FACULTY OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

Rigshospitalet

PhD thesis

Emma Maund

The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviour

Academic advisor: Peter C.Gøtzsche

Submitted: 28/06/2015. *This thesis has been submitted to the Graduate School at the Faculty of Health and Medical Sciences, University of Copenhagen*

Institutnavn:	Det Nordiske Cochrane Centre
Name of department:	The Nordic Cochrane Centre
Author:	Emma Maund
Titel og evt. undertitel:	Effekten af selektive serotoningenoptagshæmmere (SSRI) på suicidalitet og voldelig adfærd
Title / Subtitle:	The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviour
Subject description:	This thesis investigates whether SSRIs are associated with an increased risk of suicidality and violent behaviour, with a focus on the factors and limitations that need to be taken into consideration when using clinical study reports as a data source for the assessment of harms.
.	
Academic advisor:	Peter C. Gøtzsche
Submitted:	28 Juni 2015

Contents

Preface and acknowledgements

Preface

This thesis is based on work performed at the Nordic Cochrane Centre whilst I was in receipt of a clinical assistant scholarship from Rigshospitalet. This thesis is a synopsis based on the following three papers:

1. Maund E, Tendal B, Hróbjartsson A, Jørgensen KJ, Lundh A, Schroll J et al. Benefits and harms in clinical trials of duloxetine for treatment of major depressive disorder: comparison of clinical study reports, trial registries, and publications. *BMJ* 2014;348:g3510.

2. Maund E, Tendal B, Hróbjartsson A, Lundh A, Gøtzsche PC. Coding of adverse events of suicidality in clinical study reports of duloxetine for the treatment of major depressive disorder: descriptive study. *BMJ* 2014;348:g3555.

3. Maund E, Schow Jensen L, Gøtzsche PC. Benefits and suicidality related harms of duloxetine for the treatment of stress urinary incontinence: a meta-analysis of clinical study reports. Submitted.

Acknowledgements

I would like to thank the following people: my supervisor, Peter Gøtzsche, for giving me the opportunity to do a PhD at the Nordic Cochrane Centre; my co-supervisor, Britta Tendal; my co-authors: Asbjørn Hróbjartsson, Karsten Juhl Jørgensen, Andreas Lundh, Jeppe Schroll and Louise Schow Jensen; and my colleagues at the Nordic Cochrane Centre. Finally, I would like to thank my family for all their encouragement and support.

Financial support

This PhD was funded wholly by Rigshospitalets Forskningsudvalg.

English summary

It is unclear whether selective serotonin reuptake inhibitors (SSRIs) can cause suicidality and violent behavior in adults. This is mainly due to issues with coding and incorrect assignment of adverse events within trials of SSRIs, and dissemination bias and poor reporting of SSRI trial results. Clinical study reports are drug regulatory documents with detailed data on efficacy and harms, including individual patient data. These reports could be used to assess the effect of SSRIs on suicidality and violent behaviour. There is, however, very little experience of using clinical study reports in the independent research community. The limitations of using these reports are therefore unknown.

The aim of this PhD was to investigate whether SSRIs are associated with an increased risk of suicidality and violent behaviour, with a focus on the factors and limitations that need to be taken into consideration when using clinical study reports as a data source for the assessment of harms. Three studies, all using clinical study reports of duloxetine, were conducted in order to fulfil this aim. The first study was to determine if there are inconsistencies between protocols, clinical study reports, and main publicly available sources (journal articles and trial registries), and inconsistencies within clinical study reports themselves, with respect to benefits and major harms. The second study was to assess the effects of coding and coding conventions on summaries and tabulations of adverse events data on suicidality within clinical study reports. The third study was to assess the benefits of duloxetine for stress urinary incontinence, and its harms in terms of drug-induced depression, suicidality, violent behaviour and their potential precursors.

The results showed that clinical study reports contained considerably more data on harms than journal articles and trial registry reports, but that there are inconsistencies within clinical study reports in data on harms of a drug. In summary tables of adverse events data, events of suicidality were obscured due to the coding dictionary and coding conventions used. Compared to placebo, duloxetine has a statistically significant, but likely clinically insignificant, beneficial effect in the treatment of stress urinary incontinence, and an increased risk of important psychiatric harms that are possible precursors to suicidality. Evaluation of these specific harms was only possible with the use of adverse event data for individual patients, which were only contained in the appendices of clinical study reports.

My conclusions are that given the inadequacies of journal articles and trial registry reports, clinical study reports should be used as the primary data source for systematic reviews. They should, however, first be checked against protocols and within themselves for accuracy and consistency. Furthermore, due to coding dictionaries and coding conventions used, adverse events data presented in summary tables may obscure adverse events of importance. Given the possible lack of benefit but increased risk of important harms, the rationale for using duloxetine for stress urinary incontinence is questionable. The evaluation of important harms can be impossible using only summary data provided in clinical study reports. Individual patient data of adverse events contained in appendices of clinical study reports are therefore essential for a reliable assessment of drug harms

Danish summary

Det er uklart, om selektive serotoningenoptagshæmmere (SSRI) kan forårsage suicidalitet og voldelig adfærd hos voksne. Dette skyldes primært problemer med kodning og fejlagtig angivelse af skadevirkninger i forsøg med SSRI-præparater, samt rapporteringsbias og dårlig beskrivelse af forsøgsresultater. Kliniske studierapporter er dokumenter fra lægemiddelgodkendende myndigheder med detaljerede data om gavnlige og skadelige virkninger, herunder individuelle patientdata. Disse rapporter kan bruges til at vurdere effekten af SSRI-præparater på suicidalitet og voldelig adfærd. Der er dog meget lidt erfaring med at bruge kliniske studierapporter blandt uafhængige forskere. Begrænsningerne ved at anvende disse rapporter er derfor ukendt.

Formålet med denne ph.d. var at vurdere, hvorvidt SSRI-præparater er forbundet med en øget risiko for selvmord og voldelig adfærd, med fokus på de faktorer og begrænsninger, der skal tages i betragtning ved brug af kliniske undersøgelsesrapporter som datakilde ved vurderingen af skadevirkninger. Tre undersøgelser, der alle bruger kliniske studierapporter af duloxetin, blev udført for at nå dette mål. Det første studie skulle vurdere, om der er uoverensstemmelser mellem protokoller, kliniske studierapporter, og de vigtigste offentligt tilgængelige kilder (tidsskriftsartikler og forsøgsregistre), og uoverensstemmelser inden for de kliniske studierapporter selv, med hensyn til de gavnlige og skadelige virkninger. Det andet studie skulle vurdere effekten af kodning og kodningsstandarder for resuméer og beregninger af data om skadevirkninger for suicidalitet i kliniske studierapporter. Det tredje studie skulle vurdere virkningen af duloxetin ved stressinkontinens, samt skadevirkninger i form af lægemiddelinduceret depression, selvmord, voldelig adfærd og deres eventuelle forstadier.

Resultaterne viste, at kliniske studierapporter indeholdt betydeligt flere data om skadevirkninger end tidsskriftsartikler og registre over studier, men at der er uoverensstemmelser inden for de enkelte kliniske studierapporter når det gælder data om skadevirkninger af et lægemiddel. I oversigtstabeller over skadevirkninger blev tilfælde af suicidalitet sløret på grund af kodningsordbøger og kodningsstandarder. Sammenlignet med placebo har duloxetin en statistisk signifikant, men sandsynligvis klinisk insignifikant, gavnlig virkning ved behandling af stressinkontinens, og en øget risiko for vigtige psykiatriske skader, der er mulige forstadier til suicidalitet. Evaluering af disse specifikke skadevirkninger var kun mulig med brug af data for de enkelte patienter, som kun fandtes i bilagene til de kliniske studierapporter.

Mine konklusioner er, at i betragtning af manglerne i tidsskriftsartikler og rapporter fra forsøgsregistre, bør de kliniske studierapporter bruges som den primære datakilde for systematiske reviews. De skal dog først kontrolleres i forhold til studiernes protokoller og for nøjagtighed og intern konsistens. På grund af kodningsordbøger og kodningsstandarder, der anvendes, kan skadevirkninger præsenteret i oversigtstabeller skjule vigtige skadevirkninger. I betragtning af den mulige mangel på gavnlige virkninger, og øget risiko for vigtige skader, er rationalet for at bruge duloxetin til stressinkontinens tvivlsomt. Opgørelse af vigtige skadevirkninger kan være umulig, hvis man kun benytter summariske data fra kliniske studierapporter. Individuelle patientdata for skadevirkninger, der er indeholdt i bilag til de kliniske studierapporter, er derfor afgørende for en pålidelig vurdering af skadevirkninger ved medikamentel behandling

Introduction

Access to complete reports of clinical trials is essential for ethical, scientific and economic reasons.[1-3] Foremost, it would allow less biased systematic assessment of benefits and harms of pharmaceutical drugs to be performed. Provision of such assessments to patients, healthcare professionals and policymakers would allow them to make better informed decisions on receiving, prescribing or providing a drug respectively. This would have positive consequences for patient health and for the efficiency of the health care system.

Journal articles, the main publicly accessible format of reporting clinical trials, are the primary data source for systematic assessments of drugs. This is problematic due to dissemination bias and poor reporting.

Dissemination bias

Dissemination bias, especially publication bias and within study selective outcome reporting, threatens the validity of systematic assessments of drugs.[4] Publication bias arises when trials are published or not depending on their results, with trials with significant or positive results more likely to be published than those with non-significant or negative results. Evidence from systematic reviews of publication bias suggest that positive results are approximately twice as likely to be published as non-positive results.[4,5] The consequence of publication bias on meta-analyses including only published studies has been found to vary by drug and outcome, with benefits of psychiatric drugs being most likely to be overestimated.[6] In regard to harms, the evidence suggests meta-analyses including only published studies underestimate harms.[4,6,7]

Within-study selective reporting bias relates to studies that have been published. It has been defined as the selection on the basis of the results of a subset of the original variables recorded for inclusion in a publication.[8] This includes selective reporting of outcomes, types of analyses, endpoint data versus change data, timepoints or subgroups. A systematic review evaluating selective outcome reporting found statistically significant outcomes had between two to almost five times higher odds of being fully reported compared to non-significant outcomes, and that cohort studies assessing selective outcome reporting found 40 to 62% of studies had at least one primary outcome that was changed, introduced, or omitted.[8] A study of the impact of selective outcome reporting found that in a sample of meta-analyses with a statistically significant result for the primary outcome, 19%

became non-significant after adjustment for outcome reporting bias.[9] By comparing original trial documents to published journal articles, data for serious adverse events have been found to be omitted or incorrectly reported.[10,11]

Publication bias and selective reporting bias of trials are endemic throughout the medical literature, regardless of the type of intervention or indication for use.[12] These biases have a human and monetary cost. For example, a study performed prior to approval of Vioxx (rofecoxib) found the drug caused nearly a seven fold increase risk in heart attacks.[13] However, this study was never published.[14] Before its removal from the market, Vioxx is estimated to have caused heart attacks or strokes in 160 000 people, of which 30 to 40% probably died as a result.[13,15] The antiviral Tamiflu (oseltamivir), which has been on the market since 1999, is estimated to have generated sales in excess of 13 billion Euros for its manufacturer Roche.[16] A recently revised Cochrane review based on fully reported clinical trial data, however, has concluded that the benefits of Tamiflu do not outweigh its harms.[17]

Past initiatives to reduce these biases include trial registration, with summary protocol information, and posting of summary trial results on publicly accessible databases. There is, however, evidence that these measures have significant limitations.[18-24]

Poor reporting

Compounding the effects of publication bias and selective outcome reporting is the poor reporting of those trials that are published. Journal articles reporting trials should have complete, clear and transparent information on the trial's methodology, and its findings, both in terms of benefits and harms.[25] Despite the existence and use of reporting guidelines, data have shown that reporting, particularly of harms, remains suboptimal.[26-29]

The most recent initiative to overcome dissemination bias and poor reporting is allowing independent researchers access to those previously confidential drug regulatory document, which are known as "clinical study reports".

Clinical study reports

As a result of the thalidomide tragedy, since the 1960s drug regulators have required drug manufacturers to submit reports of trials pertinent to the safety and efficacy of a drug, whether the results are favourable or not, when applying for its approval for use.[30] In 1995, the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), a group made up of EU, USA, and Japanese regulatory authorities, pharmaceutical industry associations, and non-governmental organisations, approved guidelines for the content and form of these trial reports.[31,32] The purpose of these guidelines was to create one report that would be acceptable to EU, US and Japanese regulatory authorities. Trial reports meeting the requirements of these guidelines are known as "clinical study reports".

A clinical study report is a standardised account of the background, methods, and results, and conclusions of a clinical trial. These documents can be hundreds of pages in length and include significantly more detailed benefits and harms data than summary level trial results found in journal articles or summary results databases.[33] Furthermore, these documents contain appendices with important information, such as the trial protocol and any amendments, and listings of all adverse events for individual patients.

Until November 2010, trial reports and clinical study reports have remained confidential between the manufacturer and regulatory authority. There have, however, been a few exceptions, which have challenged published evidence on a given healthcare intervention or reversed conclusions based on publicly available information. For example, examination of a clinical study report, obtained through litigation, of a trial of the antidepressant paroxetine showed the drug was ineffective and seriously harmful in children, in contrast to the claims in the journal article reporting the trial.[10] Using clinical study reports obtained as part of the German Health Technology Assessment process, Eyding et al. found that, contrary to meta-analyses based on published data, the antidepressant reboxetine was no more effective than placebo but had a greater risk of harms.[7]

In 2010, after a three year fight with the European Medicines Agency (EMA), the Nordic Cochrane Centre gained access to clinical study reports held by the agency for two anti-obesity drugs.[34] This set a precedent allowing independent researchers access to clinical study reports from the EMA. Access was limited for a period, due to legal action by two pharmaceutical companies, however,

between November 2010 and April 2013, the EMA released over 1.9 million pages of documents in response to safety related requests.[35,36]

Clinical study reports, however, should not be viewed as a panacea to the problem of dissemination bias, as they come with their own limitations and potential biases. Firstly, clinical study reports are considerably larger and more detailed documents than journal articles and there is very limited experience among the independent research community in using these documents as sources of data for systematic assessment of benefits and harms of drugs.[37,38] Concerns have been expressed, by industry and industry conflicted academics, that clinical trial data contained in these documents could be re-analysed in an inexpert or irresponsible way.[39] Secondly, some researchers, whilst calling for access to full clinical study reports, have also highlighted that as some sections of the reports are prepared by the companies after data analysis, they may be subject to biased reporting.[40] Given that medical writers openly admit that the safety section of a clinical study report "presents an opportunity for spin", these concerns are not without foundation.[41 (p 6)]

Despite these limitations, the detailed data on benefits and harms in clinical study reports is an advantage over other publicly existing data sources as it would allow a considerably more detailed assessment of the harms of a drug. This is particularly important when the benefits of a drug are small and there is uncertainty around important harms, such as whether selective serotonin reuptake inhibitors induce suicidality and violence in adults.[42]

Selective serotonin reuptake inhibitors

The first antidepressant drug became available in the 1950s, after its serendipitous discovery. In 1952 doctors using a new anti-tuberculous drug, iproniazid, noticed that in addition to its anti-tuberculous effect it had a beneficial effect on mood. Soon this drug was being used to treat depression.[43] Subsequent investigations into how it exerted an antidepressant effect found that it inhibited the enzyme monoamine oxidase, consequently causing an increase of monoamine neurotransmitters, norepinephrine and serotonin, in the brain. [43-45] As a result, the "monoamine hypothesis" was put forward that depression is a result of a depletion of monoamine neurotransmitters.[43-45] There is, however, no logical foundation for the idea that the effect of a medication reveals the underlying cause of the disease it treats. Furthermore, evidence for the monoamine hypothesis is lacking, and in some cases contradictory.[46-48] Despite this, a depletion of

monoamines as the cause of depression has been presented as fact in journals, books, and advertising.[47,49]

Based on the monoamine hypothesis, antidepressants known as selective serotonin reuptake inhibitors (SSRIs) and serotonin and norepinephrine reuptake inhibitors (SNRIs) were developed (see table 1).

Table 1: SSRIs and SNRIs licensed for use in Denmark

Generic name	Brand name (Denmark)	Brand name (USA)
SSRIs		
Citalopram	Cipramil	Celexa
Escitalopram	Cipralex	Lexapro
Fluoxetine	Fontex	Prozac, Sarafem[1]
Fluvoxamine	*Fevarin*	Luvox
Paroxetine	Seroxat	Paxil
Sertraline	Zoloft	Zoloft
SNRIs		
Duloxetine	Cymbalta, Xeristar[2], Yentreve[3]	Cymbalta
Venlafaxine	Efexor	Effexor

Sources of data:
http://pro.medicin.dk/Laegemiddelgrupper/Grupper/243058; http://pro.medicin.dk/Laegemiddelgrupper/Grupper/315580
http://www.fda.gov/Drugs/DrugSafety/InformationbyDrugClass/ucm283587.htm

Footnote: [1] licensed for premenstrual dysphoric disorder; [2] licensed for generalised anxiety disorder; [3] licensed for stress urinary incontinence

These drugs first appeared on the market in the late 1980s and soon superseded pre-existing classes of antidepressants, the monamine oxidase inhibitors and tricyclic antidepressants. For the purpose of this introduction, SSRIs and SNRIs, which both inhibit the reuptake of serotonin, will be abbreviated as SSRIs in the following.

Usage of SSRIs

Licensed indications for SSRIs has increased over time and now include other psychiatric disorders, behaviour disorders, and even non-psychiatric disorders. For example, in Europe duloxetine is

approved for major depressive disorder, generalised anxiety disorder, diabetic peripheral neuropathic pain, and stress urinary incontinence.

Retrospective and prospective studies have revealed that many patients receiving antidepressants do not meet the diagnostic criteria for any psychiatric disorder.[50-52] For example, a recent prospective study by Takayanagi et al. found that 69% of patient receiving an antidepressant never met criteria for major depressive disorder and 38% never met criteria for any psychiatric disorder indicated for antidepressant use in their lifetime.[52] As only a small proportion of antidepressants are prescribed for non-psychiatric indications, this supports the view that antidepressants are being commonly prescribed for mild symptoms that do not meet the diagnostic criteria for indicated antidepressant use.[52]

Since the introduction of SSRIs, their use has increased dramatically. In Denmark, figures shows that in 2005 approximately 6% of the population could be in treatment with antidepressants, the majority of which were SSRIs.[53] In 2013 this figure had risen to 8%.[53] The increase in consumption of SSRIs over time is seen throughout the Nordic countries, the rest of Europe, the US, Australia and in developing countries.[53-56]

Benefits of SSRIs

Evidence from a meta-analysis that included unpublished data shows that compared to placebo, antidepressants, including SSRIs, have a statistically significant but small beneficial effect in the treatment of depression in adults (standardised mean difference (SMD) 0.31, 95% confidence interval 0.27 to 0.35).[57] Due to unblinding bias because of adverse events, however, even this small effect may be an *overestimate* of the beneficial effect of antidepressants. Studies have found shown that patients and health care professionals participating in randomised double blind placebo controlled trials of psychiatric drugs can determine, on the basis of adverse events, who is receiving active drug or placebo.[58] In order to control for this, an active rather than inert placebo can be used. An active placebo is a substance which does not have any therapeutic effect but mimics side effects of the active drug.[58,59] A systematic review comparing antidepressants to active placebo found a smaller beneficial effect of antidepressants (SMD 0.17, 0.00 to 0.34).[59] There is also evidence that compared to placebo, the benefit of antidepressants, including SSRIs, increases with severity of

depression symptoms. Benefits maybe minimal or nonexistent, on average, in patients with mild or moderate symptoms.[60]

Harms - SSRI-induced suicidality and violence

In 1990, shortly after SSRIs first appeared on the market, case reports of SSRI- induced suicidality (i.e. suicidal ideation, suicidal behaviour, suicide attempts, and suicide) and violence began to emerge.[61-63]

Mechanism of SSRI-induced suicidality and violence

A commonly ascribed mechanism for linking the use of antidepressants to suicide, is the "roll-back" phenomenon.[64] According to this theory, antidepressants have an early energizing effect before they have a mood lifting effect. As a result patients have the energy to follow through on suicidal impulses before their mood improves. This mechanism, however, neither accounts for suicidality observed in randomised controlled trials (RCTs) of SSRIs, given that these studies typically exclude patients experiencing suicidal ideation, nor does it explain instances of suicidality observed in healthy volunteer studies.[65-67]

Postulated iatrogenic mechanisms include akathisia (inner restlessness), emotional blunting (reduced emotional response), psychosis, and activation (e.g. insomnia, nervousness, anxiety, increased energy, agitation, akathisia, hypomania, and mania). Published evidence show an association between these mechanisms and suicidality and, in some cases violence; but these studies are arguably limited in terms of quantity and study design.[68-79] The US Food and Drug Administration's (FDA) product label for all antidepressants warns of symptoms of activation in *all* patients. Although the labelling states that a causal link between such symptoms and suicidality has not been established, it advises that patients, their families and their caregivers should be alert for such symptoms stating that: *'these may be associated with an increased risk for suicidal thinking and behavior and indicate a need for very close monitoring and possibly changes in the medication.'* [80 (p 4)]

SSRI-induced suicidality and violence in children and adolescents

In 2003, the FDA found instances of suicidality events in paediatric trials of antidepressants that had been coded by the pharmaceutical companies to both more severe terms and less severe terms.[81-83] Coding refers to the clinical trial process where events described by the original

investigators are coded to terms in a specialized, hierarchically structured, medical coding dictionary. When coding the original investigator reported term, the closest matching lowest level term in the dictionary is chosen. Synonymous lowest level terms are aggregated at the next level into preferred terms, so named because they are the preferred term for the statistical analyses and presentation of adverse event data for regulatory submissions.[84] As a result of its findings, the FDA commissioned an independent body to reclassify all events in antidepressant paediatric trials that could represent suicidality. This resulted in the development of the Columbia Classifications Algorithm of Suicide Assessment (C-CASA).[81-83,85]

Using the reclassified data, the FDA found that regardless of indication, antidepressants were associated with nearly a twofold increase in risk of suicidality compared to placebo (Risk ratio (RR) 1.95, 95% confidence interval 1.28 to 2.98). In trials of major depressive disorder, compared to placebo, SSRIs were associated with an increased risk of suicidality (RR 1.66, 1.02 to 2.68)[82] and agitation and hostility (RR 2.34, 1.24 to 4.41).[82,83]

The UK drug regulator also assessed the risk of suicidality and hostility in paediatric trials for SSRIs and reached similar conclusions.[86]

SSRI-induced suicidality and violence in adults

Case reports, analysis of spontaneous reports to regulatory agencies, and observational studies demonstrate an association between SSRIs and suicidality or violence, particularly shortly after initiating or discontinuing treatment.[67,69,87-95] Case reports and observational studies can allow the determination of whether a drug causes a specific harm.[96-99] The most reliable study design to determine causality, however, is the RCT.[100] Individual RCTs typically enroll a few hundred patients however, and therefore do not have a large enough sample size to adequately assess rare harms, such as suicidality.[99] For example, analysis by David Healy, using FDA data, showed suicide rates of 0.20% on antidepressants versus 0.04% on placebo.[63] A formal sample size calculation shows that if the placebo rate is 0.04% and a difference of 0.16% is not to be overlooked, with conventional levels for alpha=0.05 and beta =0.20, a trial would need 8600 patients in the antidepressant group and 8600 in the placebo group. In order to overcome this limitation, data from multiple trials can be combined using a statistical technique known as meta-analysis.[101]

Published meta-analyses performed by pharmaceutical companies in the 1990s found no difference in suicidal ideation between SSRIs and placebo.[102-104] The main limitation of these analyses were that none of the included trials were designed to test whether the drug was associated with suicidal ideation. Consequently, responses to suicide items from depression scales used to measure the efficacy of the drugs in the trials were used as measures of suicidal ideation. Suicide items from depression scales, however, are an insensitive measure of suicidal ideation.[105] Furthermore, a published pharmaceutical company analysis of suicidality in paroxetine trials used an incorrect number of suicide attempts in the placebo arm.[104] Evidence from FDA documents show that two suicide attempts assigned to the placebo arm in the analysis actually occurred during the placebo lead-in phase, i.e. *before* randomisation, and therefore should not have been counted.[106] Wrongful assignment of suicide attempts that occurred in the placebo lead-in phase to the placebo arm have also been found in sertraline data submitted by Pfizer to the FDA;[107] and the number of reported suicides in the placebo arms of paroxetine trials has included three events that actually occurred in the post-treatment phase.[108] In comparison to suicidality, SSRI-induced violence in adults has seldom been the subject of published studies. A published meta-analysis of fluoxetine trials performed by Eli Lilly, the manufacturer of fluoxetine, found the drug was not associated with an increased risk of violence or aggression.[109] The FDA, however, stated it was impossible to draw any conclusions from the analysis given the method of adverse event elicitation in the trials and the very small number of violent events, which suggested considerable under-reporting.[105]

Pooled analyses using summary trial data obtained from the FDA and from the Dutch drug regulator found no difference in suicide rates between SSRI and placebo.[110,111] However, the reliability of these analyses is uncertain due to methodological limitations of performing a pooled analysis rather than a meta-analysis.[112]

The three largest meta-analyses assessing whether SSRIs, compared to placebo, are associated with an increased risk of suicidality in adults have been performed by Fergusson et al., Gunnell et al., and the FDA.[64, 113-116]

Fergusson et al. conducted a systematic review of all published RCTs, regardless of indication, to evaluate the association between suicide attempts and use of SSRIs.[113] This review included 702 RCTs (87 650 patients), comparing SSRIs with placebo or other antidepressants. Four hundred and

eleven trials (18 413 patients) compared SSRI to placebo. SSRIs caused a two fold increase in odds of suicide attempt compared to placebo (odds ratio (OR) 2.28, 95% confidence interval 1.14 to 4.55). There was, however, no statistically significant difference in suicides (OR 0.95, 0.24 to 3.78). In addition to using published data only, a serious limitation of this analysis was the lack of information in 58% of publications as to whether or not any suicide attempts had occurred. On contacting a random sample of authors of trial publications where this was the case, Fergusson et al. were informed by some authors that data on suicide attempts were collected but not reported in publications, and by other authors that data on suicide attempts were not collected during the trials.

Gunnell et al. performed a meta-analysis on placebo controlled trial data provided by pharmaceutical companies to the UK drug regulator as part of its review on the safety of SSRIs.[114,115] This analysis included data from 477 RCTs (45 704 patients) and found no statistically significant difference between SSRI and placebo for suicidal ideation (OR 0.80; 95% credible interval 0.49 to 1.30), non-fatal self harm (OR 1.21, 0.87 to 1.83), or suicide (OR 0.85, 0.20 to 3.40). In all cases, however, a harmful effect of SSRIs could not be excluded. There are several limitations with this analysis. First, the authors of the analysis did not have access to trial level data. For each drug, pharmaceutical companies only provided the UK regulator with summed endpoint data across all trials and all indications for drug and placebo groups separately. Second, there was underreporting of self harm and of suicidal thoughts among the trials, which was apparent by the lower than expected ratio of suicides to self harm, and of self harm to suicidal thoughts. Third, it was uncertain whether the number of suicide attempts in the placebo group was correct.[117]

The largest analysis was carried out by the FDA using data it had requested from pharmaceutical companies specifically for this purpose. The analysis initially included 372 controlled trials of modern antidepressants (99 231 patients; 52 2960 receiving antidepressants of which 42 550 received a SSRI) for all indications.[64,116] Due to an insufficient number of suicidality events in trials for non-psychiatric indications the analysis was limited to trials of psychiatric indications only, and it was found that the risk of suicidality associated with antidepressant use was age dependent. The FDA concluded there was increased risk among young adults under 25, an overall neutral effect in adults aged 25-64, and a protective effect in adults aged 65 and over. Graphical data within the full FDA report, however, showed an increased risk of *suicidal behaviour* until 40 years of age.[64] The main limitation of the FDA analysis was that there was no independent reclassification of events.

Instead the FDA gave pharmaceutical companies the responsibility of both finding events that could potentially represent suicidality in their trial data and of reclassifying these events according to the Columbia Classifications Algorithm of Suicide Assessment.

Overall, all three meta-analyses have important limitations. Consequently, there is still considerable uncertainty as to whether SSRIs can cause suicidality in adults.

Objectives

The initial objectives of my PhD were:

1. To study the effects of SSRIs on suicidality, violent behaviour, activation syndrome symptoms, emotional blunting and disinhibition

2. To determine the degree of selective reporting of these harms by comparing the reporting of suicidality, violent behaviour, activation syndrome symptoms, emotional blunting and disinhibition, in publications with that in corresponding clinical study reports, and with what was planned according to corresponding trial protocols.

3. To assess the inter-observer variability on the coding of adverse events using the Medical Dictionary for Regulatory Activities (MedDRA).

This was to be achieved by performing systematic reviews of all trials, regardless of indication, of placebo controlled trials of SSRIs using clinical study reports submitted to the EMA (duloxetine), the Swedish Medicines Product Agency Board (citalopram, escitalopram, venlafaxine), and the Dutch Medicines Evaluation Board (fluoxetine, paroxetine, sertraline). However, we found we needed to modify these objectives for two reasons. First we could not gain access to all the data we needed. We found that the UK drug agency, the Medicines and Healthcare products Regulatory Agency (MHRA), and not the Dutch agency or the EMA held data for fluoxetine. However, the UK MHRA has destroyed the clinical study reports for adults.[118] The Swedish Medicine Products Agency stores its clinical study reports in an archive facility in Northern Sweden but was willing to transport the documents to its offices free of charge. It would, however, charge the Nordic Cochrane Centre the cost of copying the documents, which was prohibitively expensive (estimated €70 000). The Dutch Medicines Evaluation Board gave us access to clinical study reports of paroxetine and sertraline, but these documents were heavily redacted, and therefore unusable for our purposes.[119] Second, concerns have been raised by some that clinical study report data may be subject to biased reporting.[40] None of the research team were familiar with clinical study reports, and there is very limited experience of using clinical study reports in the independent research community.[37,38] It was therefore unclear what factors and limitations need to be taken into consideration when using clinical study reports as data sources for meta-analyses or systematic reviews.

The modified objectives were:

1. 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 inconsistencies within clinical study reports themselves, with respect to benefits and major harms.

2. To assess the effects of coding and coding conventions on summaries and tabulations of adverse events data on suicidality within clinical study reports.

3. To assess the benefits of duloxetine for stress urinary incontinence, and its harms in terms of depression, suicidality, violent behaviour and their potential precursors.

Description of the research project

Paper 1: comparison of protocols, clinical study reports, trial registries, and publications

In the first study we determined, 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 inconsistencies *within* clinical study reports themselves, with respect to benefits and major harms.

Research using clinical study reports of antidepressant trials have found publication bias and selective outcome reporting.[10,120] Furthermore, clinical study reports have been shown to contain significantly more data on benefits and harms than journal articles or trial registries.[33] This suggests that clinical study reports should be used as a data source for the assessment of benefits and harms of a drug. There are, however, concerns that clinical study reports themselves maybe subject to biased to reporting.[40]

We assessed clinical study reports, including protocols, (total 13 729 pages) of the nine main randomised placebo controlled trials used in the marketing authorisation of duloxetine for the treatment of major depressive disorder.[121] Journal articles were identified through relevant literature databases and contacting the manufacturer, Eli Lilly. Further, Clinicaltrials.gov and the manufacturer's online clinical trial registry were searched for trial results. Data on primary efficacy analysis and major harms (deaths (including suicides), suicide attempts, serious adverse events, and discontinuations because of adverse events) were extracted from each data source and compared.

Paper 2: coding of adverse events of suicidality

In the second study we assessed the effects of coding and coding conventions on summaries and tabulations of adverse events data on suicidality within clinical study reports of duloxetine for the treatment of major depressive disorder.

Coding of original trial investigator reported terms of adverse events is essential for the statistical analysis and evaluation of harms. Coding, however, is a little known and seldom studied clinical trial process. A recent systematic review of potential issues related to the coding of adverse events

in clinical trials, which I co-authored, found only 10 relevant studies. These dated from 1996 to 2008.[122]

We used clinical study reports, including protocols, (total 13 729 pages) of the nine main randomised placebo controlled trials used in the marketing authorisation of duloxetine for the treatment of major depressive disorder.[121] Systematic electronic searches for adverse events of suicidality were performed in tables, narratives, and listings of adverse events for individual patients within each clinical study report. Where possible, for each event we extracted the original term reported by the investigator, the term as coded by the medical coding dictionary, medical coding dictionary used, and the patient's trial identification number. Using the patient's trial identification number, we attempted to reconcile data on the same event between the different formats for presenting data on adverse events within the clinical study report.

Paper 3: benefits and harms of duloxetine for the treatment of stress urinary incontinence

In the third and final study, using data from clinical study reports, we determined the benefits and harms, including drug-induced depression, suicidality, violent behaviour and their potential precursors, of duloxetine for the treatment of stress urinary incontinence.

Suicidality and violent behaviour associated with the use of SSRIs have been reported among patients receiving treatment for psychiatric disorders and non-psychiatric disorders, and also among healthy volunteers.[69] Duloxetine is approved for use in the EU for the treatment of stress urinary incontinence. The manufacturer of duloxetine, however, withdrew its US marketing authorisation application for duloxetine for stress urinary incontinence in 2005, as the FDA was not positive to approving it.[123] In contrast to the EMA,[124,125] the FDA does not usually publish the reasons why applications are denied or withdrawn,[126] but it did say that a higher than expected rate of suicide attempts was observed in the open label extensions of the controlled studies.[127]

Data on benefits and harms were obtained from clinical study reports, including appendices of listings of adverse events for individual patient, (total 6,870 pages) of the four main randomised placebo controlled trials used in the marketing authorisation application of duloxetine for the treatment of stress urinary incontinence.[128] Benefits of interest were incontinence episode frequency

and incontinence quality of life. Harms of general interest were: serious adverse events, discontinuations because of adverse events, and number of patients experiencing at least one treatment emergent adverse event. Specific harms of interest were: suicidality, violent behaviour, and their potential precursors (akathisia, activation, emotional disturbance, and psychotic behaviour) and depression. Data on benefits were extracted from summary tables by one observer and checked by a second. Two observers independently searched all data formats of harms manually using an *a-priori* defined list of terms. Each observer independently recoded preferred terms (and if available, also verbatim terms) of each adverse event of interest using the most recent version of MedDRA (version 17.0). Meta-analyses were performed using this data.

Summary of results

Paper 1: comparison of protocols, clinical study reports, trial registries, and publications

We found all nine clinical study reports of duloxetine for major depressive disorder fully described the primary efficacy analysis and major harms. There were, however, inconsistencies *within* clinical study reports with regard to both benefits and harms. 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. There was contradictory information within clinical study reports for seven serious adverse events and eight adverse events that led to discontinuation but with no apparent bias.

There was evidence of publication bias in relation to beneficial effects. Six of the nine trials had significant results for the protocol specified primary efficacy analysis and were each published as a journal article reporting significant results. Three trials had non-significant results for the protocol specified primary efficacy analysis of which two were not published and the third was, but reported significant results according to the journal article. Its significant result was based on a different population and analytical method to that specified in the trial protocol. The journal article, however, did not mention that the analysis it presented was not the primary efficacy analysis specified in the protocol.

Harms, including serious adverse events, were generally poorly reported in both journal articles and trial registry reports. In comparison to clinical study reports, Lilly trial registry reports and journal articles reported considerably less data on harms. In each trial, a median of 166 (range 100 to 241) and 406 (range 177 to 645) treatment emergent adverse events (adverse events that emerged or worsened after study drug was started) in the randomised phase were not reported in Lilly trial registry reports and journal articles respectively.

Strengths and limitations of our study

The generalisability of our findings is unclear given that they are based on nine trials of a single drug from a single company. Furthermore, it is unclear if our findings are applicable to trial reports

that predate the 1995 ICH guidelines for clinical study reports. We do note, however, that a study published in 1980 found that trials reports of psychotropic drugs submitted to the Finnish and Swedish regulators as part of marketing authorisation applications had more data on harms than journal articles of the same trials.[129] It may be interpreted as a weakness that we only looked at inconsistencies in and completeness of reporting; we did not meta-analyse the clinical study report data to see whether the results differed from meta-analyses based on publicly available data. Although we found only a relatively small number of discrepancies in serious adverse events and adverse events that led to discontinuation within the reports we note that, within a trial, misclassifying or omitting even one adverse event can mean the difference between a statistically significant and non-statistically significant association with a drug.[130,131]

Relation of findings to those of other studies

Our findings of dissemination bias 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.[7,57,120] The poor reporting of harms, including serious adverse events, in journal articles or trial registry reports in comparison to clinical study reports, is in line with the findings of a study by Wieseler et al.[33] In that study, which included 86 clinical study reports, serious adverse events were found to be completed reported in 88% of clinical study reports but only 60% of journal articles or trial registry reports. Rodgers et al. also had similar findings in their study of Medtronic's bone implant for spinal fusion.[132]

Our findings of discrepancies in analysis agrees with the results of a recent systematic review by Dwan et al. that assessed the selective reporting of analyses in publications of clinical trials.[133] We note that none of the studies included in the review investigated evidence for bias resulting from selective reporting of analyses.

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 rejections by editors, or whether the reasons for incomplete data was constraints on word count. A recent qualitative study, involving interviews with clinical trialists, found multiple reasons for failure to publish trials including negative results, lack of time, lack of resources, and rejection by journals. Reasons for incomplete reporting include word limit constraints imposed by journals and peer reviewers not wanting all outcomes reported.[134]

Our findings support the view that journal articles are not an appropriate format to disseminate the results of clinical trials. We note, however, despite the evidence of publication bias and selective reporting, the UK Government, the British Academy of Medical Sciences and the Cochrane Collaboration Individual Patient Data Meta-analysis Method Group jointly stated dissemination of summary results in the format of a journal article would be preferable to alternative models, including summary results databases.[39] The reasons they cited included: permanence, quality control through peer review, possibility of corrections and retraction, and opportunity for post publication commentary an discussion.[39] These advantages, however, do not stand up to scrutiny as there is evidence of positive outcome bias during peer review and failure of journals to retract incorrect and misleading articles reporting trial results, despite resounding calls to do so.[135-137]

Paper 2: coding of adverse events of suicidality

Among nine trials of duloxetine for the treatment of major depressive disorder, we found six trials used the medical coding dictionary COSTART (Coding Symbols for a Thesaurus of Adverse Reaction Terms) to code original investigator reported terms of adverse events and three used the larger and more recent dictionary MedDRA (Medical Dictionary for Regulatory Activities).

In all nine clinical study reports we found tables reported coded terms only, and listings of adverse events reported verbatim terms only. Narratives, which were only for patients that experienced a serious or clinically important adverse event, or an adverse event that led to discontinuation, were the only format of adverse event data to report both verbatim and coded terms.

Suicides were clearly identifiable in all formats of adverse event data, that is tables, narratives, and listings of adverse events for individual patients, in the clinical study reports. Suicide attempts presented in summary tables included both definitive and provisional diagnoses.

Narratives containing data on suicidality were only available for six patients from three trials. In two of the three trials, COSTART was used. Narratives from these two trials showed that one event of "suicidal urges" while receiving duloxetine and two events of "suicidal ideation" while receiving paroxetine were coded as depression. There is no exact lowest level term for suicidal ideation in COSTART; the closest possible matching term is suicidal tendency, which is coded to the preferred term depression. In the third trial MedDRA was used. One patient in this trial experienced "suicidal

threat", which was coded as suicidal ideation. We found this was the closest matching term in the version of MedDRA used to code data for this trial. As a result, suicidal ideation and preparatory behaviour were obscured in some tables owing to the lack of specificity of the medical coding dictionary, especially COSTART.

Furthermore, one event of suicidal ideation described in narrative text was absent from tables and adverse event listings of individual patients. The reason for this is unclear, but may be due to the coding conventions used, i.e. to only code a diagnosis if both a diagnosis and symptoms are reported.

Strengths and limitations of our study

Our study is based on a small number of trials for a single drug manufactured by a single company. Another limitation is that, although the guideline for clinical study reports suggests that listings of adverse events for individual patients should include coded terms in addition to the verbatim terms,[32] this was not the case for the nine trials we examined. Our analysis of discrepancies in adverse events data was therefore limited to comparing data already coded in tables to those of narratives, which included verbatim and coded terms, of those patients who had adverse events that were serious, led to discontinuation of the study drug, or were non-serious but clinically important.

Relation of findings to those of other studies

Problems with terms in COSTART, including lack of specific preferred terms, were acknowledged in journal articles in the 1990s (the last version of COSTART was released in 1995).[138,139] It is therefore possible that our finding that adverse events of suicidal ideation were obscured in summary tables of COSTART coded data could also apply to other types of adverse events. Furthermore, our findings could also be applied to coded data from coding dictionaries, other than COSTART, that predate MedDRA. WHO-ART (The World Health Organisation Adverse Reaction Terminology) was predominately used throughout Europe before the introduction of MedDRA; some pharmaceutical companies modified existing dictionaries to create their own customised version e.g. GlaxoSmithKline modified COSTART; or created entirely new dictionaries, e.g. HARTS (Hoechst Adverse Reaction Terminology System) developed by Hoescht, now part of Sanofi.[140]

A study examining the difference between WHO-ART and MedDRA found that when the same verbatim text was coded with MedDRA and WHO-ART and compared, 16% of paired codes were rated as medically different. Furthermore, Brown et al. and Fescharek et al. both found that the use of different coding dictionaries can alter the apparent safety profile of a drug.[141,142]

It should also be noted that, while clinical study reports contain detailed data on adverse events, there is evidence from FDA analyses and court cases that access to case reports forms from trials can reveal discrepancies that would not be apparent from clinical study reports alone.[63,143]For example, an FDA analysis of a sample of case report forms from the RECORD trial revealed many missing cases of cardiac problems, which allowed the determination that, in contrast to the manufacturer's (GlaxoSmithKline) claims, rosiglitazone increased the risk of cardiac problems fourfold. Furthermore, as part of a court case concerning SSRIs and suicidality, examination of case report forms from clinical trials of fluoxetine found: reasons for termination were due to adverse events of activation, but were recorded as lack of efficacy; some investigators did not code events of insomnia, agitation, nervousness, restlessness as adverse events because they considered them to be symptoms of depression; on concomitant medicine pages of case report forms, treatment for adverse events were recorded but either the indication for treatment was not recorded on the adverse events section of the case report forms, or the terminology differed, with terms in the adverse events page less serious.[144] Case report forms, however, are currently unavailable to independent researchers. Should case report forms become available, any independent research using case report forms is likely to be costly, in terms of both time and money, as a case report form for a single patient can be hundreds of pages in length and require a considerable infrastructure to ensure unbiased judgements.[143]

Paper 3: benefits and harms of duloxetine for the treatment of stress urinary incontinence

Using data obtained from four placebo controlled trials of duloxetine for stress urinary incontinence (958 female patients receiving duloxetine 80mg daily and 955 female patients receiving placebo), we found that duloxetine was better than placebo for percent change in weekly incontinence episodes (mean difference (MD) -13.6% , 95% confidence interval -21.6% to -5.5%; standardized mean difference (SMD) 0.13, -0.22 to -0.04) and mean change in Incontinence Quality of Life total score (MD 3.2, 2.0 to 4.5; SMD 0.24, 0.15 to 0.33). These benefits, however, are so small they are

unlikely to be of clinical significance and could be fully explained by unblinding due to adverse events.

Duloxetine increased the risk of experiencing at least one treatment-emergent adverse event (risk ratio (RR) 1.32, 1.24 to 1.41; number needed to harm (NNH) 6, 5 to 7) and the risk of discontinuing because of an adverse event (RR 5.73 ,4.00 to 8.20; NNH 7, 6 to 8).There were no events of suicidality, violence or akathisia. Duloxetine increased the risk of experiencing emotional disturbance (RR 4.73, 1.62 to 13.85; NNH 65, 40 to 170) and the risk of experiencing a core or potential activation event (RR 4.45, 3.22 to 6.14; NNH 7, 6 to 9). There were also more core or potential psychotic events (RR 2.25, 1.06 to 4.81; NNH 80, 40 to 834) and more depression related events, but the risk was not significantly increased, RR 1.26 (0.58 to 2.71). With regard to coding of adverse events, we found there was excellent inter-observer agreement (lower level terms: kappa = 0.92; preferred terms: kappa = 0.99).

Strengths and limitations of our study

Access to individual patient data allowed us to elucidate specific harms caused by duloxetine, which would not have been possible if we only had access to summary data, either from published trial reports or clinical study reports.

Our findings, however, should be interpreted cautiously given that the data for beneficial effects, especially for incontinence episode frequency, were considerably skewed.[101] Furthermore, the effect of unblinding due to adverse events needs to be taken into consideration. A systematic review comparing antidepressants to active placebo found that the treatment effect of antidepressants was small (SMD 0.17, 0.00 to 0.34). This suggests that treatment effects are overestimated in trials that use inert placebos because of unblinding due to adverse events.[59]

There were only 958 patients on duloxetine, which means that the sample was too small to detect rare events of suicidality and violence. Furthermore, the data on adverse events were obtained through non-probing questions, which leads to underreporting of adverse events,[100,145] especially for events of a sensitive nature, such as suicidal ideation and behaviour, and violence.

Relation of findings to those of other studies

Our findings on harms are in agreement with a published pooled analysis, which is not a meta-analysis, performed by Lilly on the same four trials included in our study.[146] Our results for treatment emergent adverse events and discontinuation because of an adverse event are also in good agreement with a Cochrane review of ten trials of duloxetine for urinary incontinence, but this review did not report on suicidality.[147]

Our findings on lack of clinical relevant benefits, and important risk of harms, also are in agreement with a prospective cohort study of 228 women treated with duloxetine for stress urinary incontinence or mixed stress urinary incontinence. After four weeks of receiving duloxetine, 45% of the cohort had discontinued the drug due to adverse events, and 24% due to lack of efficacy. After 4 months, only 12% of the original cohort were still taking duloxetine.[148]

In order for a reliable evaluation of harms, coding of adverse events must be accurate and consistent. A systematic review of the challenges of coding adverse events, which I co-authored, found only one inter-observer study of coding.[122,149] In that study, using MedDRA version 3.0, two observers coded 260 verbatim terms of all types of adverse events in HIV clinical trials.[149] 12% of coded terms were found to differ between observers at the preferred term level. In our study, using MedDRA version 17.0, we coded 139 verbatim terms and found 0.7 % of coded terms differed between observers at the preferred term level. A possible reason for the discrepancy in inter-observer reliability found in the two studies is that while Tonéatti et al. coded all types of adverse events, we coded a subset of events, the vast majority of which were psychiatric in nature. The most likely explanation for the difference in findings, however, is the difference in the version of MedDRA used. This is because MedDRA version 17.0 has 25,000 more lower level terms, and twice as many preferred terms as version 3.0.[150]

All three studies forming this thesis show that individual patient data contained in appendices of clinical study reports are essential for a reliable assessment of drug harms. Such data, however, may not always accompany clinical study reports. In a sample of clinical study reports obtained by other researchers only 29 of 78 clinical study reports they reviewed had appendices of individual patient data.[151] Furthermore, the EMA recently announced that it will not publish individual anonymised patient data contained in appendices of clinical study reports in the first round of implementation of its new policy.[152] This effects clinical study reports submitted as part of new applications after 1

January 2015 and those submitted as part of extension applications after 1 July 2015. The apparent reasons for this are concerns over patient confidentiality. In accordance with current legislation, however, the data are already anonymised and the EMA's approach is inconsistent, as we can get access to the harms in the old trials in the EMA's possession.

It should be noted that some pharmaceutical companies have recently began making certain clinical study report data publicly available. For example, GlaxoSmithKline (GSK) have made clinical study reports of some of their drugs publicly available online, but with individual patient data removed. Researchers can request access to individual patient data, including statistical analysis software (SAS) datasets and case report forms, but have to submit a research proposal. This research proposal is assessed by an independent panel who decide whether or not to grant access to the data. If granted, access to data is available by viewing it onscreen only – data cannot be saved, downloaded, printed, copied, annotated, or shared in any way. This severely impedes researchers wishing to perform independent analyses, in fact, many have said it would be impossible. Researchers who have accessed data this way have described it as "unnecessarily maddening and a severe obstruction to the task."[153 (p 1)]

Conclusions and perspectives for further research

Conclusions

My overall conclusion is that, given the inadequacies of journal articles and trial registry reports, clinical study reports should be used as the primary data source for systematic reviews. They should, however, first be checked against protocols and within themselves for accuracy and consistency. Furthermore, researchers need to be aware that because of coding dictionaries and coding conventions used, adverse events data presented in summary tables may obscure adverse events of importance. Compared to placebo, duloxetine has a statistically significant, but likely clinically insignificant, beneficial effect in the treatment of stress urinary incontinence, and an increased risk of important psychiatric harms that are possible precursors to suicidality. The rationale for using duloxetine for stress urinary incontinence is therefore questionable. The evaluation of harms that were possible precursors to suicidality would have been impossible using only summary data in the clinical study report. Individual patient data of adverse events contained in appendices of clinical study reports are therefore essential for a reliable assessment of drug harms.

Implications for further research

My research has implications for both primary and secondary research.

First, in common with other researchers, we found that using clinical study reports is a time consuming process. We therefore need to develop tools and methodological approaches that will reduce the workload and allow researchers to use clinical study reports in an efficient manner. Natural language processing based text mining is already being used by pharmaceutical companies to extract and synthesise information found in unstructured text regions of clinical study reports. Given the possibility of discrepancies and potential for bias within clinical study reports, however, further research is needed to elucidate the accuracy and usefulness of such methods**.**

Second, our research highlights the need for independent researchers to have access to detailed data from trials, including original investigator reported terms of harms. Researchers with access to such

data, particularly of older trials, should recode these terms using the latest version of MedDRA, which can enable a more accurate assessment of harms to be made. Researchers with access to tables of coded data of harms only should consider whether the coding dictionary and coding conventions used affect the robustness of their assessment.

Third, research on SSRI-induced akathisia, activation, emotional disturbance and psychotic events as precursors to suicidality is sparse and largely retrospective, and antidepressant induced akathisia is underdiagnosed. Furthermore, research has shown that adverse event elicitation using open ended questions leads to underreporting of adverse events, especially for events of a sensitive nature such as suicidal ideation. The FDA has issued draft guidance for the assessment of suicidality in all clinical trials of drugs with central nervous system activity, including multiple dose phase 1 trials in healthy volunteers. This guidance does not address potential precursory events to suicidality, however, despite some of their inclusion in warnings in FDA antidepressant product labelling. I therefore suggest that future clinical trials of drugs with central nervous system activity, regardless of proposed indications for use, should prospectively assess suicidality *and* potential precursors to suicidality.

References

1. Strech D, Littmann J. Lack of proportionality. Seven specifications of public interest that override post-approval commercial interests on limited access to clinical data. *Trials* 2012;13: 100.

2. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014;383(9913): 257–66.

3. Glasziou P, Altman D, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383(9913): 267-76.

4. Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, et al. Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess* 2010;14:8.

5. Schmucker C, Schell LK, Portalupi S, Oeller P, Cabrera L, Bassler D, et al. Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries. *PLoS ONE* 2014;9(12): e114023.

6. Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *BMJ* 2012;344: d7202.

7. Eyding D, Lelgemann M, Grouven U, Harter M, Kromp M, 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.

8. Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias—An updated review. *PLoS ONE* 2013;8(7): e66844.

9. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;340: c365.

10. Jureidini JN, McHenry LB, Mansfield PR. Clinical trials and drug promotion: selective reporting of study 329. *International Journal of Risk & Safety in Medicine* 2008;20: 73-81.

11. 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 Revs* 2012;1: CD008965.

12. McGauran N, Wieseler B, Kreis J, Schüler YB, Kölsch H, Kaiser T. Reporting bias in medical research - a narrative review. *Trials* 2010;11:37.

13. United States Senate Committee on Finance. *FDA, Merck and Vioxx: Putting patient safety first. Testimony of David Graham;* 2004. http://www.finance.senate.gov/hearings/hearing/?id=48b33994-9de9-df09-dfb8-bb3599a92fbb (accessed 17 June 2015).

14. Author's reply to MM Wolfe. Rofecoxib, Merck, and the FDA. *N Engl J Med* 2004;351(27): 2877-8.

15. Topol EJ. Failing the public health-rofecoxib, Merck, and the FDA. *N Engl J Med* 2004;351(17): 1707-9.

16. Jack A. Tamiflu: A nice little earner. *BMJ* 2014;348: g2524.

17. Jefferson T, Jones M, Doshi P, Spencer EA, Onakpoya I, Heneghan CJ. Oseltamivir for influenza in adults and children: systematic review of clinical study reports and summary of regulatory comments. *BMJ* 2014;348: g2545.

18. Rasmussen N, Lee K, Bero L. Association of trial registration with the results and conclusions of published trials of new oncology drugs. *Trials* 2009;10:116.

19. Reveiz L, Chan AW, Krleza-Jerić K, Granados CE, Pinart M, Etxeandia I, et al. Reporting of methodologic information on trial registries for quality assessment: a study of trial records retrieved from the WHO search portal. *PLoS ONE* 2010;8(5): e12484.

20. Huić M, Marušić M, Marušić A. Completeness and changes in registered data and reporting bias of randomized controlled trials in ICMJE journals after trial registration policy. *PLoS ONE* 2011 6(9): e25258.

21. Viergever RF, Karam G, Reis A, Ghersi D. The quality of registration of clinical trials: still a problem. *PLoS ONE* 2014 9(1): e84727.

22. Ross JS, Mulvey GK, Hines EM, Nissen SE, Krumholz HM. Trial publication after registration in clinicaltrials.gov: a cross-sectional analysis. *PLoS Med* 2009 6(9): e1000144.

23. Prayle AP, Hurley MN, Smyth AR. Compliance with mandatory reporting of clinical trial results on ClinicalTrials.gov: cross sectional study. *BMJ* 2012;344: d737.

24. Anderson MK, Chiswell K, Peterson ED, Tasneem A, Topping J, Califf RM. Compliance with results reporting at ClinicalTrials.gov. *N Engl J Med* 2015;372(11): 1031-9.

25. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials *BMC Medicine* 2010;8:18.

26. Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev* 2012;11: MR000030.

27. Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. *Arch Intern Med* 2009;169:19: 1756-61.

28. Haidich AB, Birtsou C, Dardavessis T, Tirodimos I, Arvanitidou M. The quality of safety reporting in trials is still suboptimal: survey of major general medical journals. *J Clin Epidemiol* 2011;64(2):124-35.

29. Hodkinson A, Kirkham JJ, Tudur-Smith C, Gamble C. Reporting of harms data in RCTs: a systematic review of empirical assessments against the CONSORT harms extension. *BMJ Open* 2013;3: e003436.

30. Shah RR, Saint Raymond A. Regulation of human medicinal products in the European Union. In: Griffin JP, editor. *The textbook of pharmaceutical medicine*. 6th ed. Blackwell Publishing Ltd; 2009. p444-99.

31. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. *Organisation of ICH - Steering Committee*. http://www.ich.org/about/organisation-of-ich/steering.html (accessed 17 June 2015).

32. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. *Structure and content of clinical study reports: E3*. 1995 http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf (accessed 17 June 2015).

33. Wieseler B, Wolfram N, McGauran N, Kerekes MF, Vervölgyi 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(10): e1001526.

34. Gøtzsche PC, Jørgensen AW. Opening up data at the European Medicines Agency. *BMJ* 2011;342: d2686.

35. Dyer C. European drug agency's attempts to improve transparency stalled by legal action from two US drug companies. *BMJ* 2013;346: f3588.

36. The European Medicines Agency. *Publication of clinical data.* http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/general/general_content_000555.jsp (accessed 17 June 2015).

37. Doshi P, Jones M, Jefferson T. Rethinking credible evidence synthesis. *BMJ* 2012;344: d7898.

38. Doshi P, Jefferson T, Del Mar C. The imperative to share clinical study reports: recommendations from the Tamiflu experience. *PLoS Med* 2012;9(4): e1001201.

39. House of Commons Science and Technology Committee. *Clinical trials,* HC 104. London: The Stationery Office; 2013. http://www.publications.parliament.uk/pa/cm201314/cmselect/cmsctech/104/104.pdf (accessed 17 June 2015).

40. Chan AW. Bias, spin, and misreporting: time for full access to trial protocols and results. *PLoS Med* 2008;5(11): e230.

41. Bellevue Biomedical writing. *Writing your first clinical study report*. http://www.bellevuebiomedical.com/writing-samples.php (accessed 17 June 2015).

42. Loke YK, Price D, Herxheimer A. Systematic reviews of adverse effects: framework for a structured approach. *BMC Medical Research Methodology* 2007;7:32.

43.Healy D. *The anti-depressant era*. Harvard University Press; 1997.

44. Lyon M. Mood disorders and suicide. In: Lyon M, editor. *Abnormal Psychology: Clinical and Scientific Perspectives*. 4th ed. BVT Publishing; 2011. p293-330.

45. Hindmarch I. Expanding the horizons of depression: beyond the monoamine hypothesis. *Hum Psychopharmacol Clin Exp* 2001;16: 203-18.

46. Moncrieff J, Cohen D. Do antidepressants cure or create abnormal brain states? *PLoS Med* 2006;3(7): e240.

47. Lacasse JR, Leo J. Serotonin and depression: A disconnect between the advertisements and the scientific literature. *PLoS Med* 2005; 2(12): e392.

48. Angoa-Pérez M, Kane MJ, Briggs DI, Herrera-Mundo N, Sykes CE, Francescutti DM, et al. Mice genetically depleted of brain serotonin do not display a depression-like behavioral phenotype. *ACS Chem Neurosci* 2014;5(10): 908-19.

49. Healy D. Serotonin and depression *BMJ* 2015;350: h1771.

50. Pagura J, Katz LY, Mojtabai R, Druss BG, Cox B, Sareen J. Antidepressant use in the absence of common mental disorders in the general population. *J Clin Psychiatry* 2011;72: 494-501.

51. Wiechers IR, Leslie DL, Rosenheck RA. Prescribing of psychotropic medications to patients without a psychiatric diagnosis. *Psychiatr Serv* 2013;64: 1243–8.

52. Takayanagi Y, Spira AP, Bienvenu J, Hock RS, Carras MC, Eaton WW, et al. Antidepressant use and lifetime history of mental disorders in a community sample: results from the Baltimore Epidemiologic Catchment Area Study. *J Clin Psychiatry* 2015;76: 40-4.

53. NOMESCO. *Health statistics in the Nordic countries*. Copenhagen: Nordic Medico Statistical Committee; 2014. http://norden.diva-portal.org/smash/get/diva2:781991/FULLTEXT03.pdf (accessed 17 June 2015).

54. OECD. *Health at a glance 2013: OECD indicators*. Paris: OECD Publishing; 2013. http://dx.doi.org/10.1787/health_glance-2013-en (accessed 17 June 2015).

55. Pratt LA, Brody DJ, Gu Q. *Antidepressant use in persons aged 12 and over: United States, 2005–2008. NCHS data brief, no 76*. Hyattsville, MD: National Center for Health Statistics; 2011.

56. Gao P, Zhang Hm Xu H, Zhang C, Liu D. Increased use of antidepressants in Wuhan, China: A retrospective study from 2006 to 2012. *Medicina (Kaunas)* 2013;49: 529-34.

57. 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: 252-60

58. Fisher S, Greenberg RP. How sound is the double-blind design for evaluating psychotropic drugs? *J Nerv Ment Dis* 1993;181(6): 345-50.

59. Moncrieff J, Wessely S, Hardy R. Active placebos versus antidepressants for depression. *Cochrane Database Syst Rev* 2004;1: CD003012.

60. Fournier JC, DeRubeis RJ, Hollon SD; Dimidjiam S, Amsterdam JD, Shelton R, et al. Antidepressant drug effects and depression severity: a patient -level meta-analysis. *JAMA* 2010;303(47): 47-53.

61. Teicher MH, Glod C, Cole JO. Emergence of intense suicidal preoccupation during fluoxetine treatment. *Am J Psychiatry* 1990;147(2): 207-10.

62. Food and Drug Administration. *Psychopharmacological Drugs Advisory Committee, thirty fourth meeting proceedings, September 20, 1991.* Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; 1991. http://www.fda.gov/OHRMS/DOCKETS/AC/prozac/2443T1.PDF (accessed 17 June 2015).

63. Healy D. *Let them eat Prozac*. New York University Press; 2004.

64. Laughren TP. *Overview for December 13 Meeting of Psychopharmacologic Drugs Advisory Committee (PDAC) 2006 Nov 16*. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; 2006 http://www.fda.gov/ohrms/dockets/ac/06/briefing/2006-4272b1-01-FDA.pdf (accessed 17 June 2015).

65. Institute of Medicine. *CNS clinical trials: Suicidality and data collection: Workshop summary.* Washington, DC: The National Academies Press; 2010.

66. Healy D. Emergence of antidepressant induced suicidality. *Primary Care Psychiatry* 2000;6:1.

67. Healy D. Lines of evidence on the risks of suicide with selective serotonin reuptake inhibitors. *Psychother Psychosom* 2003;72:71-9.

68. Breggin BR. Suicidality, violence and mania caused by selective serotonin reuptake inhibitors (SSRIs): A review and analysis. *International Journal of Risk & Safety in Medicine* 2003/2004;16: 31–49.

69. Healy D, Herxheimer A, Menkes DB. Antidepressants and violence: Problems at the interface of medicine and law. *PLoS Med* 2006;3(9): e372.

70. Hansen L. A critical review of akathisia, and its possible association with suicidal behaviour. *Hum Psychopharmacol Clin Exp* 2001;16: 495-505.

71. Koliscak LP, Makela EH. Selective serotonin reuptake inhibitor-induced akathisia. *J Am Pharm Assoc* 2009;49: e28–e38.

72. Seemüller F, Riedel M, Obermeier M, Bauer M, Adli M, Mundt C, et al. The controversial link between antidepressants and suicidality risks in adults: data from a naturalistic study on a large sample of in-patients with major depressive episode. *Int J Neuropsychoph* 2009;12:181-89.

73. Price J, Cole V, Goodwin GM. Emotional side-effects of selective serotonin reuptake inhibitors: qualitative study. *BJP* 2009;195: 211-17.

74. Goldsmith L, Moncrieff J. The psychoactive effects of antidepressants and their association with suicidality. *Curr Drug Saf.* 2011; 6: 115-21.

75. Read J, Cartwright C, Gibson K. Adverse emotional and interpersonal effects reported by 1829 New Zealanders while taking antidepressants. *Psychiat Res* 2014;216: 67–73.

76. Preda A, MacLean RWm, Mazure CM, Bowers MB Jr. Antidepressant-associated mania and psychosis resulting in psychiatric admissions. *J Clin Psychiatry* 2001;62: 30-3.

77. Perlis RH, Beasley CM Jr, Wines JD Jr, Tamura RN, Cusin C, Shear D, et al. Treatment-associated suicidal ideation and adverse effects in an open, multicenter trial of fluoxetine for major depressive episodes. *Psychother Psychosom* 2007;76: 40-6.

78. Sinclair LI, Christmas DM, Hood SD, Potokar JP, Robertson A, Isaac A, et al. Antidepressant-induced jitteriness/anxiety syndrome: systematic review. *BJP* 2009;194: 483-90.

79. Harada T, Inada K, Yamada K, Sakamoto K, Ishigooka J. A prospective naturalistic study of antidepressant induced jitteriness/anxiety syndrome. 2014. *Neuropsychiatr Dis Treat* 2014;10: 2115–21.

80. Food and Drug Administration. *Revisions to product labelling.* Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research http://www.fda.gov/downloads/Drugs/DrugSafety/InformationbyDrugClass/UCM173233.pdf (accessed 17 June 2015).

81. Laughren TP. *Background Comments for February 2, 2004 Meeting of Psychopharmacological Drugs Advisory Committee (PDAC) and Pediatric Subcommittee of the Anti-Infective Drugs Advisory Committee (Peds AC).* Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; 2004 http://www.fda.gov/ohrms/dockets/ac/04/briefing/4006B1_03_Background%20Memo%2001-05-04.pdf (accessed 17 June 2015).

82. Hammad T. *Relationship between psychotropic drugs and pediatric suicidality.* Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; 2004 http://www.fda.gov/ohrms/dockets/ac/04/briefing/2004-4065b1-10-tab08-hammads-review.pdf (accessed 17 June 2015)

83. Hammad TA, Laughren T, Racoosin J. Suicidality in pediatric patients treated with antidepressant drugs. *Arch Gen Psychiatry* 2006;63: 332-9.

84. White CA. A preliminary assessment of the impact of MEDDRA on adverse event reports

and product labelling. *Drug Inf J* 1998;32: 347-62.

85. Posner K, Oquendo MA, Gould M, Stanley B, Davies M. Columbia Classification Algorithm of Suicide Assessment (C-CASA): Classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants. *Am J Psychiatry* 2007;164(7): 1035–43.

86. United Kingdom Committee of Safety of Medicines. *Report of the CSM Expert Working Group on the safety of selective serotonin reuptake inhibitor antidepressants*. 2004 http://webarchive.nationalarchives.gov.uk/20141205150130/http://mhra.gov.uk/home/groups/pl-p/documents/drugsafetymessage/con019472.pdf (accessed 17 June 2015).

87. Medawar C, Herxheimer A. A comparison of adverse drug reaction reports from professionals and users, relating to risk of dependence and suicidal behaviour with paroxetine. *International Journal of Risk & Safety in Medicine* 2003-2004;16: 5–19

88. Sansone RA, Sansone LA. Duloxetine-related acute dysphoria. *Psychiatry* 2007;4: 65-8.

89. Parikh AR, Thatcher BT, Tamano EA, Liskow BI. Suicidal ideation associated with duloxetine use: a case series. *J Clin Psychopharmacol* 2008;28: 101-2.

90. Salem BA, Karam EG. Duloxetine and suicide attempts: a possible relation. *Clin Pract Epidemiol Ment Health* 2008;4: 18.

91. Vilhelmsson A, Svensson T, Meeuwisse A, Carlsten A. What can we learn from consumer reports on psychiatric adverse drug reactions with antidepressant medication? Experiences from reports to a consumer association. *BMC Clinical Pharmacology* 2011;11: 16.

92. Thomson KH, Martin RM, Potokar J, Pirmohamed M, Gunnell D. Reporting of drug induced depression and non-fatal suicidal behaviour in the UK from 1998 to 2011. *BMC Pharmacology and Toxicology* 2014;15: 54.

93. Moore TJ, Glenmullen J, Furberg CD. Prescription drugs associated with reports of violence towards others. *PLoS ONE* 2010;5(12): e15337.

94. Björkenstam C, Möller J, Ringbäck G, Salmi P, Hallqvist J, Ljung R. An association between initiation of selective serotonin reuptake inhibitors and suicide - a nationwide register-based case-crossover study. *PLoS ONE* 2013;8(9): e73973.

95. Tiihonen J, Lehti M, Aaltonen M, Kivivuori J, Kautiainen H, Virta LJ, et al. Psychotropic drugs and homicide: a prospective cohort study from Finland. *World Psychiatry* 2015;14(2): 245-7.

96. Howick J, Glasziou P, Aronson JK. The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute? *J R Soc Med* 2009;102: 186–94.

97. Vandenbroucke JP, Psaty BM. Benefits and risks of drug treatments how to combine the best evidence on benefits with the best data about adverse effects. *JAMA* 2008;300(20): 2417-9.

98. Loke YK, Price D, Herxheimer A. Chapter 14: Adverse effects. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions.* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

99. Chou R, Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, et al. *Assessing harms when comparing medical interventions.* In: Agency for Healthcare Research and Quality. Methods Reference Guide for Comparative Effectiveness Reviews [posted November 2008]. Rockville, MD. http://www.effectivehealthcare.ahrq.gov/tasks/sites/ehc/assets/File/MethodsGuide_Chou_Assessing_Harms_When_Comparing.pdf (accessed 17 June 2015).

100. Friedman LM., Furberg CD, DeMets D. *Fundamentals of clinical trials*. 4th ed. New York: Springer-Verlag; 2012.

101. Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

102. Beasley CM Jr, Dornseif BE, Bosomworth JC, Sayler ME, Rampey AH Jr, Heilgenstein JH. Fluoxetine and suicide: a meta-analysis of controlled trials of treatment of depression. *BMJ* 1991;303: 685-92.

103. López-Ibor Jr JJ. Reduced suicidality with paroxetine. *Eur Psychiatry* 1993;8: Suppl 1.

104. Montgomery SA, Dunner DL, Dunbar GC. Reduction of suicidal thoughts with paroxetine in comparison with reference antidepressants and placebo. *European Neuropsychopharm* 1995;5: 5-13.

105. Graham D. *Sponsor's ADR submission on fluoxetine dated July 17, 1990.* Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; 1990. http://www.healyprozac.com/Trials/CriticalDocs/memo110990.htm (accessed 17 June 2015).

106. Brecher M. *Review and evaluation of clinical data. Original NDA 20-031. Paroxetine (Aropax) safety review. 1992 Oct 5*. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; 1992. http://www.healyprozac.com/Trials/CriticalDocs/parox1992.pdf (accessed 17 June 2015).

107. Lee H. *Statistical review and evaluation (addendum 2) NDA 19-839/drug class 1C. 1991 Jan 31.* Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; 1991. http://www.healyprozac.com/Trials/CriticalDocs/sert1991.pdf (accessed 17 June 2015).

108. Healy D. Did regulators fail over selective serotonin reuptake inhibitors? *BMJ* 2006; 333(7558): 92-5.

109. Heiligenstein JH, Beaseley CM Jr, Potvin JH. Fluoxetine not associated with increased aggression in controlled clinical trials. *Intl Clin Psychopharm* 1993;8: 277-80.

110. Khan A, Warner HA, Brown WA: Symptom reduction and suicide risk in patients treated with placebo in antidepressant clinical trials. *Arch Gen Psychiatry* 2000; 57: 311-7.

111. Storosum JG, van Zweiten BJ, van den Brink W, Gersons BPR, Broekmans AW: Suicide risk in placebo-controlled studies of major depression. *Am J Psychiatry* 2001;158: 1271-5.

112. Bravata DM, Olkin I. Simple pooling versus combining in a meta-analysis. *Eval Health Prof.* 2001;24: 218-30.

113. Fergusson D, Doucette S, Glass KC, Shapiro S, Healy D, Hebert P, et al. Association between suicide attempts and selective serotonin reuptake inhibitors: systematic review of randomised controlled trials. *BMJ* 2005;330(7488): 396.

114. Gunnell D, Saperia J, Ashby D. Selective serotonin reuptake inhibitors (SSRIs) and suicide in adults: meta-analysis of drug company data from placebo controlled, randomised controlled trials submitted to the MHRA's safety review. *BMJ* 2005;330(7488): 385.

115. Saperia J, Ashby D, Gunnell D. Suicidal behaviour and SSRIs: updated meta-analysis. *BMJ* 2006;332(7555): 1453.

116. Stone M, Laughren T, Jones ML, Levenson M, Holland PC, Hughes A, et al. Risk of suicidality in clinical trials of antidepressants in adults: analysis of proprietary data submitted to US Food and Drug Administration. *BMJ* 2009;339: b2880.

117. Healy D. Letters to the editor: risk of suicide. *BMJ* 2005;330(7500): 1148.

118. Gøtzsche PC. UK drug regulator destroys all evidence after 15 years. *BMJ* 2011;343: d4203.

119. Gøtzsche PC. *Deadly medicines and organised crime: How big pharma has corrupted health care*. London: Radcliffe Publishing; 2013.

120. Melander H. Evidence b(i)ased medicine—selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 2003;326: 1171-3.

121. The European Medicines Agency. *Scientific Discussion.* 2005. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000572/WC500036776.pdf

122. Schroll JB, Maund E, Gøtzsche PC. Challenges in coding adverse events in clinical trials: a systematic review. *PLoS ONE* 2012;7(7): e41174.

123. Eli Lilly and Company. *Lilly and Boehringer Ingelheim jointly announce the recision of U.S. FDA application for duloxetine for treatment of stress urinary incontinence.* 2005. https://investor.lilly.com/releasedetail.cfm?ReleaseID=154081 (accessed 17 June 2015).

124. The European Medicines Agency. *Procedural advice on publication of information on negative opinions and refusals of marketing authorisation applications for human medicinal products*. 2013.

http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004188.pdf (accessed 17 June 2015).

125. The European Medicines Agency. *Procedural advice on publication of information on withdrawals of applications related to the marketing authorisation of human medicinal products*. 2013. http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500005061.pdf (accessed 17 June 2015).

126. Food and Drug Administration Centre for Drug Evaluation and Research. Personal communication. 26 November 2014.

127. Food and Drug Administration. *Historical Information on Duloxetine hydrochloride (marketed as Cymbalta).* http://google2.fda.gov/search?q=www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm114970.htm&client=FDAgov-archive&proxystylesheet=FDAgov&output=xml_no_dtd&site=FDAgov-archive&requiredfields=archive:Yes&sort=date:D:L:d1&filter=1 (accessed 17 June 2015).

128. The European Medicines Agency. *Scientific discussion*. 2005. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000545/WC500045641.pdf (accessed 17 June 2015).

129. Hemminki E. Study of information submitted by drug companies to licensing authorities. *Br Med J* 1980;280: 833-6.

130. Peters PH, Gravenstein S, Norwood P, De Bock V, Van Couter A, Gibbens M, et al. Long-term use of oseltamivir for the prophylaxis of influenza in a vaccinated frail older population. *J Am Geriatr Soc* 2001;49:1025-31. [Correction in *J Am Geriatr Soc* 61:478.]

131. Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *BMJ* 2012;344: e1119.

132. 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.

133. Dwan K, Altman DG, Clarke M, Gamble C, Higgins JPT, Sterne JAC, et al. Evidence for the selective reporting of analyses and discrepancies in clinical trials: a systematic review of cohort studies of clinical trials. *PLoS Med* 2014;11(6): e1001666.

134. Smyth R, Jacoby A, Altman DG, Gamble C, Williamson PR. The natural history of conducting and reporting clinical trials: interviews with trialists. *Trials* 2015;16:16.

135. Emerson GB, Warme WJ, Wolf FM, Heckman JD, Brand RA, Seth S, et al. Testing for the presence of positive-outcome bias in peer review: a randomized controlled trial. *Arch Intern Med* 2010;170(21): 1934-39.

136. Newman M. The rules of retraction. *BMJ* 2010;341: c6985.

137. Doshi P. Putting GlaxoSmithKline to the test over paroxetine. *BMJ* 2013;347: f6754.

138. Fizames C. How to improve the medical quality of the coding reports based in Who-art and Costart use. *Drug Inf J* 1997;31: 85-92.

139. Matsumoto C. Implementation and use of the Costart dictionary at Lilly: would we do it again. *Drug Inf J* 1991;25: 201-3.

140. Kornai A, Stone L. Automatic translation to controlled medical vocabularies. In: Abraham A, Jain LC, van der Zwaag BJ (eds). *Innovation in intelligent systems*. 1st ed. Springer; 2004

141. Brown DR, Brown EG, Moulvad TB. A comparison of two medical terminologies in coding and analysing clinical trial safety data. *International Journal of Pharmaceutical Medicine* 1997;11: 85–9.

142. Fescharek R, Kubler J, Elsasser U, Frank M, Guthlein P. Medical dictionary for regulatory activities (MedDRA): Data retrieval and presentation. *International Journal of Pharmaceutical Medicine* 2004;18: 259–69.

143. Parks MH. *Advisory Committee Meeting for NDA 21071 Avandia (rosiglitazone maleate tablet) July 13 and 14, 2010.* Department of Health and Human Services, Food and Drug Administration; 2010.

144. McBride JK. *Trial testimony of Dr Nancy Lord on 24 October 1994 in Fentress et al. versus Shea Communications et al*. http://www.healyprozac.com/Trials/Fentress/Transcripts/b10-24-%201994.txt (accessed 17 June 2015).

145. Montejo-Gonzalez AL, Llorca G, Izquierdo JA, Ledesma A, Bousono M, Calcedo A, et al. SSRI-induced sexual dysfunction: fluoxetine, paroxetine, sertraline, and fluvoxamine in a prospective, multicenter, and descriptive clinical study of 344 patients. *J. Sex Marital Ther.* 1997;23: 176–94.

146. Hurley DJ, Turner CL, Yalcin I, Viktrup L, Baygani SK. Duloxetine for the treatment of stress urinary incontinence in women: an integrated analysis of safety. *Eur J Obstet Gynecol Reprod Biol*. 2006 1;125: 120-8.

147. Mariappan P, Alhasso AA, Grant A, N'Dow JMO. Serotonin and noradrenaline reuptake inhibitors (SNRI) for stress urinary incontinence in adults. *Cochrane Database Syst Rev* 2005;3: CD004742.

148. Vella M, Duckett J, Basu M. Duloxetine 1 year on: the long term outcome of a cohort of women prescribed duloxetine. *Int Urogynecol J* 2008;19: 961–4.

149. Tonéatti C, Saïdi Y, Meiffrédy V, Tangre P, Harel M, Eliette V, et al. Experience using MedDRA for global events coding in HIV clinical trials. *Contemp Clin Trials* 2006;27: 13–22.

150. MedDRA Maintenance and Support Services Organization. Personal communication. 12 June 2015.

151. Doshi P, Jefferson T. Clinical study reports of randomised controlled trials: an exploratory review of previously confidential industry reports. *BMJ Open* 2013;3: e002496.

152. The European Medicines Agency. *Questions and answers on the European Medicines Agency policy on publication of clinical data for medicinal products for human use.* 2015. http://www.ema.europa.eu/docs/en_GB/document_library/Report/2014/10/WC500174378.pdf (accessed 17 June 2015).

153. Jureidini J, Nardo J. Inadequacy of remote desktop interface for independent reanalysis of data from drug trials. *BMJ* 2014;349: g4353.

Paper 1: comparison of protocols, clinical study reports, trial registries, and publications

BMJ

BMJ 2014;348:g3510 doi: 10.1136/bmj.g3510 (Published 4 June 2014)

RESEARCH

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

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

[1]Nordic Cochrane Centre, Rigshospitalet Dept 7811, Copenhagen, Denmark; [2]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,[1] and the other half is often published selectively,[2] 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,[3 4] which in one study led to an overall 32% overestimation of the treatment effect.[4] 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[5]; 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.[6] Furthermore, a recent study found that clinical study reports were more complete in their reporting of outcomes than published articles and trial registries combined.[7] It should be

Correspondence to: E Maund 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.[8 9]

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.[10] 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.[11-13]

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.[14] 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⇓). 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),[15] 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%.[16]

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/). 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"[33]
- 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"[33]
- 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"[33]
- 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"[10]
- 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[10 34 35]
- 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[10]

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 ≥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⇓). 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, 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⇓ 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 ($HAM\text{-}D_{17}$) 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.[17-25] 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.[17] 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_{17} 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⇓ 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.[26] 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.[3-5] 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.[27]

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.[28] 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.[29] 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.[30] 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.[31]

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.[32]

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 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.

1 Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev* 2007;2:MR000005.
2 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.
3 Melander H. Evidence b(i)ased medicine—selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 2003;326:1171-3.
4 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.
5 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.
6 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.
7 Wieseler B, Wolfram N, McGauran N, Kerekes MF, Vervölgyi 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.
8 Chan AW. Bias, spin, and misreporting: time for full access to trial protocols and results. *PLoS Med* 2008;5:e230.
9 Jefferson T, Doshi P, Thompson M, Heneghan C. Ensuring safe and effective drugs: who can do what it takes? *BMJ* 2011;342:c7258.
10 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.
11 Gøtzsche PC, Jørgensen AW. Opening up data at the European Medicines Agency. *BMJ* 2011;342:d2686.
12 Doshi P. Transparency interrupted: the curtailment of the European Medicines Agency's policy on access to documents. *JAMA Intern Med* 2013;173:2009-11.
13 Dyer C. European drug agency's attempts to improve transparency stalled by legal action from two US drug companies. *BMJ* 2013;346:f3588.
14 Scientific Discussion EMA 2005. www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000572/WC500036776.pdf.
15 Central authorisation of medicines. www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000109.jsp.
16 ABBYY FineReader 11 professional edition—full feature list. http://finereader.abbyy.com/professional/full_feature_list/.
17 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.
18 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.
19 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.
20 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.
21 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.
22 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.
23 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.
24 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.
25 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.
26 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.
27 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.
28 Haddad PM, Anderson IM. Recognising and managing antidepressant discontinuation symptoms. *Adv Psychiatr Treat* 2007;13:447-57.
29 Smith R. Medical journals are an extension of the marketing arm of pharmaceutical companies. *PLoS Med* 2005;2:e138.
30 Agreement reached in Europe on Clinical Trials Regulation. www.alltrials.net/2013/agreement-reached-europe-clinical-trials-regulation/.
31 Access to clinical trial information and the stockpiling of Tamiflu. Thirty-fifth report of session 2013-4. House of Commons Committee of Public Accounts.
32 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.
33 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.
34 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.
35 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.

Accepted: 05 May 2014

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

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[17]	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[18]	16 Nov '04	1196	11	23
5	HMAYa	15 Feb '00	30 Mar '00	16 Nov '00	15 Jan '03	Yes†	2004[19]	23 Jun '05	2211	14	71
6	HMAYb	15 Feb '00	30 Mar '00	31 Oct '00	15 Jan '03	Yes‡	May '06[20]	27 Jul '06	2095	12	33
7	HMBHa	19 Jul '00	None	09 Nov '00	27 Jul '01	Yes†	Apr '02[21]	15 Nov '04	916	8	20
8	HMBHb	19 Jul '00	None	13 Nov '00	6 Aug '01	Yes*	2002[22]	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,[25]¶ Fava,[23]**Hudson[24]††	17 Jul '06	3369	Perahia[25] 8; Fava[23] 9, Hudson[24] 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.

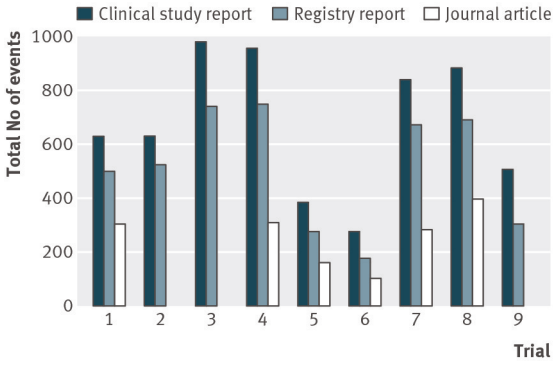
§Additional telephone call to patients from one investigative site.

¶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.

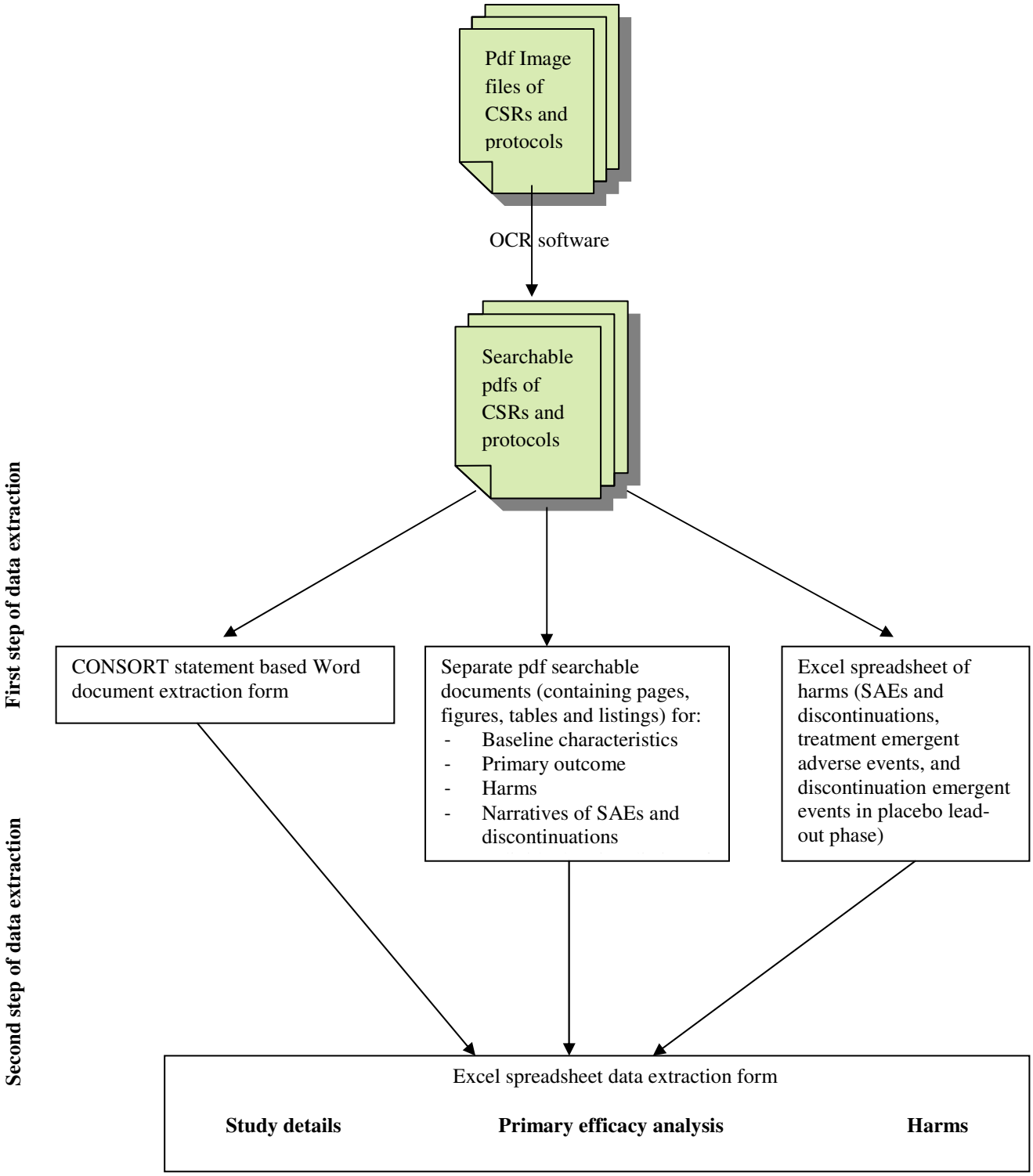
Figure



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

Appendix 1: Flowchart of data extraction methods [posted as supplied by author]



Appendix 2: Supplementary table [posted as supplied by author]

Proportion of trials with full, partial and unreported harms, and proportion of fully reported trials showing inconsistencies with respect to data in the clinical study report.

	CSR		Publically available data			
			Journal article		Lilly trial registry report	
	Reported (proportion)	**Inconsistency within CSR (proportion)**	**Reported (proportion)**	**Inconsistency between article and CSR (proportion)**	**Reported (proportion)**	**Inconsistency between registry report and CSR (proportion)**
Trials with occurrence of death in any phase	Full: 3/3	1/3 [a]	Full: 2/3 No article: 1/3	None	Full: 2/3 Incomplete: 1/3[a]	None
Trials with non-occurrence of death in all phases	Full: 6/6	None	Full: 1/6 Incomplete: 1/6 Unreported: 3/6 No article: 1/6	None	Full: 4/6 Incomplete: 2/6	None
SAEs - randomised phase (number of patients)	Full: 9/9	4/9 [b]	Full: 3/9 Unreported: 4/9 No article: 2/9	None	Full: 3/9 Incomplete: 6/9	None
SAEs - randomised phase (number of events)	Full: 9/9	4/9 [b]	Full: 3/9 Unreported: 4/9 No article: 2/9	1/3 [c]	Full: 3/9 Incomplete: 5/9 Unreported: 1/9	None
Discontinuations due to AEs - randomised phase (number of patients)	Full: 9/9	1/9 [d]	Full: 7/9 No article: 2/9	1/7 [e]	Full: 7/9 Incomplete: 2/9	None
Discontinuations due to AEs - randomised phase (number of events)	Full: 9/9	1/9 [d]	Full: 1/9 Unreported: 6/9 No article: 2/9	None	Full: 1/9 Incomplete: 1/9 Unreported: 7/9	None
TEAEs- randomised phase (number of patients)	Full: 9/9	None	Incomplete: 1/9 Unreported: 6/9	Not applicable	Full: 9/9	None

				No article: 2/9			
TEAEs- randomised phase (number of events)	Full:	9/9	None	Incomplete: 7/9 No article: 2/9	Not applicable	Incomplete: 9/9	Not applicable
DEAEs (number of patients)	Full:	9/9	1/9 [a]	Incomplete: 1/9 Unreported: 6/9 No article: 2/9	Not applicable	Unreported: 9/9	Not applicable
DEAEs (number of events)	Full:	9/9	1/9 [a]	Incomplete: 6/9 Unreported: 1/9 No article: 2/9	Not applicable	Incomplete: 2/9 Unreported: 7/9	Not applicable
Trials with occurrence of suicide in any phase	Full:	2/2	None	Full: 2/2	None	Full: 2/2	None
Attempted suicide - any phase (number of patients)	Full:	1/1	None	Incomplete: 1/1[f]	None	Unreported: 1/1 [g]	Not applicable
Attempted suicide - any phase (number of events)	Full:	1/1	None	Unreported: 1/1[f]	Not applicable	Unreported: 1/1[g]	Not applicable

AEs=adverse events; CSR=clinical study report; DEAEs=discontinuation emergent adverse events; p=page number; SAEs=serious adverse events; TEAEs=treatment emergent adverse events

[a] Trial 3 (HMATa); [b] Trials 3,7,8,9 (HMATa, HMBHa, HMBHb, HMBC); [c] Trial 5 (HMAYa); [d] Trial 9 (HMBC);[e] Trial 8 (HMBHb), [f] Perahia 2006 and Hudson 2007 [24,25]; [g] unreported because suicide attempts were below the incidence threshold of ≥2% used in the tables of the Lilly trial registry reports.

Paper 2: coding of adverse events of suicidality

BMJ

BMJ 2014;348:g3555 doi: 10.1136/bmj.g3555 (Published 4 June 2014)

RESEARCH

Coding of adverse events of suicidality in clinical study reports of duloxetine for the treatment of major depressive disorder: descriptive study

Emma Maund *PhD student*[1], Britta Tendal *postdoctoral researcher*[1], Asbjørn Hróbjartsson *senior researcher*[1], Andreas Lundh *physician*[1 2], Peter C Gøtzsche *professor*[1]

[1]Nordic Cochrane Centre, Rigshospitalet Department 7811, Copenhagen, Denmark ; [2]Department of Infectious Diseases, Hvidovre University Hospital, Hvidovre, Denmark

Abstract

Objective To assess the effects of coding and coding conventions on summaries and tabulations of adverse events data on suicidality within clinical study reports.

Design Systematic electronic search for adverse events of suicidality in tables, narratives, and listings of adverse events in individual patients within clinical study reports. Where possible, for each event we extracted the original term reported by the investigator, the term as coded by the medical coding dictionary, medical coding dictionary used, and the patient's trial identification number. Using the patient's trial identification number, we attempted to reconcile data on the same event between the different formats for presenting data on adverse events within the clinical study report.

Setting 9 randomised placebo controlled trials of duloxetine for major depressive disorder submitted to the European Medicines Agency for marketing approval.

Data sources Clinical study reports obtained from the EMA in 2011.

Results Six trials used the medical coding dictionary COSTART (Coding Symbols for a Thesaurus of Adverse Reaction Terms) and three used MedDRA (Medical Dictionary for Regulatory Activities). Suicides were clearly identifiable in all formats of adverse event data in clinical study reports. Suicide attempts presented in tables included both definitive and provisional diagnoses. Suicidal ideation and preparatory behaviour were obscured in some tables owing to the lack of specificity of the medical coding dictionary, especially COSTART. Furthermore, we found one event of suicidal ideation described in narrative text that was absent from tables and adverse event listings of individual patients. The reason for this is unclear, but may be due to the coding conventions used.

Conclusion Data on adverse events in tables in clinical study reports may not accurately represent the underlying patient data because of the medical dictionaries and coding conventions used. In clinical study reports, the listings of adverse events for individual patients and narratives of adverse events can provide additional information, including original investigator reported adverse event terms, which can enable a more accurate estimate of harms.

Introduction

A proper assessment of the benefits and harms of a medical intervention requires accurate data on harms. An assessment of the harms of an intervention in a randomised clinical trial is more difficult than an assessment of the benefits, as harms can be unpredictable and harms events may be rare.

In a classic drug trial, run and financed by the producer of the drug (the sponsor), doctors interacting with patients (the trial investigators) describe in case report forms those adverse events occurring in each patient, and the sponsor then codes them and enters them in clinical safety databases. The coded data are used for production of summaries of product characteristics and clinical study reports. Clinical study reports comprise detailed information on efficacy and adverse events data from a single trial and can be hundreds of pages in length. These clinical study reports form part of the marketing authorisation application submitted to regulatory authorities, and they should also be used as the primary data source for systematic reviews of drugs.[1] This has been most aptly illustrated by the Cochrane review of neuraminidase inhibitors for preventing and treating influenza in healthy adults and children, where the review based on clinical study reports on oseltamivir (Tamiflu) gave much more modest results than those based on published reports.[2]

In a clinical study report, data on adverse events are presented in various summaries and tabulations, including listings of all

Correspondence to: E Maund em@cochrane.dk

Extra material supplied by the author (see http://www.bmj.com/content/348/bmj.g3555?tab=related#webextra)

Supplementary table 1: adverse events of suicidal ideation and preparatory behaviour
Appendices 1a-c: examples of data formats

adverse events and pre-existing medical conditions in individual patients; narratives of clinically important adverse events (including serious adverse events or discontinuations of the study drug as a result of adverse events), which also include data on pre-existing medical conditions; and summary tables of treatment emergent adverse events (events that occurred or worsened after the study drug was started) or adverse events that emerged after discontinuation of the study drug (see box and supplementary appendices 1a-c for examples of each format).[3]

There are important differences between these three data formats that are related to the coding procedures. The narratives—and in some cases also the listings of data on individual patients—contain the investigator's description of the adverse event on the case report form (commonly referred to as the "verbatim" description). In summary tables, the events appear as coded terms. This is necessary to analyse rates of occurrence because investigators may use different terms to describe the same type of events. The grouping of similar events is achieved by coding verbatim terms to the most closely matching lowest level term in a hierarchically structured medical coding dictionary (tables 1⇓ and 2⇓). Similar lowest level terms are aggregated at the next level into a preferred term, so named because it is a favoured term for use in submissions to regulatory authorities, which are presented in summary tables of adverse events.[10]

Historically, the most widely used dictionaries have been the US Food and Drug Administration's Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART) and the World Health Organization Adverse Reaction Terminology (WHO-ART). These dictionaries were introduced in 1969 in response to increased regulation of the pharmaceutical industry after the thalidomide scandal.[11] However, they had limitations, including lack of specificity of lowest level terms, and in 1999 the International Conference on Harmonisation (ICH) launched the Medical Dictionary for Regulatory Activities (MedDRA).[12] MedDRA has more terms and they are more specific than those in the earlier dictionaries (see tables 1 and 2). For example, a study found that MedDRA contained exact or acceptable matches for 90% of verbatim terms but that COSTART contained only 62%.[13]

MedDRA cannot, however, solve all problems. Firstly, data in summary tables and in listings of adverse events for individual patients may differ from those presented in narratives because of the coding conventions used. For example, the preferred coding convention for a definitive diagnosis with symptoms, such as "anaphylactic reaction, rash, dyspnea, hypotension, and laryngospasm" is to code the diagnosis only as, for example, anaphylactic reaction.[14] In contrast, the preferred coding convention for a provisional diagnosis with symptoms, such as "Possible myocardial infarction with chest pain, dyspnea, diaphoresis" is to code the provisional diagnosis and symptoms as, for example, myocardial infarction, chest pain, dyspnoea, diaphoresis.[14] Secondly, coding can be inconsistent. For example, when the FDA wanted to analyse the risk of suicidality (ideation, behaviour, suicide attempts, and suicide) in paediatric trials of selective serotonin reuptake inhibitors (SSRIs), they found instances of suicidality events coded to both more severe terms and less severe terms. The FDA found that any conclusion based on such data would be unreliable and might lead to either an unwarranted restriction of the drugs or an underestimation of their dangers.[15] Unsurprisingly, research in other areas has shown that misclassifying or omitting even one adverse event can mean the difference between a statistically significant and non-statistically significant association with a drug.[16 17]

We assessed the effects of coding and coding conventions on adverse events data within clinical study reports and compared three different data formats. We used the nine main placebo controlled trials submitted to the European Medicines Agency in the marketing authorisation application of duloxetine for the treatment of major depressive disorder in adults.[18]

Methods

The nine clinical study reports on duloxetine date from September 2000 to September 2003 and total 13 729 pages. We obtained these documents in May 2011 as part of a wider request of access to reports on SSRIs and serotonin norepinephrine reuptake inhibitors. Duloxetine was the only centrally approved product (whereby a single application to the EMA can lead to a European Union wide marketing authorisation for a drug),[19] which is why we focused on this drug. We specifically chose to assess the coding of adverse events of suicidality (ideation, behaviour, attempts, and suicide) within these reports, given the FDA's findings of inconsistency in coding of suicidality in trials of SSRIs in young people, and ongoing public concern and scientific debate about suicidality in adults.[15 20]

One researcher used optical character recognition software to make searchable the 47 PDF documents, comprising the nine clinical study reports. Adobe Acrobat was used for all text portions. ABBYY Finereader enabled the efficient conversion of tables of harms into Excel spreadsheets; according to its manufacturer this software has an accuracy rate of 99.8%.[21]

Two observers, one an experienced medical coder (EM), independently did electronic searches in the clinical study reports of summary tables of coded adverse events, narratives of serious adverse events, narratives of discontinuations of the study drug as a result of adverse events, and listings of adverse events in individual patients. The box describes each data format, and table 3⇓ provides key features of narratives and of individual patient adverse event listings. (See supplementary appendices 1a-c for examples of each data format.)

Search terms included those that the FDA requested pharmaceutical companies to use when searching company databases for events of suicidality in paediatric trials ("suic", "overdos, "attempt", "cut", "gas", "hang", "hung", "jump", "mutilate", "overdos", "self damage", "self harm", "self inflict", "self injur", "shoot", and "slash").[15 22] We additionally used the terms "poi", "emot", "labi", "hos", "vio", "agg", "thought", and "think". We were only interested in adverse events that met the definition in the international statistical principles for clinical trials guideline of a treatment emergent adverse event—that is, an event that occurred or worsened after the study drug was started,[23] or adverse events that emerged after discontinuation of the study drug.

In an Excel spreadsheet we recorded the results of the searches, including which data format the term was found in, which study arm (investigational drug, active comparator, or placebo) the suicidality event occurred in, whether the term found was a verbatim or coded term, and the medical coding dictionary used in the trial. We then compared the extracted data and resolved discrepancies by consensus.

When a verbatim term was reported, one researcher (EM) consulted the medical coding dictionary used in the study and chose the closest matching lowest level term, and then the preferred term that was used in summary tables. As the lowest level terms were not available in the reports, we could not compare our choices with these, but we checked the preferred terms. We accessed COSTART version 5, which was released in 1995 and was the last version of COSTART, through the

Glossary of clinical study report related terms

- 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"[4]
- ICH (International Conference on Harmonisation) E3: ICH guidelines on the structure and content of clinical study reports[3]
- 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"[4]
- 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"[4]
- Summary tables: in a CSR "All adverse events occurring after the initiation of study treatment . . . should be displayed in summary tables . . . The tables should list each adverse event, the number of patients in each treatment groups in whom the event occurred, and the rate of occurrence"[3]
- 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."[3] Narratives are based on extracted data from source files (for example, case report forms). They are written by medical writers. Narratives can be written before data are finalised, but updates are required based on the final data[5]
- Appendices: CSRs include appendices on study information (for example, protocol and protocol amendments, sample case report forms, a list of institutional review boards/ethics committees, a 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 on 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 amendments to the protocol[3 6 7]
- 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[3]

website http://purl.bioontology.org/ontology/CST; MedDRA versions 2.1 to 16.0 were accessed electronically through an academic subscription.

For all trials we attempted to reconcile each suicidality event in the three data formats. Firstly, using the patient's trial identification number we were able to reconcile data reported in the patient listings with those in the narrative. Secondly, using data (treatment assignment, coded term, and timing of event) from the patient listings and narratives, we were able to reconcile data from these two formats with the data in summary tables.

Results

Six trials (1586 patients) used the coding dictionary COSTART (version number not provided) and three trials (1292 patients) used MedDRA (version 5.0 or 6.0). Adverse events listings for individual patients were available for all nine trials (1672 patients receiving duloxetine, 777 receiving placebo, 70 receiving fluoxetine, and 359 receiving paroxetine). These listings provided data on individual adverse events experienced by each patient in the trial and included the verbatim term, severity of the event, if the adverse event was serious or led to discontinuation of the study drug, and whether the adverse event was considered to be related to the study drug (see table 3). The listings did not, however, provide the preferred term of the event. Narratives were the only data format to provide both verbatim and preferred terms. We were therefore only able to compare verbatim terms to coded terms for those patients who had a narrative—that is, patients who experienced a serious adverse event, discontinued the study drug as a result of an adverse event, or had a clinically significant non-serious adverse event. A median of 11% of patients in each trial had a narrative. We also noted that the listings contained no information on action taken with the study drug in response to the adverse event—for example, dose reduction, the date that the study drug was stopped, whether the adverse event resolved on reducing the dose or stopping the drug, or whether the patient received any treatment for the adverse event.

Within the clinical study report, summary tables of adverse events for some of the trial phases were presented; the lead-in phase of 3-10 days without drugs was always missing. If patients experienced a specific adverse event in the randomised phase, its incidence for each arm was reported in the table. All events were presented as the preferred term.

Individual patient listings versus narratives

The listings of adverse events for individual patients described three suicides and three definitive suicide attempts, which were serious adverse events. The listings also showed a "possible suicide attempt," which was mild, non-serious, and did not result in the patient discontinuing the study drug.

There were narratives for the suicides and definitive suicide attempts, as these were all serious adverse events. From the narratives it could be discerned that verbatim terms were coded to identical terms.

No narrative was present for the patient who experienced a "possible suicide attempt," because the patient did not experience any serious or clinically important adverse events and did not discontinue the study drug as a result of an adverse event. Furthermore, there were no events, such as "overdose" mentioned for this patient in the individual patient listings that could possibly constitute a suicide attempt. Therefore, we did not have any information as to what the possible suicide attempt comprised.

The patient listings described 10 patients who experienced events relating to suicidal ideation (six receiving duloxetine, two receiving placebo, two receiving paroxetine), one patient receiving placebo who experienced "increased suicidality," and one patient receiving duloxetine who experienced "suicide threat" (see supplementary table 1).

Narratives were only available for six patients from three trials (as the other six patients did not experience any serious or clinically important adverse events or adverse events that led to discontinuation). In two of the three trials, COSTART was used. Narratives from these two trials showed that one event of "suicidal urges" while receiving duloxetine and two events of "suicidal ideation" while receiving paroxetine were coded as depression. According to the definitions of the International Conference on Harmonisation, adverse events can include pre-existing conditions that worsen after starting the study drug.[23] One of the two patients receiving paroxetine had a mild baseline "suicidal ideation" that worsened in severity in the randomised phase of the trial (see supplementary table 1), therefore meeting the criterion of an adverse event. However, "suicidal ideation" was only recorded in the narrative as a pre-existing condition, not as an adverse event. Furthermore, there was no mention in the narrative text of a worsening of the severity of suicidal ideation. The only suicidality preferred term in the last version of COSTART (version 5) is suicide attempt. There is no exact lowest level term for suicidal ideation in COSTART; the closest possible matching term is suicidal tendency, which is coded to the preferred term depression.

In the third trial, MedDRA version 6.0 was used, and two events of suicidal ideation were coded as the preferred term suicidal ideation. The event "suicidal threat" (the patient threatened to harm herself while in possession of a knife) was coded as suicidal ideation. Although more recent versions of MedDRA have an appropriate term (lowest level term preparatory actions towards imminent suicidal behaviour, which codes to the preferred term suicidal behaviour), version 6.0 did not.

We also found a case of suicidal ideation (see supplementary table 1) that did not appear in the patient listings but in the narrative text only, in a patient receiving duloxetine who experienced "worsened depression." This finding agrees with the common coding convention of only coding a definitive diagnosis and not its symptoms.

Individual patient listings and narratives versus tables

The three suicides could clearly be identified in the coded data presented in summary tables of adverse events in the clinical study reports. The three definitive suicide attempts and one "possible suicide attempt" came from one trial, and its summary table reported four suicide attempts.

Summary tables showed important loss of information on adverse events. Two of the 10 events related to suicidal ideation were coded as suicidal ideation using MedDRA version 6.0. We were only able to reconcile verbatim terms to coded terms for three of the nine other events in summary tables of trials using the COSTART dictionary.

In all three cases the original term reported by the investigator was coded to the COSTART preferred term depression. Two of these cases (one patient receiving duloxetine and the other receiving paroxetine) occurred in the randomised phase of one trial, and the summary table reported them as depression. The third case (patient receiving paroxetine) occurred in the randomised phase of a different trial where the summary table for the randomised phase reported depression while receiving paroxetine.

The event of "suicidal threat," where a patient receiving duloxetine threatened to harm herself while in possession of a knife, was coded to the preferred term suicidal ideation using MedDRA version 6.0, which was also the term used in the summary table.

We also found instances where events of suicidal ideation were present in patient listings but were absent from summary tables, and vice versa. In the patient listings of one trial there was an adverse event of suicidal ideation in a patient receiving paroxetine that met the criteria of a treatment emergent adverse event in the patient listings, but was only shown as a pre-existing condition, coded to depression, in the narratives. In the summary table there were zero adverse events of depression in the paroxetine arm. Furthermore, in one trial, which used the coding dictionary MedDRA, summary tables of coded data for the open label single arm run-in phase reported three events of suicidal ideation. From the patient listings and narratives, however, we could only identify two adverse events of suicidal ideation.

Discussion

We wanted to assess the effects of coding and coding conventions on summaries and tabulations of adverse events data on suicidality within clinical study reports. From the small number of suicidal events that we were able to reconcile, coding was both accurate, given the constraints of the dictionaries used, and consistent. The suicides were clearly identifiable in all formats of adverse events data whereas, in line with common coding conventions, suicide attempts in tables included both definitive and provisional diagnoses. However, some events of suicidal ideation and preparatory behaviour were obscured in tables owing to the lack of specificity in the coding dictionary used. Instances of suicidal ideation events were present in patient listings but were absent from summary tables, and vice versa. One event of suicidal ideation appeared in the narrative text only. This may result from the common coding convention that if symptoms and a definitive diagnosis are both provided, only the diagnosis is coded.

Strengths and limitations of this study

Our study is based on a small number of trials for a single drug manufactured by a single company.

Another limitation is that, although the guideline for clinical study reports suggests that listings of adverse events for individual patients should include coded terms in addition to the verbatim terms,[3] this was not the case for the nine trials we examined. Our analysis of discrepancies in adverse events data was therefore limited to comparing data already coded in tables to those of narratives, which included verbatim and coded terms, of those patients who had adverse events that were serious, led to discontinuation of the study drug, or were non-serious but clinically important.

Comparisons with other studies

Problems with terms in COSTART, including lack of specific preferred terms, were acknowledged in journal articles in the 1990s (the last version of COSTART was released in 1995).[24 25] It is therefore possible that our finding that adverse events of suicidal ideation were obscured in summary tables of COSTART coded data could also apply to other types of adverse events.

Conclusion and implications for researchers and clinicians

Our study has shown that researchers and clinicians need to be aware that because of coding dictionaries and coding conventions used, adverse events data presented in summary tables may obscure adverse events of importance. Furthermore, important data, in particular the verbatim terms of adverse events, can be presented in the patient listings in the appendices

of clinical study reports. To obtain a more accurate estimate of the incidence of specific adverse events, the verbatim terms should be recoded with the latest version of MedDRA. This is in agreement with informal advice from the FDA and from the MedDRA Maintenance and Support Services Organization. However, researchers contemplating using clinical study reports as sources of data for adverse events need to be aware that access to verbatim terms may not be possible, as the individual patient listings of all adverse events, in contrast with serious adverse events, is not a mandatory part of the submission to the EMA.[6] Furthermore, patient listings may not contain information on certain events (owing to coding conventions), or important information, such as action taken with the study drug and treatments given in response to adverse events. Important adverse events data may therefore only be available in the narratives of patients who experienced adverse events that were serious, led to discontinuation of the study drug, or were non-serious but clinically important.

It should also be noted that, while clinical study reports contain detailed data on adverse events, there is evidence from FDA analyses and court cases that access to case reports forms reveal discrepancies that would not be apparent from clinical study reports alone.[26 27] For example, an FDA analysis of a sample of case report forms from the RECORD trial revealed many missing cases of cardiac problems, which allowed the determination that, in contrast to the manufacturer's (GlaxoSmithKline) claims, rosiglitazone increased the risk of cardiac problems fourfold.[26] Furthermore, case report forms are sometimes unavailable to, or rarely used by, academic authors of journal articles reporting industry sponsored trials. Readers of journal articles should therefore be aware that academic authors often only use data files of coded data or coded data from clinical study reports to perform or check analyses presented in journal articles.[28] Case report forms are currently unavailable to independent researchers. Should case report forms become available, any independent research using case report forms is likely to be costly, in terms of both time and money, as a case report form for a single patient can be hundreds of pages in length and require a considerable infrastructure to ensure unbiased judgments.[26]

In conclusion, adverse event data in tables in clinical study reports may not accurately represent the underlying patient data owing to medical coding dictionaries and coding conventions used. In clinical study reports, the individual patient listings of harms and narratives of adverse events can provide important additional data, including the original terms for adverse events reported by the investigators, which can enable a more accurate estimate of harms.

We thank Julie Borring, Kristine Rasmussen, and Trine Gro Saida for assistance with initial data extraction; the EMA for providing the material and for responding to queries related 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 and BT contributed to the study concept and design. EM, BT, AH and AL 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 and 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 (available on request from the corresponding author) and declare: this study is part of a PhD funded by Rigshospitalets Forskningsudvalg; 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.

Data sharing: The clinical study reports we used can be obtained from the authors (em@cochrane.dk).

Transparency: 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.

1 Doshi P, Jefferson T, Del Mar C. The imperative to share clinical study reports: recommendations from the Tamiflu experience. *PLoS Med* 2012;9:e1001201.
2 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.
3 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Structure and content of clinical study reports: E3. 1995. www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf.
4 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Guideline for good clinical practice: E6 (R1). 2013. www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6_R1/Step4/E6_R1__Guideline.pdf.
5 Challenges of writing patient/ safety narratives—part 1. 2014. http://info.quanticate.com/bid/56497/Challenges-of-Writing-Patient-Safety-Narratives-Part-1.
6 Committee for medicinal products for human use (CHMP). Note for guidance on the inclusion of appendices to clinical study reports in marketing authorisation applications (CHMP/EWP/2998/03 Final). 2004. www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003638.pdf.
7 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. 2013. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:311:0067:0067:EN:PDF.
8 Chow S-H, Lui J-P. Design and analysis of clinical trials: concepts and methodologies. 2nd ed. Wiley-Interscience, 2004.
9 Brown EG. Dictionaries and coding in pharmacovigilance. In: Talbot J, Waller P, eds. Stephens detection of new adverse drug reaction. 5th ed. Wiley, 2004:533-57.
10 White CA. A preliminary assessment of the impact of MEDDRA on adverse event reports and product labelling. *Drug Inf J* 1998;32:347-62.
11 Kornai A, Stone L. Automatic translation to controlled medical vocabularies. In: Abraham A, Jain LC, van der Zwaag BJ, eds. Innovation in intelligent systems. Springer, 2004:413-34.
12 MedDRA. History. 2013. www.meddra.org/about-meddra/history.
13 Brown DR, Brown EG, Moulvad TB. A comparison of two medical terminologies in coding and analysing clinical trial safety data. *Int J Pharm Med* 1997;11:85-9.
14 ICH Secretariat. MedDRA term selection: points to consider release 4.6. 2014. www.meddra.org/sites/default/files/guidance/file/9491-1610_termselptc_r4.6_sep2013.pdf.
15 Posner K, Oquendo MA, Gould M, Stanley B, Davies M. Columbia Classification Algorithm of Suicide Assessment (C-CASA): classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants. *Am J Psychiatry* 2007;164:1035-43.
16 Peters PH, Gravenstein S, Norwood P, De Bock V, Van Couter A, Gibbens M, et al. Long-term use of oseltamivir for the prophylaxis of influenza in a vaccinated frail older population. *J Am Geriatr Soc* 2001;49:1025-31. [Correction in *J Am Geriatr Soc* 61:478.]
17 Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *BMJ* 2012;344:e1119.
18 Scientific Discussion EMA 2005. 2013. www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000572/WC500036776.pdf.
19 Central authorisation of medicines. 2014. www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000109.jsp.
20 Selective serotonin reuptake inhibitors and serotonin and noradrenaline reuptake inhibitors information for healthcare professionals. 2014. www.mhra.gov.uk/Safetyinformation/Generalsafetyinformationandadvice/Product-specificinformationandadvice/index.htm.
21 ABBYY FineReader 11 Professional edition—full feature list. 2013. http://finereader.abbyy.com/professional/full_feature_list/.
22 Hammad TA, Laughren T, Racoosin J. Suicidality in pediatric patients treated with antidepressant drugs. *Arch Gen Psychiatry* 2006;63:332-9.
23 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Statistical principles for clinical trials: E9. 1998. www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf.
24 Fizames C. How to improve the medical quality of the coding reports based in Who-Art and Costart use. *Drug Inf J* 1997;31:85-92.
25 Matsumoto C. Implementation and use of the Costart dictionary at Lilly: would we do it again. *Drug Inf J* 1991;25:201-3.
26 US FDA. Advisory Committee Meeting for NDA 21071 Avandia (rosiglitazone maleate tablet). 2010.
27 Healy D. Let them eat Prozac. New York University Press, 2004.
28 Lundh A, Krogsbøll LT, Gøtzsche PC. Access to data in industry-sponsored trials. *Lancet* , 2011;378:1995-6.

What is already known on this topic

For statisticians to analyse adverse events recorded in a clinical trial, it is necessary that events described by the original investigators are coded to terms in a specialised medical coding dictionary

Miscoding of harms can prevent an accurate risk assessment of harms

Extensive coded data on adverse events, in different summaries and tabulations, and provision of original investigator reported terms, can be found in clinical study reports submitted in drug licensing applications to the regulatory authorities

What this study adds

The use of coding dictionaries and coding conventions may inadvertently obscure events that are important in summary tables

Individual patient listings of harms and narratives of adverse events can provide important additional data, including original investigator reported descriptions of the adverse events, which can enable a more accurate estimate of harms

Accepted: 5 May 2014

Cite this as: *BMJ* 2014;348:g3555

Tables

Table 1| MedDRA hierarchical structure

No of terms	Level of term	Example
26	System organ class (SOC)	Psychiatric disorder
>330	High level group terms (HLGT)	Suicidal and self injurious behaviour NEC
>1700	High level terms (HLT)	Suicidal and self injurious behaviour
>20 000	Preferred terms (PT)	Suicidal ideation
>70 000	Lowest level terms (LLT)	Suicidal tendency, active suicidal ideation, death wishes, life weariness, passive suicidal ideation, suicidal ideation, suicidal intention, suicidal plans

MedDRA=Medical Dictionary for Regulatory Activities; NEC=not elsewhere classified.

Table 2| COSTART hierarchical structure[8 9]

No of terms	Level of term	Example
12	Body system	Nervous system
	Mid-level classification	Depression
About 1200	Coding symbol (analogous to preferred term)	Depression
>6000	Glossary terms (used to assist in selection of coding symbol, analogous to lowest level terms of MedDRA)	Suicidal tendency, depression agitated, depression mental, melancholia, depression worsened, depressed reaction, oppression, dejection emotional, moroseness, depression reactive, dysphoria, depression aggravated, depression functional, depression psychic, depression neurotic, anhedonia, depressed state

COSTART=Coding Symbols for a Thesaurus of Adverse Reaction Terms.
Examples using http://bioportal.bioontology.org/ontologies/COSTART.

Table 3| Variables in individual patient listings for adverse events, and in narratives

Variables	ICH E3 specified variables for IP AE listing*	Variables in DLX CSR IP AE listing	ICH E3 specified variables for narrative*	Variables in DLX CSR narratives
Investigator	Yes	Yes	—	—
Treatment group	Yes	Yes		Yes
Patient identifier	Yes	Yes	Yes	Yes
Age, race, sex, weight	Yes	—	Yes	Yes
Location of CRFs, if provided	Yes	—	—	—
The adverse event (preferred term, reported term)	Yes	Reported term only	—	Yes (preferred and reported terms)
Duration of adverse event	Yes	Yes*	—	Yes
Severity	Yes	Yes	—	Only for adverse events that led to discontinuation
Seriousness (serious/non-serious)	Yes	Yes	—	Yes
Action taken (none, dose reduced, treatment stopped, specific treatment instituted, etc)	Yes	—	—	Yes†
Outcome (for example, CIOMS format)	Yes	Start and stop dates of event provided	—	Yes*
Causality assessment (for example, related/not related)	Yes	Yes	Yes	Yes
Date of onset or date of clinic visit at which event was discovered	Yes	Yes	—	Yes
Timing of onset of adverse event in relation to last dose of test drug/investigational product (when applicable)	Yes	Yes	—	Yes
Study treatment at time of event or most recent study treatment taken	Yes	—	—	Yes
Test drug/investigational product dose in absolute amount, mg/kg or mg/m² at time of event	Yes	—	—	Yes
Drug concentration (if known)	Yes	—	—	—
Duration of test drug/investigational product treatment	Yes	—	—	Yes
Concomitant treatment during study	Yes	—	—	Yes
Date study drug started	—	Yes	Yes	Yes
Relevant concomitant/previous illnesses with details of occurrence/duration	—	Yes	Yes	Yes
Relevant concomitant/previous drug with details of dosage	—	—	Yes	Yes
Relevant laboratory measurements	—	—	Yes	Yes

CIOMS=Council for International Organizations of Medical Sciences guidelines; CRF=case report form; CSR=clinical study report; DLX=duloxetine; ICH E3=International Conference on Harmonization of technical requirements for registration of pharmaceuticals for human use: structure and content of clincal study reports: E3; IP AE=individual patient adverse event.

*Can be calculated from information in data format.

†Apparent from information in text.

Supplementary table 1: Adverse events of suicidal ideation and preparatory behaviour

Study ID	Dictionary	Drug	Phase	IP AE listings ("verbatim term", severity)	Narrative ("verbatim term" and *coded term*)	Summary tables[a] (*coded term*)	Issue
HMATa	COSTART	DLX 20mg BID	Acute (randomised)	"suicidal urges", severe	"suicidal urges" coded as *depression* Excerpt from narrative text: "Suicidal urges began [number of days] after randomization. Dissociative feelings began [number of days] after randomization. Patient stated that dissociative feelings contributed to [their] suicidal thoughts."	DLX 20mg BID: 2 events of *depression*	Loss of information in summary data, as cannot determine from summary table alone that one event of depression was actually an event of "suicidal urges"
		PRX	Acute (randomised)	"suicidal ideation", mild	"suicidal ideation" coded as *depression* Excerpt from narrative text: "The patient expressed suicidal ideation that lasted one day on [date], [number of days] after randomization to paroxetine"	PRX: 1 event of *depression*	Loss of information in summary data, as cannot determine from summary table alone that the single event of depression in the paroxetine arm was actually suicidal ideation
HMATb	COSTART	PRX	Acute (randomised)	"suicidal ideation", moderate (mild "suicidal ideation" was also listed as pre-existing condition)	"suicidal ideation" coded as *depression* Date of onset shows suicidal ideation as a pre-existing condition. Excerpt from narrative	PRX: 0 events of *depression*	TEAE in IP AE listing is not found in narratives or summary tables

					text: no mention of suicidal ideation in narrative text.		
		DLX 20mg BID	Placebo lead-out	"passive suicidal ideation", mild	No narrative[b]	[c]	[c]
HMBHa	COSTART	DLX	Acute (randomised)	"suicidal ideation", moderate	No narrative[b]	[c]	[c]
		PBO	Acute (randomised)	"increased suicidality", moderate	No narrative[b]	[c]	[c]
HMBHb	COSTART	DLX	Acute (randomised)	"fleeting suicidal thoughts", moderate	No narrative[b]	[c]	[c]
		PBO	Acute (randomised)	"suicidal ideation", moderate (mild suicidal ideation listed as pre-existing condition)	No narrative[b]	[c]	[c]
		PBO	Acute (randomised) and placebo lead-out	"life is not worth living thoughts", moderate (mild "life is not worth living thoughts" listed as pre-existing condition.)	No narrative[b]	[c]	[c]
HMAYb	MedDRA version 5.0	DLX 60mg BID	Continuation (non-randomised)	"worsening of depression", severe (No events of "suicidal ideation "and "labile mood" listed)	"worsening of depression" coded as *depression aggravated* Excerpt from narrative text: "On [date], the patient was hospitalized for the worsening of depressive	DLX 60mg BID: 1 event of *depression aggravated*	Non-coding of labile mood and suicidal ideation agrees with common coding conventions, i.e. symptoms not coded as adverse events because a definitive diagnosis

					symptoms...while hospitalized the patient was treated with [list of medications]. The patient's mood was reportedly labile but the depression improved and suicidal ideations disappeared during hospitalization".		was also provided. Example that due to coding conventions some adverse events may be found in narratives only.
HMBC	MedDRA version 6.0	DLX	Open label single arm run-in phase	"suicidal ideation", severe	"suicidal ideation" coded as *suicidal ideation* Excerpt from narrative text: "[the patient] was hospitalized in the inpatient mental health unit for suicidal ideation....".	DLX: 3 events of *suicidal ideation*	Narratives and IP AE listing only provide information on 2 events of suicidal ideation during open label single arm run-in phase.
		DLX	Open label single arm run-in phase	"psychiatric hospitalization for suicidal ideation", severe	"psychiatric hospitalization for suicidal ideation" coded as *suicidal ideation* Excerpt from narrative text: "the patient was hospitalized for suicidal ideation....The patient stated [their] suicidal ideation was based upon anger and frustration with [their] family"	DLX: 3 events of *suicidal ideation*	Narratives and IP AE listings only provide information on 2 events of suicidal ideation during run-in open label single arm phase
		DLX	Continuation (randomised)	"suicide threat", severe	"suicide threat" coded as *suicidal ideation* Excerpt from narrative text: "the patient	DLX: 1 event of *suicidal ideation*	Coding was accurate with MedDRA version available at that time. However, contemporary recoding of event with latest version of

				experienced auditory hallucinations, increased depressive symptoms, and made suicidal threats. Early that AM, the patient threatened to harm [themselves] while in the possession of a knife....”		MedDRA (as advised by regulators and MSSO when reanalysing legacy data) would code this event as *preparatory actions toward imminent suicidal behaviour*

BID=twice daily; CSR=clinical study report; DLX=duloxetine; IP AE=individual patient adverse event; MedDRA=Medical Dictionary for Regulatory Activities; MSSO=MedDRA Maintenance and Support Services Organization; PBO=placebo; PRX=paroxetine; TEAE=treatment-emergent adverse event (events that occurred or worsened after the study drug was started)

[a] Summary tables that show frequency of adverse events for the relevant trial phase e.g. acute (randomised) phase;

[b] There was no narrative as the patient did not experience any adverse events that were considered serious or clinically significant, or that led to discontinuation from the trial;

[c] There was no narrative for this patient and therefore it was not possible to definitively determine what the verbatim term was coded to. Consequently, it was not possible to definitively reconcile data in individual patient adverse event listings to that in summary tables.

Appendices 1a-c

Examples of different summaries and tabulations of harms data within clinical study reports

Appendix 1a

1 page excerpt from individual patient adverse event listing clinical study report appendix

Note: There were no redactions in the material received from the European Medicines Agency. Redactions were made by the authors for publication purposes.

Listing of Adverse Events and Pre-Existing Conditions
F1J-MC-HMAQ(A)
All Randomized Subjects

Protocol: F1J-MC-HMAQ

Inv	Pat	Therapy Group	Therapy Start Date	Visit	Actual Term	Event Start Date	Event Stop Date	Start Rel. Day of Ther.	Stop Rel. Day of Ther.	Event Severity	Event Cause Disc.	Serious Criteria	Poss. Rel. to Study Drug
[redacted]					[redacted]			.	.	Mild	NO	NONE	
					STOMACH "BURNING"					Mild	NO	NONE	No
					STOMACH "BURNING"					Mild	NO	NONE	No
					IRRITABILITY					Moderate	YES	NONE	Yes
					TENSENESS					Moderate	NO	NONE	Yes
					INCREASED DEPRESSION					Moderate	NO	NONE	Yes
					DRY MOUTH					Moderate	NO	NONE	No
					INCREASED CRYING					Moderate	NO	NONE	Yes
					S KNEE PAIN			.	.	Moderate	NO	NONE	
					S BACK PAIN			.	.	Moderate	NO	NONE	
					HYPERTENSION					Moderate	NO	NONE	Yes
					HYPERTENSION					Moderate	NO	NONE	Yes
					HYPERTENSION					Moderate	NO	NONE	Yes
					HYPERTENSION					Mild	NO	NONE	Yes
					HYPERTENSION					Moderate	NO	NONE	Yes
					HYPERTENSION					Moderate	NO	NONE	Yes
					HYPERTENSION					Moderate	NO	NONE	Yes
					S SEASONAL ALLERGIES			.	.	Moderate	NO	NONE	
					S TENSION HEADACHES			.	.	Moderate	NO	NONE	
					S INSOMNIA			.	.	Severe	NO	NONE	
					S MIGRAINE HEADACHES			.	.	Severe	NO	NONE	

Output stored as RMP.F1JO.HMAQ.FINALA(AE35100Z)
Data from RMP.SAS.F1JM.MCHMAQSW.STUDY1
An 'S ' preceding the Event Term (Class. or Actual) denotes that the record is a secondary condition.
Code: CA-Congenital Anomaly CN-Cancer DI-Died HO-Hospitalized LT-Life-threatening OD-Overdose PD-Permanently Disabled
OTH-Other Serious Criteria
XAEL0003

Appendix 1b

1 page excerpt of a narrative from a clinical study report

Note: There were no redactions in the material received from the European Medicines Agency. Redactions were made by the authors for publication purposes.

Protocol: F1J-MC-HMAQ Discontinued Due To Adverse Event
Patient : [redacted]

Study Medication: Randomized on [redacted]

Age: [redacted] years Sex: [redacted]
Origin: [redacted] Weight: [redacted] kg

Event: Actual Term - IRRITABILITY
Class Term - AGITATION
Severity - Moderate
Serious? - No

Onset from initial therapy: [redacted]
Days of therapy: [redacted]

Historical Diagnoses
[redacted]

Secondary Conditions
[redacted]

Abnormal Labs

Analyte	Result		Units	Lo Limit	Hi Limit	Sample Date
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Adverse Events

Classification Term (Actual)	Onset	Relative Onset	Duration
DYSPEPSIA (STOMACH "BURNING")	[redacted]	[redacted]	[redacted]
AGITATION (IRRITABILITY)	[redacted]	[redacted]	
ANXIETY (TENSENESS)	[redacted]	[redacted]	
DEPRESSION (INCREASED DEPRESSION)	[redacted]	[redacted]	
DRY MOUTH (DRY MOUTH)	[redacted]	[redacted]	[redacted]
EMOTIONAL LABILITY (INCREASED CRYING)	[redacted]	[redacted]	

* denotes event was classified as serious.

Concomitant Medications

Drug	Start Date	Stop Date
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Summary Paragraph

Patient [redacted], a [redacted] year old [redacted] was discontinued on [redacted] from the study due to the adverse event irritability. Patient had the adverse events irritability, tenseness, increased crying, and increased depression [redacted] days after randomization to [redacted] therapy. Patient was taking [redacted] at time of event. Last dose of study drug was on [redacted] [redacted]days after randomization to therapy. Patient discontinued participation in the study to begin clinical treatment. Principal investigator considered the event to be related to study drug.

Appendix 1c

1 page excerpt from a table of treatment-emergent adverse events presented in a clinical study report

Table HMAQa.12.3. Treatment-Emergent Adverse Events By Decreasing Frequency All Randomized Patients Acute Therapy Phase

	Placebo N=70	Dulox N=70	Fluox N=33	Total N=173	p-Values* Overall	(1) vs.(2)	(1) vs.(3)	(2) vs.(3)
PATIENTS WITH ANY TESS	58(82.9)	62(88.6)	30(90.9)	150(86.7)	.538	.469	.376	1.00
HEADACHE	22(31.4)	14(20)	11(33.3)	47(27.2)	.205	.175	1.00	.149
DRY MOUTH	12(17.1)	21(30)	7(21.2)	40(23.1)	.196	.110	.599	.477
RHINITIS	12(17.1)	11(15.7)	5(15.2)	28(16.2)	1.000	1.00	1.00	1.00
DIARRHEA	7(10)	10(14.3)	10(30.3)	27(15.6)	.039	.606	.020	.066
SOMNOLENCE	7(10)	13(18.6)	7(21.2)	27(15.6)	.202	.227	.135	.793
NAUSEA	9(12.9)	9(12.9)	6(18.2)	24(13.9)	.745	1.00	.553	.553
INSOMNIA	5(7.1)	14(20)	3(9.1)	22(12.7)	.070	.046	.709	.255
SWEATING	6(8.6)	13(18.6)	3(9.1)	22(12.7)	.213	.137	1.00	.258
ASTHENIA	3(4.3)	12(17.1)	5(15.2)	20(11.6)	.037	.026	.107	1.00
DIZZINESS	5(7.1)	11(15.7)	2(6.1)	18(10.4)	.180	.183	1.00	.216
CONSTIPATION	4(5.7)	8(11.4)	5(15.2)	17(9.8)	.265	.366	.141	.751
DYSPEPSIA	7(10)	4(5.7)	5(15.2)	16(9.2)	.258	.532	.515	.141
PHARYNGITIS	4(5.7)	6(8.6)	3(9.1)	13(7.5)	.803	.745	.677	1.00
ANOREXIA	3(4.3)	7(10)	2(6.1)	12(6.9)	.428	.326	.654	.715
ABDOMINAL PAIN	5(7.1)	3(4.3)	3(9.1)	11(6.4)	.513	.718	.709	.382
AMBLYOPIA	3(4.3)	4(5.7)	3(9.1)	10(5.8)	.631	1.00	.382	.677
PAIN	2(2.9)	4(5.7)	3(9.1)	9(5.2)	.445	.681	.324	.677
PALPITATION	5(7.1)	3(4.3)	1(3)	9(5.2)	.739	.718	.661	1.00
ANXIETY	4(5.7)	3(4.3)	1(3)	8(4.6)	1.000	1.00	1.00	1.00

(1) = Placebo, (2) = Dulox, (3) = Fluox

(Continued)

Paper 3: benefits and harms of duloxetine for the treatment of stress urinary incontinence

Benefits and suicidality related harms of duloxetine for the treatment of stress urinary incontinence: a meta-analysis of clinical study reports.

(Submitted)

Nordic Cochrane Centre, Rigshospitalet Dept 7811, Copenhagen, Denmark
Emma Maund
PhD student

Nordic Cochrane Centre, Rigshospitalet Dept 7811, Copenhagen, Denmark
Louise Schow Jensen
Medical student

Nordic Cochrane Centre, Rigshospitalet Dept 7811, Copenhagen, Denmark
Peter C Gøtzsche
Professor

Correspondence to: E Maund em@cochrane.dk

Abstract

Objective: To determine the benefits of duloxetine for stress urinary incontinence, and its harms in terms of depression, suicidality, violent behaviour and their potential precursors.

Design: Meta-analysis of the four randomised placebo controlled trials of duloxetine submitted to the European Medicines Agency (EMA) for marketing approval for stress urinary incontinence in women. We obtained clinical study reports, including protocols and adverse event listings for individual patients as appendices (6,870 pages) from the EMA.

Participants: 958 women receiving duloxetine 80mg daily and 955 women receiving placebo.

Main outcomes: Benefits were incontinence episode frequency and incontinence quality of life. Harms in general were serious adverse events, discontinuations because of adverse events, and number of patients experiencing at least one treatment emergent adverse event. Specific harms of interest were suicidality, violent behaviour, and their potential precursors (akathisia, activation, emotional disturbance, and psychotic behaviour) and depression.

Results: Duloxetine was better than placebo for percent change in weekly incontinence episodes (N = 1738, mean difference (MD) -13.6% (95% confidence interval -21.6% to -5.5%) and mean change in Incontinence Quality of Life total score (MD 3.2, 2.0 to 4.5). However, the benefits were so small, standardized mean differences of -0.13 (-0.22 to -0.04) and 0.24 (0.15 to 0.33), respectively, that they were not clinically relevant and could be fully explained by unblinding bias. Duloxetine increased the risk of experiencing at least one adverse event, risk ratio (RR) 1.32 (1.24 to 1.41) and the risk of discontinuing because of an adverse event, RR 5.73 (4.00 to 8.20). There were no events of suicidality, violence or akathisia. Duloxetine increased the risk of experiencing emotional disturbance, RR 4.73 (1.62 to 13.85) and the risk of experiencing a core or potential activation event, RR 4.45 (3.22 to 6.14), number needed to harm 7 (6 to 9). There were also more core or potential psychotic events, RR 2.25 (1.06 to 4.81), number needed to harm 80 (40 to 834) and more depression related events, but the risk was not significantly increased, RR 1.26 (0.58 to 2.71).

Conclusions: Given its unlikely clinically meaningful benefit and its many important harms, we question the rationale for using duloxetine for stress urinary incontinence. Individual patient data contained in appendices of clinical study reports are essential for a reliable assessment of drug harms.

Introduction

In 1990, the first case reports of selective serotonin reuptake inhibitors (SSRIs) induced suicidality in adults were published.[1] Subsequently, suicidality and violent behaviour associated with the use of SSRIs have been reported among patients receiving treatment for psychiatric disorders and non-psychiatric disorders, and also among healthy volunteers.[2] Putative causes for such behaviour include SSRI induced akathisia (extreme restlessness), activation (e.g. insomnia, nervousness, anxiety, agitation), emotional blunting and psychotic events.[2,3]

In 2006, the FDA carried out a large meta-analysis and found that the risk of suicidality declined with increasing age.[4] However, the risk may be far greater than that found by the FDA as drug companies have underrecorded suicide attempts, non-fatal self harm and suicidal ideation,[5,6] and as the methods the FDA used underestimated the harms, e.g. by only including events that occurred during randomisation and the first 24 hours after the patients stopped taking the drug.[7] The FDA has warned that patients of *all* ages should be monitored for "clinical worsening, suicidality, and unusual changes in behavior".[8]

The serotonin-norepinephrine reuptake inhibitor duloxetine was granted marketing authorization for the treatment of stress urinary incontinence in the EU in 2004 under the trade name Yentreve.[9] Eli Lilly withdrew its US application in 2005, as the FDA was not positive to approving it.[10] In contrast to the EMA,[11,12] the FDA does not usually publish its reasons why applications are denied or withdrawn,[13] but it did say that a higher than expected rate of suicide attempts was observed in the open label extensions of the controlled studies.[14]

Drug agencies require companies to submit clinical study reports of all completed and ongoing studies of the drug in proposed and non-proposed indications.[15] These documents contain considerably more data on harms than journal articles and online registry reports.[16] We assessed the benefits and harms of duloxetine in the four such reports of stress urinary incontinence submitted to the EMA for marketing authorization. With respect to harms, we focused on depression, suicidality, violent behaviour, and their potential precursors (akathisia, activation, emotional disturbance, and psychotic behaviour).

Methods

Data

In 2011, we obtained from the EMA the clinical study reports for four randomised placebo controlled trials of duloxetine for the treatment of stress urinary incontinence in women.[17] The reports were from 2001 and 2002 and totaled 6,870 pages, including protocols. The documents were provided as non-searchable pdfs but we made them searchable using Adobe Acrobat Pro XI.

Outcomes

Benefits

The *a priori* outcomes specified in our protocol were the primary outcomes of each trial. We had also planned to look at Patient Global Impression of Improvement (PGI-I) as an outcome, but dropped it as we felt it would not add anything to the two outcomes we focused on.

Harms

The harms data specified *a priori* were divided into general harms, and harms related to suicidality and violence.

General harms were: deaths, non-fatal serious adverse events (any adverse event that was life threatening, required initial or prolonged inpatient hospitalization, caused severe or permanent disability, congenital anomaly, or were significant for other reasons), discontinuations because of adverse events, and number of patients experiencing at least one treatment-emergent adverse event.

Suicidality and violence related harms were: suicidality (ideation, behaviour, suicide attempts, suicide), violent behaviour and their potential precursors (akathisia, emotional disturbance, psychotic events, activation), depression or worsening of depression.

Search terms

Terms for suicidality were those the FDA requested the pharmaceutical companies to use when searching company databases (see table 1).[4] For violence, the terms were those used in a study to determine the association of prescription drugs with violence using FDA Adverse Event Reporting System data (see table 1).[18]

We focused on akathisia, emotional disturbance and psychotic events because these events are known as the "psychotropic suicidogenic triumvirate", and can predispose to suicidality and violence. [2,3] We also recorded activation symptoms (including akathisia) as these are included in warnings in FDA product labelling.[8]

FDA activation terms were obtained from approved product labels;[8] terms for other potential precursors to suicidality and violence were obtained from the literature.[3,19,20] A systematic review has shown there is lack of consensus on what the symptoms of activation are,[19] and we were uncertain about some events, e.g. whether nightmares, which can be a prelude to a psychotic event, should be treated as psychotic events.[21] We therefore consulted with a professor in psychiatry and when we were uncertain, we preliminarily categorised the events as "potential". We also created a second category of activation events (see table 1).

Data extraction

Benefits

Data were extracted from summary tables. For each study arm, one observer (LSJ) extracted the number of patients randomised and analysed, means and standard deviations. Extracted data were checked by a second observer (EM).

Harms

Data extraction was blinded. All drug names were redacted from all data formats of harms including pre-existing conditions in individual patients by one observer (EM) using the white redaction tool in Adobe Acrobat Pro XI. Additionally, narrative texts were placed in Word documents, replacing all drug names (including dosages) and also placebo by the generic term “drug X” to avoid the possibility that the identity of the drug could be guessed from the number of missing characters.

Two observers (EM and LSJ) independently searched all data formats of harms manually, using the terms listed in table 1. Each observer recorded patient identification number, date randomised,

adverse event term and data format (e.g. in a listing of all adverse events), onset and stop date of event, severity, whether the event was serious or led to discontinuation, and whether the term was the original investigator reported term ("verbatim term") or Medical Dictionary for Regulatory Activities (MedDRA) preferred term. MedDRA is a hierarchical medical terminology used to standardise data entry, retrieval, analysis and display of adverse events data.[22,23] Verbatim terms are coded to the closest matching lowest level terms in MedDRA. These lowest level terms are aggregated at the next level into preferred terms, which are the favoured terms for use in submissions to regulatory authorities.[24]

Two observers (EM and LSJ) independently recoded preferred terms (and if available, also verbatim terms) using the most recent version of MedDRA (version 17.0). Inter-observer agreement was calculated, and discrepancies were resolved by consensus. To ensure we had identified all the relevant terms, we carried out electronic searches on all the blinded documents using all the terms we had identified. The data were then unblinded.

Post hoc decisions

We moved the activation event of tension from the "potential" subcategory to the core subcategory of activation, as it belongs to the MedDRA high level term of anxiety symptoms, and anxiety was a core event of activation. We also added feeling abnormal, which wasn't included in any of the original categories, to emotional disturbance on the basis of the original investigator reported term, e.g. fuzzy feeling. Finally, events of dysthymic disorder and depressed mood were added as potential depression events.

Statistical analysis

For each outcome, data were combined in a meta-analysis. For binary outcomes, we calculated risk ratios and risk differences, and for continuous outcomes mean differences, with 95% confidence intervals using a fixed effect model, as it gives more weight to large trials. If heterogeneity was substantial (I^2 above 50%), the reasons were explored in sensitivity analyses. Meta-analyses were performed in RevMan, which adds 0.5 to cells with zero events.[25] This does not cause bias if the study arms are of equal size.[26]

For harms, only treatment emergent adverse events, i.e. those that began or worsened during the randomised phase, were of interest. Where categories of suicidality and violence related harms consisted or core and potential events, sensitivity analyses were performed by only using core events. For categories which contained insomnia, sensitivity analyses were performed by excluding events that were not definitively insomnia, e.g. poor quality sleep and sleep disorder.

For benefits, as specified *a priori* in the protocol, we determined if the results were clinically relevant by comparing with the minimum clinically important difference for the primary outcome as stated in the literature.

Post hoc, for continuous outcomes of benefits, we also calculated standardized mean differences, and for binary outcomes of harms, we calculated number needed to harm (NNH).

Results

Overall, 958 women with stress urinary incontinence were randomised to duloxetine 80 mg and 955 to placebo (in one trial, the starting dose was 40 mg) (weighted average age 52 years). In all trials,

use of antidepressants within 14 days prior to trial entry or during the trial was an exclusion criterion. During two weeks without medication, the patients completed daily diaries about voluntary and involuntary urination. Patients who completed the diaries and met the inclusion criteria, entered a placebo lead-in period of two weeks, followed by a 12-week randomised phase.

The clinical study reports contained trial protocols, summary tables of adverse events, listings and narratives of serious adverse events or discontinuations because of adverse events, and adverse event listings for individual patients as appendices. There were no examples of case report forms.

Both protocols and clinical study reports specified that adverse events data would be collected at randomisation, and at study visits every four weeks thereafter. However, none of the sources specified *how* they would be ascertained. In the publications of the trials, it was stated that adverse events were ascertained through non-probing questions.[27-30]

All formats of harms data presented MedDRA preferred terms. Narratives were the only format to report verbatim terms.

Benefits

Data were only shown for patients with baseline and at least one post-baseline value, and the method used for missing values was Last Observation Carried Forward.

Duloxetine was significantly better than placebo for percent change from baseline in weekly incontinence episodes (N = 1738, mean difference (MD) -13.6%, 95% CI: -21.6% to -5.5%, I^2 = 42%) and for number of weekly incontinence episodes (MD -2.9, -3.9 to -1.8, I^2 = 27%) (see figures 1 and 2). In both instances, however, the effects sizes were small, standardized mean difference (SMD) -0.13, -0.22 to -0.04, I^2 = 64%, and -0.26, -0.35 to -0.16, I^2 = 0% (figures 3 and 4).

Duloxetine was also better than placebo for mean change in Incontinence Quality of Life total score, MD 3.2, 2.0 to 4.5, I^2 = 5% (figure 5), but again, the effect size was small, SMD 0.24, 0.15 to 0.33, I^2 = 0% (figure 6).

Harms

Inter-observer agreement

In the four trials, an average of 22% of duloxetine patients and 5% of placebo patients experienced serious adverse events or discontinued because of adverse events, and therefore had narratives. A total of 96 patients with 139 adverse events of our interest had a narrative, and when two of us (EM and LSJ) independently recoded the verbatim terms using MedDRA, there was excellent inter-observer agreement (lower level terms: kappa = 0.92; preferred terms: kappa = 0.99).

Adverse events in general

On cross referencing summary tables, line listings and narratives we found one serious adverse event in a patient receiving duloxetine and two in a patient receiving placebo, and adverse events that led to discontinuations in six patients in the duloxetine group and four in the placebo group, which began prior to randomisation and did not worsen in severity. These events therefore did not meet the criteria for treatment emergent adverse events and were excluded from our analyses. 727 patients receiving duloxetine had one or more adverse events compared to 548 receiving placebo,

RR 1.32 (1.24 to 1.41, I^2=51%), risk difference (RD) 18.5% (14.4% to 22.7%; NNH 6, 5 to 7 (figures 7 and 8). The risk of discontinuing because of an adverse event was over five times higher (RR 5.73, 4.00 to 8.20, I^2=26%; RD 16.4%, 13.6% to 19.1%; NNH 7, 6 to 8 (figures 9 and 10), and the risk of experiencing a non-fatal serious adverse event was also higher (RR 1.77; 0.79 to 3.98, I^2=0%) (figure 11) although the difference was not statistically significant.

Deaths, violence and suicidality

One patient who received duloxetine died after a cerebrovascular accident. There were no events of violence that met our prespecified criteria but one patient receiving duloxetine experienced mild hostility that began two days after randomisation. There were no reports of suicidality either, but eight patients (four in the duloxetine group and four in the placebo group) experienced injuries or burns. There were no narratives for these patients and therefore no information regarding the context and nature of these events.

Harms predisposing to violence and suicidality

According to our criteria, two patients receiving duloxetine experienced five serious adverse events potentially predisposing to suicidality or violence, which were severe depression, panic attacks and severe anxiety. The many events that were not serious are described next.

Activation events

Core or potential activation events were experienced by 187 patients in the duloxetine group and 42 patients in the placebo group (table 2), RR 4.45, 3.22 to 6.14, I^2=0%; RD 15.1%, 12.3% to 18.0%; NNH 7, 6 to 9 (figures 12 and 13). The result was similar after exclusion of patients who only experienced sleep problems that were not definitively insomnia, RR 4.96, 3.47 to 7.09, I^2=0% (figure 14). The risk of experiencing a core event was over three times greater with duloxetine (RR 3.59, 2.04 to 6.32, I^2=0%; RD 4.1%, 2.4% to 5.7%; NNH 25, 18 to 42) (figures 15 and 16). Twenty eight patients discontinued due to events of activation (27 in the duloxetine group and 1 in the placebo group). The most frequently reported core event was anxiety (18 patients in the duloxetine group and 6 in the placebo group).

The results were similar for FDA defined activation events (table 2, see figures 17 and 18). Twenty eight patients in the duloxetine group and one in the placebo group experienced more than one such event (range 2 to 4 events). The most frequently occurring event was insomnia (120 patients in the duloxetine group and 19 in the placebo group, RR 6.30, 3.92 to 10.13, I^2=0%; RD 10.5%, 8.3% to 12.9%; NNH 10, 8 to 13 (figures 19 and 20).

Akathisia, emotional disturbance, psychosis and depression

No events of akathisia were reported, whereas 18 patients in the duloxetine group and 3 in the placebo group experienced emotional disturbance (table 3), RR 4.73, 1.62 to 13.85, I^2=0%; RD 1.6%, 0.6% to 2.5%; NNH 65, 40 to 170 (figures 21 and 22), and 3 in the duloxetine group and 1 in the placebo group discontinued because of emotional disturbance. The most frequently reported event was feeling abnormal (8 patients in the duloxetine group and 1 in the placebo group).

Thirty patients (21 in the duloxetine group and 9 in the placebo group) experienced a core or potential psychotic event (table 2), RR 2.25, 1.06 to 4.81, I^2=0%; RD 1.3%, 0.1% to 2.4%; NNH 80, 40 to 834 (figures 23 and 24). The risk of experiencing a core event was similar (RR 2.49, 0.78 to 7.89, I^2=0%, figure 25), but not statistically significant. The most frequently reported core event was disorientation (4 patients in the duloxetine group and 1 in the placebo group). One patient receiving duloxetine discontinued because of a confusional state.

Depression related events were similar on duloxetine and placebo (RR 1.26, 0.58 to 2.71, I^2=26%, figure 26).

Discussion

We wished to assess the clinical effect of duloxetine and its harms in terms of suicidality and violence and their possible precursors in women with stress urinary incontinence based on the four trials used for marketing authorisation in the EU.

The effects we found on incontinence were so small that it is unlikely that they are clinically relevant. For incontinence quality of life, we found a difference of 3.2 between duloxetine and placebo. Eli Lilly staff has suggested that the minimum clinically important difference is 2.5, but this could be a post hoc suggestion, as it was based on two of the trials we reviewed.[31] It is hard to believe that such a small difference on a scale that goes up to 100 can be relevant. For depression, for example, the Hamilton scale ranges from 0 to 52, and the smallest effect that can be perceived is 5-6,[32] which is less than what is clinically relevant. SSRI trials are not adequately blinded because of the many side effects these drugs have, and the small effects we found could be fully explained by unblinding bias.[33,34]

We did not find any reported events of suicidality or violence, but many patients experienced unpleasant events that predispose to suicidality and violence, e.g. the number needed to harm was only 7 for activation. Furthermore, one in eight patients receiving duloxetine developed insomnia and one in a hundred developed a core or potential psychotic event because of duloxetine.

Strengths and limitations of our study

Access to individual patient data allowed us to elucidate the harms caused by duloxetine in meta-analyses, which wouldn't have been possible if we had only had access to published trial reports or to the summary data provided in clinical study reports.

It is a limitation that the data for the beneficial effects, especially for incontinence episode frequency, were considerably skewed.[35]Another reason why these results should be interpreted cautiously is the unblinding due to adverse events.[33,34]

There were only 958 patients on duloxetine, which means that the sample was too small to detect rare events of suicidality and violence. Furthermore, the data on adverse events were obtained through non-probing questions, which leads to underreporting of adverse events,[22,36] especially for events of a sensitive nature,[37] such as suicidal ideation and behaviour, and violence. Finally, suicidality events have been much underreported , also in clinical trials and observational studies conducted by Eli Lilly.[7]

Comparisons with other studies

In its scientific discussion about Lilly's application for Yentreve, the EMA reported on a suicide attempt occurring in the open label extension phase of one of the trials we examined.[17] These extension phases were ongoing when the marketing authorization was granted in the EU but the results are now available in Lilly's trial registry. We found reports of two suicide attempts on duloxetine in one of the trials. Lilly's trial registry report stated that the overall rates of suicidal behaviour for patients in duloxetine stress urinary incontinence studies (190/100,000 patient years)

"was within the range described for women in the general population in epidemiological survey studies (149 to 897/100,000 patient years)." [38] In contrast, the FDA stated that suicide attempts were 2.6 times higher than for other women of similar age based on a suicide attempt rate among middle-aged U.S. women in published studies of 150 to 160 per 100,000 person years.[14] We find it very worrying that the FDA reported that 11 of 9,400 women receiving duloxetine in the stress urinary incontinence studies had a suicide attempt in the open label extensions when they all received duloxetine.[14]

It is seriously misleading to report suicidality in relation to patient years, as Eli Lilly did.[7] Those patients who continue with the drug for a long time after the randomised phase is over are those who tolerate it, which means that person years are added to the drug group "for free" in terms of suicidality. The effect of this survivorship bias can be dramatic.[7] In the regulatory submission for paroxetine, for example, suicide attempts on the drug decreased by 52% by using patient years, but increased by 25% using patients.[39] Our four clinical study reports did not contain any such extension phase.

Our findings on harms are in agreement with a published pooled analysis, which is not a meta-analysis performed by Lilly on the same trials.[40] Our results for treatment-emergent adverse events and discontinuation because of an adverse event are also in good agreement with a Cochrane review of ten trials of duloxetine for urinary incontinence,[41] but this review did not report on suicidality.

Our findings also are in agreement with a prospective cohort study of 228 women treated with duloxetine for SUI or mixed SUI. After four weeks of receiving duloxetine, 45% of the cohort had discontinued the drug due to adverse events, and 24% due to lack of efficacy.[42] After 4 months, only 12% of the original cohort were still taking duloxetine.[42]

Prospective assessment of potential precursors to suicidality and violence

Research on SSRI induced akathisia, activation, emotional disturbance and psychotic events as precursors to suicidality is sparse and largely retrospective, and antidepressant induced akathisia is underdiagnosed.[43,44] The FDA has issued draft guidance for the assessment of suicidality in all clinical trials of drugs with central nervous system activity, including multiple dose phase 1 trials in healthy volunteers.[45] This guidance does not address potential precursory events to suicidality, however, despite their inclusion in warnings in antidepressant product labelling. It would be a missed opportunity not to include prospective assessment of these symptoms also in the draft guidance.

EMA changes its announcements about access to trial data

Despite earlier promises, the EMA recently announced that it will not publish individual anonymised patient data contained in appendices of clinical study reports in the first round of implementation of its new policy,[46] which came into effect on 1 January 2015. The EMA's excuse is that it needs to find a reliable way to anonymise the data. However, in accordance with current legislation, the data are already anonymised, and the EMA's approach is inconsistent, as we can get access to the harms in the old trials in the EMA's possession. As we have shown here and earlier,[16] individual patient data contained in appendices of clinical study reports are essential for a reliable assessment of drug harms.

The interpretation of drug harms differ not only between drug agencies but also within drug agencies where scientific staff are sometimes overruled by their superiors, who tend to downplay

issues of patient safety.[47,48] The assessment of the balance between benefits and harms is subjective and qualitative,[49] and we therefore believe that the FDA, like the EMA, should be held accountable to the public it serves by explaining why an application was refused or withdrawn.[50]

Conclusions and implications

Given its unlikely clinically meaningful benefit and its many important harms, we question the rationale for using duloxetine for stress urinary incontinence.

References

1. Teicher MH, Glod C, Cole JO. Emergence of intense suicidal preoccupation during fluoxetine treatment. *Am J Psychiatry* 1990;147:207-210.

2. Healy D, Herxheimer A, Menkes DB. Antidepressants and violence: Problems at the interface of medicine and law. *PLoS Med* 2006;3:e372.

3. Breggin BR. Suicidality, violence and mania caused by selective serotonin reuptake inhibitors (SSRIs): A review and analysis. *International Journal of Risk & Safety in Medicine* 2003/2004;16:31–49.

4. Laughren TP. Overview for December 13 Meeting of Psychopharmacologic Drugs Advisory Committee (PDAC). 2006 Nov 16. www.fda.gov/ohrms/dockets/ac/06/briefing/2006-4272b1-01-FDA.pdf

5. Gunnell D, Saperia J, Ashby D. Selective serotonin reuptake inhibitors (SSRIs) and suicide in adults: meta-analysis of drug company data from placebo controlled, randomised controlled trials submitted to the MHRA's safety review. *BMJ* 2005;330:385.

6. Fergusson D, Doucette S, Glass KC, Shapiro S, Healy D, Hebert P et al. Association between suicide attempts and selective serotonin reuptake inhibitors: systematic review of randomised controlled trials. *BMJ* 2005;330:396.

7. Gøtzsche PC. Deadly psychiatry and organised denial. Copenhagen: People's Press 2015 (in press).

8. FDA. *Revisions to product labelling.* www.fda.gov/downloads/Drugs/DrugSafety/InformationbyDrugClass/UCM173233.pdf

9. EMA. *Yentreve Authorisation details.* www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000545/human_med_001164.jsp&mid=WC0b01ac058001d124

10. Eli Lilly and Company. *Lilly and Boehringer Ingelheim Jointly Announce the Recision of U.S. FDA Application for Duloxetine for Treatment of Stress Urinary Incontinence.* 2005. investor.lilly.com/releasedetail.cfm?ReleaseID=154081

11. EMA. *Procedural advice on publication of information on negative opinions and refusals of marketing authorisation applications for human medicinal products*. 2013. www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004188.pdf

12. EMA. *Procedural advice on publication of information on withdrawals of applications related to the marketing authorisation of human medicinal products*. 2013. www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500005061.pdf

13. FDA Centre for Drug Evaluation and Research. Personal communication. 26 November 2014

14. FDA. *Historical Information on Duloxetine hydrochloride (marketed as Cymbalta).* http://google2.fda.gov/search?q=www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm114970.htm&client=FDAgov-archive&proxystylesheet=FDAgov&output=xml_no_dtd&site=FDAgov-archive&requiredfields=archive:Yes&sort=date:D:L:d1&filter=1

15. International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use. *The common technical document for the registration of pharmaceuticals for human use efficacy–M4E(R1*). 2002. www.ich.org/fileadmin/Public_Web_Site/ICH_Products/CTD/M4__R1__Efficacy/M4E__R1_.pdf

16. Maund E, Tendal B, Hróbjartsson A, Jørgensen KJ, Lundh A, Schroll J et al. Benefits and harms in clinical trials of duloxetine for treatment of major depressive disorder: comparison of clinical study reports, trial registries, and publications. *BMJ* 2014;348:g3510.

17. EMA. *Scientific discussion.* 2005. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000545/WC500045641.pdf

18. Moore TJ, Glenmullen J, Furberg CD. Prescription drugs associated with reports of violence towards others. *PLoS ONE* 2010;5:e15337.

19. Sinclair LI, Christmas DM, Hood SD, Potokar JP, Robertson A, Isaac A et al. Antidepressant-induced jitteriness/anxiety syndrome: systematic review. *Br J Psychiatry* 2009;194:483-90.

20. Price J, Cole V, Goodwin GM. Emotional side-effects of selective serotonin reuptake inhibitors: qualitative study. *Br J Psychiatry* 2009;195:211-7.

21. Levin R, Daly RS. Nightmares and psychotic decompensation: a case study. *Psychiatry* 1998;61:217-22.

22. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of clinical trials*. 4th ed. Springer; 2010

23. ICH Secretariat. *Understanding MedDRA.* 2013. http://www.meddra.org/sites/default/files/page/documents/meddra2013.pdf

24. White CA. A preliminary assessment of the impact of MEDDRA on adverse event reports and product labelling. *Drug Inf J* 1998;32:347-62.

25. Higgins JPT, Deeks JJ, Altman DG (editors). Chapter 16: Special topics in statistics. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions.* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

26. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in Medicine* 2004;23:1351-1375.

27. Norton PA, Zinner NR, Yalcin I, Bump RC. Duloxetine versus placebo in the treatment of stress urinary incontinence. *Am J Obstet Gynecol* 2002;187:40–8.

28. Dmochowski RR, Miklos JR, Norton PA, Zinner NR, Yalcin I, BumpRC. Duloxetine versus placebo for the treatment of North American women with stress urinary incontinence. *J Urol* 2003;170:1259–63.

29. van Kerrebroeck P, Abrams P, Lange R, Slack M, Wyndaele JJ, Yalcin I et al. Duloxetine versusplacebo in the treatment of European and Canadian women with stress urinary incontinence. *BJOG* 2004;111:249–57.

30. Millard RJ, Moore K, Rencken R, Yalcin I, Bump RC. Duloxetine vs.placebo in the treatment of stress urinary incontinence: a four-continent randomized clinical trial. *BJU Int* 2004;93:311–8.

31. Yalcin I, Patrick DL, Summers K, Kinchen K, Bump RC. Minimal clinically important differences in Incontinence Quality-of-Life scores in stress urinary incontinence. *Urology* 2006;67:1304-8.

32. Leucht S, Fennema H, Engel R, Kaspers-Janssen M, Lepping P, Szegedi A. What does the HAMD mean? *J Affect Disord* 2013;148:243-8.

33. Fisher S, Greenberg RP. How sound is the double-blind design for evaluating psychotropic drugs? *J Nerv Ment Dis* 1993;181;345-350.

34. Moncrieff J, Wessely S, Hardy R. Active placebos versus antidepressants for depression. *Cochrane Database of Systematic Reviews* 2004, Issue 1. Art. No.: CD003012. DOI: 10.1002/14651858.CD003012.pub2.

35. Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org

36. Spielmans GI, Kirsch I. Drug approval and drug effectiveness. *Annu Rev Clin Psychol* 2014;10:741-66.

37. Montejo-Gonzalez AL, Llorca G, Izquierdo JA, Ledesma A, Bousono M, Calcedo A et al. SSRI-induced sexual dysfunction: fluoxetine, paroxetine, sertraline, and fluvoxamine in a prospective, multicenter, and descriptive clinical study of 344 patients. *J. Sex Marital Ther 1997* 23:176–94.

38. Eli Lilly and Company. *Summary ID# 5309. Clinical Study Summary: Study F1J-MC-SBBM. Long-Term Monitoring of Safety in Subjects Treated with Duloxetine for Stress Urinary Incontinence*. http://lillytrials.com/results/Yentreve.pdf

39. FDA. *Review and evaluation of clinical data. Original NDA 20-031. Paroxetine (Aropax) safety review*. 1992 Oct 5.

40. Hurley DJ, Turner CL, Yalcin I, Viktrup L, Baygani SK. Duloxetine for the treatment of stress urinary incontinence in women: an integrated analysis of safety. *Eur J Obstet Gynecol Reprod Biol* 2006;125:120-8.

41. Mariappan P, Alhasso AA, Grant A, N'Dow JMO. Serotonin and noradrenaline reuptake inhibitors (SNRI) for stress urinary incontinence in adults. Cochrane Database of Systematic Reviews 2005, Issue 3. Art. No.: CD004742. DOI: 10.1002/14651858.CD004742.pub2.

42. Vella M, Duckett J, Basu M. Duloxetine 1 year on: the long term outcome of a cohort of women prescribed duloxetine. *Int Urogynecol J* 2008;19:961–964.

43. Lipinski JF, Mallya G, Zimmerman P, Pope H. Fluoxetine-induced akathisia. Clinical and theoretical implications. *J Clin Psychiatry* 1989;50:339-342.

44. Baldassano CF, Truman CJ, Nierenberg A, Ghaemi SN, Sach GS. Akathisia: a review and case report following paroxetine treatment. *Comprehensive Psychiatry* 1996;37:122-129.

45. Department of Health and Human Services, Food and Drug Administration Centre for Drug Evaluation and Research. *Guidance for Industry: Suicidal ideation and behaviour: Prospective assessment of occurrence in clinical trials - draft guidance*. 2012. www.fda.gov/drugs/guidancecomplianceregulatoryinformation/guidances/ucm315156.htm

46. EMA. *Questions and answers on the European Medicines Agency policy on publication of clinical data for medicinal products for human use.* 2015. http://www.ema.europa.eu/docs/en_GB/document_library/Report/2014/10/WC500174378.pdf

47. Gøtzsche PC. Deadly medicines and organised crime: How big pharma has corrupted health care. London: Radcliffe Publishing; 2013.

48. EMA Benefit-Risk Methodology Project Team. *Benefit-risk methodology project: Work package 1 report: description of the current practice of benefit-risk assessment for centralised procedure products in the EU regulatory network.* 2011. www.ema.europa.eu/docs/en_GB/document_library/Report/2011/07/WC500109478.pdf

49. Committee for Medicines Products for Human Use. *Reflection paper on benefit-risk assessment methods in the context of the evaluation of marketing authorisation applications of Medicinal products for human use*. 2008. www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2010/01/WC500069634.pdf

50. Wood AJ. Progress and deficiencies in the registration of clinical trials. *N Engl J Med*. 2009, 19;360:824-30.

Tables and figures

Table 1: *A priori* defined adverse events in suicidality and violence related adverse event categories

Adverse event category	Core adverse events	Potential adverse events[1]
Suicidality	accident-, attempt, burn, cut, drown, gas, gun, hang, hung, immolat,injur-, jump, monoxide, mutilat-, overdos-, self damag-, self harm, self inflict, self injur-, shoot, slash, suic-", poison, asphyxiation, suffocation, firearm	
Violent behaviour	Homicide, physical assault, physical abuse, homicidal ideation, violence related symptoms (e.g. criminal behaviour, anti-social behaviour)	
Depression	depression	
Emotional disturbance	Anhedonia, apathy, depersonalisation, derealisation, disinhibition, emotionally detached, emotional lability, flat effect, impulsivity, lack of empathy	
Psychotic behaviour	Abnormal thinking (intrusive thoughts, unusual thoughts), confusion (disorientation, incoherent thoughts), delirium, delusions, hallucinations, hysteria, manic reaction, paranoia, psychosis	Abnormal dreams, nightmares
Activation	Agitation (aggression, hostility), akathisia, anxiety, increased energy (euphoria, irritability, jitteriness, mania[2]), restlessness (hyperactivity), shakiness	Insomnia, panic, tension[3], tremor
FDA defined activation symptoms	anxiety, agitation, panic attacks, insomnia, irritability, hostility, aggressiveness, impulsivity, akathisia (psychomotor restlessness), hypomania, and mania	

[1]Potential events were those events were there was lack of consistency in the literature, or uncertainty over whether an event was relevant. The effect of including these events was explored in sensitivity analyses.
[2]Mania was reported as both an activation event and as a psychotic event, as patients can report being "manic", when they are describing being more active than usual i.e. experiencing activation.
[3]Tension was originally categorised as a potential activation event however, tension codes to the higher level term of symptoms of anxiety in MedDRA. Tension was therefore considered a core event in the main analyses. Sensitivity analyses were performed to evaluate the effect of this decision

Table 2: Adverse events of suicidality and violence related adverse event categories reported among four placebo controlled trials of duloxetine for stress urinary incontinence

Adverse event category	Core adverse events reported among the four trials[1,2]	Potential adverse events reported among the four trials[1,2]
Activation	anxiety, central nervous system stimulation, energy increased, euphoric mood, feeling jittery, hostility, irritability, mania, nervousness psychomotor hyperactivity, restlessness, stress, tension	insomnia (including initial and middle insomnia), panic attack, panic disorder, poor quality sleep, restless leg syndrome, sleep disorder and tremor.
FDA defined activation symptoms	Agitation, anxiety, insomnia (including initial and middle insomnia),mania, nervousness[3], panic attack,poor quality sleep, sleep disorder, stress[3], tension[3]	
Emotional disturbance	feeling abnormal (verbatims included "feeling drugged", "foggy in the head", "fuzzy feeling"), apathy, emotional disorder, cognitive disorder ("lack of awareness"), emotional poverty ("emotionless"), listless, mood altered ("be moody").	
Psychotic behaviour	disorientation, confusional state, euphoric mood, mania, and mental disorder (verbatim "nervous breakdown").	abnormal dreams and nightmares
Depression	Depression	Depressed mood, dysthymic disorder

[1] These adverse events occurred in either the duloxetine arm or the placebo arm, or in both.

[2] The data presented in this table are the MedDRA 17.0 preferred terms that we used when we recoded the original preferred terms (and if available, also verbatim terms) of adverse events provided in the clinical study reports of the 4 trials of duloxetine for stress urinary incontinence.

[3] Nervousness, stress and tension are not explicitly mentioned in FDA defined activation. Anxiety is stated as a FDA defined activation event and nervousness, stress and tension all code to the higher level term of symptoms of anxiety in MedDRA. We therefore included these 3 types of events in our analyses of FDA defined activation.

Figure 1: forest plot (mean difference) of percentage change from baseline in weekly incontinence episodes

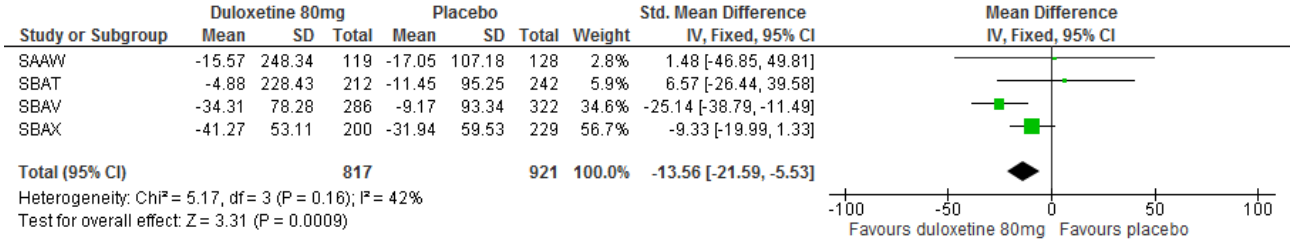


Figure 2: forest plot (mean difference) of numerical change from baseline in weekly incontinence episodes

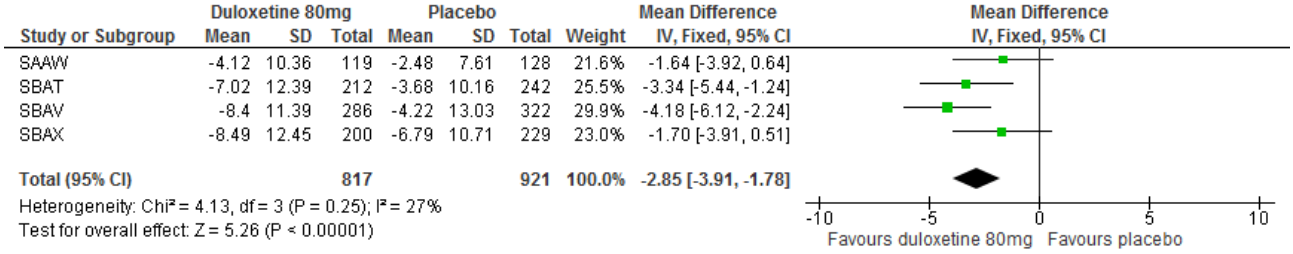


Figure 3: forest plot (standardised mean difference) of percentage change from baseline in weekly incontinence episodes

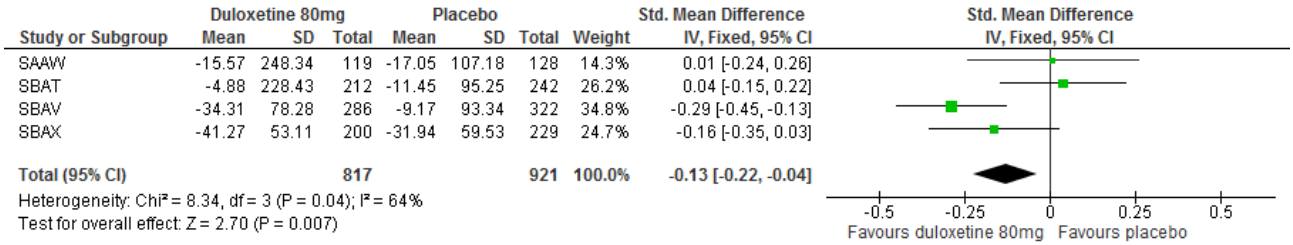


Figure 4: forest plot (standardised mean difference) of numerical change from baseline in weekly incontinence episodes

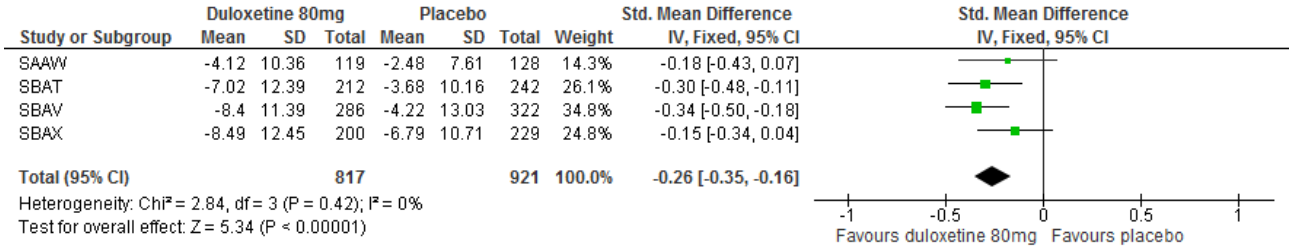


Figure 5: forest plot (mean difference) of change in Incontinence Quality of Life total score

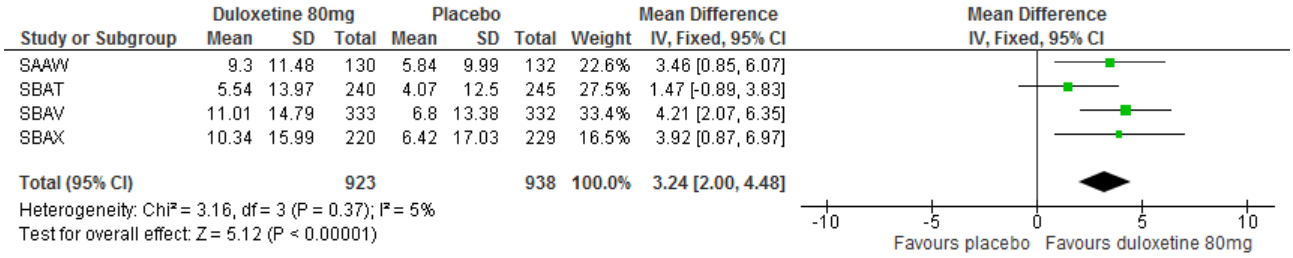


Figure 6: forest plot (standardised mean difference) of change in Incontinence Quality of Life total score

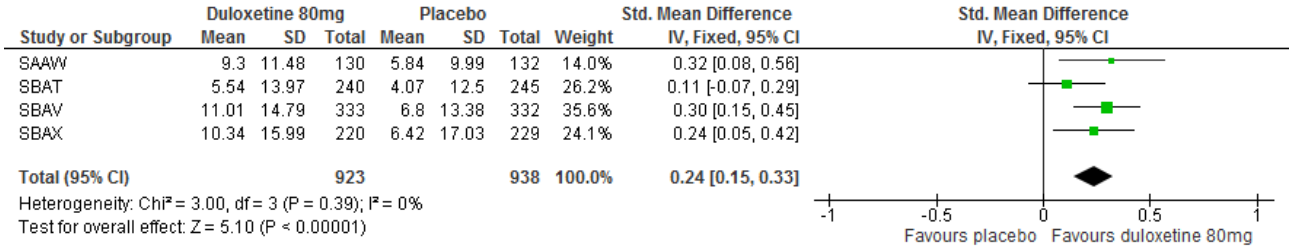


Figure 7: forest plot (risk ratio) of experiencing at least one treatment emergent adverse event

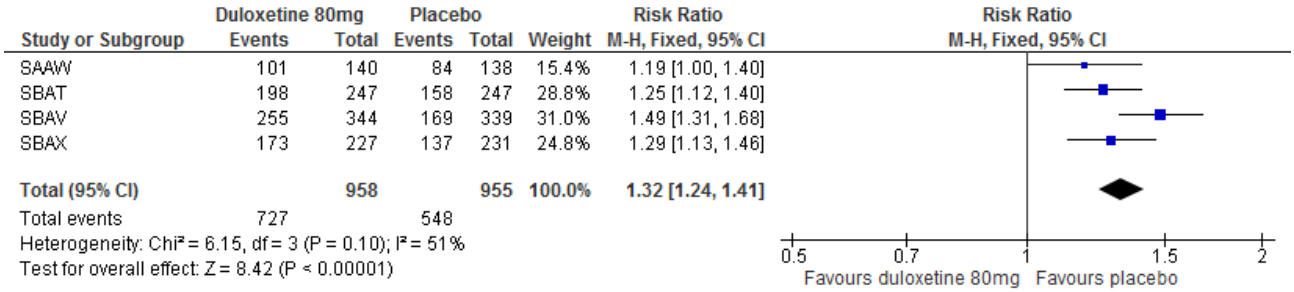


Figure 8: forest plot (risk difference) of experiencing at least one treatment emergent adverse event

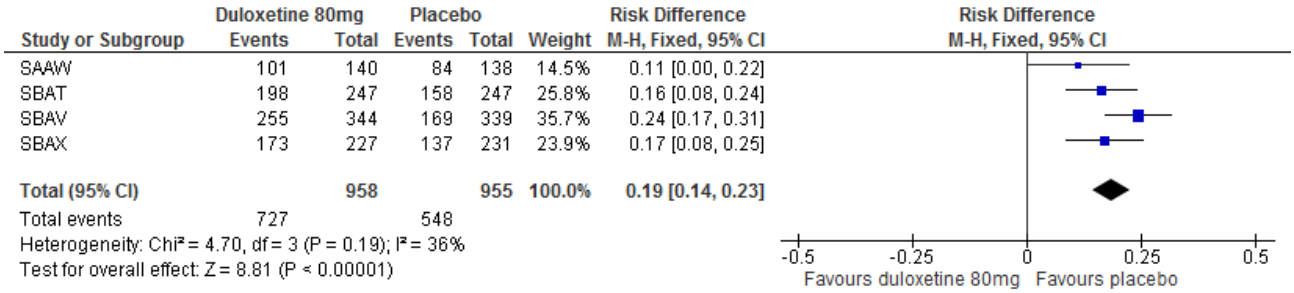


Figure 9: forest plot (risk ratio) of discontinuing because of adverse events

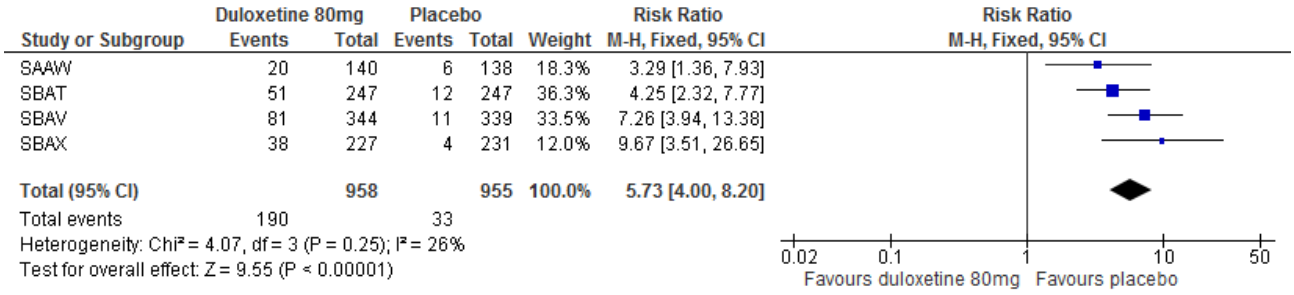


Figure 10: forest plot (risk difference) of discontinuing because of adverse events

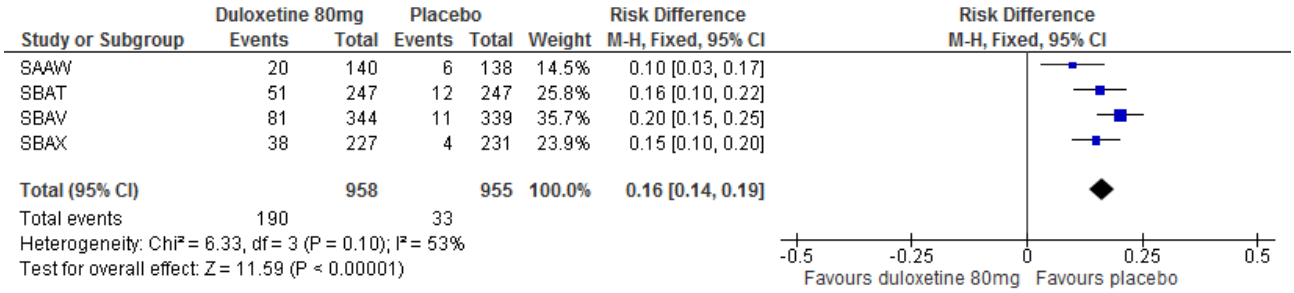


Figure 11: forest plot (risk ratio) of experiencing at least one non-fatal serious adverse event

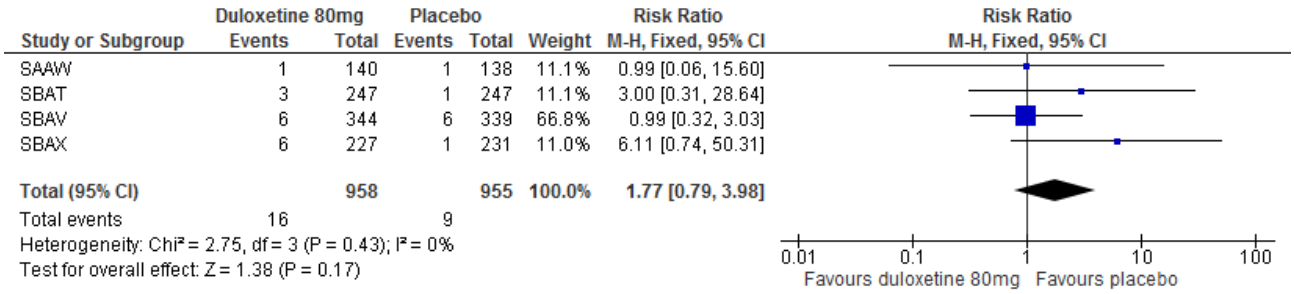


Figure 12: forest plot (risk ratio) of experiencing at least one core or potential activation event

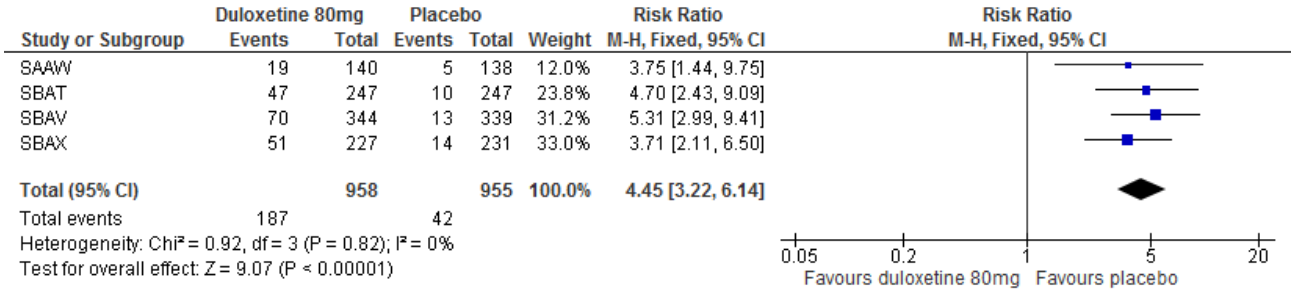


Figure 13: forest plot (risk difference) of experiencing at least one core or potential activation event

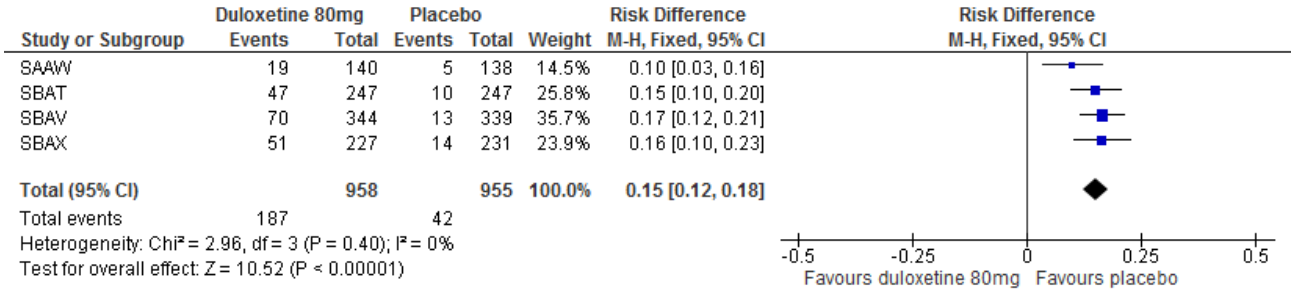


Figure 14: forest plot (risk ratio) of experiencing at least one core or potential activation event, excluding patients who experienced solely sleep disorder or poor quality sleep

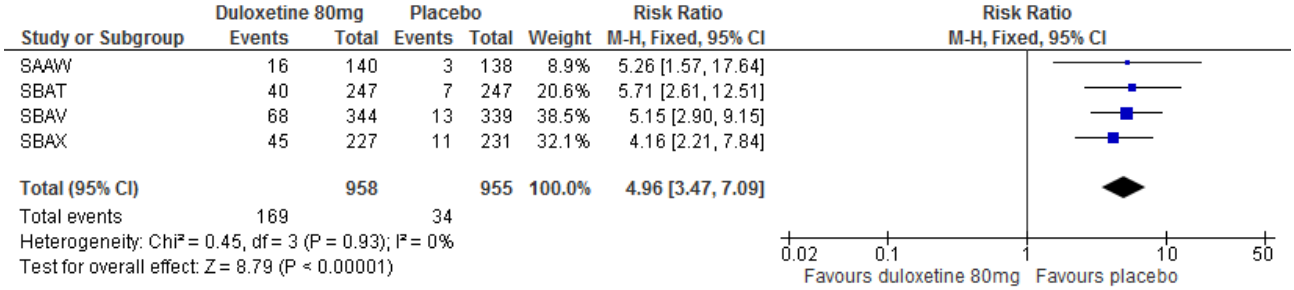


Figure 15: forest plot (risk ratio) of experiencing at least one core activation event

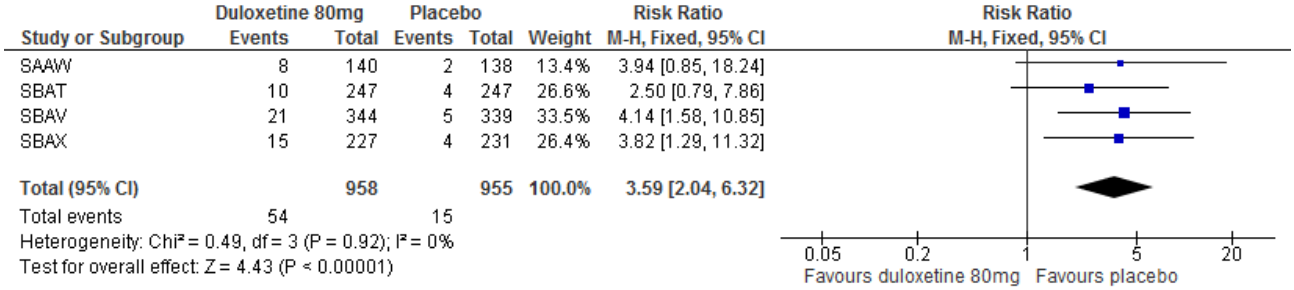


Figure 16: forest plot (risk difference) of experiencing at least one core activation event

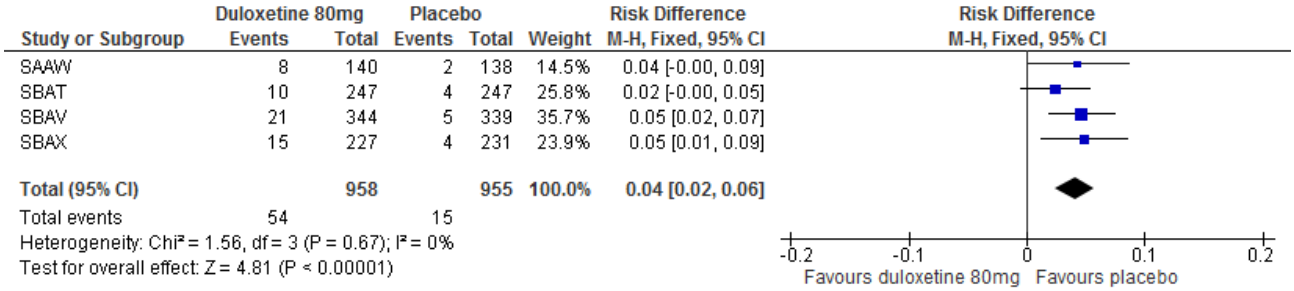


Figure 17: forest plot (risk ratio) of experiencing at least one FDA defined activation event

Study or Subgroup	Duloxetine 80mg Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
SAAW	16	140	5	138	12.9%	3.15 [1.19, 8.37]
SBAT	40	247	9	247	23.1%	4.44 [2.20, 8.96]
SBAV	64	344	12	339	31.0%	5.26 [2.89, 9.56]
SBAX	48	227	13	231	33.0%	3.76 [2.09, 6.74]
Total (95% CI)		**958**		**955**	**100.0%**	**4.30 [3.07, 6.02]**
Total events	168		39			

Heterogeneity: Chi² = 1.03, df = 3 (P = 0.79); I² = 0%
Test for overall effect: Z = 8.50 (P < 0.00001)

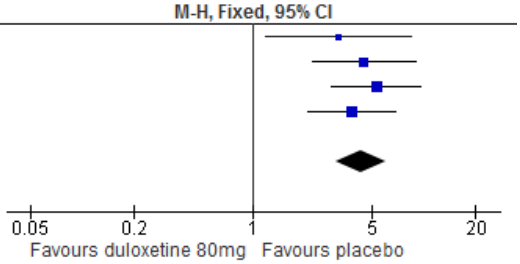


Figure 18: forest plot (risk ratio) of experiencing at least one FDA defined activation event, excluding patients who experienced solely sleep disorder or poor quality sleep

Study or Subgroup	Duloxetine 80mg Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
SAAW	13	140	3	138	9.7%	4.27 [1.24, 14.66]
SBAT	32	247	6	247	19.3%	5.33 [2.27, 12.53]
SBAV	61	344	12	339	39.0%	5.01 [2.75, 9.13]
SBAX	42	227	10	231	32.0%	4.27 [2.20, 8.31]
Total (95% CI)		**958**		**955**	**100.0%**	**4.77 [3.27, 6.94]**
Total events	148		31			

Heterogeneity: Chi² = 0.23, df = 3 (P = 0.97); I² = 0%
Test for overall effect: Z = 8.14 (P < 0.00001)

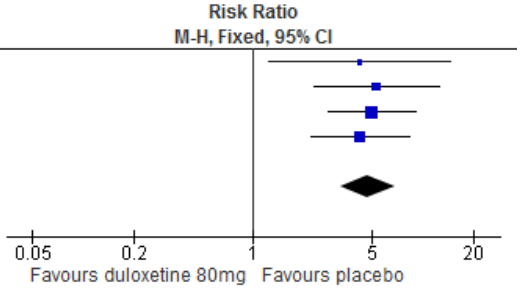


Figure 19: forest plot (risk ratio) of experiencing at least one event of insomnia (including initial and middle insomnia)

Study or Subgroup	Duloxetine 80mg Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
SAAW	10	140	1	138	5.3%	9.86 [1.28, 75.97]
SBAT	29	247	3	247	15.8%	9.67 [2.98, 31.32]
SBAV	49	344	9	339	47.7%	5.37 [2.68, 10.75]
SBAX	32	227	6	231	31.3%	5.43 [2.31, 12.73]
Total (95% CI)		**958**		**955**	**100.0%**	**6.30 [3.92, 10.13]**
Total events	120		19			

Heterogeneity: Chi² = 1.02, df = 3 (P = 0.80); I² = 0%
Test for overall effect: Z = 7.60 (P < 0.00001)

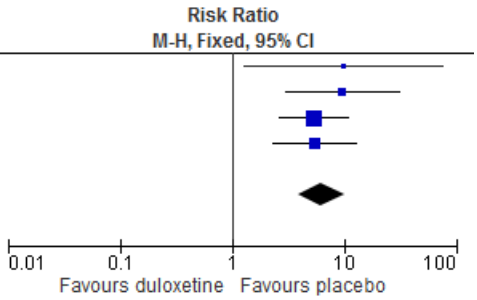


Figure 20: forest plot (risk difference) of experiencing at least one event of insomnia (including initial and middle insomnia)

Study or Subgroup	Duloxetine 80mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
SAAW	10	140	1	138	14.5%	0.06 [0.02, 0.11]
SBAT	29	247	3	247	25.8%	0.11 [0.06, 0.15]
SBAV	49	344	9	339	35.7%	0.12 [0.08, 0.16]
SBAX	32	227	6	231	23.9%	0.11 [0.07, 0.16]
Total (95% CI)		**958**		**955**	**100.0%**	**0.11 [0.08, 0.13]**
Total events	120		19			

Heterogeneity: Chi² = 3.63, df = 3 (P = 0.30); I² = 17%
Test for overall effect: Z = 9.10 (P < 0.00001)

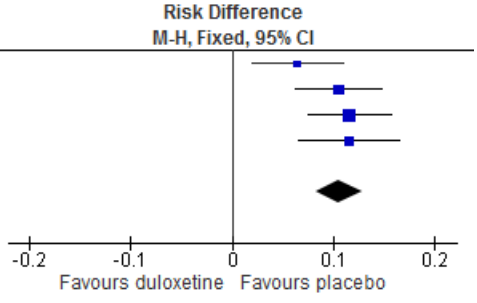


Figure 21: forest plot (risk ratio) of experiencing at least one event of emotional disturbance

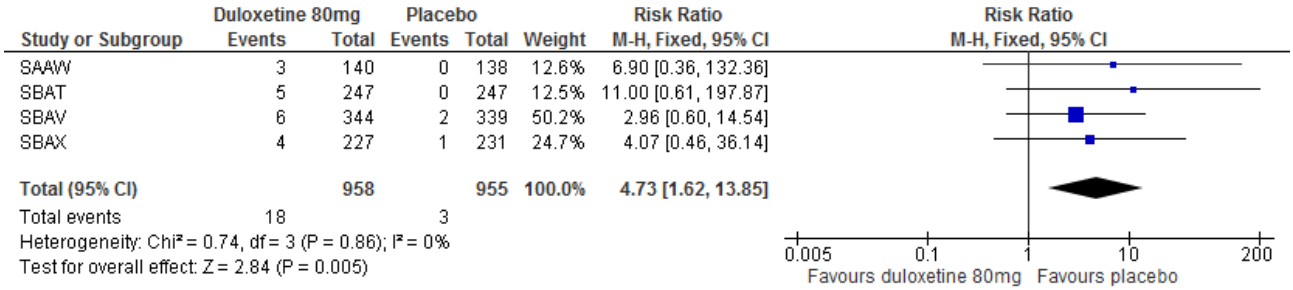


Figure 22: forest plot (risk difference) of experiencing at least one event of emotional disturbance

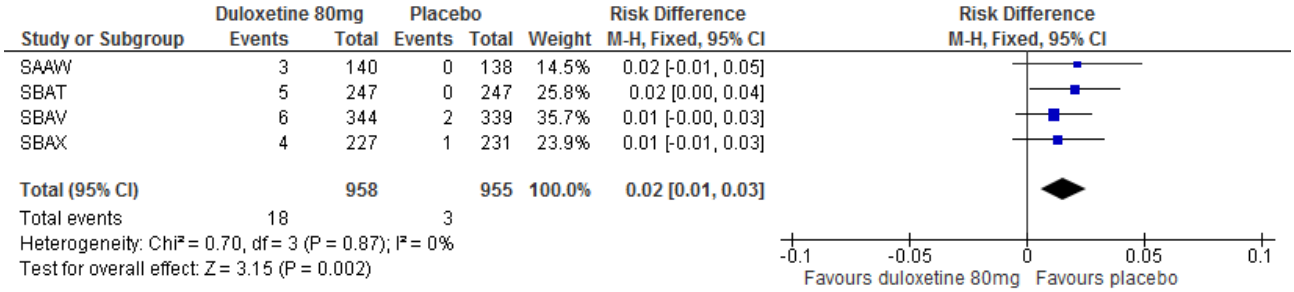


Figure 23: forest plot (risk ratio) of experiencing at least one core or potential psychotic event

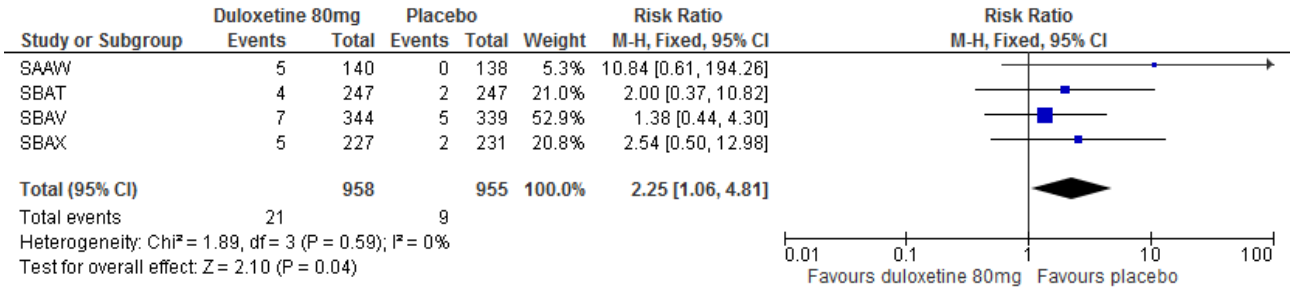


Figure 24: forest plot (risk difference) of experiencing at least one core or potential psychotic event

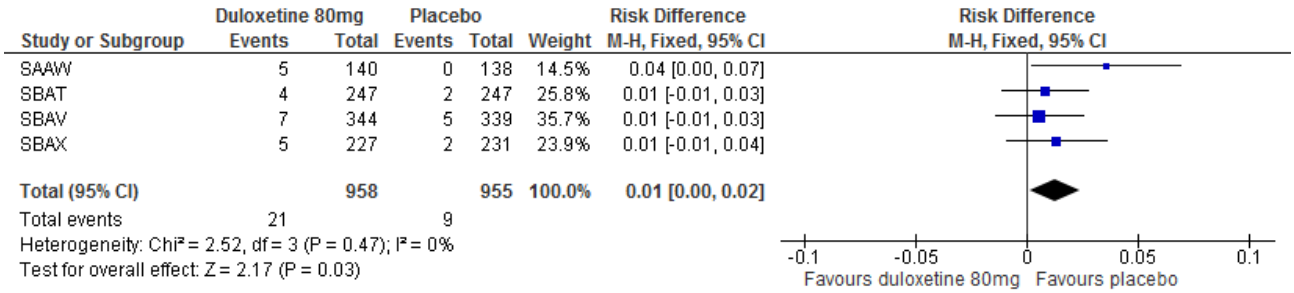


Figure 25: forest plot (risk ratio) of experiencing at least one core psychotic event

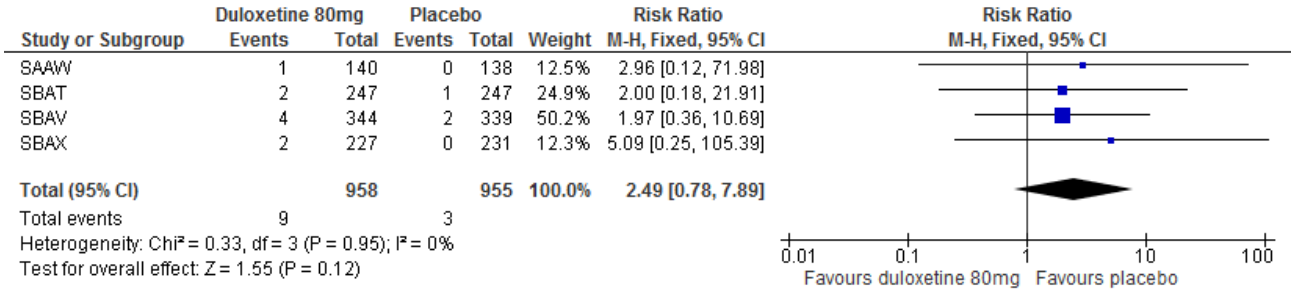
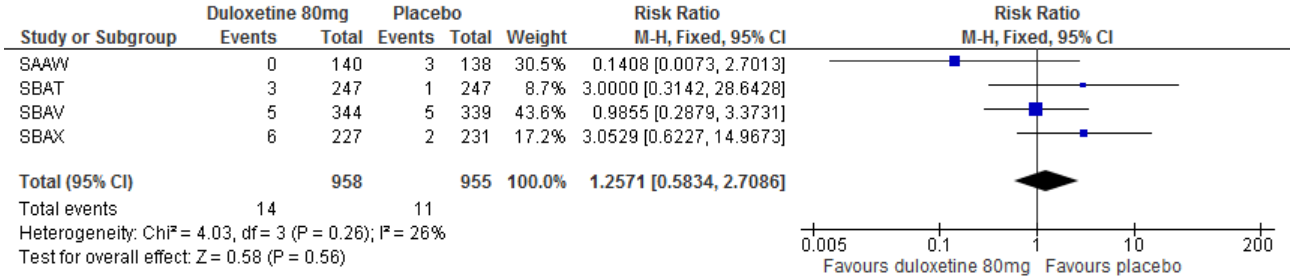


Figure 26: forest (risk ratio) of experiencing at least one core or potential depression related event



Signed co-authorship forms

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviour

This declaration concerns the following article:
Benefits and harms in clinical trials of duloxetine for treatment of major depressive disorder: comparison of clinical study reports, trial registries, and publications

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	**(A,B,C)**
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	B
2. Planning of the experiments and methodology design, including selection of methods and method development	B

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
27.11.14	Britta Tendal	PhD	
	Asbjørn Hróbjartsson	PhD	
	Karsten Juhl Jørgensen	DrMedSci	

Signature of the PhD student and the principal supervisor:	
Date: 23 JUNE 2015	**Date:** 23 JUNE 2015
PhD student:	**Principal supervisor:**

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviour

This declaration concerns the following article:
Benefits and harms in clinical trials of duloxetine for treatment of major depressive disorder: comparison of clinical study reports, trial registries, and publications

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	(A,B,C)
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	B
2. Planning of the experiments and methodology design, including selection of methods and method development	B

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
	Britta Tendal	PhD	
8/1 2015	Asbjørn Hróbjartsson	PhD	
	Andreas Lundh	PhD	

Signature of the PhD student and the principal supervisor:	
Date: 23 JUNE 2015	**Date:** 23 JUNE 2015
PhD student:	**Principal supervisor:**

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviour

This declaration concerns the following article:
Benefits and harms in clinical trials of duloxetine for treatment of major depressive disorder: comparison of clinical study reports, trial registries, and publications

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	(A,B,C)
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	B
2. Planning of the experiments and methodology design, including selection of methods and method development	B

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
	Britta Tendal	PhD	
	Asbjørn Hróbjartsson	PhD	
8.1.15	Karsten Juhl Jørgensen	DrMedSci	

Signature of the PhD student and the principal supervisor:	
Date: 23 JUNE 2015	**Date:** 23 JUNE 2015
PhD student:	**Principal supervisor:**

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviour

This declaration concerns the following article:
Benefits and harms in clinical trials of duloxetine for treatment of major depressive disorder: comparison of clinical study reports, trial registries, and publications

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	(A,B,C)
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	B
2. Planning of the experiments and methodology design, including selection of methods and method development	B

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
	Britta Tendal	PhD	
	Asbjørn Hróbjartsson	PhD	
14/1-15	Andreas Lundh	PhD	Andreas Lund

Signature of the PhD student and the principal supervisor:	
Date: 23 JUNE 2015	**Date:** 23 JUNE 2015
PhD student:	**Principal supervisor:**

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviour

This declaration concerns the following article:
Benefits and harms in clinical trials of duloxetine for treatment of major depressive disorder: comparison of clinical study reports, trial registries, and publications

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	**(A,B,C)**
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	B
2. Planning of the experiments and methodology design, including selection of methods and method development	B

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
	Andreas Lundh	PhD	
14/1-15	Jeppe Schroll	PhD student	
	Peter C Gøtzsche	Professor	

Signature of the PhD student and the principal supervisor:	
Date: 23 JUNE 2015	**Date:** 23 JUNE 2015
PhD student:	**Principal supervisor:**

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviour

This declaration concerns the following article:
Benefits and harms in clinical trials of duloxetine for treatment of major depressive disorder: comparison of clinical study reports, trial registries, and publications

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	**(A,B,C)**
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	B
2. Planning of the experiments and methodology design, including selection of methods and method development	B

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
	Andreas Lundh	PhD	
	Jeppe Schroll	PhD student	
7/1 - 15	Peter C Gøtzsche	Professor	

Signature of the PhD student and the principal supervisor:	
Date: 7/1/2015 **PhD student:**	**Date:** 7/1 - 15 **Principal supervisor:**

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviourr

This declaration concerns the following article:
Coding of adverse events of suicidality in clinical study reports of duloxetine for the treatment of major depressive disorder: descriptive study

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	**(A,B,C)**
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	B
2. Planning of the experiments and methodology design, including selection of methods and method development	C

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
27.11.17	Britta Tendal	PhD	BT
	Asbjørn Hróbjartsson	PhD	
	Andreas Lundh	PhD	

Signature of the PhD student and the principal supervisor:	
Date: 23 JUNE 2015	**Date:** 23 JUNE 2015
PhD student:	**Principal supervisor:**

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviourr

This declaration concerns the following article:
Coding of adverse events of suicidality in clinical study reports of duloxetine for the treatment of major depressive disorder: descriptive study

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	**(A,B,C)**
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	B
2. Planning of the experiments and methodology design, including selection of methods and method development	C

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
	Britta Tendal	PhD	
8/2015	Asbjørn Hróbjartsson	PhD	
	Karsten Juhl Jørgensen	DrMedSci	

Signature of the PhD student and the principal supervisor:	
Date: 23 JUNE 2015 **PhD student:**	**Date:** 23 JUNE 2015 **Principal supervisor:**

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviourr

This declaration concerns the following article:
Coding of adverse events of suicidality in clinical study reports of duloxetine for the treatment of major depressive disorder: descriptive study

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	(A,B,C)
1. **Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments**	B
2. **Planning of the experiments and methodology design, including selection of methods and method development**	C

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
14/1-15	Andreas Lundh	PhD	
	Jeppe Schroll	PhD student	
	Peter C Gøtzsche	Professor	

Signature of the PhD student and the principal supervisor:	
Date: 23 JUNE 2015	**Date:** 23 JUNE 2015
PhD student:	**Principal supervisor:**

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

About this form

- Declarations of co-authorship must be included in your PhD thesis if your thesis contains or is based on co-authored manuscripts or published articles.
- You must submit a declaration of co-authorship for all articles/manuscripts included.
- The declarations are important to the assessment committee's work as they must be able to identify and assess your contribution to the article(s).
- If you have more co-authors for an article than there is space for in the table, simply just complete another declaration for the same article with the remaining co-authors.
- Remember that you and your principal supervisor must sign the declarations.
- Please submit the co-author declarations together with your thesis. Either insert the declarations in the back of the thesis or create a joint PDF which you submit in the email together with the thesis.

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviourr

This declaration concerns the following article:
Coding of adverse events of suicidality in clinical study reports of duloxetine for the treatment of major depressive disorder: descriptive study

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	**(A,B,C)**
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	B
2. Planning of the experiments and methodology design, including selection of methods and method development	C

3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	C

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	**Name:**	**Title:**	**Signature:**
7/1-15	Peter C Gøtzsche	Professor	

Signature of the PhD student and the principal supervisor:	
Date: 7/11 2015 **PhD student:**	**Date:** 7/1-15 Same **Principal supervisor:**

GRADUATE SCHOOL OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

DECLARATION OF CO-AUTHORSHIP

Information on PhD student:	
Name of PhD student	Emma Maund
E-mail	em@cochrane.dk
Date of birth	02/03/1977
Work place	The Nordic Cochrane Centre
Principal supervisor	Professor Peter C Gøtzsche

Title of PhD thesis:
The effect of selective serotonin reuptake inhibitors (SSRIs) on suicidality and violent behaviour

This declaration concerns the following article:
Benefits and suicidality related harms of duloxetine for the treatment of stress urinary incontinence: a meta-analysis of clinical study reports

The PhD student's contribution to the article: *(please use the scale (A,B,C) below as benchmark*)*	(A,B,C)
1. Formulation/identification of the scientific problem that from theoretical questions need to be clarified. This includes a condensation of the problem to specific scientific questions that is judged to be answerable by experiments	C
2. Planning of the experiments and methodology design, including selection of methods and method development	C
3. Involvement in the experimental work	B
4. Presentation, interpretation and discussion in a journal article format of obtained data	B

**Benchmark scale of the PhD student's contribution to the article*		
A. refers to:	*Has contributed to the co-operation*	*0-33 %*
B. refers to:	*Has contributed considerably to the co-operation*	*34-66 %*
C. refers to:	*Has predominantly executed the work independently*	*67-100 %*

Signature of the co-authors:			
Date:	Name:	Title:	Signature:
18/6-15	Louise Schow Jensen	Dr	Louise Schow Jensen
23/6/15	Peter C Gøtzsche	Professor	

Signature of the PhD student and the principal supervisor:	
Date: 23 JUNE 2015 PhD student:	Date: 23 JUNE 2015 Principal supervisor: