be explained by the difference in study materials. All the studies in our

### What is already known on this topic

Duplicate publication, selective publication, and selective reporting are likely to introduce bias in systematic literature reviews and meta-analyses

Several reports have provided evidence of duplicate publication and selective publication as well as the tendency to publish only studies with significant findings

### What this study adds

Access to full documentation of all studies (published and unpublished) made it possible to investigate the relative impact of the different sources of bias

Selective reporting (tendency to publish the more favourable per protocol results only) was a major cause for bias

A sponsor in control of all studies does not seem to improve the situation with respect to duplicate publication, selective publication, and selective reporting

investigation were initiated by the sponsor, and the investigators were usually clinical practitioners for whom academic research was not the primary interest. Hence, the decision on how and whether a study should be published was probably left entirely to the sponsor. The studies in the other investigations were more heterogeneous with respect to funding (public funding, no external funding, etc) and the study sponsors probably took a less active part in the reporting of the studies.

### Conclusions

The outcome of our investigation should not be used to dispute the value of systematic literature reviews and meta-analyses in general. However, for anyone who relies on published data alone to choose a specific drug, our results should be a cause for concern. Without access to all studies (positive as well as negative, published as well as unpublished) and without access to alternative analyses (intention to treat as well as per protocol), any attempt to recommend a specific drug is likely to be based on biased evidence. The probable choice of a specific selective serotonin reuptake inhibitor based on a pooled analysis of publicly available data is not likely to be supported by an analysis considering the total body of evidence.

Contributors: HM, JA-R, and BB designed the study. GM performed the search for publications based on the data submitted to the Swedish drug regulatory authority. HM and JAR reviewed the submitted reports and the publications, and extracted the data. HM performed the statistical analysis. HM, JAR, and BB interpreted the results. All authors contributed to writing the paper. HM is guarantor for the study.

Funding: None.

Competing interests: None declared.

- Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;263:1385-9.
- Dickersin K, Chalmers I. Under-reporting research is scientific misconduct. *JAMA* 1990;263:1405-8.
- Göttsche PC. Multiple publication of reports of drug trials. *Eur J Clin Pharmacol* 1989;36:429-32.

- Tramèr MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ* 1997;315:635-40.
- Davidoff F, DeAngelis C, Drazen J, Hoet J, Højgaard L, Horton R, et al. Sponsorship, authorship, and accountability [commentary]. *Lancet* 2001;358:854-6.
- Simes RJ. Publication bias: the case for an international registry for clinical trials. *J Clin Oncol* 1986;4:1529-41.
- Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H. Publication bias and clinical trials. *Control Clin Trials* 1987;8:343-53.
- Dickersin K, Min YI. NIH clinical trials and publication bias. *Online J Curr Clin Trials* 1993; Apr 28; Doc No 50.
- Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results: follow-up of applications submitted to two institutional review boards. *JAMA* 1992;263:374-8.
- Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;337:867-72.
- Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ* 1997;315:640-5.
- DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88.
- Amin MM, Ananth JV, Coleman BS, Darcourt G, Farkas T, Goldstein B, et al. Fluvoxamine: antidepressant effects confirmed in a placebo-controlled international study. *Clin Neuropharmacol* 1984;7(suppl 1):580-1.
- Bech P, Cialdella P. Citalopram in depression—meta-analysis of intended and unintended effects. *Int Clin Psychopharmacol* 1992;6(suppl 5):45-54.
- Byerley WF, Reimherr FW, Wood DR, Grosser BI. Fluoxetine, a selective serotonin uptake inhibitor, for the treatment of outpatients with major depression. *J Clin Psychopharmacol* 1988;8:112-5.
- Claghorn JL, Kiev A, Rickels K, Smith WT, Dunbar GC. Paroxetine versus placebo: a double-blind comparison in depressed patients. *J Clin Psychiatry* 1992;53:434-8.
- Claghorn JL. A double-blind comparison of paroxetine and placebo in the treatment of depressed outpatients. *Int Clin Psychopharmacol* 1992;6(suppl 4):25-30.
- Claghorn JL. The safety and efficacy of paroxetine compared with placebo in a double-blind trial of depressed outpatients. *J Clin Psychiatry* 1992;53(suppl):33-5.
- Cohn JB, Wilcox C. A comparison of fluoxetine, imipramine, and placebo inpatients with major depressive disorder. *J Clin Psychiatry* 1985;46:26-31.
- Cohn JB, Wilcox CS. Paroxetine in major depression: a double-blind trial with imipramine and placebo. *J Clin Psychiatry* 1992;53(suppl):52-6.
- Dominguez RA, Goldstein BJ, Jacobson AF, Steinbook RM. A double-blind placebo-controlled study of fluvoxamine and imipramine in depression. *J Clin Psychiatry* 1985;46:84-7.
- 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.
- Dunbar GC, Claghorn JL, Kiev A, Rickels K, Smith WT. A comparison of paroxetine and placebo in depressed outpatients. *Acta Psychiatr Scand* 1993;87:302-5.
- Dunbar GC, Cohn JB, Fabre LF, Feigner JP, Fieve RR, Mendels J, et al. A comparison of paroxetine, imipramine and placebo in depressed outpatients. *Br J Psychiatry* 1991;159:394-8.
- Edwards JG, Goldie A. Placebo-controlled trial of paroxetine in depressive illness. *Hum Psychopharmacol* 1993;8:203-9.
- Fabre LF, Crismon L. Efficacy of fluoxetine in outpatients with major depression. *Curr Ther Res* 1985;37:115-23.
- Fabre LF. A 6-week, double-blind trial of paroxetine, imipramine, and placebo in depressed outpatients. *J Clin Psychiatry* 1992;53(suppl):40-3.
- Feighner JP, Boyer WF, Merideth CH, Hendrickson GG. A double-blind comparison of fluoxetine, imipramine and placebo in outpatients with major depression. *Int Clin Psychopharmacol* 1989;4:127-34.
- Feighner JP, Boyer WF, Merideth CH, Hendrickson GG. A placebo-controlled inpatient comparison of fluvoxamine maleate and imipramine in major depression. *Int Clin Psychopharmacol* 1989;4:239-44.
- Feighner JP, Boyer WF. Paroxetine in the treatment of depression: a comparison with imipramine and placebo. *Acta Psychiatr Scand* 1993;80(suppl 350):125-9.
- Feighner JP, Boyer WF. Paroxetine in the treatment of depression: a comparison with imipramine and placebo. *J Clin Psychiatry* 1992;53(suppl):44-7.
- Feighner JP, Cohn JB, Fabre LF, Fieve RR, Mendels J, Shrivastava RK, et al. A study comparing paroxetine, placebo and imipramine in depressed patients. *J Affect Dis* 1993;28:71-9.
- Feighner JP. A double-blind comparison of paroxetine, imipramine and placebo in depressed outpatients. *Int Clin Psychopharmacol* 1992;6(suppl 4):31-5.
- Itil TM, Shrivastava RK, Mukherjee S, Coleman BS, Michael ST. A double-blind placebo-controlled study of fluvoxamine and imipramine in outpatients with primary depression. *Br J Clin Pharmacol* 1983;15:433-8S.
- Kiev A. A double-blind, placebo-controlled study of paroxetine in depressed outpatients. *J Clin Psychiatry* 1992;53(suppl):27-9.
- LaPierre YD, Browne M, Horn E, Oyewumi LK, Sarantidis D, Roberts N, et al. Treatment of major affective disorder with fluvoxamine. *J Clin Psychiatry* 1987;48:65-8.
- Mendels J, Kiev A, Fabre LF. Double-blind comparison of citalopram and placebo in depressed outpatients with melancholia. *Depress Anxiety* 1999;9:54-60.
- Miller SM, Naylor GJ, Murtahg M, Winslow G. A double-blind comparison of paroxetine and placebo in the treatment of depressed patients in a psychiatric outpatient clinic. *Acta Psychiatr Scand* 1993;80(suppl 350):143-4.
- Montgomery SA, Rasmussen JGC, Lyby K, Connor P, Tanghøj P. Dose response relationship of citalopram 20 mg, citalopram 40 mg and

- placebo in the treatment of moderate and severe depression. *Int Clin Psychopharmacol* 1992;6(suppl 5):65-70.
- 40 Norton KRW, Sireling LI, Bhat AV, Rao B, Paykel ES. A double-blind comparison of fluvoxamine, imipramine and placebo in depressed patients. *J Affect Dis* 1984;7:297-308.
- 41 Olie JP, Gunn KO, Katz E. A double-blind placebo-controlled multicentre study of sertraline in the acute and continuation treatment of major depression. *Eur Psychiatry* 1997;12:34-41.
- 42 Ottevanger EA. The efficacy of fluvoxamine in patients with severe depression. *Prog Neuropsychopharmacol Biol Psychiatry* 1994;18:731-40.
- 43 Peselow ED, Filippi AM, Goodnick P, Barouche F, Fieve RR. The short- and long-term efficacy of paroxetine HCl: A Data from a 6-week double-blind parallel design trial vs. imipramine and placebo. *Psychopharmacol Bull* 1989;25:267-71.
- 44 Reimherr FW, Chouinard G, Cohn CK, Cole JO, Itil TM, LaPierre YD, et al. Antidepressant efficacy of sertraline: a double-blind, placebo- and amitriptyline controlled, multicenter comparison study in outpatients with major depression. *J Clin Psychiatry* 1990;51(suppl B):18-27.
- 45 Rickels K, Amsterdam J, Clary C, Fox I, Schweizer E, Weise C. A placebo-controlled, double-blind, clinical trial of paroxetine in depressed outpatients. *Acta Psychiatr Scand* 1993;80(suppl 350):117-23.
- 46 Rickels K, Amsterdam J, Clary C, Fox I, Schweizer E, Weise C. The efficacy and safety of paroxetine compared with placebo in outpatients with major depression. *J Clin Psychiatry* 1992;53(suppl):30-2.
- 47 Rickels K, Amsterdam JD, Avallone ME. Fluoxetine in major depression: a controlled study. *Curr Ther Res* 1986;39:559-63.
- 48 Shrivastava RK, Shrivastava SH, Overweg N, Blumhardt CL. A double-blind comparison of paroxetine, imipramine, and placebo in major depression. *J Clin Psychiatry* 1992;53(suppl):48-51.
- 49 Smith WT, Glaudin V. A placebo-controlled trial of paroxetine in the treatment of major depression. *J Clin Psychiatry* 1992;53(suppl):36-9.
- 50 Stark P, Hardison CD. A review of multicenter controlled studies of fluoxetine vs. imipramine and placebo in outpatients with major depressive disorder. *J Clin Psychiatry* 1985;46:53-8.
- 51 Dickersin K. How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997;9(suppl A):15-21.

(Accepted 6 March 2003)