ABSTRACT

Objectives To assess the benefits and harms of reboxetine versus placebo or selective serotonin reuptake inhibitors (SSRIs) in the acute treatment of depression, and to measure the impact of potential publication bias in trials of reboxetine.

Design Systematic review and meta-analysis including unpublished data.

Data sources Bibliographic databases (Medline, Embase, PsycINFO, BIOSIS, and Cochrane Library), clinical trial registries, trial results databases, and regulatory authority websites up until February 2009, as well as unpublished data from the manufacturer of reboxetine (Pfizer, Berlin).

Eligibility criteria Double blind, randomised, controlled trials of acute treatment (six weeks or more) with reboxetine versus placebo or SSRIs in adults with major depression.

Outcome measures Remission and response rates (benefit outcomes), as well as rates of patients with at least one adverse event and withdrawals owing to adverse events (harm outcomes).

Data extraction and data synthesis The procedures for data extraction and assessment of risk of bias were always conducted by one person and checked by another. If feasible, data were pooled by meta-analyses (random effects model). Publication bias was measured by comparing results of published and unpublished trials.

Results We analysed 13 acute treatment trials that were placebo controlled, SSRI controlled, or both, which included 4098 patients. Data on 74% (3033/4098) of these patients were unpublished. In the reboxetine versus placebo comparison, no significant differences in remission rates were shown (odds ratio 1.17, 95% confidence interval 0.91 to 1.51; P=0.216). Substantial heterogeneity (I²=67.3%) was shown in the meta-analysis of the eight trials that investigated response rates for reboxetine versus placebo. A sensitivity analysis that excluded a small inpatient trial showed no significant difference in response rates between patients receiving reboxetine and those receiving placebo (OR 1.24, 95% CI 0.98 to 1.56; P=0.071; I²=42.1%). Reboxetine was inferior to SSRIs (fluoxetine, paroxetine, and citalopram) for remission rates (OR 0.80, 95% CI 0.67 to 0.96; P=0.015) and response rates (OR 0.80, 95% CI 0.67 to 0.95; P=0.01). Reboxetine was inferior to placebo for both harm outcomes (P<0.001 for both), and to fluoxetine for withdrawals owing to adverse events (OR 1.79, 95% CI 1.06 to 3.05; P=0.031). Published data overestimated the benefit of reboxetine versus placebo by up to 11.5% and reboxetine versus SSRIs by up to 23%, and also underestimated harm.

Conclusions Reboxetine is, overall, an ineffective and potentially harmful antidepressant. Published evidence is affected by publication bias, underlining the urgent need for mandatory publication of trial data.

INTRODUCTION

Reboxetine, the first selective norepinephrine (noradrenaline) reuptake inhibitor used in the treatment of depression,1 mainly acts by binding to the norepinephrine transporter and blocking reuptake of extracellular norepinephrine.2 The drug is “indicated for the acute treatment of depressive illness or major depression and for maintaining the clinical improvement in patients initially responding to treatment.”24 Reboxetine has been approved for marketing in many European countries (for example, the United Kingdom and Germany) since 1997. In the United States, however, the application for approval was ultimately rejected after preliminary acceptance.24

Compared with the overall amount of antidepressants prescribed, reboxetine’s share is relatively small. For example, of 974 million defined daily doses of antidepressants prescribed in Germany in 2008, reboxetine accounted for 6.7 million defined daily doses.3 The average cost of reboxetine per defined daily dose was €1.87 (£1.54; $2.39) for Edronax (Pfizer,
Berlin) to €2.09 for Solvex (Merz, Frankfurt), compared with €0.52 for selective serotonin reuptake inhibitors (SSRIs), the most commonly prescribed antidepressants.5

Although reboxetine has been claimed to show superior efficacy to placebo and similar efficacy to other antidepressants,6–10 the clinical relevance of the drug has been queried. A recent systematic review by Cipriani et al11 included a network meta-analysis of active controlled trials and found that reboxetine was not only significantly less effective than the other newer antidepressants investigated, but was also the drug with the highest dropout rates.

The German Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen [IQWiG]) conducted a health technology assessment of the short term and long term benefits and harms of reboxetine, bupropion, and mirtazapine in placebo controlled and active controlled trials of adult patients with major depressive disorder. Both published and previously unpublished data were considered. The full German language report and an English summary are available on the institute’s website.12,13 The responsibilities and methodological approach of IQWiG are described in its methods paper online.14

This publication presents the main findings of the reboxetine trials with the aim of assessing the benefits (remission and response rates) and harms (rates of patients with at least one adverse event and rates of withdrawals owing to adverse events) of reboxetine versus placebo or SSRIs in the acute treatment of major depressive disorder. In addition, for the present paper we assessed potential publication bias by comparing results from published and unpublished trials of reboxetine.

METHODS

We developed and followed a standardised protocol for all steps of the review.15

Eligibility criteria

The health technology assessment report that formed the basis of this publication included both published and unpublished trials of reboxetine that had the following characteristics:

- Double blind, randomised controlled design
- Investigation of adult patients with major depressive disorder as their primary diagnosis according to the International Classification of Diseases, the Diagnostic and Statistical Manual of Mental Disorders, or the Research Diagnostic Criteria
- Acute treatment (at least six weeks duration) or long term treatment (at least six months [relapse] or 12 months [recurrence]) for prevention of relapse or recurrence
- Comparison of reboxetine with placebo or any antidepressant (including St John’s wort); treatment according to approval status in Germany
- Evaluation of at least one prespecified patient relevant outcome (in this context, the term “patient relevant” refers to “how a patient feels, functions, or survives”16)
- Publication in English, German, or French (or any other language if the trial was classified as potentially relevant according to the English title or abstract)
- Availability of a full text document (for example, journal article or clinical study report).

This publication is limited to acute treatment trials of reboxetine versus placebo or SSRIs. The outcomes presented are restricted to the most commonly reported outcomes in depression trials. Benefit outcomes were remission and response rates. Harm outcomes were rates of patients with at least one adverse event (any adverse event according to the definitions used in the primary trials) and rates of withdrawals owing to adverse events (any adverse event according to the definitions used in the primary trials). Harms were further described by the overall rates of patients with serious adverse events (any serious adverse event according to the definitions used in the primary trials).

According to the review protocol, response and remission data were analysed on the basis of the definitions and instruments used in the primary trials. All trials applied the Hamilton depression rating scale and 10 trials additionally applied the Montgomery-Åsberg depression rating scale. We primarily considered the response and remission outcomes on the Hamilton depression rating scale. In all trials, response was defined as a reduction in the score on the Hamilton depression rating scale of 50% or more from baseline to end of study, and remission was defined as a reduction in the score on the Hamilton depression rating scale to below an absolute threshold at end of study (score ≤10 in all trials except in one trial where the score threshold was ≤8).

Search strategy and study selection

We searched for relevant primary and secondary publications (systematic reviews and health technology assessment reports) in Medline, Embase, PsycINFO, BIOSIS, and the Cochrane Library published up until February 2009. The full search strategy, including the search terms used for the various databases, has been described elsewhere.12

We scrutinised the reference lists of the primary and secondary publications retrieved to identify further trials. In addition, clinical trial registries and trial results databases available on the internet were screened, as were the websites of the European Medicines Agency and the US Food and Drug Administration.

In order to obtain the most complete dataset possible, we asked the manufacturer of reboxetine (Pfizer) to supply unpublished trials and additional unpublished data from published trials. As a prerequisite for the use of unpublished data, IQWiG asked the manufacturer to sign an agreement requiring: (1) submission of a list...
of all sponsored published and unpublished trials investigating reboxetine; (2) submission of documents (generally the clinical study reports) compliant with the CONSORT criteria for all relevant trials selected by IQWiG; and (3) permission for publication of all previously unpublished relevant data. This procedure was required to avoid bias through selective provision of data. Finally, people and parties who had submitted comments on the preliminary version of the health technology assessment report at the public hearing in July 2009 were asked to provide any additional relevant trials.

Two reviewers independently screened titles and abstracts of the retrieved citations to identify potentially eligible primary and secondary publications. In a first broad screening step, citations were excluded if clearly irrelevant; that is, if a primary publication was not a clinical trial in humans with depression, or if a secondary publication of eligible trials was not a systematic review. In a second screening step, the full set of eligibility criteria was applied. Potentially relevant articles were then screened as full texts. Disagreement was resolved by consensus.

Data extraction and assessment of risk of bias
The individual steps of the data extraction and assessment of risk of bias were always conducted by one person and checked by another. Details of the trials were extracted using standardised tables. Information and data from publications were supplemented by clinical study reports provided by the manufacturer. We always extracted data from the intention to treat populations. Clinical study reports were always considered the primary source in instances of conflict with the publication. Disagreement was resolved by consensus.

Information was extracted from each included trial on:

- Study characteristics, including citation, study design, setting (inpatient or outpatient), inclusion and exclusion criteria, length of follow-up, sample size, location, number of centres, and year of completion
- Characteristics of the study participants, including age, gender, and disease severity at baseline
- Characteristics of the test and control interventions, including dose
- Outcomes and type of outcome measures (outcomes as presented above; measurement tools as used in the individual trials)
- Risk of bias items.

The risk of bias at the study level was assessed on the basis of the adequacy of the following criteria: randomisation; allocation concealment; blinding of patients and investigators; and complete and non-selective results reporting. The risk of bias at the outcome level was assessed on the basis of the adequacy of: application of the intention to treat principle; blinding of the outcome assessor; statistical evaluation; and complete and non-selective results reporting. Trials and outcomes were categorised into those with a low risk of bias and those with a high risk.

Data analysis
If feasible and meaningful, data were pooled by means of meta-analyses. Effect measures were reported as odds ratios (ORs) and 95% confidence intervals (CIs) for binary data. A random effects model was used to calculate a pooled effect estimate. Statistical significance was assumed for P<0.05. Heterogeneity of effect sizes was assessed by using the I² statistic; pooled estimates were not calculated if substantial heterogeneity was observed (I²>50%). If heterogeneity with I²>50% was shown, sensitivity analyses were conducted, when appropriate, to assess possible sources of heterogeneity across the trials included. The review protocol prespecified potential effect modifiers, including gender and trial setting (inpatient or outpatient). These factors were investigated by means of random effects meta-regression analyses based on aggregate study data.17
To assess publication bias, effect sizes in the published, unpublished, and overall dataset were compared. In addition, the differences in effect sizes between published and unpublished data, and between published and overall data, were expressed as the ratio of odds ratios (OR). The magnitude of the overestimation or underestimation of effect sizes in published versus overall data (publication bias) was expressed as percentage changes.

Meta-analyses were performed using SAS version 9.1.3. If meta-analyses were not possible, the results of the individual trials were assessed.

RESULTS

Study selection

The process of study selection is presented in figure 1. The search in bibliographic databases yielded 2596 citations, of which 713 were classified as potentially relevant and subjected to strict eligibility assessment. A total of 13 citations (10 trials) met the inclusion criteria; however, two of these 13 citations were publications on subgroups of otherwise unpublished trials,18 19 and one was the only available publication on the total population being studied but did not report the main outcomes.20 In the assessment of publication bias, we considered these three trials to be “unpublished.” No trials were identified in clinical trial or trial results registers or in the European Medicines Agency or FDA websites.

The retrieval of previously unpublished trials was hampered by the fact that during preparation of the preliminary health technology assessment report, the manufacturer of reboxetine did not provide a complete list of unpublished trials as requested by IQWiG.21 22 Secondary publications clearly indicated that further potentially relevant unpublished trials existed.68 As the preliminary report showed that reboxetine had been tested in at least 16 trials including about 4600 patients, but data on almost two thirds of these patients were not

<table>
<thead>
<tr>
<th>Trial</th>
<th>Year of completion</th>
<th>Primary publication available?</th>
<th>Clinical study report available?*</th>
</tr>
</thead>
<tbody>
<tr>
<td>014</td>
<td>Before 1996</td>
<td>Refs 42-44</td>
<td>Ref 45†</td>
</tr>
<tr>
<td>015</td>
<td>1992</td>
<td>None, only a pooled analysis (ref 6)</td>
<td>Ref 46</td>
</tr>
<tr>
<td>016</td>
<td>1993</td>
<td>Ref 47</td>
<td>Ref 48</td>
</tr>
<tr>
<td>032</td>
<td>2001</td>
<td>None</td>
<td>Ref 49</td>
</tr>
<tr>
<td>043</td>
<td>2001</td>
<td>Ref 50</td>
<td>Ref 51</td>
</tr>
<tr>
<td>045</td>
<td>1999</td>
<td>None</td>
<td>Ref 52</td>
</tr>
<tr>
<td>046</td>
<td>2000</td>
<td>None</td>
<td>Ref 53</td>
</tr>
<tr>
<td>047</td>
<td>2000</td>
<td>Ref 19, although the data for the full study population were not reported</td>
<td>Ref 54</td>
</tr>
<tr>
<td>049</td>
<td>1998</td>
<td>None</td>
<td>Ref 55</td>
</tr>
<tr>
<td>050</td>
<td>1999</td>
<td>Ref 20, although only data on sexual dysfunction and other adverse events were reported</td>
<td>Ref 56</td>
</tr>
<tr>
<td>052</td>
<td>2000</td>
<td>Ref 18, although the data for the full study population were not reported</td>
<td>Ref 57</td>
</tr>
<tr>
<td>091</td>
<td>1990</td>
<td>Refs 58 and 59</td>
<td>Ref 60</td>
</tr>
<tr>
<td>Berlanga and Flores-Ramos 2006</td>
<td>2003</td>
<td>Ref 61</td>
<td>No</td>
</tr>
</tbody>
</table>

*As a matter of principle, the German Institute for Quality and Efficiency in Health Care requests documents compliant with the CONSORT criteria from manufacturers on all relevant trials selected. If cooperative, manufacturers usually provide the full clinical study report; that is, a written description of the study that follows the guidelines of the International Conference on Harmonisation.62
†Only addendum.

![Fig 2](http://www.bmj.com/) Forest plot showing meta-analyses of remission and response rates for trials that compared reboxetine with placebo. Empty boxes show published studies and filled boxes show unpublished studies. Study 091 is not included in the pooled analysis of response of reboxetine versus placebo because of high heterogeneity (see text for details). CI, confidence interval; n, number of patients with event; N, number of patients in treatment group.
Table 2 | Trial characteristics and baseline demographics

<table>
<thead>
<tr>
<th>Trial</th>
<th>Treatments</th>
<th>Dose (mg/d)</th>
<th>Proportion of maximum approved daily dose (%)</th>
<th>Number of patients randomised</th>
<th>Duration of active medication (weeks)</th>
<th>Number of centres (locations)*</th>
<th>Setting</th>
<th>Baseline demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>014</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>156</td>
<td>8</td>
<td>33 (Europe, South America)</td>
<td>Inpatient and outpatient</td>
<td>40 (12)</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Fluoxetine 20-40</td>
<td>50-55</td>
<td>128</td>
<td>44 (12)</td>
<td>54</td>
<td>27.6 (3.4)</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo —</td>
<td>—</td>
<td>34 (North America, Europe, Australia)</td>
<td>Inpatient and outpatient</td>
<td>46 (13)</td>
<td>63</td>
<td>27.5 (5.1)</td>
<td>21</td>
</tr>
<tr>
<td>016</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>79</td>
<td>8</td>
<td>16 (Europe, South America, Australia)</td>
<td>Inpatient and outpatient</td>
<td>44 (11)</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Fluoxetine 20-40</td>
<td>50-55</td>
<td>89</td>
<td>44 (12)</td>
<td>72</td>
<td>27.4 (4.1)</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo —</td>
<td>—</td>
<td>43 (12)</td>
<td>48</td>
<td>27.1 (5.3)</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>032</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>43</td>
<td>8</td>
<td>5 (Asia)</td>
<td>Inpatient and outpatient</td>
<td>47 (15)</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Fluoxetine 20-40</td>
<td>30-50</td>
<td>42</td>
<td>36 (13)</td>
<td>62</td>
<td>28.3 (5.3)</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>043</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>183</td>
<td>24</td>
<td>23 (Europe)</td>
<td>Outpatient</td>
<td>47 (13)</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Citalopram 20-40</td>
<td>33-67</td>
<td>176</td>
<td>42 (12)</td>
<td>60</td>
<td>27.4 (3.9)</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>045</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>89</td>
<td>6</td>
<td>48 (Europe, Asia)</td>
<td>Inpatient and outpatient</td>
<td>42 (11)</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Placebo —</td>
<td>—</td>
<td>41 (11)</td>
<td>70</td>
<td>26.4 (2.6)</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>046</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>265</td>
<td>8</td>
<td>94 (North America)</td>
<td>N/A</td>
<td>40 (11)</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>Paroxetine 20-40</td>
<td>40-80</td>
<td>265</td>
<td>40 (12)</td>
<td>69</td>
<td>22.8 (5.4)</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo —</td>
<td>—</td>
<td>39 (12)</td>
<td>70</td>
<td>23.0 (5.2)</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>047</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>258</td>
<td>8</td>
<td>68 (North America)</td>
<td>N/A</td>
<td>39 (12)</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Paroxetine 20-40</td>
<td>40-80</td>
<td>262</td>
<td>40 (11)</td>
<td>72</td>
<td>23.9 (5.4)</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo —</td>
<td>—</td>
<td>37 (11)</td>
<td>82</td>
<td>23.7 (4.8)</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>049</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>107</td>
<td>6</td>
<td>9 (North America)</td>
<td>Outpatient</td>
<td>40 (12)</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Placebo —</td>
<td>—</td>
<td>40 (11)</td>
<td>58</td>
<td>25.3 (3.0)</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>050</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>150</td>
<td>8</td>
<td>24 (North America)</td>
<td>Outpatient</td>
<td>40 (11)</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Fluoxetine 20-40</td>
<td>50-50</td>
<td>150</td>
<td>41 (11)</td>
<td>66</td>
<td>26.0 (3.3)</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo —</td>
<td>—</td>
<td>40 (11)</td>
<td>60</td>
<td>25.3 (3.3)</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>052</td>
<td>Reboxetine 8-10</td>
<td>67-83</td>
<td>159</td>
<td>8</td>
<td>41 (Europe)</td>
<td>N/A</td>
<td>42 (12)</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Paroxetine 20-40</td>
<td>40-80</td>
<td>166</td>
<td>45 (11)</td>
<td>62</td>
<td>24.1 (3.4)</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>091</td>
<td>Reboxetine 10</td>
<td>80-133</td>
<td>33-67</td>
<td>6</td>
<td>3 (North America, South America)</td>
<td>Inpatient</td>
<td>42 (N/A)</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Citalopram 20-40</td>
<td>33-67</td>
<td>55</td>
<td>40 (N/A)</td>
<td>50</td>
<td>35.1 (N/A)</td>
<td>57</td>
<td></td>
</tr>
</tbody>
</table>

*Details on individual countries are provided in web table A.
†To comply with the intention to treat principle, missing data from discontinued patients were imputed by using the last observation carried forward method.
N/A, not available.

accessible, the institute initially concluded that no meaningful assessment of reboxetine was possible.21 22

After the publication of the preliminary report, the manufacturer decided to cooperate and provided most of the missing data (one venlafaxine controlled trial23 was not available as a full publication). Thus, an additional 10 previously unpublished or incompletely published reboxetine trials were considered in the final health technology assessment report.12 Two trials with tricyclic antidepressants as active controls and two relapse prevention trials were excluded from the present analysis.

Of the remaining 13 eligible acute treatment trials, three were placebo controlled, five were active controlled, and five had both placebo and active controlled arms (one of which had a tricyclic antidepressant arm that was not considered). A total of 4098 patients were analysed: 2256 in the reboxetine versus placebo comparisons and 2641 in the reboxetine versus SSRI comparisons.

Study characteristics
The characteristics of the pool of 13 acute treatment trials that were placebo controlled, SSRI controlled, or both are presented in tables 1 and 2. All trials were sponsored by predecessors of Pfizer (Pharmacia, and Pharmacia & Upjohn), except for Berlanga and Flores-Ramos 2006 (sponsored by Lundbeck), and included adult patients with major depressive disorder according to the third edition, revised or the fourth edition of the Diagnostic
and Statistical Manual of Mental Disorders. In the four fluoxetine controlled trials and in one citalopram controlled trial, the SSRIs were potentially underdosed compared with reboxetine (according to doses standardised on the basis of the maximum approved dose; see table 2). The trials were well balanced between treatment arms with respect to patient baseline characteristics.

There were no major differences between trials in terms of dosage and mean patient age. However, there were differences in setting (inpatient, outpatient, or both) and baseline severity of depression as measured by the Hamilton depression rating scale. For more details on trial characteristics see web table A.

Risk of bias

The overall methodological quality of the trials was good (table 3). At the trial level, the risk of bias was low in all but one study, which had a high risk of bias at the trial level owing to unclear allocation concealment and blinding. At the outcome level, the risk of bias was low for all four benefit and harm outcomes in nine out of the 13 trials. Three trials had a high risk of bias at the outcome level owing to an inadequate intention to treat analysis. Analyses excluding the outcomes at high risk of bias did not alter the conclusions (data not shown). As no clear dose-response relationship has been shown for fluoxetine and citalopram, the potential underdosing of these agents in five trials did not affect the risk of bias.

Owing to the availability of a comprehensive set of the relevant data on reboxetine versus placebo and SSRIs, the risk of publication bias on the results of the final analysis was minor.

Effects of interventions

In this text, the terms “superior” and “inferior” refer to statistically significant differences between treatment groups (P<0.05).

Meta-analyses of remission and response rates

The Hamilton depression rating scale was used in the meta-analyses of remission and response rates. No statistically significant difference between reboxetine and placebo was noted in the meta-analysis of remission rates (OR 1.17, 95% CI 0.91 to 1.51; P=0.216; fig 2).

Substantial heterogeneity (I²=67.3%; p=0.003) was shown in the meta-analysis of response rates including all eight trials that compared reboxetine with placebo, and consequently no point estimate was calculated. The only known inpatient trial—trial 091 (n=52), which had an OR of 11.43 (95% CI 3.10 to 42.12)—was obviously a statistical outlier (figure 2).

In the sensitivity analysis using meta-regression analysis, setting had an effect on the outcome response. Patients who received reboxetine in an inpatient setting were more likely to show a good response compared with placebo than were patients who received reboxetine in an outpatient setting (P=0.001 inpatients v outpatients; trials 091 v 049 and 050). In a second scenario, the proportion of inpatients was used as the independent variable. This analysis also included trials 014 and 015, for which the proportion of inpatients was available from Montgomery et al 2003.7 This scenario confirmed the influence of setting (P<0.001). The meta-analysis of response rates in the outpatient only trials (049 and 050) showed no statistically significant difference between reboxetine and placebo (OR 1.05, 95% CI 0.73 to 1.50; P=0.796 I²=0%). These findings indicate that patient setting was the most probable effect modifier. After exclusion of trial 091, the meta-analysis of response rates in the seven remaining trials showed no statistically significant difference between reboxetine and placebo (OR 1.24, 95% CI 0.98 to 1.56, P=0.071, I²=42.1%; figure 2).

Reboxetine was inferior to SSRIs in the meta-analysis of remission rates (OR 0.80, 95% CI 0.67 to 0.96; P=0.015; fig 3). A similar, although non-significant, trend in remission rates was shown when reboxetine was compared with the individual SSRIs (fluoxetine, paroxetine, and citalopram). However, if remission rates according to the Montgomery-Åsberg depression rating scale rather than the Hamilton depression rating scale were analysed from trials using this instrument as the primary scale (trials 046 and 047), reboxetine was inferior to paroxetine (OR 0.72, 95% CI 0.56 to 0.93). In the long term acute treatment trial (trial 043), reboxetine was inferior to citalopram (OR 0.51, 95% CI 0.32 to 0.83). However, the intention to treat principle was violated in this trial, so a worst case analysis was conducted in which the difference in remission rate compared with citalopram was no longer statistically significant.

Reboxetine was also inferior to SSRIs in the meta-analysis of response rates (OR 0.80, 95% CI 0.67 to 0.95; P=0.010). A similar trend was shown when reboxetine was compared with the individual SSRIs, where the trend reached statistical significance in the comparison of reboxetine and paroxetine (OR 0.79,
95% CI 0.64 to 0.99; P=0.04). In the long term acute treatment trial, no statistically significant difference was shown between reboxetine and citalopram (OR 0.60, 95% CI 0.35 to 1.03).

The overall findings were also reflected in the subset of trials that were both placebo controlled and SSRI controlled (n=4), which are suited to demonstrating assay sensitivity. The SSRIs were superior to placebo and reboxetine in this analysis, but no statistically significant difference was shown between reboxetine and placebo (see web figure A).

Meta-analyses of adverse events and withdrawals owing to adverse events
Reboxetine was inferior to placebo (that is, it was associated with higher event rates) in the meta-analyses of the rates of patients with at least one adverse event and in the meta-analysis of the rates of withdrawals owing to adverse events (OR 2.14, 95% CI 1.59 to 2.88; P<0.001 and OR 2.21, 95% CI 1.45 to 3.37; P<0.001, respectively; fig 4).

The rates of patients with at least one adverse event did not differ significantly between patients treated with reboxetine and those who received an SSRI (OR 1.06, 95% CI 0.82 to 1.36; P=0.667; fig 5). The same finding was seen for patients on reboxetine versus those treated with individual SSRIs. A meta-regression analysis showed a gender effect in the comparison of reboxetine with fluoxetine (P=0.022 for the interaction test): in men reboxetine was inferior to fluoxetine in the meta-analysis of patients with at least one adverse event (OR 2.76, 95% CI 1.28 to 5.93), whereas no significant difference was shown in women (OR 0.90, 95% CI 0.51 to 1.59). In the long term acute treatment trial, reboxetine was inferior to citalopram (OR 1.57, 95% CI 1.03 to 2.38).

Fig 3 | Forest plot showing meta-analyses of remission and response rates for trials that compared reboxetine with selective serotonin reuptake inhibitors (SSRIs; fluoxetine, paroxetine, and citalopram). Empty boxes show published studies and filled boxes show unpublished studies. Empty diamonds show subtotals (individual SSRIs) and filled diamonds show overall totals (all SSRIs). CI, confidence interval; n, number of patients with event; N, number of patients in treatment group
other hand, the comparison between reboxetine and paroxetine concerning rates of withdrawals owing to adverse events was inferior to citalopram (OR 1.79, 95% CI 1.06 to 3.05; P = 0.031). On the other hand, the comparison between reboxetine and paroxetine showed substantial heterogeneity (I² = 84.2%), but the sensitivity analysis did not identify a potential effect modifier. We therefore concluded that there was no proof of a difference between reboxetine and paroxetine concerning rates of withdrawals owing to adverse events. In the long term acute treatment trial, reboxetine was inferior to citalopram (OR 4.61, 95% CI 2.15 to 9.89).

Further information on adverse events

For both harm outcomes, the addition of unpublished data changed the superiority of reboxetine versus placebo shown in published data to a non-significant difference and also changed the non-significant difference between reboxetine and SSRIs to an inferiority of reboxetine (fig 6). Comparison of the published data with the full dataset (published and unpublished) showed that the published data overestimated the beneficial effect of reboxetine compared with placebo by 99-115% and of reboxetine compared with SSRIs by 19-23%.

For both harm outcomes, the addition of unpublished data changed the non-significant difference between reboxetine and placebo shown in published data to an inferiority of reboxetine. For rates of withdrawals owing to adverse events, the addition of unpublished data changed the non-significant difference in the comparisons of reboxetine to an inferiority of reboxetine; this was primarily owing to the increased power of the analysis rather than to major differences in withdrawal rates between published and unpublished data. For patients with at least one adverse event, no significant impact of unpublished data was shown in the comparison between reboxetine and SSRIs.

DISCUSSION

To our knowledge, this is the first systematic review of a comprehensive evidence base of published and unpublished acute treatment trials of reboxetine versus...
placebo or SSRIs in adults with major depressive disorder. We found that, overall, reboxetine was ineffective as an antidepressant because it showed no benefit over placebo and was inferior to SSRIs for remission and response rates. A benefit of reboxetine (higher response rates) was shown in a placebo controlled trial in inpatients; however, this trial was too small to draw general conclusions on the effect of reboxetine in this patient population. Reboxetine was inferior to placebo for both harm outcomes and to fluoxetine for rates of withdrawals owing to adverse events.

Given the potential underdosing of fluoxetine and citalopram in five trials, our findings on reboxetine might be considered conservative. At the same time, the advantages of SSRIs concerning harm might be overestimated. However, as stated, no clear dose-response relationship has been shown for fluoxetine and citalopram. Furthermore, in our test of assay sensitivity that included two of the four potentially underdosed fluoxetine arms, even the lower fluoxetine dose showed a clear benefit compared with placebo (OR 1.98, 95% CI 1.19 to 3.28, I²=53.8%), thus qualifying the effect of dosing in treatment for depression.

Data on 74% of the patients included in our analysis was unpublished, indicating that the published evidence on reboxetine so far has been severely affected by publication bias. Our comparison of published and unpublished trials confirmed this assumption: the positive benefit-risk ratio of reboxetine in the published literature was changed to a negative ratio if unpublished trials were added to the analysis.

Comparison with other reviews

The results of our review largely contradict the findings of previous systematic reviews and analyses of reboxetine versus placebo[7,9] and reboxetine versus active comparators. The solely placebo controlled analyses by Ferguson et al[6] and Montgomery et al[7] both found greater efficacy (including higher response rates) for reboxetine compared with placebo, and Ferguson et al also found comparable harms. However, both reviews included only three of the eight studies considered in our review (plus the inpatient trial by Ban et al[198]). These two reviews also included only one unpublished trial[105], even though the relevant unpublished trials had been included in our analysis. Therefore, our results are more conservative and provide a more reliable basis for judging the efficacy and safety of reboxetine for treatment of depression.

### Table

<table>
<thead>
<tr>
<th>Patients with adverse events</th>
<th>Odds ratio (95% CI)</th>
<th>Weight (%)</th>
<th>Odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluoxetine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>014</td>
<td>1.90 (1.14 to 1.71)</td>
<td>13.1</td>
<td>1.09 (0.52 to 1.88)</td>
</tr>
<tr>
<td>016</td>
<td>1.09 (0.52 to 1.88)</td>
<td>13.1</td>
<td>1.90 (1.14 to 1.71)</td>
</tr>
<tr>
<td>050</td>
<td>1.01 (0.60 to 1.82)</td>
<td>7.3</td>
<td>1.01 (0.60 to 1.82)</td>
</tr>
<tr>
<td>032</td>
<td>1.01 (0.60 to 1.82)</td>
<td>7.3</td>
<td>1.01 (0.60 to 1.82)</td>
</tr>
<tr>
<td>Fluoxetine subtotal</td>
<td>1.25 (0.89 to 1.76)</td>
<td>12.5</td>
<td>1.25 (0.89 to 1.76)</td>
</tr>
<tr>
<td>Paroxetine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>046</td>
<td>1.16 (0.68 to 2.23)</td>
<td>11.8</td>
<td>1.16 (0.68 to 2.23)</td>
</tr>
<tr>
<td>047</td>
<td>1.16 (0.68 to 2.23)</td>
<td>11.8</td>
<td>1.16 (0.68 to 2.23)</td>
</tr>
<tr>
<td>052</td>
<td>1.16 (0.68 to 2.23)</td>
<td>11.8</td>
<td>1.16 (0.68 to 2.23)</td>
</tr>
<tr>
<td>Paroxetine subtotal</td>
<td>1.25 (0.89 to 1.76)</td>
<td>12.5</td>
<td>1.25 (0.89 to 1.76)</td>
</tr>
<tr>
<td>Total</td>
<td>1.14 (0.68 to 1.91)</td>
<td>11.2</td>
<td>1.14 (0.68 to 1.91)</td>
</tr>
<tr>
<td>Total heterogeneity: I²=0.0%, P=0.565; subtotal effect: P=0.192</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Figure 5

Forest plot showing meta-analyses of rates of patients with at least one adverse event and rates of withdrawals owing to adverse events for trials that compared reboxetine with selective serotonin reuptake inhibitors (SSRIs; fluoxetine and paroxetine). Empty boxes show published studies and filled boxes show unpublished studies. Empty diamonds show subtotals (individual SSRIs) and filled diamonds show overall totals (all SSRIs). CI, confidence interval; n, number of patients with event; N, number of patients in treatment group.
completed before publication of these analyses and both reviews were co-written by a sponsor employee. The meta-analysis by Chuluunkhuu et al \(^9\) concluded that reboxetine showed superior efficacy to placebo and found no difference in efficacy of reboxetine compared with SSRIs and other antidepressants. However, this analysis considered only published data.

Although the meta-analysis by Papakostas et al \(^8\) identified and included a large body of unpublished studies that used SSRIs as the control (the same set as we used), they found no significant difference in response rates between SSRIs and reboxetine (risk ratio 1.08, 95% CI 0.98 to 1.19). Their analysis included the long term acute treatment trial 043, which we analysed separately. In contrast, our analysis showed that reboxetine was inferior to SSRIs, even if trial 043 was included (recalculated according to Papakostas: risk ratio (SSRI vs reboxetine) 1.10, 95% CI 1.03 to 1.17; \(P=0.003\)). The reason for this discrepancy is unclear, because Papakostas et al reported only point estimates and CIs and did not report the number of actual events or the corresponding populations.

The widely discussed systematic review by Cipriani et al \(^11\), which assessed 12 new generation antidepressants in a network meta-analysis and ranked reboxetine last, had similar findings to those of our review. These authors found significantly lower response rates for reboxetine than for all SSRIs investigated, as well as significantly higher dropout rates versus fluoxetine, citalopram, escitalopram, and sertraline. However, despite the similarity in findings, the evidence base of the Cipriani review differed markedly from that in our review because placebo controlled trials were omitted and trials that were not double blind, which carry a higher risk of bias, were considered. In addition, unpublished trials of reboxetine on file at the manufacturers were not considered, even though significant publication bias has been
Table 4 | Examples of publication bias and industry sponsorship bias in trials of antidepressants

<table>
<thead>
<tr>
<th>Source</th>
<th>Study type</th>
<th>Antidepressant type</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turner et al 2008</td>
<td>Comparison of FDA reviews and matching publications</td>
<td>SSRI, SNRI, NDRIS, TCAs, and atypical antidepressants</td>
<td>Among 74 FDA registered studies, 31%, accounting for 3449 study participants, were not published... A total of 37 studies viewed by the FDA as having positive results were published... Studies viewed by the FDA as having negative or questionable results were, with 3 exceptions, either not published (22 studies) or published in a way that, in our opinion, conveyed a positive outcome (11 studies). According to the published literature, it appeared that 94% of the trials conducted were positive. By contrast, the FDA analysis showed that 51% were positive... the increase in effect size ranged from 11% to 69% for individual drugs and was 32% overall.*</td>
</tr>
<tr>
<td>Kirsch et al 2008</td>
<td>Meta-analysis of data submitted to the FDA</td>
<td>SSRI, SNRI, and atypical antidepressants</td>
<td>“[The FDA public disclosure did not include mean changes for nine trials that were deemed adequate and well controlled but that failed to achieve a statistically significant benefit for drug over placebo... Specifically, four sertraline trials involving 486 participants and one citalopram trial involving 274 participants were reported as having failed to achieve a statistically significant drug effect, without reporting mean Hamilton rating scale of depression scores. We were unable to find data from these trials on pharmaceutical company websites or through our search of the published literature. These omissions represent 38% of patients in sertraline trials and 23% of patients in citalopram trials.”</td>
</tr>
<tr>
<td>Whittington et al 2004</td>
<td>Systematic review of published versus unpublished data</td>
<td>SSRI, SNRI</td>
<td>“Data for two published trials suggest that fluoxetine has a favourable risk-benefit profile, and unpublished data lend support to this finding. Published results from one trial of paroxetine and two trials of sertraline suggest equivocal or weak positive risk-benefit profiles. However, in both cases, addition of unpublished data indicates that risks outweigh benefits. Data from unpublished trials of citalopram and venlafaxine show unfavourable risk-benefit profiles.”</td>
</tr>
<tr>
<td>Melander et al 2003</td>
<td>Analysis of industry sponsored studies in new drug applications</td>
<td>SSRI</td>
<td>“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.”</td>
</tr>
<tr>
<td>Jureidini et al 2008</td>
<td>Case report on selective reporting</td>
<td>Paroxetine</td>
<td>“The published report of study 329 of paroxetine in adolescents sponsored by GlaxoSmithKline claims that paroxetine is generally well tolerated and effective for major depression in adolescents. By contrast, documents obtained during litigation reveal that study 329 was negative for efficacy on all eight protocol specified outcomes and positive for harm.”</td>
</tr>
<tr>
<td>Tungaraza et al 2007</td>
<td>Analysis of influence of industry authorship and funding</td>
<td>Not specified*</td>
<td>“Independent studies were more likely to report negative findings than were industry funded studies. However, the involvement of a drug company employee had a much greater effect on study outcome than financial sponsorship alone.”</td>
</tr>
<tr>
<td>Perlis et al 2005</td>
<td>Analysis of influence of industry funding and financial conflict of interest</td>
<td>Not specified*</td>
<td>“Among the 162 randomised, double blind, placebo controlled studies examined, those that reported conflict of interest were 4.9 times more likely to report positive results; this association was significant only among the subset of pharmaceutical industry funded studies.”</td>
</tr>
<tr>
<td>Kelly et al 2006</td>
<td>Analysis of influence of industry funding</td>
<td>Not specified*</td>
<td>“Favourable outcomes were significantly more common in studies sponsored by the drug manufacturer (78%) than in studies without industry sponsorship (48%) or sponsored by a competitor (28%).”</td>
</tr>
</tbody>
</table>

*Findings also refer to other psychiatric drugs. All analyses examined drug trials reported in psychiatric journals. No separate results for antidepressants were reported.

Our findings that reboxetine was superior [higher response rates] to placebo in a small trial in inpatients and that patient setting was a probable effect modifier are supported by the four week active controlled and placebo controlled inpatient trial by Ban et al (n=169 in the reboxetine and placebo arms), which we excluded owing to its short duration. Ban et al also found a statistically significant higher response rate in inpatients who received reboxetine compared with those who received placebo (60% v 35%; OR 2.70, 95% CI 1.45 to 5.03 [own calculation]).

Strengths and limitations of the review

The main strength of our review is the inclusion of a large amount of previously unpublished data. As we made extensive efforts to identify unpublished trials, we are optimistic that we analysed the vast majority or even all of the placebo controlled and SSRI controlled double blind randomised trials of reboxetine in adults with major depression.

Our review also has a number of limitations. We only had access to aggregated data. To assess the impact of effect modifiers, meta-analysis of individual patient data would be needed to determine the setting in studies with mixed settings and to test our hypothesis that the setting was the effect modifier explaining the substantial heterogeneity in the meta-analysis of response rates in placebo controlled trials.

Our results are further limited by the fact that they only refer to acute treatment trials, only one of which lasted more than eight weeks. However, six to eight weeks is the standard study duration in trials investigating the acute treatment of depression. The long term acute treatment trial showed similar, though not always statistically significant, trends to the short term trials. Other long term outcomes in depression, such as prevention of relapse or recurrence, were not the focus of this paper.

Finally, except for a subgroup analysis for gender and setting, we assessed total populations of patients with major depressive disorder. No analyses were performed in other subgroups of patients (for example, patients with severe disease or specific major depressive...
WHAT IS ALREADY KNOWN ON THIS TOPIC

Reboxetine has been approved for the treatment of major depression in many European countries, but the application for approval was rejected in the United States. Doubts have been raised about the efficacy of reboxetine. Research into antidepressants is particularly affected by publication bias.

WHAT THIS STUDY ADDS

Overall, reboxetine is not effective for the treatment of major depressive disorder. We found a higher rate of patients affected by adverse events than with placebo and higher withdrawal rates owing to adverse events than with placebo and fluoxetine. This meta-analysis provides a striking example of publication bias, in which the previously favourable risk-benefit profile of reboxetine shown in published trials is reversed by the addition of unpublished data. Post-approval regulatory decisions (for example, reimbursement decisions based on the findings of health technology assessment reports) might be affected by publication bias. Our findings underline the need for mandatory publication of clinical trial results in the treatment of depression. In our opinion, this conclusion can no longer be upheld.

The ongoing problem of publication bias shows that unbiased decision making in health care requires mandatory public disclosure of all clinical trial data. The US FDA Amendments Act of 2007 solves the problem in part by requiring protocol information and study results for clinical trials to be made public on the clinicaltrials.gov website (www.clinicaltrials.gov; please see accompanying comment (doi:10.1136/bmj.c4942) for further details). Similar legislation is also being introduced in Europe, with the mandatory public disclosure of data from the clinical trials database EudraCT (eudraact.ema.europa.eu) but the date of implementation is not yet clear.

As the full assessment reports on reboxetine prepared by regulatory authorities are not publicly available, it is not clear as to how the comprehensive body of evidence (including that on efficacy outcomes) generated after reboxetine was approved in Europe in the late 1990s has been analysed by these authorities. The reason for the difference in approval status of reboxetine between Europe and the US thus remains unclear.

Conclusions and policy implications

Our analysis of a comprehensive evidence base of published and unpublished trials of reboxetine compared with placebo or SSRIs in adults with major depressive disorder indicates that reboxetine is, overall, an ineffective and potentially harmful antidepressant. Published evidence on reboxetine has been substantially affected by publication bias, underlining the urgent need for mandatory publication of clinical trial data, including data on older agents.

We thank Katharina Quitmann, Yvonne-Beatrice Schülker, and Volker Verheugen for general support, and Natalie McGauran for editorial support. We also thank Elke Hausner for developing the search strategy and performing the literature search. All persons listed above are employed by the Institute for Quality and Efficiency in Health Care (IQWiG), Cologne, Germany.

Contributors: DE coordinated and BW supervised the review. DE, ML, UG, MH, TK, MG, and BW designed the protocol. UG and MK planned and did the statistical analyses. DE, ML, and MG selected studies. MFK did the data extraction. All authors (except MFK) assessed studies. DE wrote the first draft of the paper. All authors contributed to data interpretation and to critical revision of the paper, and have seen and approved the final version. BW is the guarantor.

Funding: This work was supported by IQWiG. The original health technology assessment report was commissioned by the German Federal Joint Committee. MG, MH, and ML acted as consultants for the preparation of the original health technology assessment report. For this they were reimbursed by IQWiG.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any company for the submitted work; DE was employed by H Lundbeck A/S, Copenhagen, between January 2006 and April 2007, MH received remuneration from Boehringer Ingelheim and Lilly Pharma for three talks on depression guidelines in 2008, and UG, MK, TK, MFK, and BW are employees of IQWiG. DE is a former employee of IQWiG. In order to produce unbiased health technology assessment reports, the institute depends on access to all of the relevant data on the topic under investigation. The authors therefore support the mandatory worldwide establishment of trial registries and study results databases. ML and MH were involved in the development of the German Disease Management Guideline on Depression. Ethical approval: Not required.
Data sharing: The full (German) version of the health technology assessment report (including the search strategy) and the clinical study reports on reboxetine are available on the IQWiG website (www.iqwig.de).


45 Pharmacia Limited. Multicentre, multinational double-blind study of the activity and tolerability of reboxetine vs fluoxetine in patients suffering from major depressive episodes (phase III): results of patient self-evaluation assessment instrument; addendum to final


Pharmacia & Upjohn. Comparison of placebo and three fixed doses of reboxetine in a population of patients with major depression: a phase II, double-blind, randomized, parallel group, multicenter study of 3 fixed doses of reboxetine or placebo, given orally twice daily to adult patients with major depressive disorder; final report of the trial 95-CRBX-045. 2001. www.iqwig.de/download/Studienbericht_zu_Studie_045.pdf.

Pharmacia & Upjohn. Comparison of placebo and three fixed doses of reboxetine in a population of patients with major depression: a phase II, double-blind, randomized, parallel group, multicenter study of 3 fixed doses of reboxetine or placebo, given orally twice daily to adult patients with major depressive disorder; final report of the trial 95-CRBX-045. 2001. www.iqwig.de/download/Studienbericht_zu_Studie_045.pdf.


Accepted: 16 August 2010