

Considering benefits and harms of duloxetine for treatment of stress urinary incontinence: a meta-analysis of clinical study reports

Emma Maund MSc PhD, Louise Schow Guski MD, Peter C. Gøtzsche MD

■ Cite as: *CMAJ* 2017 February 6;189:E194-203. doi: 10.1503/cmaj.151104

See related article at www.cmaj.ca/lookup/doi/10.1503/cmaj.161088

ABSTRACT

BACKGROUND: The European Medicines Agency makes clinical study reports publicly available and publishes reasons for not approving applications for marketing authorization. Duloxetine has been approved in Europe for the treatment of stress urinary incontinence in women. The reported adverse effects of duloxetine include mental health problems and suicidality. We obtained clinical study reports from the European Medicines Agency concerning use of this drug for stress urinary incontinence.

METHODS: We performed a meta-analysis of 4 randomized placebo-controlled trials of duloxetine (involving a total of 1913 patients) submitted to the European Medicines Agency for marketing approval for the indication of stress urinary incontinence in women. We used

data from the clinical study reports (totalling 6870 pages and including individual patient data) to assess benefits (including frequency of incontinence and changes in quality-of-life scores, such as Patient Global Impression of Improvement rating) and harms (both general harms, including discontinuation because of adverse events, and harms related to suicidality, violent behaviour and their potential precursors, such as akathisia and activation [stimulating effects such as insomnia, anxiety and agitation]).

RESULTS: Duloxetine was significantly better than placebo in terms of percentage change in weekly incontinence episodes (mean difference -13.56%, 95% confidence interval [CI] -21.59% to -5.53%) and change in Incontinence

Quality of Life total score (mean difference 3.24, 95% CI 2.00 to 4.48). However, the effect sizes were small, and a sensitivity analysis (with removal of one trial) showed that the number needed to treat for a Patient Global Impression of Improvement rating of “much better or very much better” was 8 (95% CI 6 to 13). The numbers needed to harm were 7 (95% CI 6 to 8) for discontinuing because of an adverse event and 7 (95% CI 6 to 9) for experiencing an activation event. No suicidality, violence or akathisia events were noted.

INTERPRETATION: Although duloxetine is effective for stress urinary incontinence in women, the rates of associated harm were high when individual patient data were analyzed, and the harms outweighed the benefits.

Clinical study reports are detailed reports of the design, conduct and results of clinical trials. These reports form part of the applications that pharmaceutical companies submit to drug agencies for marketing authorization for new drugs.¹ The European Medicines Agency (EMA) makes these reports publicly available upon request. Drug agencies in Canada and the United States, by contrast, keep this information confidential.² Clinical study reports contain considerably more data on harms than do journal articles and trial registry reports, and therefore should be used, when available, as a primary data

source for systematic reviews.³ Various North American organizations, including the Canadian Medical Association and the American Medical Association, support the AllTrials initiative, which calls for clinical study reports, with individual patient data redacted, to be made publicly available.⁴ However, access to individual patient data in the reports would allow for more complex analyses of harms, such as when a harm of interest comprises multiple symptoms.

Stress urinary incontinence is the involuntary leakage of urine on exertion, sneezing or coughing.⁵ It is estimated that 1 in every

10 adult women suffer from this condition,⁶ which causes substantial impairment of quality of life and a considerable economic burden.⁷ Historically, treatment has consisted of conservative measures such as training of pelvic floor muscles or surgery. Duloxetine is a drug that has been approved in Canada, the US and Europe for the treatment of major depressive disorder. In 2004, Eli Lilly submitted applications for duloxetine for the treatment of stress urinary incontinence in women to the Canadian, US and European drug agencies,^{8,9} but this indication was approved only in Europe.^{9,10} Currently, unlike the situation for the EMA,^{11,12} reasons why marketing authorization applications are withdrawn or denied is not published by either the Canadian¹³ or the US drug agency (US Food and Drug Administration [FDA], Center for Drug Evaluation and Research: personal communication, Nov. 26, 2014). However, the FDA has said that a higher-than-expected rate of suicide attempts was observed in the open-label extensions of controlled trials of duloxetine for stress urinary incontinence.¹⁴

Given the FDA's statement about the rate of suicide attempts,¹⁴ we wanted to determine whether duloxetine increased the risk of suicidality, violence or their possible precursors (drug-induced akathisia, an extreme type of restlessness; activation, which consists of stimulating effects such as insomnia, anxiety and agitation; emotional disturbance, such as depersonalization and derealization; or psychotic events, such as delusions and hallucinations) in the randomized phases of the trials.^{15,16} We therefore assessed the benefits and harms of duloxetine in stress urinary incontinence using clinical study reports, including individual patient data, of the 4 main trials submitted to the EMA.¹⁷

Methods

Data sources

In 2011, in response to a wider request for access to clinical study reports of antidepressants, we obtained from the EMA reports concerning duloxetine for various indications. In our first research project, which was about depression, we found that the listings of adverse events for individual patients and the narratives of adverse events allowed a more accurate estimate of harms.³

The 4 reports that we received of placebo-controlled trials in stress urinary incontinence each had a unique identifier (SAAW, SBAT, SBAV, SBAX). The reports dated from 2001 and 2002 and totalled 6870 pages, including the protocols. The documents were provided as nonsearchable pdfs, but we made them searchable using Adobe Acrobat Pro XI.

Outcomes

The a priori benefits specified as outcomes in our protocol were the primary outcomes of each trial and the Patient Global Impression of Improvement.¹⁸

We divided the harms data (which were specified a priori) into general harms and harms related to suicidality and violence.

The general harms were deaths, nonfatal serious adverse events (any adverse event that was life-threatening, required initial or prolonged inpatient hospitalization, caused severe or permanent disability, caused congenital anomaly or was important

for other reasons) and discontinuation because of adverse events. We also determined the number of patients experiencing at least one treatment-emergent adverse event.

The harms related to suicidality and violence 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 that the FDA asked pharmaceutical companies to use when searching their own databases (Table 1).¹⁹ For violence, the terms were those used in a study to determine the association of prescription drugs with violence using data from the FDA Adverse Event Reporting System (Table 1).²⁰

We focused on akathisia, emotional disturbance and psychotic events because these events, known as the “psychotropic suicidogenic triumvirate,” can predispose to suicidality and violence.^{15,16} We also recorded activation symptoms (including akathisia), which we obtained from the warnings in FDA product labelling for antidepressants.²¹ We obtained terms for other potential precursors to suicidality and violence from the literature.^{16,22,23} A systematic review has shown a lack of consensus about what the symptoms of activation are,²² and we were uncertain about whether some events (e.g., nightmares, which can be a prelude to a psychotic event) should be treated as psychotic events.²⁴ We therefore consulted a professor in psychiatry. Core activation events were those that the psychiatrist considered as activation, according to his knowledge and clinical experience. When we were uncertain, we preliminarily categorized the events as “potential.” Because of differences in symptoms, we kept psychiatrist-defined activation as a category separate from FDA-defined activation (Table 1).

Data extraction

Data on benefits were extracted from summary tables. For each study arm, one observer (L.S.G.) extracted the number of patients included in randomization and subsequent analyses, the means and standard deviations for benefits and the number of patients in each category of the Patient Global Impression of Improvement. Extracted data were checked by a second observer (E.M.).

Two observers (E.M., L.S.G.) independently searched all data formats of harms manually, using the terms listed in Table 1. For one observer (L.S.G.), the study materials were blinded for data extraction, as follows. The other observer (E.M.) used the white redaction tool in Adobe Acrobat Pro XI to redact all drug names from all data formats of harms, including pre-existing conditions in individual patients. In addition, narrative texts were placed in Word documents, and all drug names (including dosages) and mentions of placebo were replaced 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.

For the manual search of harms data, each observer recorded the patient identification number, date of random assignment, adverse event term and data format (e.g., listing of all adverse events), onset and stop date of the event, severity, whether the

event was serious or led to discontinuation, and whether the term was the original investigator-reported term (“verbatim term”) or the preferred term from the Medical Dictionary for Regulatory Activities (MedDRA). The MedDRA is a hierarchic medical terminology system used to standardize entry, retrieval, analysis and display of adverse events data.^{25,26} Within MedDRA, verbatim terms are coded to the closest matching lowest-level terms. 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.²⁷

The two observers (E.M., L.S.G.) independently recoded preferred terms (and, if available, the verbatim terms as well) using the most recent version of MedDRA (version 17.0). Interobserver agreement was calculated, and discrepancies were resolved by consensus. To ensure that we had identified all of the relevant terms, we carried out electronic searches on all of the blinded documents using all of the terms identified. The data were then unblinded.

Overall, this process took an average of 3 months per observer.

Post hoc decisions

We moved the activation event of tension from the potential subcategory to the core subcategory of activation, because 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 was not included in any of the original categories, to emotional disturbance, on the basis of the verbatim terms (e.g., fuzzy feeling). Finally, the events of dysthymic disorder and

depressed mood were added as potential depression events. Sensitivity analyses were performed to determine the effect of these decisions.

Statistical analysis

For each outcome, we combined the data in a meta-analysis. For binary outcomes, we calculated risk ratios (RRs) and risk differences, and for continuous outcomes we calculated mean differences with 95% confidence intervals (CIs) using a fixed-effect model, because this approach gives more weight to large trials. If the heterogeneity was substantial ($I^2 > 50\%$), we explored the reasons in sensitivity analyses. We performed the meta-analyses in RevMan, which adds 0.5 to cells with zero events.²⁸ This adjustment does not cause bias if the study arms are of equal size.²⁹

For harms, only treatment-emergent adverse events (i.e., those that began or worsened during the randomized phase) were of interest. Where categories of suicidality and violence-related harms consisted of both core and potential events, sensitivity analyses were performed using only the core events. For categories that included 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 whether the results were clinically relevant by comparing with the minimum clinically important difference for the primary outcome, as stated in the literature.

In post hoc analyses, we also calculated the number needed to treat and the number needed to harm for binary outcomes and

Table 1: Terms for adverse events, defined a priori, in the suicidality and violence-related adverse event categories

Adverse event category	Terms for core adverse events*	Terms for potential adverse events†
Suicidality	Accident-, attempt, burn, cut, drown, gas, gun, hang, hung, immolat-, injur-, jump, monoxide, mutilat-, overdos-, self damage-, 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, antisocial behaviour)	
Depression	Depression	
Emotional disturbance	Anhedonia, apathy, depersonalization, derealization, disinhibition, emotional detachment, emotional lability, flat affect, 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§), restlessness (hyperactivity), shakiness	Insomnia, panic, tension,¶ tremor
FDA-defined activation symptoms	Anxiety, agitation, panic attacks, insomnia, irritability, hostility, aggressiveness, impulsivity, akathisia (psychomotor restlessness), hypomania, mania	

Note: FDA = US Food and Drug Administration.

*Core adverse events were those that had been used as search terms in the published research or that were considered relevant by expert opinion.

†Potential adverse events were events for which there was a lack of consistency in the literature or uncertainty over whether they were relevant. The effect of including potential events was explored in sensitivity analyses.

‡Activation refers to stimulating effects, such as insomnia, anxiety and agitation.

§Mania was reported as both an activation event and a psychotic event, because patients can report being “manic” when they are describing being more active than usual (i.e., experiencing activation).

¶Tension was originally categorized as a potential activation event; however, tension codes to the higher-level term of “anxiety symptoms” in the Medical Dictionary for Regulatory Activities. Tension was therefore considered a core event in the main analyses. A sensitivity analysis was performed to evaluate the effect of this decision.

the standardized mean differences for continuous outcomes of benefits. We considered a standardized mean difference of 0.2 a small effect, 0.5 a moderate effect and 0.8 a large effect.³⁰

Results

Overall, 958 women with stress urinary incontinence were randomly assigned to receive duloxetine 80 mg, and 955 were assigned to receive placebo; in one trial, the starting dose of duloxetine was 40 mg and titrated upward. The weighted average age of women in the trials was 52 years. In all trials, use of antidepressants within 14 days before trial entry or during the trial was an exclusion criterion. Apart from substance abuse, there were no exclusion criteria pertaining to psychiatric disorders.

All arms in individual trials were comparable in terms of baseline characteristics of pre-existing or historical diagnoses of psychiatric symptoms and disorders, with the exception of one trial (SBAV), in which more women in the placebo arm had pre-existing depression, relative to the duloxetine arm (18 v. 6, $p = 0.01$). During 2 weeks without medication, the participants 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 2 weeks, followed by randomization and 12 weeks of treatment with either duloxetine or placebo.

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 the protocols and the clinical study reports specified that adverse event data would be collected at the time of randomization and at study visits every 4 weeks thereafter. However, none of the sources specified how these data would be ascertained. The published articles from these trials stated that adverse events were ascertained through nonprobing questions.³¹⁻³⁴

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

Benefits

The protocol-specified primary outcomes were percentage change from baseline in frequency of incontinence episodes and mean change from baseline in Incontinence Quality of Life total score (range of scores 0 [worst] to 100 [best]).³⁵ Data were shown only for patients with a baseline value and at least one post-baseline value; the method used for missing values was last observation carried forward.

The weighted mean baseline value for weekly frequency of incontinence episodes was 16.8. Duloxetine was significantly better than placebo in terms of percentage change from baseline in weekly incontinence episodes ($n = 1738$ patients, mean difference -13.56% , 95% CI -21.59% to -5.53% , $I^2 = 42\%$) and change in weekly number of incontinence episodes (mean difference -2.85 , 95% CI -3.91 to -1.78 , $I^2 = 27\%$) (Figure 1). We did not find any published minimum clinically important differences for this outcome. Effect sizes were small for both percentage change

from baseline in weekly incontinence episodes (standardized mean difference -0.13 , 95% CI -0.22 to -0.04 , $I^2 = 64\%$; sensitivity analysis: standardized mean difference -0.05 , 95% CI -0.16 to 0.07 , $I^2 = 19\%$) and change in the number of weekly incontinence episodes (standardized mean difference -0.26 , 95% CI -0.35 to -0.16 , $I^2 = 0\%$) (Appendix 1, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.151104/-/DC1).

The weighted mean baseline value for Incontinence Quality of Life total score was 64.0. Duloxetine was better than placebo in terms of mean change in this score (mean difference 3.24, 95% CI 2.00 to 4.48, $I^2 = 5\%$; Figure 2). However, given that the published minimum clinically important difference for the Incontinence Quality of Life total score is 2.5,³⁵ this difference may not be clinically important. Furthermore, the effect size for this outcome was small (standardized mean difference 0.24, 95% CI 0.15 to 0.33, $I^2 = 0\%$; Appendix 1).

The Patient Global Impression of Improvement rating is a globally validated measure of the improvement that a patient perceives since starting treatment. The trial protocols specified that responses would be grouped into 3 categories as a secondary analysis variable: “a little worse or very much worse,” “no change or a little better” and “much better or very much better”; we considered the last of these categories to represent a response to treatment. When all 4 trials were included in a meta-analysis, there was considerable heterogeneity ($I^2 = 68\%$; Appendix 1). Removal of 2 trials, one at a time, from the analysis reduced the heterogeneity to below 50% (Appendix 1). The greatest reduction in heterogeneity was seen with removal of trial SBAX (RR 1.51, 95% CI 1.29 to 1.77, $I^2 = 36\%$; Appendix 1), with a number needed to treat of 8 (95% CI 6 to 13; Appendix 1).

Harms

Summary tables and line listings provided only coded terms of adverse events. Narratives, which were only for patients who experienced a serious adverse event or discontinued participation because of an adverse event, contained the verbatim terms for the adverse events that the patients experienced and the corresponding coded terms.

Adverse events in general

On cross-referencing the summary tables, line listings and narratives, we found 1 serious adverse event in a patient receiving duloxetine and 2 in a patient receiving placebo, as well as adverse events that led to discontinuations in 6 patients receiving duloxetine and 4 receiving placebo, that began before randomization and did not worsen in severity. These events did not meet the criteria for treatment-emergent adverse events and were excluded from our analyses. There was one death (in a patient receiving duloxetine, who died after a cerebrovascular accident), whereas 727 patients receiving duloxetine and 548 receiving placebo had one or more adverse events (RR 1.32, 1.24 to 1.41, $I^2 = 51\%$ [Figure 3]; number needed to harm 6, 95% CI 5 to 7 [Appendix 1]; sensitivity analysis RR 1.25, 95% CI 1.16 to 1.35, $I^2 = 0\%$; number needed to harm 7, 95% CI 5 to 10 [Appendix 1]). The risk of discontinuing because of an adverse event was more than 5 times higher among patients receiving duloxetine (RR

5.73, 4.00 to 8.20, $I^2 = 26\%$ [Figure 3]; number needed to harm 7, 95% CI 6 to 8 [Appendix 1], and the risk of experiencing a nonfatal serious adverse event was also higher (RR 1.77, 95% CI 0.79 to 3.98, $I^2 = 0\%$; Appendix 1), although the difference was not statistically significant.

Suicidality and violence

There were no events of violence that met our prespecified criteria, but one patient receiving duloxetine experienced mild hostility that began 2 days after randomization. There were no reports of suicidality, but 8 patients (4 in the duloxetine group and 4 in the placebo group) experienced injuries or burns. There were no narratives for these patients, and therefore no information about the context or nature of these events.

Harms predisposing to suicidality or violence

According to our criteria, 2 patients receiving duloxetine experienced a total of 5 serious adverse events potentially predisposing to suicidality or violence; these were severe depression, panic attacks and severe anxiety. The many events that were not considered serious are described in the following section.

Activation events

Core or potential activation events were experienced by 187 patients in the duloxetine group and 42 patients in the placebo group. Among these patients, 46 (41 in the duloxetine group and 5 in the placebo group) experienced more than one event of interest (range 2 to 6 events). The risk of experiencing a core or potential activation event was more than 4 times higher in the

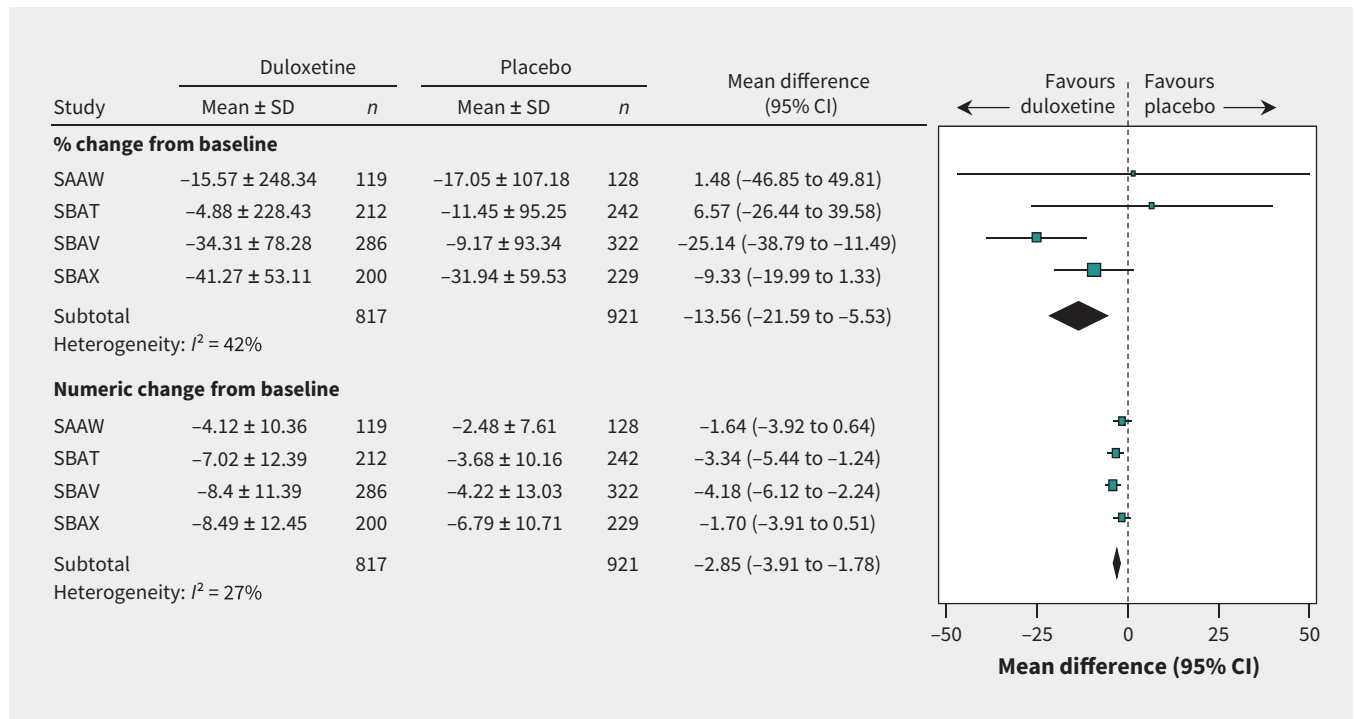


Figure 1: Benefits of duloxetine (80-mg dose), in terms of change in weekly incontinence episodes. CI = confidence interval; SAAW, SBAT, SBAV, SBAX = trial designations in clinical study reports; SD = standard deviation.

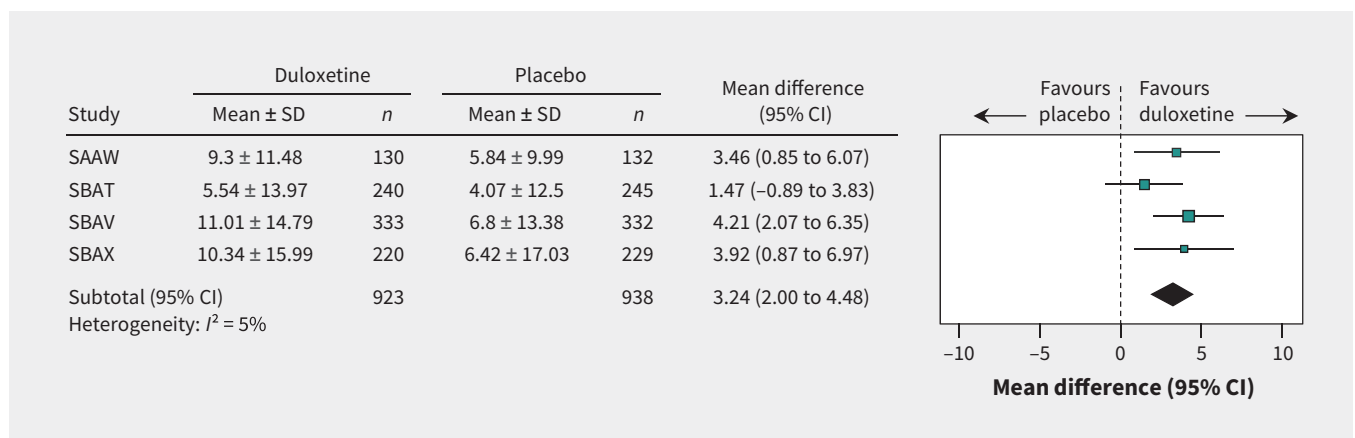


Figure 2: Benefits of duloxetine (80-mg dose), in terms of the Incontinence Quality of Life total score. CI = confidence interval; SAAW, SBAT, SBAV, SBAX = trial designations in clinical study reports; SD = standard deviation.

duloxetine group than the placebo group (RR 4.45, 3.22 to 6.14, $I^2 = 0\%$ [Figure 3]; number needed to harm 7, 95% CI 6 to 9 [Appendix 1]). The result was similar after exclusion of patients who experienced only sleep problems that were not definitively insomnia (RR 4.96, 3.47 to 7.09, $I^2 = 0\%$; Appendix 1). 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, $I^2 = 0\%$; number needed to harm 10, 95% CI 8 to 13; Appendix 1). The risk of experiencing a core event was more than 3 times greater with duloxetine (RR 3.59, 2.04 to 6.32, $I^2 = 0\%$;

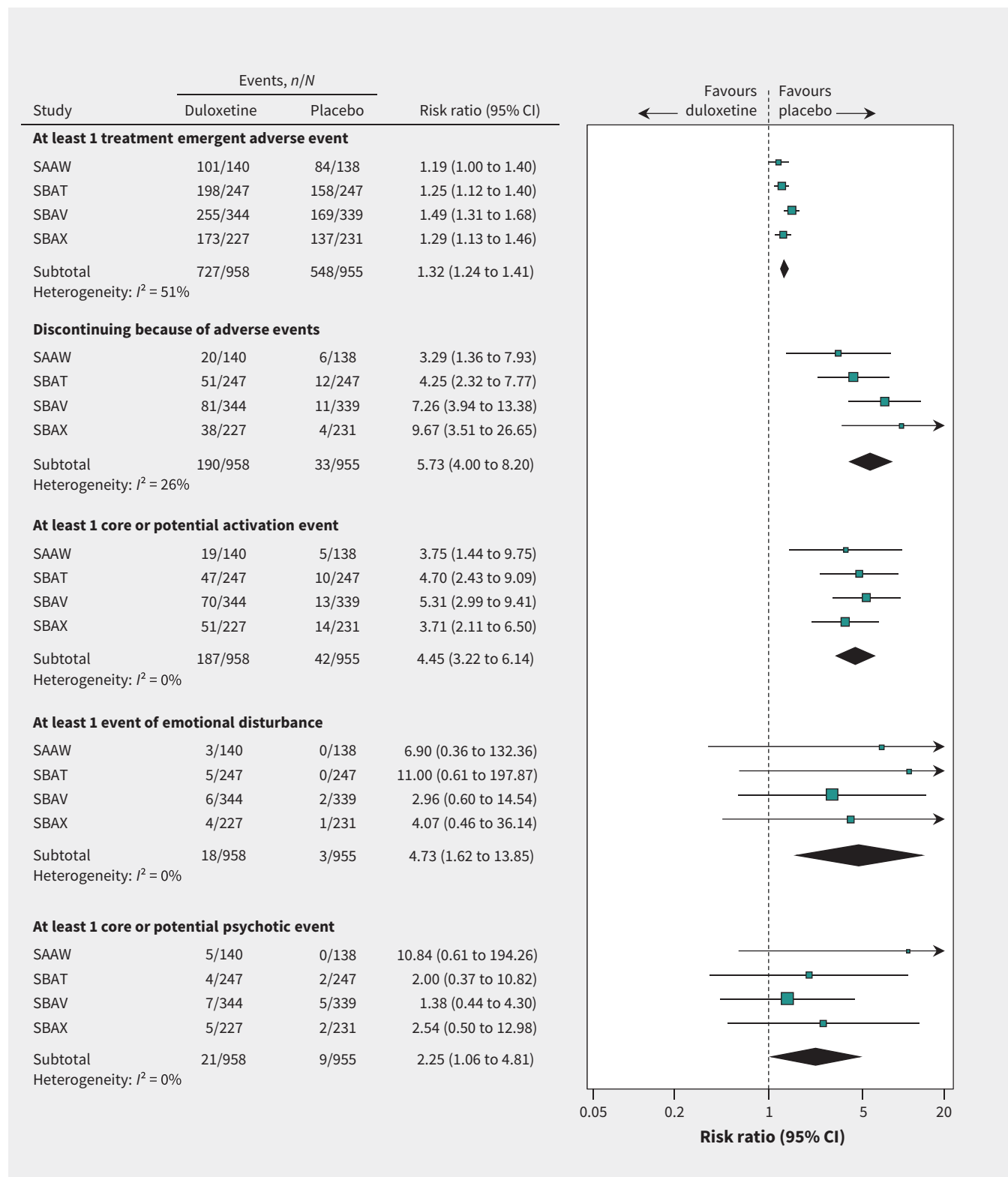


Figure 3: Harms of duloxetine (80-mg dose). CI = confidence interval; SAAW, SBAT, SBAV, SBAX = trial designations in clinical study reports.

number needed to harm 25, 95% CI 18 to 42; Appendix 1). Twenty-eight patients (27 in the duloxetine group and 1 in the placebo group) discontinued participation because of activation events. 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 and Appendix 1).

Akathisia, emotional disturbance, psychosis and depression

No akathisia events were reported, whereas 18 patients in the duloxetine group and 3 in the placebo group experienced emotional disturbance (Table 2) (RR 4.73, 1.62 to 13.85, $I^2 = 0\%$ [Figure 3]; number needed to harm 65, 95% CI 40 to 170 [Appendix 1]). In addition, 3 patients in the duloxetine group and 1 in the placebo group discontinued participation 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 (as defined in Table 2) (RR 2.25, 1.06 to 4.81, $I^2 = 0\%$ [Figure 3]; number needed to harm 80, 95% CI 40 to 834 [Appendix 1]). The risk of experiencing a core event was similar (RR 2.49, 0.78 to 7.89, $I^2 = 0\%$; Appendix 1), 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 participation because of a confusional state.

Depression-related events were similar for patients receiving duloxetine and placebo (RR 1.26, 0.58 to 2.71, $I^2 = 26\%$; Appendix 1).

Interobserver agreement

In the 4 trials, an average of 22% of patients receiving duloxetine and 5% of those receiving placebo experienced serious adverse events or discontinued participation because of adverse events (and therefore had narratives). A total of 96 patients with 139 adverse events of interest to the current analysis had a narrative. When 2 of us (E.M., L.S.G.) independently recoded the verbatim terms using MedDRA, there was excellent interobserver agreement (for lower-level terms, $\kappa = 0.92$; for preferred terms, $\kappa = 0.99$).

Events identifiable only from narratives

There were 4 events of interest, involving 4 patients (all receiving duloxetine), that were obscured by the coded term used but of which we became aware through the verbatim terms provided in the narratives. For example, one patient had a “nervous breakdown,” which was coded as mental disorder, and another patient reported “feeling drugged,” which was coded as somnolence. In addition, 5 patients, all receiving duloxetine, experienced a total of 8 events that were mentioned only in the narrative text.

Interpretation

We found a statistically significant difference between duloxetine and placebo in terms of percentage change in frequency of incontinence episodes and change in the Incontinence Quality of Life total score. However, the effect sizes for both of these benefits were small. Furthermore, given that the confidence interval for the effect size for Incontinence Quality of Life total score crossed the

Table 2: Adverse events in the suicidality and violence-related adverse event categories, as reported in 4 placebo-controlled trials of duloxetine for stress urinary incontinence

Adverse event category	Type of adverse event; specific events* reported in trial†	
	Core adverse events‡	Potential adverse events§
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,¶ panic attack, poor-quality sleep, sleep disorder, stress,¶ tension¶	
Emotional disturbance	Feeling abnormal (verbatim terms 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, mental disorder (verbatim term “nervous breakdown”)	Abnormal dreams and nightmares
Depression	Depression	Depressed mood, dysthymic disorder

Note: FDA = US Food and Drug Administration, MedDRA = Medical Dictionary for Regulatory Activities.

*These adverse events occurred in either the duloxetine arm or the placebo arm, or in both arms.

†The data presented in this table are the preferred terms from the MedDRA (version 17.0) that were used for recoding of the original preferred terms (and also the verbatim terms, if available) of adverse events provided in the clinical study reports of the 4 trials of duloxetine for stress urinary incontinence.

‡Core adverse events were those that had been used as search terms in the published research or that were considered relevant by expert opinion.

§Potential adverse events were events for which there was a lack of consistency in the literature or uncertainty over whether they were relevant. The effect of including potential events was explored in sensitivity analyses.

¶Nervousness, stress and tension are not explicitly mentioned in FDA-defined activation. Anxiety is categorized as an FDA-defined activation event, and nervousness, stress and tension all code to the higher-level term of “anxiety symptoms” in the MedDRA. These 3 types of events were therefore included in the analyses of FDA-defined activation.

published value for the minimum clinically important difference, the difference for this outcome may not be of clinical significance.

We also found a statistically significant difference for Patient Global Impression of Improvement, but the number needed to treat for this outcome was not less than the number needed to harm in terms of discontinuations due to adverse events or in terms of core or potential activation events. This finding suggests that the benefits of duloxetine for stress urinary incontinence do not outweigh its harms. We did not find any adverse events of suicidality or violence, but many of the patients experienced unpleasant events that might have predisposed them to suicidality and violence (e.g., the number needed to harm was 7 for a core or potential activation event). It was possible for us to analyze harms consisting of multiple symptoms, such as activation, only because we had access to individual patient data contained in the clinical study reports.

Systematic reviews assessing the benefits and harms of duloxetine for stress urinary incontinence have been performed by the Cochrane Collaboration³⁶ and by the US Agency for Healthcare Research and Quality (AHRQ).³⁷ Neither of these reviews included meta-analyses of percentage or numeric change in weekly frequency of incontinence episodes, and neither reported on suicidality. Furthermore, for the outcomes below, none of the analyses in either review included data from all 4 trials that were included in our review.

The Cochrane review³⁶ found a slightly larger change in Incontinence Quality of Life total score (RR 4.50, 95% CI 2.83 to 6.18). However, its analysis included data from 3 trials, only 2 of which were among our 4 trials. Furthermore, the dose of duloxetine used in the third trial was higher (120 mg daily) than that used in our trials. The AHRQ review³⁷ noted only that Incontinence Quality of Life total score was inconsistent among trials.

Both of the reviews assessed the Patient Global Impression of Improvement. The Cochrane review³⁶ included patients who felt “very much better, much better, or a little better,” and found a relative treatment effect slightly lower than ours (RR 1.25, 95% CI 1.14 to 1.36). The AHRQ review³⁷ used the same Patient Global Impression of Improvement variable as we did, but its analysis consisted of 4 trials, only one of which was included in our analysis. It found a larger number needed to treat (13, 95% CI 7 to 143) than we did.

Our results for any treatment-emergent adverse events were similar to those of both the Cochrane review³⁶ (RR 1.31, 95% CI 1.24 to 1.39) and the AHRQ review³⁷ (RR 1.36, 95% CI 1.28 to 1.44). Both the AHRQ review³⁷ and our own review found that the relative risk of individual harms was much larger than the relative risk of the outcome of any treatment-emergent adverse event. Our results for discontinuation due to adverse events were slightly higher than those reported in both the Cochrane review³⁶ (RR 4.50, 95% CI 3.44 to 5.89) and the AHRQ review³⁷ (RR 4.4, 95% CI 3.24 to 5.86). Our results for discontinuation due to adverse events were similar to those of a published pooled analysis (i.e., not a meta-analysis) performed by Eli Lilly using data from the 4 trials that we included.³⁸ That analysis found adverse event discontinuation rates of 20.5% for duloxetine and 3.9% for placebo ($p < 0.001$). That study also found no reported cases of suicide or attempted suicide among patients taking duloxetine.³⁸

Our findings are also in line with evidence from a cohort study, in which 228 women with stress urinary incontinence were given duloxetine as an alternative to surgery.³⁹ Two-thirds of the patients (68%) discontinued the drug within the first 4 weeks of treatment, mainly because of adverse events. By the end of 1 year, only 9% of the cohort were still taking duloxetine, with 82% having decided to undergo surgery.

Notably, the UK National Institute for Health and Care Excellence guideline states that duloxetine should not be used as a first-line treatment or routinely offered as a second-line treatment for stress urinary incontinence, given that pelvic floor muscle training is more effective and less costly than duloxetine and that surgery is more cost-effective than duloxetine.⁴⁰

Some of our harms of interest (e.g., activation) consisted of multiple possible symptoms. It would have been impossible to perform meta-analyses of these harms if we had had access only to journal articles or to the summary data provided in clinical study reports, and not to individual patient data. According to a blog post, the authors of the Cochrane review of oseltamivir for influenza also found that narratives and line listings contained in clinical study reports were essential for their review.⁴¹ Despite earlier promises, the EMA recently announced that it will not publish individual anonymized patient data contained in the appendices of clinical study reports in the first round of implementation of its new policy,⁴² which came into effect on Jan. 1, 2015. The EMA’s reason is that it needs to find a reliable way to anonymize the data. However, in accordance with current legislation,⁴³ the data are already anonymized, and the EMA’s approach is inconsistent, given that researchers can get access to the harms in the old trials that are in the EMA’s possession. As we have shown here and previously,³ individual patient data contained in appendices of clinical study reports are essential for reliable assessment of drug harms. Furthermore, we did not find any deficiencies in the anonymization of the individual patient data that we received from the EMA for the current study.

Limitations

One limitation of this study was that the data for the beneficial effects of duloxetine, especially for frequency of incontinence episodes, were considerably skewed.⁴⁴ Another reason why these results should be interpreted with caution is the unblinding due to adverse effects.^{45,46} The effects on incontinence that we found, in terms of changes in frequency of incontinence episodes and the Incontinence Quality of Life total score, were small. Anti-depressant trials have been shown to be inadequately blinded because of the adverse effects that these drugs have.^{45,46} Therefore, the small effects we found could be fully explained by unblinding bias.

Only 958 patients were receiving 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 nonprobing questions, an approach that leads to underreporting of adverse events,⁴⁷ especially for events of a sensitive nature⁴⁸ (such as suicidal ideation, suicidal behaviour and violence). Suicidality events have been much underreported in clinical trials and observational studies, including those conducted by Eli Lilly.⁴⁹

We may also have underestimated harms because we did not have access to the completed case report forms (data collection forms) of the trials. Le Noury and associates⁵⁰ recently reanalyzed a trial of the antidepressants paroxetine and imipramine versus placebo, using data from the clinical study report (including individual patient data) and a sample of the trial's case report forms. They found adverse events recorded on case report forms that were not transcribed into patient-level listings of adverse events. The most common adverse events not transcribed were psychiatric in nature, occurring among patients who received paroxetine. Furthermore, relying on adverse event listings in individual patient data, instead of case report forms, caused underestimation of adverse events by between 7% and 14%.

A further limitation of our study was that data extraction of benefits was performed by one person and checked by a second, rather than being performed in duplicate.

Conclusion

Given the uncertainty as to whether duloxetine leads to clinically significant improvement in quality of life, and given that improvements measured as Patient Global Impression of Improvement did not outweigh discontinuations due to adverse effects, or core or potential activation events, we question the rationale for using duloxetine for stress urinary incontinence. Evaluation of harms that were possible precursors to suicidality would have been impossible using only summary data in the clinical study reports. Individual patient data of adverse events contained in clinical study reports are therefore essential for a reliable assessment of drug harms.

References

1. ICH harmonised tripartite guideline: the common technical document for the registration of pharmaceuticals for human use. Efficacy — M4E(R1). Clinical overview and clinical summary of module 2. Module 5: Clinical study reports. Geneva (Switzerland): International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2002. Available: www.ich.org/fileadmin/Public_Web_Site/ICH_Products/CTD/M4_R1_Efficacy/M4E_R1.pdf (accessed 2012 Mar. 6).
2. Persaud N, Doshi P. North American regulatory agencies can and should make clinical trial data publicly available. *CMAJ* 2016;188:96-7.
3. Maund E, Tendal B, Hróbjartsson A, 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.
4. AllTrials: Supporters — organisation list. London (UK): AllTrials; 2016. Available: www.alltrials.net/supporters/supporters-organisation-list/ (accessed 2016 May 2).
5. Bettez M, Tu le M, Carlson K, et al. 2012 update: guidelines for adult urinary incontinence collaborative consensus document for the Canadian Urological Association. *Can Urol Assoc J* 2012;6:354-63.
6. Ebbesen MH, Hunskaar S, Rortveit G, et al. Prevalence, incidence and remission of urinary incontinence in women: longitudinal data from the Norwegian HUNT study (EPINCONT). *BMC Urol* 2013;13:27.
7. *Incontinence: the Canadian perspective*. Peterborough (ON): Canadian Continence Foundation; 2014. Available: www.canadiancontinence.ca/pdfs/en-incontinence-a-canadian-perspective-2014.pdf (accessed 2016 Mar. 8).
8. Emerging drug list: duloxetine for major depressive disorder and stress urinary incontinence. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 2004. Available: www.cadth.ca/media/pdf/108_No61_duloxetine_edrug_e.pdf (accessed 2016 Jan. 27).
9. Yentreve (duloxetine): authorisation details. London (UK): European Medicines Agency; 2004. Available: www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000545/human_med_001164.jsp&mid=WC0b01ac058001d124 (accessed 2014 Nov. 18).
10. Lilly and Boehringer Ingelheim jointly announce the decision of U.S. FDA application for duloxetine for treatment of stress urinary incontinence [media release]. Indianapolis (IN) and Ridgefield (CT): Eli Lilly and Company and Boehringer Ingelheim Pharmaceuticals Inc.; 2005. Available: <https://investor.lilly.com/releasedetail.cfm?releaseid=154081> (accessed 2014 Oct. 30).
11. Procedural advice on publication of information on negative opinions and refusals of marketing authorisation applications for human medicinal products. London (UK): European Medicines Agency; 2013. Available: www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004188.pdf (accessed 2015 Jan. 27).
12. Procedural advice on publication of information on withdrawals of applications related to the marketing authorisation of human medicinal products. London (UK): European Medicines Agency; 2013. Available: www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500005061.pdf (accessed 2015 Jan. 27).
13. Regulatory transparency and openness — 2015–2016 activities. Ottawa: Health Canada; 2016. Available: www.hc-sc.gc.ca/home-accueil/rto-tor/aap-pa/2015-16-eng.php (accessed 2016 Jan. 27).
14. Historical information on duloxetine hydrochloride (marketed as Cymbalta). Silver Spring (MD): Food and Drug Administration (US); 2013. Available: www.fda.gov/drugs/drugsafety/postmarketdrugsafetyinformationforpatientsandproviders/ucm114970.htm (accessed 2016 May 30).
15. Healy D, Herxheimer A, Menkes DB. Antidepressants and violence: problems at the interface of medicine and law. *PLoS Med* 2006;3:e372.
16. Breggin PR. Suicidality, violence and mania caused by selective serotonin reuptake inhibitors (SSRIs): a review and analysis. *Int J Risk Saf Med* 2003-2004;16:31-49.
17. *Scientific discussion* [Yentreve]. London (UK): European Medicines Agency; 2005. Available: www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Scientific_Discussion/human/000545/WC500045641.pdf (accessed 2013 July 25).
18. Yalcin I, Bump RC. Validation of two global impression questionnaires for incontinence. *Am J Obstet Gynecol* 2003;189:98-101.
19. Laughren TP. Overview for December 13 meeting of Psychopharmacologic Drugs Advisory Committee (PDAC) [memorandum]. Silver Spring (MD): Food and Drug Administration (US); 2006 Nov. 16. Available: www.fda.gov/ohrms/dockets/ac/06/briefing/2006-4272b1-01-FDA.pdf (accessed 2012 June 28).
20. Moore TJ, Glenmullen J, Furberg CD. Prescription drugs associated with reports of violence towards others. *PLoS One* 2010;5:e15337.
21. Revisions to product labeling. Silver Spring (MD): Food and Drug Administration (US); 2007. Available: www.fda.gov/downloads/Drugs/DrugSafety/InformationbyDrugClass/UCM173233.pdf (accessed 2014 Oct. 2).
22. Sinclair LI, Christmas DM, Hood SD, et al. Antidepressant-induced jitteriness/anxiety syndrome: systematic review. *Br J Psychiatry* 2009;194:483-90.
23. Price J, Cole V, Goodwin GM. Emotional side-effects of selective serotonin reuptake inhibitors: qualitative study. *Br J Psychiatry* 2009;195:211-7.
24. Levin R, Daly RS. Nightmares and psychotic decompensation: a case study. *Psychiatry* 1998;61:217-22.
25. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of clinical trials*. 4th ed. New York: Springer-Verlag New York; 2010.
26. *Understanding MedDRA: the medical dictionary for regulatory activities*. Geneva (Switzerland): International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2013. Available: www.meddra.org/sites/default/files/page/documents/meddra2013.pdf (accessed 2014 June 8).
27. White CA. A preliminary assessment of the impact of MedDRA on adverse event reports and product labelling. *Drug Inf J* 1998;32:347-62.
28. Higgins JPT, Deeks JJ, Altman DG. Chapter 16: Special topics in statistics. In: Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0. Oxford (UK): The Cochrane Collaboration; 2011. Available: handbook.cochrane.org (accessed 2011 July 10).
29. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004;23:1351-75.
30. Cohen J. *Statistical power analysis in the behavioral sciences*. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.
31. Norton PA, Zinner NR, Yalcin I, et al. Duloxetine versus placebo in the treatment of stress urinary incontinence. *Am J Obstet Gynecol* 2002;187:40-8.
32. Dmochowski RR, Miklos JR, Norton PA, et al. Duloxetine versus placebo for the treatment of North American women with stress urinary incontinence. *J Urol* 2003;170:1259-63.
33. van Kerrebroeck P, Abrams P, Lange R, et al. Duloxetine versus placebo in the treatment of European and Canadian women with stress urinary incontinence. *Br J Obstet Gynaecol* 2004;111:249-57.
34. Millard RJ, Moore K, Rencken R, et al. Duloxetine vs. placebo in the treatment of stress urinary incontinence: a four-continent randomized clinical trial. *BJU Int* 2004;93:311-8.

35. Yalcin I, Patrick DL, Summers K, et al. Minimal clinically important differences in Incontinence Quality-of-Life scores in stress urinary incontinence. *Urology* 2006;67:1304-8.
36. Mariappan P, Alhasso AA, Grant A, et al. Serotonin and noradrenaline reuptake inhibitors (SNRI) for stress urinary incontinence in adults. *Cochrane Database Syst Rev* 2005; (3):CD004742.
37. Shamliyan T, Wyman J, Kane RL. *Comparative effectiveness review no. 36: Nonsurgical treatments for urinary incontinence in adult women: diagnosis and comparative effectiveness*. Rockville (MD): Agency for Healthcare Research and Quality; 2012. Available: <https://www.effectivehealthcare.ahrq.gov/ehc/products/169/834/urinary-incontinence-treatment-report-130909.pdf> (accessed 2014 Mar. 26).
38. Hurley DJ, Turner CL, Yalcin I, et al. 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.
39. Vella M, Duckett J, Basu M. Duloxetine 1 year on: the long term outcome of a cohort of women prescribed duloxetine. *Int Urogynecol J Pelvic Floor Dysfunct* 2008;19:961-4.
40. *Urinary incontinence in women: the management of urinary incontinence in women: NICE clinical guideline CG171*. London (UK): National Institute for Health and Care Excellence; 2013. Available: <http://guidance.nice.org.uk/CG171> (accessed 2014 Mar. 14).
41. El Emam K, Jefferson T, Doshi P. The release of regulatory documents under EMA policy 0070: now you see them, now you don't [blog]. London (UK): BMJ Publishing Group Limited; 2015 Sept. 15. Available: <http://blogs.bmj.com/bmj/2015/09/15/the-release-of-regulatory-documents-under-ema-policy-0070-now-you-see-them-now-you-dont/> (accessed 2016 Jan. 27).
42. Questions and answers on the European Medicines Agency policy on publication of clinical data for medicinal products for human use. London (UK): European Medicines Agency; 2015. Available: www.ema.europa.eu/docs/en_GB/document_library/Report/2014/10/WC500174378.pdf (accessed 2015 June 10).
43. *HMA/EMA guidance document on the identification of commercially confidential information and personal data within the structure of the marketing authorisation (MA) application — release of information after the granting of a marketing authorisation*. London (UK): HMA/EMA Working Group on Transparency; 2012. Available: www.ema.europa.eu/docs/en_GB/document_library/Other/2012/03/WC500124536.pdf (accessed 2012 May 3).
44. Deeks JJ, Higgins JPT, Altman DG. 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. Oxford (UK): The Cochrane Collaboration; 2011. Available: handbook.cochrane.org (accessed 2011 July 10).
45. Fisher S, Greenberg RP. How sound is the double-blind design for evaluating psychotropic drugs? *J Nerv Ment Dis* 1993;181:345-50.
46. Moncrieff J, Wessely S, Hardy R. Active placebos versus antidepressants for depression. *Cochrane Database Syst Rev* 2004;(1):CD003012.
47. Rief W, Nestoriuc Y, von Lilienfeld-Toal A, et al. Differences in adverse effect reporting in placebo groups in SSRI and tricyclic antidepressant trials: a systematic review and meta-analysis. *Drug Saf* 2009;32:1041-56.
48. Montejo-González AL, Llorca G, Izquierdo JA, 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.
49. Gøtzsche PC. *Deadly psychiatry and organised denial*. Copenhagen (Denmark): People's Press; 2015.
50. Le Noury J, Nardo JM, Healy D, et al. Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. *BMJ* 2015;351:h4320.

Competing interests: None declared.

This article has been peer reviewed.

Affiliation: Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark

Contributors: Emma Maund and Peter Gøtzsche contributed to the study concept and design. Emma Maund and Louise Schow Guski contributed to the acquisition of data. All of the authors contributed to the analysis and interpretation of data and to the drafting and revising of the manuscript. All of the authors approved the final version for publication and agreed to act as guarantors of the work.

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

Accepted: July 19, 2016

Early release: Nov. 14, 2016

Correspondence to: Emma Maund, em@cochrane.dk