

## RESEARCH

# Benefits and harms in clinical trials of duloxetine for treatment of major depressive disorder: comparison of clinical study reports, trial registries, and publications

 OPEN ACCESS

Emma Maund *PhD student*<sup>1</sup>, Britta Tendal *postdoctoral researcher*<sup>1</sup>, Asbjørn Hróbjartsson *senior researcher*<sup>1</sup>, Karsten Juhl Jørgensen *senior researcher*<sup>1</sup>, Andreas Lundh *physician*<sup>1,2</sup>, Jeppe Schroll *PhD student*<sup>1</sup>, Peter C Gøtzsche *Professor*<sup>1</sup>

<sup>1</sup>Nordic Cochrane Centre, Rigshospitalet Dept 7811, Copenhagen, Denmark; <sup>2</sup>Department of Infectious Diseases, Hvidovre University Hospital, Kettegårds Allé 30, 2650 Hvidovre, Denmark

## Abstract

**Objective** To determine, using research on duloxetine for major depressive disorder as an example, if there are inconsistencies between protocols, clinical study reports, and main publicly available sources (journal articles and trial registries), and within clinical study reports themselves, with respect to benefits and major harms.

**Design** Data on primary efficacy analysis and major harms extracted from each data source and compared.

**Setting** Nine randomised placebo controlled trials of duloxetine (total 2878 patients) submitted to the European Medicines Agency (EMA) for marketing approval for major depressive disorder.

**Data sources** Clinical study reports, including protocols as appendices (total 13 729 pages), were obtained from the EMA in May 2011. Journal articles were identified through relevant literature databases and contacting the manufacturer, Eli Lilly. Clinicaltrials.gov and the manufacturer's online clinical trial registry were searched for trial results.

**Results** Clinical study reports fully described the primary efficacy analysis and major harms (deaths (including suicides), suicide attempts, serious adverse events, and discontinuations because of adverse events). There were minor inconsistencies in the population in the primary efficacy analysis between the protocol and clinical study report and within the clinical study report for one trial. Furthermore, we found contradictory information within the reports for seven serious adverse events and eight adverse events that led to discontinuation but with no apparent bias. In each trial, a median of 406 (range 177-645) and 166 (100-241) treatment emergent adverse events (adverse events that emerged or worsened after study drug was started) in the randomised phase were not reported

in journal articles and Lilly trial registry reports, respectively. We also found publication bias in relation to beneficial effects.

**Conclusion** Clinical study reports contained extensive data on major harms that were unavailable in journal articles and in trial registry reports. There were inconsistencies between protocols and clinical study reports and within clinical study reports. Clinical study reports should be used as the data source for systematic reviews of drugs, but they should first be checked against protocols and within themselves for accuracy and consistency.

## Introduction

About half of all randomised clinical trials are never published,<sup>1</sup> and the other half is often published selectively,<sup>2</sup> in both cases depending on the direction of the results.

Researchers who had access to unpublished clinical study reports at drug agencies have found that reporting biases were common in trials of antidepressants,<sup>3,4</sup> which in one study led to an overall 32% overestimation of the treatment effect.<sup>4</sup> Other researchers have found that, contrary to meta-analyses based on published data only, meta-analysis of published and unpublished data showed that the antidepressant reboxetine was no more effective than placebo but caused greater harm<sup>5</sup>; and an analysis of company documents obtained through litigation found that paroxetine was ineffective and seriously harmful in children, in contrast with the claims in the journal article reporting the trial.<sup>6</sup> Furthermore, a recent study found that clinical study reports were more complete in their reporting of outcomes than published articles and trial registries combined.<sup>7</sup> It should be

Correspondence to: E Maund [em@cochrane.dk](mailto:em@cochrane.dk)

Extra material supplied by the author (see <http://www.bmj.com/content/348/bmj.g3510?tab=related#webextra>)

**Appendix 1:** Supplementary figure

**Appendix 2:** Supplementary table

noted, however, that clinical study reports can also be subject to biased reporting.<sup>8,9</sup>

Since 1995, clinical study reports submitted to the regulatory authorities in Europe, the United States, and Japan are expected to follow the International Conference of Harmonisation (ICH) E3 guideline.<sup>10</sup> These reports can be thousands of pages in length and include detailed information on efficacy and harms in various formats (see box). For example, data on harms can be presented in summary tables; both narratives and line listings (with data for each adverse event listed in a separate row) can provide information on serious adverse events, discontinuations because of adverse events and non-serious clinically relevant adverse events; and there can be individual patient listings of all adverse events and pre-existing medical conditions.

Until recently, independent researchers had limited access to these reports. From 2010 until legal action by two pharmaceutical companies in April 2013, however, the European Medicines Agency (EMA) released nearly two million pages of documents to academics, media, legal entities, and the pharmaceutical industry.<sup>11-13</sup>

In May 2011, before the EMA began to limit access to clinical study reports, we obtained such reports for the nine placebo controlled trials submitted in the marketing authorisation application of duloxetine for the treatment of major depressive disorder.<sup>14</sup> These reports, which included protocols as appendices, comprised 47 non-searchable pdf documents totalling 13 729 pages with no redactions. These documents were obtained as part of a wider request of access to reports on selective serotonin reuptake inhibitors (SSRIs) and serotonin norepinephrine reuptake inhibitors (SNRIs). While there were no redactions within the reports, they were incomplete because certain appendices were missing for all trials (table 1).

Duloxetine was the only centrally approved product (whereby a single application to the European Medicines Agency can lead to an EU-wide marketing authorisation for a drug),<sup>15</sup> which is why we focused on this drug.

We determined inconsistencies between protocols, clinical study reports, and publicly available sources, and within clinical study reports themselves, with respect to the primary efficacy analysis and major harms (deaths (including suicides), suicide attempts, serious adverse events, and discontinuations because of adverse events).

## Methods

We assessed clinical study reports, including protocols, and the main sources of publicly available data (published journal articles describing a single trial only and results posted on trial registries) of the nine randomised placebo controlled trials of duloxetine to determine whether there was evidence of inconsistencies in the primary efficacy analysis between protocols and clinical study reports; inconsistencies in the primary efficacy analysis and data on harms within clinical study reports; publication bias; and inconsistencies and incomplete reporting of the primary efficacy analysis and data on harms between the clinical study report and publicly available sources.

One researcher made the 47 pdfs, comprising the nine clinical study reports, searchable using optical character recognition software. Adobe Acrobat was used for all text portions. ABBYY Finereader was used to enable the efficient conversion of tables of harms into Excel spreadsheets; according to its manufacturer this software has an accuracy rate of 99.8%.<sup>16</sup>

For each of the nine trials, we identified journal articles that described a single trial only and not several trials or pooled analyses of two or more trials. We searched PubMed (final search 5 February 2013) and Cochrane Central Register of Controlled Trials (final search 12 March 2013) and contacted the manufacturer (Eli Lilly). One researcher (EM) identified relevant trials based on study ID, indication, sample size, study duration, and dose groups. When there was doubt as to whether a paper should be included, consensus was sought with a second researcher (BT).

One researcher (EM) searched for trial results on Clinicaltrials.gov (<http://clinicaltrials.gov/>). We obtained a pdf of trial registry reports for duloxetine from the manufacturer because we could not open the relevant links in their clinical trial registry website ([www.lillytrials.com/](http://www.lillytrials.com/)). We were interested in data on primary efficacy and major harms as they are especially pertinent for assessment of the efficacy and safety of the drug. The data of interest we specified a priori in our protocol were:

- Primary efficacy analysis:
  - Scale, effect size (group means/medians or differences)
  - Measure of precision or variability (confidence intervals, standard deviation, or standard error; interquartile range or other range for medians; precise P value)
  - Time point, type of analysis, and analysis population (for example, intention to treat, per protocol)
- Major harms (for each phase of the trial, such as randomised phase and placebo lead-out phase), number of patients and events in each arm: deaths (including suicides), attempted suicides, serious adverse events, and discontinuations because of adverse events.

Before data extraction, we chose treatment emergent adverse events (adverse events that emerged or worsened after study drug was started), and adverse events that emerged on discontinuation of study drug as additional harms of interest.

One set of independent observers used a two step data extraction process to extract data on these outcomes from protocols and clinical study reports (see appendix 1) and from published articles by a second set of observers. Data from trial registry reports were extracted by one observer and checked by a second. Any discrepancies were resolved by discussion and referral to the source documents, within each set of observers. A third opinion was sought when necessary.

For each trial, we compared data for each outcome between the protocol and the clinical study report, within the clinical study report, and between the clinical study report and publicly available data (journal trial report or trial registry report, or both), for consistency and, when applicable, completeness of reporting.

One researcher (EM) assessed completeness of reporting. The primary efficacy analysis was considered to be fully reported if scale, effect size for each group, measure of precision or variability, time point, type of analysis, and analysis population were provided, as described above. For major harms, the reporting of the number of patients and number of events was considered separately. In both instances, the phase (for example, randomised phase, placebo lead-out phase) and the number for each group needed to be reported. For treatment emergent adverse events, and for adverse events that emerged on discontinuation of study drug, the number of patients and number of events were also considered separately. The number of patients experiencing at least one treatment emergent adverse event, the number of patients experiencing at least one adverse

### Glossary of terms in clinical study reports

- Clinical study report (CSR): "A written description of a trial/study of any therapeutic, prophylactic, or diagnostic agent conducted in human subjects, in which the clinical and statistical description, presentations, and analyses are fully integrated into a single report"<sup>63</sup>
- ICH E3: ICH Guideline for Structure and Content of Clinical Study Reports
- Adverse event (AE): "Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have a causal relationship with this treatment"<sup>63</sup>
- Serious adverse event (SAE): "Any untoward medical occurrence that at any dose: results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect"<sup>63</sup>
- Narratives: In a CSR "There should be brief narratives describing each death, each other serious adverse event, and those of the other significant adverse events that are judged to be of special interest because of clinical importance. These narratives can be placed either in the text of the report or in section 14.3.3, depending on their number. Events that were clearly unrelated to the test drug/investigational product may be omitted or described very briefly. In general, the narrative should describe the following: the nature and intensity of event, the clinical course leading up to event, with an indication of timing relevant to test drug/investigational product administration; relevant laboratory measurements, whether the drug was stopped, and when; countermeasures; post mortem findings; investigator's opinion on causality, and sponsor's opinion on causality, if appropriate"<sup>10</sup>
- Appendices: CSRs include appendices on study information (for example, protocol and protocol amendments, sample case report forms, list of institutional review boards/ethics committees, list of investigators) and patient data listings (discontinued patients, protocol deviations, patients excluded from the efficacy analysis, individual efficacy response data, adverse event listings, individual laboratory measurements listings). Under directive 2001/83/EC and ICH E3, these appendices do not necessarily have to be submitted to the EMA as part of the regulatory submission for marketing authorisation, but the sponsor must make these available to the EMA upon request. The "note for guidance on the inclusion of appendices to clinical study reports in marketing authorisation applications" lists the appendices required to be submitted to the EMA with each CSR. These appendices include the protocol and protocol amendments<sup>10 34 35</sup>
- Individual patient adverse event listings: All adverse events for each patient, including the same event on several occasions, should be available as an appendix of the CSR. ICH E3 suggests the variables, such as patient identifier, the adverse event (preferred term and reported term), duration of the adverse event, severity (for example, mild, moderate, severe), seriousness (serious/non-serious), action taken (none, dose reduced, treatment stopped, etc), and outcome, that should be included in the listing<sup>10</sup>

event that emerged on discontinuation, and the frequency of each named event needed to be reported. Outcomes were considered incompletely reported if any of the aforementioned elements was missing, if only adverse events that met a threshold (such as an incidence of  $\geq 5\%$ ) were reported, or if only a qualitative statement was provided. If there were no data, either qualitative or quantitative, the outcome was considered to be unreported. We define publication bias as preferential publication of trial reports with positive findings.

## Results

The nine trials we included were all placebo controlled. To avoid confusion, we have called them trials 1 to 9, as their official names were similar (table 1). A total of 2878 patients entered the trials; the largest one (trial 9) recruited 533 patients.

The randomised phase lasted eight to nine weeks in all trials, apart from one in which it lasted 26 weeks (trial 9). Trials 1-6 had a one week placebo lead-in phase before randomisation, trials 7 and 8 had no lead-in, and in trial 9 all patients were openly treated with duloxetine in a 12 week lead-in phase. In trials 5 and 6, those who responded continued taking the randomised treatment for another 26 weeks. The dates of protocols, including any amendments, for eight trials were before the reported date of enrolment of the first patient (table 1). For trial 9 there was one minor protocol amendment (an additional telephone call) made after the first patient was enrolled.

Our searches of the published literature identified 1578 unique references. For six trials, we found one journal article for each trial; for trial 9, we found three articles that reported on the different phases of this trial, and for trials 2 and 3, we found no journal articles.

Only trial 9 was registered on [www.clinicaltrials.gov](http://www.clinicaltrials.gov), but no results were posted. All trials had a report on Lilly's publicly available clinical trial registry, and, with one exception, these reports had a later approval date than the publication date of the journal article reporting the trial.

The table 1 shows the characteristics of the clinical study reports (mean length 1525 pages), journal articles (mean 10 pages), and Lilly trial registry reports (mean 33 pages).

## Inconsistencies between protocols and clinical study reports

All protocols specified only one primary outcome and type of analysis. For two trials (trials 1 and 2), the protocols did not provide information on the analysis population (for example, intention to treat, subgroup, per protocol) but stated that this information was described in appendices, which, however, were not in the possession of the EMA.

The primary outcome and type of analysis used were consistent between protocols and clinical study reports. For one of seven trials with information in the protocol on analysis population, there were inconsistencies between the protocol (subgroup of patients with negative results on urine drug screen) and the clinical study report, which itself was internally inconsistent (intention to treat versus all randomly assigned patients with at least one follow-up after baseline). We considered these inconsistencies to be minor, however, as the number of patients differed by only 0.3% and 3%, respectively.

## Inconsistencies within clinical study reports

There were no inconsistencies in relation to the primary outcome (total score on 17 item Hamilton depression scale (HAM-D<sub>17</sub>) in eight trials) and analytical method used. All clinical study reports reported fully on the primary efficacy analysis and, for each phase of the trial, the harms of interest.

For harms there were no inconsistencies in suicides or attempted suicides (appendix 2). We did, however, find inconsistencies for seven serious adverse events and eight adverse events leading to discontinuation: adverse event listings from narratives, tables, and individual patient data were inconsistent with the safety conclusion as to whether the events occurred in the randomised phase of treatment or in the subsequent placebo lead-out; narratives clearly describing serious adverse events that began before randomisation and did not worsen in severity—that is, did not meet the clinical study report defined criteria for a treatment emergent adverse event but were listed in tables as if they had occurred while the patients received study drug; and events that appeared in summary tables and narratives but were

missing from the relevant line listings. There was no bias in these inconsistencies.

### Publication bias and inconsistencies between clinical study reports and publicly available sources

Six trials had significant results for the primary efficacy analysis specified in the protocol as defined in the clinical study report, and these trials' results were each published as a trial report reporting significant results in a journal article.<sup>17-25</sup> As noted above, two of the nine trials (trials 2 and 3) were not published and both had non-significant results for the primary efficacy analysis. The third trial (trial 1) had a non-significant result according to the clinical study report but significant results according to the journal article.<sup>17</sup> Its significant result for the primary efficacy analysis was based on patients with post-baseline efficacy data, whereas the result in the clinical study report was based on those patients who had a decrease in HAM-D<sub>17</sub> total score of at least 30% in the one week placebo lead-in phase plus a score of at least 14 at randomisation plus at least one score after randomisation. Furthermore, the analytical method used in the article (likelihood based mixed models repeated measures approach) was added after completion of the protocol, but the journal article did not mention that the analysis it presented was not the primary efficacy analysis specified in the protocol.

In regard to harms, we found inconsistencies between clinical study reports and journal articles for two trials. For trial 5, the journal article reported on only one serious adverse event in the randomised phase, but the clinical study report stated that two serious adverse events occurred in one patient taking paroxetine. For trial 8, the article reported that four patients in the placebo group discontinued because of adverse events in the randomised phase, while the clinical study report stated it was six patients. The Lilly trial registry reports for both trials were consistent with the clinical study reports (see appendix 2).

### Reporting of harms in publicly available sources

Harms were generally poorly reported in journal articles and Lilly trial registry reports.

#### *Deaths, including suicides*

Five deaths (four in the duloxetine group and one in the placebo group) including three suicides (two in the duloxetine group and one in the placebo group) were reported in three clinical study reports. For two of the trials (trials 5 and 9), the journal articles accurately reported the deaths and suicides. There was no journal article for the third trial (trial 3), in which the clinical study report stated that a patient taking duloxetine died after cardiopulmonary arrest. The Lilly trial registry report provided this information but did not state which phase of the trial the death occurred in. As described above, the clinical study report was internally inconsistent as the death was reported to have occurred both in the randomised treatment phase and in the placebo lead-out phase.

For six trials, the clinical study reports stated that no deaths occurred but only two of the five journal articles stated that no patients had died. One of the two articles, however, reported only that no patients had died in the acute or continuation phases. Neither the article, nor the Lilly trial registry report, stated whether any patients had died in the placebo lead-out phase. The articles reporting three trials (trials 4, 7, and 8) did not mention whether there were any deaths in these trials. From the

Lilly trial registry report for trial 4 it was unclear if anyone had died as it stated "no patients died during this study" but this appears in a section entitled "Safety-acute therapy phase." The Lilly trial registry reports for trials 7 and 8, and for one trial with no journal article (trial 2), fully reported on deaths (see appendix 2).

#### *Attempted suicides*

Summary tables of treatment emergent adverse events reported four suicide attempts in trial 9, in the open lead-in duloxetine phase. Data from the individual patient listings of harms and narratives of the clinical study report showed that three of the suicide attempts were definitive and serious and led to the patients being withdrawn from the trial. The fourth suicide attempt was reported in the individual patient listings of harms as a "possible suicide attempt" and was reported as being neither serious nor leading to the patient being withdrawn from the trial. Only the three definitive suicide attempts were reported in two of three journal articles for this trial, as serious adverse events or as reasons for patients being withdrawn from the trial. There was no mention of suicide attempts in the Lilly trial registry report, either in the text or tables.

All Lilly trial registry reports we examined reported only adverse events that had a total incidence of at least 2%, and suicide attempts were below this threshold. We did not find reports of suicide attempts in the clinical study reports of the other eight trials (see appendix 2).

#### *Serious adverse events other than death*

Serious adverse events were mentioned in eight of nine clinical study reports (one trial mentioned there were none).

Three journal articles (trials 4, 7, and 8) did not report on the occurrence or non-occurrence of serious adverse events. Lilly trial registry reports for these three trials, and for a trial with no journal article (trial 2), correctly gave the number of patients in each arm who experienced serious adverse events but either did not report which phase the events occurred in or did not report, or were unclear, as to how many events there were.

None of the three journal articles for trial 9 reported the occurrence of serious adverse events for the randomised phase of the trial, and the Lilly trial registry report gave only a qualitative statement that there was no significant difference between the groups.

For the trial without a journal article (trial 3), the Lilly trial registry report mentioned a patient who wasn't randomised when the serious adverse event occurred and another who experienced an exacerbation of asthma while taking placebo. According to the clinical study report, however, the patient had the exacerbation of asthma before randomisation and it did not worsen during the trial, which means that the event cannot be considered a serious adverse event for analytical purposes (see appendix 2).

#### *Discontinuations because of adverse events*

All seven trials with a journal article reported the number of patients who discontinued because of an adverse event; for the two other trials (trials 2 and 3) it was reported in the Lilly trial registry reports. In trial 2, however, the data were not divided per group, as only the total number of patients was reported (see appendix 2).

## Treatment emergent adverse events

Treatment emergent adverse events were poorly reported. None of the journal articles reported the number of patients that experienced such events in the randomised phase. The journal article for the largest trial (trial 9) gave a qualitative statement that there was no significant difference in the rates between duloxetine and placebo, and the number of patients affected was given only for a post-randomisation so called rescue phase, when they are not particularly relevant. Six journal articles reported events in the randomised phase only if they met a specified threshold—for example, had an incidence of greater than 10% for duloxetine patients.

The Lilly trial registry reports for all nine trials provided the number of patients who had experienced at least one treatment emergent adverse event, but they reported only the number and types of adverse events that had a total incidence of at least 2% in each trial. The figure<sup>1</sup> shows the number of treatment emergent adverse events reported in clinical study reports, Lilly trial registry reports, and journal articles for the randomised phase of each trial. Because of the use of reporting thresholds, data on between a median of 406 (range 177-645) and 166 (100-241) treatment emergent adverse events in the randomised phase per trial were not reported in journal articles and Lilly trial registry reports, respectively.

## Adverse events that emerged on discontinuation

There was little information about what happened when the treatment was stopped according to plan. Only articles reporting the results of two trials provided numbers of adverse events that emerged on discontinuation in the treated groups, but reporting was incomplete (see appendix 2). For four other trials, the journal articles mentioned either a specific event, or the events were not quantified but reported in brief general terms. The article for another trial (trial 8) did not report on such events at all.

The events were even less frequently reported in the Lilly trial registry reports. Information was available for only two of the nine trials (trial 3, for which there was no journal article, and trial 9), and the information was only qualitative.

## Discussion

### Principal findings

In this comparison of clinical study reports, trial registries, and publications on duloxetine for treatment of major depressive disorder, we found minor inconsistencies between the protocol and clinical study report and within the clinical study report for the primary efficacy analysis for one trial. More importantly, we found inconsistencies in the harms data within some of the clinical study reports—for example, serious adverse events were reported as if they had occurred during the randomised phase of the study though they started before randomisation, and events that were presented in tables were absent from line listings. There was, however, no apparent bias in these inconsistencies.

We found evidence of publication bias. All six trials with a significant result on the primary efficacy analysis specified in the protocol were published as a journal article. Two of the three trials with non-significant results were not published, and the third had significant results when published because of the use of a different analysis population and statistical method than those specified in the protocol.

Harms were generally poorly reported in both journal articles and in Lilly trial registry reports. In both formats, cut points for

the incidence of treatment emergent adverse events necessary to report them were arbitrary.

## Strengths and limitations of the study

The generalisability of our findings is unclear given that they are based on nine trials of a single drug from a single company. Another limitation of our study is that we looked only at inconsistencies in and completeness of reporting; we did not meta-analyse the clinical study report data to see whether the results were different from meta-analyses based on publicly available data only, but we plan to do this in a larger sample of trials.

## Comparisons with other studies

We do not know whether Lilly's failure to publish certain trials as trial reports in journals was because of non-submission of manuscripts or rejection by editors or whether the reason for incomplete data in publications was constraints on word count by journals. We do note, however, that previous research has shown that publication rates for submitted manuscripts with non-significant results are similar to those with significant results.<sup>26</sup> Furthermore, we studied only the primary efficacy analysis and the major harms, and there is no valid excuse for not publishing results for these outcomes.

Our findings of biased publications agree with previous studies that have compared publications with clinical study reports or other types of comprehensive data sources, including those studies that focused on depression.<sup>3-5</sup> Furthermore, the poor or missing reporting even of serious adverse events in journal articles is in line with the findings of a recent study of Medtronic's bone implant for spinal fusion.<sup>27</sup>

## Conclusions and implications for clinicians and researchers

Prescribers and patients need all the pertinent information on benefits and harms of a treatment, including information on any effects of withdrawal, to make an informed decision about treatment. It has been known for many years that serious discontinuation symptoms can occur on withdrawal from tricyclic antidepressants, monamine oxidase inhibitors, and SSRIs, including psychiatric symptoms that can be misdiagnosed as a recurrence of depression.<sup>28</sup> The lack of reporting of data on adverse events that emerge on discontinuation in journal articles and also in trial registry reports of an antidepressant was therefore disappointing. The use of reporting thresholds in the reporting of treatment emergent adverse events in journal articles and trial registries is problematic as important, but rare, events such as suicidal thoughts, behaviour, and attempts usually fall below the threshold.

Data on harms were often incompletely reported or were absent from publicly available sources but were fully reported in clinical study reports. Our findings support the view that journal articles are not an appropriate format to disseminate the results of clinical trials. Instead of publishing trials, journals could concentrate on discussing their merits and implications.<sup>29</sup> Furthermore, the incomplete reporting of harms in trial registry reports highlights that access to these reports is not an adequate alternative to access to clinical study reports. Clinical study reports should therefore be the primary data source for systematic reviews of drugs. This requires public access to these documents. Recently the committee of representatives from every EU member state government agreed with the text of the Clinical Trials Regulation, which includes the proposal of a publicly accessible EU database, set up and run by the EMA,

containing clinical study reports for new trials, when applicable, starting from 2014, used in a marketing authorisation request.<sup>30</sup> Furthermore, the UK Public Accounts Committee has recently recommended that National Institute for Health and Care Excellence (NICE) should ensure that it obtains full methods and results on all trials for all treatments that it reviews, including clinical study reports when necessary, and that it makes all this information available to the medical and academic community for independent scrutiny.<sup>31</sup>

As we found inconsistencies between protocols and clinical study reports, and even between different summaries and tabulations of harms data within clinical study reports, clinical study reports should be checked against protocols and within themselves for accuracy and consistency. Furthermore, clinical study reports are extremely lengthy documents and represent a considerable challenge to researchers. There is a need to develop tools and methodological approaches that will reduce the workload and still allow researchers to use them in an accurate and efficient manner.<sup>32</sup>

In conclusion, we found that clinical study reports contained extensive data on major harms that were not available in journal articles and in trial registry reports. There were minor inconsistencies in primary efficacy analysis population between protocols and clinical study reports and within clinical study reports. There were also inconsistencies between different summaries and tabulations of harms data within clinical study reports. Clinical study reports should be used as the data source for systematic reviews of drugs, but they should first be checked against protocols and within themselves for accuracy and consistency.

We thank Julie Borring, Kristine Rasmussen, Trine Gro Saida, and Louise Schow Jensen for assistance with data extraction; Eli Lilly Medical Information Department for their response to a request for a list of publications for the nine trials and providing the pdf of duloxetine Lilly trial registry reports; the EMA for providing the material and for responding to queries relating to the material; and Jesper Krogh for sharing material he obtained from the EMA.

**Contributors:** All authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. EM, BT, and PCG contributed to the study concept and design. EM, BT, AH, KJ, AL, and JS contributed to the acquisition of data. All authors contributed to the analysis and interpretation of data, and drafts of manuscripts. All the authors critically reviewed the manuscript for publication. PCG provided administrative, technical, and material support, and was the study supervisor. PCG is guarantor.

**Funding:** This study is part of a PhD (EM) funded by Rigshospitalets Forskningsudvalg. The funding source had no role in the design and conduct of the study; data collection, management, analysis, and interpretation; preparation, review, and approval of the manuscript; or the decision to submit the paper for publication.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

**Ethical approval:** Not required.

**Transparency declaration:** the manuscript's guarantor affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

**Data sharing:** The clinical study reports we used can be obtained from us.

- Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev* 2007;2:MR000005.
- Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for the selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457-65.
- Melander H. Evidence based medicine—selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 2003;326:1171-3.
- Turner EH, Mathews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;358:252.
- Eyding D, Lelgermann M, Grouven U, Harter M, Krompt M, Kaiser T, et al. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. *BMJ* 2010;341:c4737.
- Juriedini JN, McHenry LB, Mansfield PR. Clinical trials and drug promotion: selective reporting of study 329. *Int J Risk Safety Med* 2008;20:73-81.
- Wieseler B, Wolfram N, McGauran N, Kerekes MF, Vervögyi V, Kohlepp P, et al. Completeness of reporting of patient-relevant clinical trial outcomes: comparison of unpublished clinical study reports with publicly available data. *PLoS Med* 2013;10:e1001526.
- Chan AW. Bias, spin, and misreporting: time for full access to trial protocols and results. *PLoS Med* 2008;5:e230.
- Jefferson T, Doshi P, Thompson M, Heneghan C. Ensuring safe and effective drugs: who can do what it takes? *BMJ* 2011;342:c7258.
- Structure and content of clinical study reports: E3. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1995 [www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Efficacy/E3/E3\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf).
- Gøtzsche PC, Jørgensen AW. Opening up data at the European Medicines Agency. *BMJ* 2011;342:d2686.
- Doshi P. Transparency interrupted: the curtailment of the European Medicines Agency's policy on access to documents. *JAMA Intern Med* 2013;173:2009-11.
- Dyer C. European drug agency's attempts to improve transparency stalled by legal action from two US drug companies. *BMJ* 2013;346:f3588.
- Scientific Discussion EMA 2005. [www.ema.europa.eu/docs/en\\_GB/document\\_library/EPAR\\_Scientific\\_Discussion/human/000572/WC500036776.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_Scientific_Discussion/human/000572/WC500036776.pdf).
- Central authorisation of medicines. [www.ema.europa.eu/ema/index.jsp?curl=pages/about\\_us/general/general\\_content\\_000109.jsp](http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000109.jsp).
- ABBY FineReader 11 professional edition—full feature list. [http://finereader.abbyy.com/professional/full\\_feature\\_list/](http://finereader.abbyy.com/professional/full_feature_list/).
- Goldstein DJ, Mallinckrodt C, Lu Y, Demitrack MA. Duloxetine in the treatment of major depressive disorder: a double-blind clinical trial. *J Clin Psychiatry* 2002;63:225-31.
- Goldstein DJ, Lu Y, Detke MJ, Wiltse C, Mallinckrodt C, Demitrack MA. Duloxetine in the treatment of depression: a double-blind placebo-controlled comparison with paroxetine. *J Clin Psychopharmacol* 2004;24:389-99.
- Detke MJ, Wiltse CG, Mallinckrodt CH, McNamara RK, Demitrack MA, Bitter I. Duloxetine in the acute and long-term treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *Eur Neuropsychopharmacol* 2004;14:457-70.
- Perahia DG, Wang F, Mallinckrodt CH, Walker DJ, Detke MJ. Duloxetine in the treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *Eur Psychiatry* 2006;21:367-78.
- Detke MJ, Lu Y, Goldstein DJ, Hayes JR, Demitrack MA. Duloxetine, 60 mg once daily, for major depressive disorder: a randomized double-blind placebo-controlled trial. *J Clin Psychiatry* 2002;63:308-15.
- Detke MJ, Lu Y, Goldstein DJ, McNamara RK, Demitrack MA. Duloxetine 60 mg once daily dosing versus placebo in the acute treatment of major depression. *J Psychiatr Res* 2002;36:383-90.
- Fava M, Detke MJ, Balestrieri M, Wang F, Raskin J, Perahia D. Management of depression relapse: re-initiation of duloxetine treatment or dose increase. *J Psychiatr Res* 2006;40:328-36.
- Hudson JI, Perahia DG, Gilaberte I, Wang F, Watkin JG, Detke MJ. Duloxetine in the treatment of major depressive disorder: an open-label study. *BMC Psychiatry* 2007;7:43.
- Perahia DG, Gilaberte I, Wang F, Wiltse CG, Huckins SA, Clemens JW, et al. Duloxetine in the prevention of relapse of major depressive disorder: double-blind placebo-controlled study. *Br J Psychiatry* 2006;188:346-53.
- Lee KP, Boyd EA, Holroyd-Leduc JM, Bacchetti P, Bero LA. Predictors of publication: characteristics of submitted manuscripts associated with acceptance at major biomedical journals. *Med J Aust* 2006;184:621-6.
- Rodgers MA, Brown JV, Heirs MK, Higgins JP, Mannion RJ, Simmonds MC, et al. Reporting of industry funded study outcome data: comparison of confidential and published data on the safety and effectiveness of rhBMP-2 for spinal fusion. *BMJ* 2013;346:f3981.
- Haddad PM, Anderson IM. Recognising and managing antidepressant discontinuation symptoms. *Adv Psychiatr Treat* 2007;13:447-57.
- Smith R. Medical journals are an extension of the marketing arm of pharmaceutical companies. *PLoS Med* 2005;2:e138.
- Agreement reached in Europe on Clinical Trials Regulation. [www.alltrials.net/2013/agreement-reached-europe-clinical-trials-regulation/](http://www.alltrials.net/2013/agreement-reached-europe-clinical-trials-regulation/).
- Access to clinical trial information and the stockpiling of Tamiflu. Thirty-fifth report of session 2013-4. House of Commons Committee of Public Accounts.
- Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. *Cochrane Database Syst Rev* 2012;1:CD008965.
- Guideline for good clinical practice E6 (R1). International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2002. [www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Efficacy/E6\\_R1/Step4/E6\\_R1\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6_R1/Step4/E6_R1_Guideline.pdf).
- Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ.L:2001:311:0067:0067:EN:PDF>.
- Committee for medicinal products for human use (CHMP). Note for guidance on the inclusion of appendices to clinical study reports in marketing authorisation applications

**What is already known on this topic**

On average, meta-analyses of randomised clinical trials based on published articles overestimate the benefits and underestimate the harms of drugs, including antidepressants

A more reliable source of data for meta-analyses are clinical study reports—detailed reports on the design, conduct, and results of clinical trials that are submitted in marketing authorisation applications to the regulatory authorities

**What this study adds**

There can be inconsistencies in harms between different summaries and tabulations of harms data within clinical study reports

Authors of systematic reviews should check clinical study reports for accuracy and consistency whenever possible

(CHMP/EWP/2998/03 Final). [www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500003638.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003638.pdf).

**Accepted:** 05 May 2014

Cite this as: *BMJ* 2014;348:g3510

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

## Table

**Table 1 | Characteristics of clinical study reports, protocols, and publicly available data sources**

Trial No used in our study	Original Trial ID	Date of protocol	Dates of protocol amendments	Date first patient enrolled	Date CSR approved	Appendices missing from CSR	Date of single trial journal article	Approval date of Lilly trial registry report	No of pages		
									CSRs	Journal article	Lilly trial registry report
1	HMAQa	20 Oct '98	4 Feb '99	11 Feb '99	2 Oct '01	Yes*	Mar '02 <sup>17</sup>	16 Nov '04	951	7	31
2	HMAQb	20 Oct '98	4 Feb '99	10 Mar '99	2 Oct '01	Yes*	No publication	16 Nov '04	859	NA	26
3	HMATa	16 Dec '99	None	10 Mar '00	17 Sep '01	Yes*	No publication	16 Nov '04	1221	NA	24
4	HMATb	16 Dec '99	None	9 Mar '00	17 Sep '01	Yes*	Aug '04 <sup>18</sup>	16 Nov '04	1196	11	23
5	HMAYa	15 Feb '00	30 Mar '00	16 Nov '00	15 Jan '03	Yes†	2004 <sup>19</sup>	23 Jun '05	2211	14	71
6	HMAyb	15 Feb '00	30 Mar '00	31 Oct '00	15 Jan '03	Yes‡	May '06 <sup>20</sup>	27 Jul '06	2095	12	33
7	HMBHa	19 Jul '00	None	09 Nov '00	27 Jul '01	Yes†	Apr '02 <sup>21</sup>	15 Nov '04	916	8	20
8	HMBHb	19 Jul '00	None	13 Nov '00	6 Aug '01	Yes*	2002 <sup>22</sup>	23 Nov '04	911	8	21
9	HMBC	25 Sep '01	3 Oct '01, 20 Dec '01, 27 Feb '02, 12 Jun '02§	11 Mar '02	23 Sep '03	Yes*	Perahia, <sup>25</sup> ¶ Fava, <sup>23</sup> **Hudson <sup>24</sup> ††	17 Jul '06	3369	Perahia <sup>25</sup> 8; Fava <sup>23</sup> 9, Hudson <sup>24</sup> 11	50

CSR=clinical study report; NA=not applicable.

\*Sample case report forms; list of investigators (names and addresses provided but additional pages missing); ethics review board documents; listing of patients receiving treatment from specific batches; audit certificates; documentation of laboratory methods; additional statistical methods; individual patient listings other than adverse events.

†All appendices listed in \* and additionally appendix of errors to locked database.

‡All appendices listed in \* with exception of listing of patients receiving treatment from specific batches, and documentation of laboratory methods.

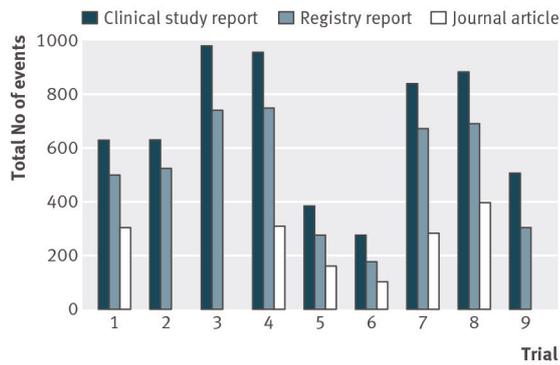
§Additional telephone call to patients from one investigative site.

¶Reported on continuation (randomised) phase, and additionally on some adverse events in open label duloxetine lead-in phase, and in follow up phase.

\*\*Reported on data from rescue phase only.

††Reported on open label duloxetine lead-in phase only.

## Figure



Registry reports were from manufacturer's (Eli Lilly) clinical trial registry; journal articles were those describing single trial only. Trials 2 and 3 had no journal article. TEAEs in trial 9 were reported qualitatively in relevant journal article

**Total number of treatment emergent adverse events (TEAEs) reported in randomised phase in different sources of trial data**